Lecture notes for Calculus and Analysis

Dr A. Tonks, a.tonks@gmail.com

Original authors: Dr K. Leschke, Dr A. Cangiani

last updated: March 7, 2022

Contents

Chapter 1

Precalculus

Let us start by reviewing some terminology, notation and formulas of elementary mathematics.

1.1 Sets

A *set* S is a collection of distinct objects, called *elements* of S.

Notations:

$x \in S$	:	x is element of the set S
$x \notin S$	:	x is not element of the set S
$\{x : P\}$	:	the set of all objects x satisfying the condition P.

Relations and operations:

$A \subset B$	:	A is a *subset* of B, that is, for all $x \in A$ we have also $x \in B$. We write for short: $\forall x \in A : x \in B$
$A \cup B$	$=$	$\{x : x \in A \text{ or } x \in B\}$ is the *union* of the sets A and B.
$A \cap B$	$=$	$\{x : x \in A \text{ and } x \in B\}$ is the *intersection* of the sets A and B.
$A \setminus B$	$=$	$\{x : x \in A \text{ but } x \notin B\}$ is the *difference* of the sets A and B.

For the phrase "there exists" we also write for short $\exists$.

1.2 Number systems

Natural numbers The set of all positive integers

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

Some mathematicians and textbook writers denote this by $\mathbb{N}_0$, as by default they *do not include* 0 in the set of natural numbers $\mathbb{N}$.

Integers

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

Rational numbers

$$\mathbb{Q} = \{x : x = \frac{p}{q}, p \in \mathbb{Z}, q \in \mathbb{Z}, q \neq 0\}$$

Irrational numbers Numbers which cannot represented as a quotient of integers, e.g. $\sqrt{2}, \pi, \ldots$.

Real numbers

$$\mathbb{R} = \{x : x \text{ is rational or irrational}\}$$

Real numbers can be represented as decimals. e.g. $x = \frac{3}{5}$ is represented $x = 0.6$, and $x = \frac{1}{3}$ as $x = 0.333\ldots$. For irrational numbers this process never terminates or repeats,

$$\text{e.g. } \sqrt{2} = 1.41421356237\ldots.$$

1.3 Ordered fields

The real numbers, and the rational numbers, are example of special sets which are called *fields*. This just means that the real numbers, and also the rational numbers, have operations (addition $x + y$ and multiplication $x \cdot y$) satisfying the field axioms.

Axiom 1.1 (Field axioms)**.** *A field is a set F together with addition and multiplication operations, and elements $0 \in F$ and $1 \in F - \{0\}$, satisfying the field axioms:*

Name:	**additive**	**multiplicative**
Identity	$\forall x \in F,\ x + 0 = x$	$\forall x \in F,\ x \cdot 1 = x$
Inverse	$\forall x \in F,\ \exists(-x),\ x + (-x) = 0$	$\forall x \in F - \{0\},\ \exists x^{-1},\ x \cdot x^{-1} = 1$
Associativity	$\forall x, y, z \in F,\ x + (y + z) = (x + y) + z$	$\forall x, y, z \in F,\ x(yz) = (xy)z$
Commutativity	$\forall x, y \in F,\ x + y = y + x$	$\forall x, y \in F,\ xy = yx$
Distributivity	$\forall x, y, z \in F,\ x(y + z) = xy + xz$	

Associativity means we can sometimes leave out some brackets. Also, we usually abbreviate $x + (-y)$ to $x - y$ and abbreviate $(-x) + y$ to $-x + y$, but the latter is not the same as $-(x + y)$.

From the field axioms we can derive (that is, prove) a lot of basic properties, such as $0x = 0$, and $(-1)x = -x$, and also $xy = 0$ implies $x = 0$ or $y = 0$.

Geometrically, we represent real numbers as points on the real line. In particular, we can give an *order* on the real numbers.

Axiom 1.2 (Order axioms)**.** *The real numbers $\mathbb{R}$ are an ordered field, that is, the field $\mathbb{R}$ has a relation $<$ so that for all $a, b, c \in \mathbb{R}$:*

Total order *Either $a < b$, $b < a$ or $a = b$.*

Transitivity *If $a < b$ and $b < c$ then $a < c$.*

Compatibility *If $a < b$ then $a + c < b + c$. If $a < b$ and $c > 0$ then $ac < bc$.*

Obviously, we define "$a > b$" by saying that this means that "$b < a$", and "$a \leq b$" by saying that this means that either "$a < b$" or "$a = b$".

Remark 1.3. *The real numbers and the rational numbers are an ordered field, but the complex numbers are not.*

Let us formulate some conclusions from these axioms.

Corollary 1.4. *(i)* $x, y > 0 \implies xy > 0$

(ii) $-x < 0 \iff x > 0$

(here "$\iff$" means "if and only if", that is, $-x < 0 \implies x > 0$ and $x > 0 \implies -x < 0$).

(iii) $x \neq 0 \implies x^2 > 0$.

(iv) $1 > 0$.

(v) $0 < x < y \implies 0 < y^{-1} < x^{-1}$.

1.3.1 Intervals and absolute value

Using the order relation on the real numbers we can define *intervals* and the *absolute value*.

We denote by

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

the *open interval* from a to b, by

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

the *closed interval* from a to b, and by

$$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}, \qquad [a, b) = \{x \in \mathbb{R} : a \leq x < b\}.$$

the half–open intervals. Moreover, we write

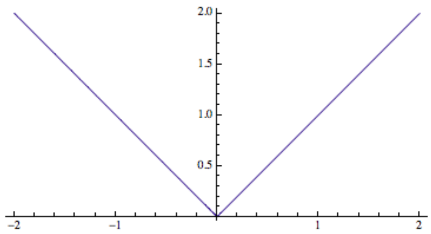
$$(-\infty, a) = \{x \in \mathbb{R} : x < a\}, \qquad (a, \infty) = \{x \in \mathbb{R} : x > a\}.$$

An open interval does not contain the *endpoints* whereas the endpoints a, b are elements of the closed interval $[a, b]$. We call a point c an *interior point* of an interval, if it is not an endpoint.

The *absolute value* is an important function related to the order on $\mathbb{R}$,

Definition 1.5. The *absolute value* of $a \in \mathbb{R}$ is defined by

$$|a| = \max\{a, -a\} = \begin{cases} a, & a \geq 0 \\ -a, & a < 0 \end{cases}$$



The absolute value measures the distance of a to 0, therefore $|a-b|$ is the distance between a and b.

Properties: For all $a, b \in \mathbb{R}$:

(i) $|a| = 0 \iff a = 0$

(ii) $|-a| = |a|$

(iii) $|ab| = |a| \cdot |b|$, $|a^2| = |a|^2$

Also we can prove that

$$|a+b| \leq |a| + |b| \qquad (\textit{triangle inequality})$$

and

$$|a-b| \geq ||a| - |b||.$$

The following, which we will meet a lot later, illustrates two definitions of two sets of real numbers, either with the use of interval notation or with the absolute value function:

$$\{x \in \mathbb{R} \, : \, |x-c| < \delta\} \;=\; (c-\delta, c+\delta)$$

$$\{x \in \mathbb{R} \, : \, 0 < |x-c| < \delta\} \;=\; (c-\delta, c+\delta) - \{c\} \;=\; (c-\delta, c) \cup (c, c+\delta)$$

1.4 Least Upper Bound Axiom

Intuitively, the *completeness axiom*, often called the *LUB axiom*, implies that there are no “gaps” in the real numbers. This contrasts with the rational numbers, whose corresponding number line has a “gap” at each irrational value.

Definition 1.6. A non–empty subset $S \subset \mathbb{R}$ of the real numbers is said to be

(i) *bounded above* if there exists $M \in \mathbb{R}$ so that $x \leq M$ for all $x \in S$; in this case M is called an *upper bound* of S;

(ii) *bounded below* if there exists $m \in \mathbb{R}$ so that $m \leq x$ for all $x \in S$; in this case m is called a *lower bound* of S;

(iii) *bounded* if S is both bounded above and below.

Note that if M is an upper bound for S then every $\tilde{M} \in [M, \infty]$ is also an upper bound for S.

So is there a best possible upper bound and how do we find it? This is easy, if we know that the upper bound M is an element of the set S; then we call it a *maximum*, denoted by $M = \max S$. Similarly, a lower bound m of S with $m \in S$ is called the *minimum* of S, denoted by $m = \min S$.

Consider the set $S = \{x \mid x = \frac{1}{n}, n \in \mathbb{N}\}$. We know that S is bounded below by 0 since $\frac{1}{n} > 0$ for all n. But is there a bigger lower bound? Assume that $m > 0$ is a lower bound of S then $\frac{1}{n} \geq m$ for all n. However, consider the number $\frac{1}{m} < \infty$. Then there exists $K \in \mathbb{N}$ which is bigger than $\frac{1}{m}$. But then $\frac{1}{K} < m$. Since $\frac{1}{K} \in S$ this gives a contradiction to m being a lower bound. Thus, 0 is the best lower bound, we call it the greatest lower bound.

Definition 1.7. Let $S \subset \mathbb{R}$ be non–empty. Then

(i) $M \in \mathbb{R}$ is called the *least upper bound* (or, also frequently used, *supremum*) of S, written as $\mathrm{lub}(S)$ (or $\sup(S)$), if

  (a) M is an upper bound of S, and

  (b) if $\tilde{M}$ is an upper bound of S then $M \leq \tilde{M}$.

(ii) $m \in \mathbb{R}$ is called the *greatest lower bound* (or, also frequently used, *infimum*) of S, written as $\mathrm{glb}(S)$ (or $\inf(S)$), if

  (a) m is a lower bound of S, and

  (b) if $\tilde{m}$ is a lower bound of S then $\tilde{m} \leq m$.

Example 1.8. The set $S = (0, \sqrt{2})$ does not have a least upper bound in $\mathbb{Q}$. However, it has a least upper bound in $\mathbb{R}$: $\sqrt{2} \in \mathbb{R}$.

And here is the property of the real numbers that guarantees that there are no gaps in the real line:

Axiom 1.9 (LUB–axiom)**.** *Every non–empty subset of the real numbers which has an upper bound has a least upper bound.*

Remark 1.10. *The least upper bound axioms together with the field axioms and the order axioms make the set of real numbers essentially unique; every other set satisfying these axioms is isomorphic to* $\mathbb{R}$*.*

Here is a property of the least upper bound we will use frequently:

Theorem 1.11. *Let* $S \subset \mathbb{R}$*. If* $M = \operatorname{lub} S$ *and* $\varepsilon > 0$ *then there exists* $s \in S$ *with*

$$M - \varepsilon < s \leq M\,.$$

Similar statements as for the least upper bound also hold for the greatest lower bound.

1.5 Elementary functions

Here is a quick introduction to functions, just a reminder to help set out the notation you should use in Calculus and Analysis.

Let $D \subset R$. Then f is a function on D if f assigns to each $x \in D$ a value $f(x) \in \mathbb{R}$. We often write

$$f : D \to \mathbb{R}, \quad x \mapsto f(x)\,.$$

We call $D = \operatorname{dom}(f)$ the *domain* of f and

$$\operatorname{range}(f) = \{f(x) : x \in D\}$$

the *range* of f, that is, all values f attains on D. We also write for a set $S \subset D$:

$$f(S) = \{y : y = f(x) \text{ for some } x \in S\}\,.$$

Thus, $\operatorname{range}(f) = f(\operatorname{dom}(f))$.

The most common domains we will look at are intervals: either a closed interval like $[a, b]$ or an open interval like (a, b) or (a, ∞) or $\mathbb{R} = (-\infty, \infty)$.

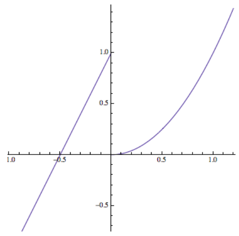
If not noted otherwise, we will usually use the biggest possible domain for the function, e.g, for

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

we use $\operatorname{dom}(f) = \mathbb{R} \setminus \{1\}$. For a lot of applications it is important to keep track of the domain; the behaviour of a function might change on a bigger or smaller domain.

A function can also be defined piecewise:

$$f(x) = \begin{cases} 2x+1, & x < 0 \\ x^2, & x \geq 0 \end{cases}\,.$$



A function is called *even* if $f(-x) = f(x)$ for all $x \in \mathrm{dom}(f)$. (Symmetric about the y–axis). A function is called *odd* if $f(-x) = -f(x)$ for all $x \in \mathrm{dom}(f)$. (Symmetric about origin).

Here are some functions you probably are already familiar with:

Polynomials: A function

$$P(x) = a_n x^n + \ldots + a_1 x + a_0, \quad x \in \mathbb{R},$$

is called a *polynomial of degree* n with coefficients $a_0, \ldots, a_n \in \mathbb{R}$, $a_n \neq 0$. If $P(x_0) = 0$ for some $x_0 \in \mathbb{R}$ then $x = x_0$ is called a *zero* or *root* of $P(x)$, or a solution of $P(x) = 0$, and we can factorise

$$P(x) = (x - x_0)Q(x)\,,$$

where Q is a polynomial of degree $n-1$. A polynomial of degree n has at most n zeros. A polynomial of odd degree has at least one zero.

Rational Functions: A *rational function* is given as a quotient of two polynomials P, Q,

$$R(x) = \frac{P(x)}{Q(x)}\,.$$

The domain of R is given by $\mathrm{dom}(R) = \{x : Q(x) \neq 0\}\,.$

Trigonometric Functions: The sine function is an odd function, the cosine is even, and the tangent $\tan(x) = \frac{\sin x}{\cos x}$ is odd. The domain of tan is $\mathrm{dom}(\tan) = \{x : \cos x \neq 0\} = \mathbb{R} \setminus \{\pm\frac{\pi}{2} + k\pi : k \in \mathbb{Z}\}$. The range of tan is $\mathrm{range}(\tan) = \mathbb{R}$.

From the Pythagorean theorem we have

$$\sin^2(x) + \cos^2(x) = 1 \quad \forall x \in \mathbb{R}\,.$$

We also have

$$\sin(x+y) = \sin x \cos y + \sin y \cos x, \quad \cos(x+y) = \cos x \cos y - \sin x \sin y\,.$$

To construct new functions out of given functions f, g on a common domain D we have the algebraic operations:

$$(f \pm g)(x) = f(x) \pm g(x), \quad (fg)(x) = f(x)g(x), \quad \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

where in the last case $\mathrm{dom}\left(\frac{f}{g}\right) = \{x \in D : g(x) \neq 0\}$.

Composition: If $f : D_f \to \mathbb{R}, g : D_g \to \mathbb{R}$ such that $\operatorname{range}(g) \subset D_f = \operatorname{dom}(f)$ then we can define the composition $f \circ g$ on $D_g = \operatorname{dom}(g)$ by setting

$$(f \circ g)(x) = f(g(x)), \quad x \in D_g\,.$$

Note: the composition is not commutative, that is, in general $f \circ g$ is not the same as $g \circ f$.

Example 1.12. $f(x) = x + 3, g(x) = x^2$, $x \in \mathbb{R}$. The

$$(f \circ g)(x) = f(x^2) = x^2 + 3$$

whereas

$$(g \circ f)(x) = g(x + 3) = (x + 3)^2\,.$$

Since $(f \circ g)(1) = 4 \neq 16 = (g \circ f)(1)$ we see that $f \circ g \neq g \circ f$.

Remark 1.13. *To prove that two functions f, g on a common domain are equal, you have to verify $f(x) = g(x)$ for* every *$x \in \operatorname{dom}(f) = \operatorname{dom}(g)$. However, to show that two functions are not the same, it is enough to show that $f(x_0) \neq g(x_0)$ for one $x_0 \in \operatorname{dom}(f)$.*

Since this is a common beginner error, let us emphasise: to prove a statement it is not enough to verify it for a special example.

For example, consider the statement "$x^2 = 1$ for all $x \in \mathbb{R}$".

Here is a "proof": for $x = 1$ we have $x^2 = 1^2 = 1$ thus the statement is correct.

Of course, in this simple example we see immediately that this is rubbish!

Never prove a statement by giving an example!

Definition 1.14. A function f is said to be *one–to–one* (or *injective*) if for all $x_1, x_2 \in \operatorname{dom}(f)$:

$$f(x_1) = f(x_2) \implies x_1 = x_2\,.$$

Thus, if f is one–to–one then $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.

Common errors when showing that f is one–to–one:

(i) To show that $f(x_1) = f(x_2)$ if $x_1 = x_2$ (no marks: after all this is true for any function!)

(ii) To show that $f(x_1) \neq f(x_2)$ for some x_1, x_2. (No marks: recall that it is never a good idea to show a statement by an example!)

Example 1.15. The function $f(x) = \frac{2x+1}{x-1}$, $x \neq 1$, is one-to-one: Let $x_1, x_2 \in \mathbb{R} \setminus \{1\}$. Then

$$\begin{aligned} f(x_1) = f(x_2) &\implies \frac{2x_1+1}{x_1-1} = \frac{2x_2+1}{x_2-1} \\ &\implies (2x_1+1)(x_2-1) = (2x_2+1)(x_1-1) \\ &\implies x_2 - 2x_1 = x_1 - 2x_2 \implies x_1 = x_2 \,. \end{aligned}$$

Thus, f is one-to-one.

Theorem 1.16. *If f is one–to–one, then there exists a unique function $g : \mathrm{range}(f) \to \mathrm{dom}(f)$ such that*

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

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

Definition 1.17. The function g in the previous theorem is called the *inverse function* of f, and is denoted by $g = f^{-1}$.

Remark 1.18. *Note that*

$$f^{-1} \neq \frac{1}{f} \,.$$

For example, $f(x) = x$ has inverse function $f^{-1}(x) = x$ but $\frac{1}{f(x)} = \frac{1}{x}$.

Example 1.19. The function $f(x) = \frac{2x+1}{x-1}$, $x \neq 1$, is one-to-one. Put $y = f^{-1}(x)$. Then $f(y) = x$ can be solved for y:

$$x = \frac{2y+1}{y-1} \implies x(y-1) = 2y+1 \implies y(x-2) = 1+x \implies y = \frac{1+x}{x-2} \,.$$

Thus, $f^{-1}(x) = \frac{1+x}{x-2}$, $x \neq 2$. The range of f is thus $\mathbb{R} \setminus \{2\}$.
To verify our result we compute

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

as it should be.

The graph of the inverse function f^{-1} is the graph of f reflected in the line $x = y$. From this geometric picture, it is intuitively clear that the inverse of a function is also one–to–one, and its inverse is the original function.

Lemma 1.20. *Let f be one–to–one and $g = f^{-1}$ its inverse function. Then*

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

for all $x \in \mathrm{dom}(f)$. In particular, f^{-1} is one–to–one and the inverse function of f^{-1} is $(f^{-1})^{-1} = f$.

1.6 Mathematical induction

There are various axioms needed to describe the set of real numbers. For the set of natural numbers, we follow a rather informal, intuitive view; the main axiom we impose is the axiom of induction.

Axiom 1.21 (Axiom of induction (see also Peano axioms))**.** *Let* $S \subset \mathbb{N}$. *If*

(i) $1 \in S$

(ii) $k \in S \implies k+1 \in S$

then $S = \mathbb{N}$

Think about domino theory: to get a chain of dominoes to fall, you have to kick the first one (base case) but it only works if every domino's fall guarantees the next one is falling too (induction step).

Mathematical induction: Let $P(n)$ be a statement depending on a natural number n. We want to show that $P(n)$ holds for all $n \in \mathbb{N}$. Thus, put $S = \{n \in \mathbb{N} : P(n) \text{ is true}\}$. If we can prove the *base case*, i.e. $1 \in S$, and show the *induction step*, that is, $(k \in S \implies k+1 \in S)$, then $S = \mathbb{N}$ by the axiom of induction. Therefore, in this case, $P(n)$ is true for all $n \in \mathbb{N}$.

Theorem 1.22. *For all* $n \in \mathbb{N}$*:*

$$1 + 2 + \ldots + n = \frac{n(n+1)}{2}$$

We introduce the $\sum$ notation for sums: if $a_m, \ldots, a_n \in \mathbb{R}, m \leq n, m, n \in \mathbb{N}$, then we denote by

$$\sum_{l=m}^{n} a_l = a_m + a_{m+1} + \ldots + a_n\,.$$

For example,

$$\sum_{l=1}^{n} l = 1 + 2 + \ldots + n$$

Similarly, we define the $\prod$ notation for products: if $a_m, \ldots, a_n \in \mathbb{R}, m \leq n, m, n \in \mathbb{N}$, then we denote by

$$\prod_{l=m}^{n} a_l = a_m \cdot a_{m+1} \cdot \ldots \cdot a_n\,.$$

For example,

$$\prod_{l=1}^{n} l = 1 \cdot 2 \cdot \ldots \cdot n = n!$$

Thus, the previous theorem reads as

Theorem 1.23. *For all $n \in \mathbb{N}$:*

$$\sum_{l=1}^{n} l = \frac{n(n+1)}{2}$$

Theorem 1.24. $\forall x \geq -1$:

$$(1+x)^n \geq 1 + nx \quad \forall n \in \mathbb{N}\,.$$

Common errors:

- The base case is not verified (heavy mark reduction).
- In the induction step the order of argument is wrong (mild to heavy mark reduction, depending on the error); in particular, showing that the induction claim implies a true statement.

  Consider the argument:

  $1 = 0$ implies $0 = 0 \cdot 1 = 0 \cdot 0$

  Even though the last statement "$0 = 0$" is correct, obviously this does not prove that $1 = 0$.
- A final sentence referring to the principle of induction (e.g., "from (1) and (2) we see...") is missing (mild mark reduction).

Example 1.25 (Quantum Numbers)**.** One can prove by induction that the sum of the first n odd numbers is n^2, that is,

$$\forall n \in \mathbb{N} : \quad \sum_{l=0}^{n-1} (2l+1) = n^2\,.$$

This equation appears in connection with quantum numbers: an electron has

- main quantum number n
- orbital angular momentum quantum number l which can attain the values $0, 1, \ldots, n-1$
- magnetic quantum number with possible values from $-l$ to l, that is, with $2l+1$ possible values.

Thus, by given n, the number of possible combinations of orbital angular momentum quantum number and magnetic quantum number is given by n^2.

Chapter 2

Limits and Continuity

The notion of a limit is a "can't afford to miss" topic — if you do, it will come back to haunt you.

Why is it so important? Limits are at the core of calculus and analysis: every single notion is a limit in one sense or another!

Examples, where limits are used, include the slope of the tangent line, the length of a curve, the area under a graph.

2.1 Definition of a limit

Let us start with the idea of a limit (and this needs to be refined later so that we can actually work with limits). If you are struggling with the topic, going to the online resource and investing time and work is particularly important; you cannot ignore the problem.

Given a function $f : D \to \mathbb{R}$ on an open interval, and $c \in \mathbb{R}$ then we mean by

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

that

"as x approaches c, $f(x)$ approaches L"

or, formulated differently,

"$f(x)$ is close to L whenever $x \neq c$ is close to c."

By the one–sided limit

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

we mean

"as x approaches c from the left, $f(x)$ approaches L"

and by

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

we mean

"as x approaches c from the right, $f(x)$ approaches L".

Note that even if both one–sided limits exist, the limit might not: the left and right limits could have different values.

Let us now give a precise definition of a limit. Firstly, for saying $f(x)$ is close to L for all $x \neq c$ close to c, we need $f(x)$ to makes sense for $x \neq c$ near c (but not for c itself).

That is, we require that there is a (possibly very small) $p > 0$ such that the open intervals $(c - p, c), (c, c + p)$ are contained in the domain of f.

$f(x)$ is close to L means that the distance $|f(x) - L|$ is small, that is, as small as we choose it to be. So what we want is that given $\varepsilon > 0$ then $|f(x) - L| < \varepsilon$ if x is close enough to c.

But x is close enough to c, $x \neq c$, means that there exists a $\delta > 0$ such that $0 < |x-c| < \delta$.

Voilà, that's our definition:

Definition 2.1 (ε, δ–definition)**.** Let $f : D \to \mathbb{R}$ be a function with $(c-p, c) \cup (c, c+p) \subset D$ for some $p > 0$. Then we say that the *limit* of f at c is L, that is

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

if for all $\varepsilon > 0$ there exists $\delta > 0$ such that :

$$0 < |x - c| < \delta \implies |f(x) - L| < \varepsilon\,.$$

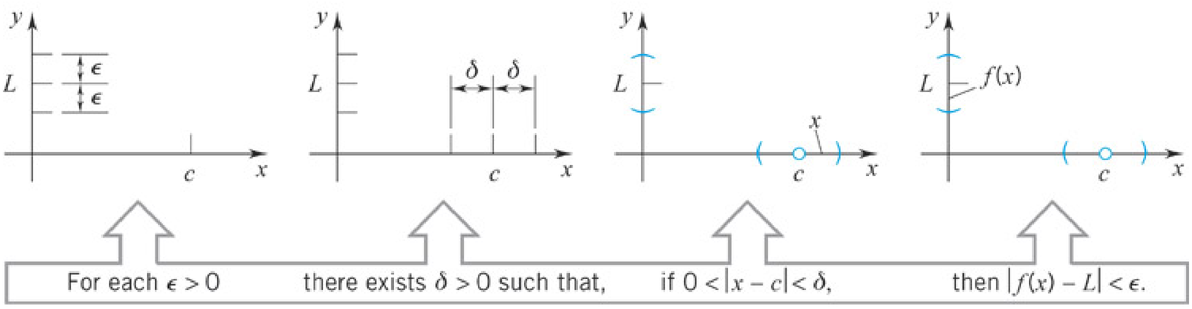


Figure 2.1: A picture of the definition of limit

Note that the order of ε and δ is crucial: we first choose how close we want to come to L and then have to find how close we have to be to c for this to be true.

Thus, δ depends on the choice of ε.

The definition of a limit is, as mentioned before, on the core of calculus and analysis. It may appear confusing in the beginning, so you should practise this notion with loads of examples.

The standard method to use the ε, δ–argument has two parts:

"Find δ" scratch work to find δ

"δ works" verify that your candidate for δ indeed satisfies the conditions.

Example 2.2. Let $f(x) = 2x - 1$. Then $\lim_{x\to 2} f(x) = 3$:

"Find δ" Scratch work to find δ: For $\varepsilon > 0$ want $\delta > 0$ so that if x has distance at most δ from 2, i.e., $0 < |x - 2| < \delta$, then $f(x)$ has distance at most ε from 3, i. e., $|f(x) - 3| < \varepsilon$. In our case $|f(x) - 3| = |2x - 1 - 3| = 2|x - 2|$ so we choose $\delta = \frac{\varepsilon}{2}$.

"δ works" We now verify that your candidate for δ indeed satisfies the conditions: Let $\varepsilon > 0$ and let $\delta = \frac{\varepsilon}{2}$. Then $\delta > 0$ and for x with $0 < |x - 2| < \delta$ we have

$$|f(x) - 3| = 2|x - 2| < 2\delta = \varepsilon\,.$$

Thus, indeed $\lim_{x\to 2} f(x) = 3$.

Example 2.3. Let $f(x) = \frac{x^2-1}{x-1}$. Then $\lim_{x\to 1} f(x) = 2$:

"Find δ" Scratch work to find δ: For $\varepsilon > 0$ want $\delta > 0$ so that if x has distance at most δ from c, i.e., $0 < |x - 1| < \delta$, then $f(x)$ has distance at most ε from 2, i. e., $|f(x) - 2| < \varepsilon$. In our case

$$|f(x) - 2| = \left|\frac{x^2 - 1}{x - 1} - 2\right| = \left|\frac{x^2 - 1 - 2(x - 1)}{x - 1}\right| = \left|\frac{x^2 - 2x + 1)}{x - 1}\right| = \left|\frac{(x - 1)^2)}{x - 1}\right| \overset{x\neq 1}{=} |x - 1|$$

so we choose $\delta = \varepsilon$.

"δ works" We now verify that our candidate for δ indeed satisfies the conditions: Let $\varepsilon > 0$ and let $\delta = \varepsilon$. Then $\delta > 0$ and for x with $0 < |x - 1| < \delta$ we have

$$|f(x) - 2| = \left|\frac{x^2 - 1}{x - 1} - 2\right| = \left|\frac{x^2 - 1 - 2(x - 1)}{x - 1}\right| = \left|\frac{x^2 - 2x + 1)}{x - 1}\right| = \left|\frac{(x - 1)^2)}{x - 1}\right| \overset{x\neq 1}{=} |x - 1| < \delta = \varepsilon.$$

Thus, indeed $\lim_{x\to 1} f(x) = 2$.

Note: we couldn't just plug in $x = 1$: in this case $x - 1 = 0$, and thus, f is not defined at $x = 1$. However, the limit is, as we have seen.

To get an idea of the value of the limit, observe that $f(x) = \frac{x^2-1}{x-1} = \frac{(x-1)(x+1)}{x-1} = x + 1$ for $x \neq 1$.

Common errors: You cannot plugin the value of the function to find the limit. Let us consider the function

$$f(x) = \begin{cases} 2 & x \neq 2 \\ 3 & x = 2 \end{cases}$$

We have $\lim_{x\to 2} f(x) \overset{x\neq 2}{=} \lim_{x\to 2} 2 \neq 3 = f(2)$.

Of course, we don't want to use the ε, δ–criterion every single time. So, we try to find some general rules which allow to derive a limit, once we know simpler ones. Here are some simple examples we can use later:

Example 2.4.

(i) $\forall c \in \mathbb{R} : \lim\limits_{x\to c} x = c$

(ii) $\forall c \in \mathbb{R} : \lim\limits_{x\to c} |x| = |c|$

(iii) If $f(x) = k$ is the constant function with constant $k \in \mathbb{R}$ then $\forall c \in \mathbb{R}$: $\lim\limits_{x\to c} f(x) = k$.

Could you explain these in terms ε, δ-arguments if you needed to in the exam?

Here are some other ways to verify that a number is the limit of f at c which will come handy in various situations.

Theorem 2.5 (Equivalent formulations)**.** *The following are equivalent:*

(i) $\lim\limits_{x\to c} f(x) = L$

(ii) $\lim\limits_{x\to c} (f(x) - L) = 0$

(iii) $\lim\limits_{h\to 0} f(c + h) = L$

(iv) $\lim\limits_{x\to c} |f(x) - L| = 0.$

In the same spirit the one–sided limits can be introduced, and corresponding statements proved. We will only touch quickly on this subject and leave it to you to fill in the details.

Definition 2.6. (i) Let $f : D \to \mathbb{R}$ be a function with $(c - p, c) \subset D$ for some $p > 0$. Then we say that the *left–sided limit* of f at c is L, that is

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

if for all $\varepsilon > 0$ there exists $\delta > 0$ such that :

$$c - \delta < x < c \implies |f(x) - L| < \varepsilon\,.$$

(ii) Let $f : D \to \mathbb{R}$ be a function with $(c, c+p) \subset D$ for some $p > 0$. Then we say that the *right–sided limit* of f at c is L, that is

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

if for all $\varepsilon > 0$ there exists $\delta > 0$ such that :

$$c < x < c + \delta \implies |f(x) - L| < \varepsilon\,.$$

The one–sided limits are connected to the limit via

Theorem 2.7. *Let $f : D \to \mathbb{R}$, $c \in \mathbb{R}$. Then*

$$\lim_{x \to c} f(x) = L \iff \lim_{x \to c^+} f(x) = L \text{ and } \lim_{x \to c^-} f(x) = L\,.$$

Of course it would not make sense for the same function at the same point to have two different limits, but this is not obvious from the definition so we should prove it:

Theorem 2.8. *If $\lim_{x \to c} f(x) = L$ and $\lim_{x \to c} f(x) = M$ then $L = M$.*

Proof. Suppose L and M were different, and try and prove this is imposible. For example, we could let $\varepsilon = |L - M|/4$, then if L and M we different we would have $\varepsilon > 0$ and by the definition of limit:

$$\exists \delta_L > 0 : 0 < |x - c| < \delta_L \Rightarrow |f(x) - L| < \varepsilon$$
$$\exists \delta_M > 0 : 0 < |x - c| < \delta_M \Rightarrow |f(x) - M| < \varepsilon$$

Now choose some x with $0 < |x - c| < \min(\delta_L, \delta_M)$, and by the triangle inequality

$$|L - M| = |(f(x) - M) - (f(x) - L)| \le |f(x) - M| + |f(x) - L| < \varepsilon + \varepsilon = |L - M|/2.$$

We have a contradiction to our hypothesis that $L \neq M$, so we must have $L = M$. □

2.2 The pinching theorem for limits

Theorem 2.9 (Pinching Theorem). *Let $f, g, h : D \to \mathbb{R}$. Suppose that for all $x \in D$*

$$h(x) \le f(x) \le g(x)\,.$$

If $\lim_{x\to c} h(x) = L$ *and* $\lim_{x\to c} g(x) = L$ *then*

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

Proof. Let $\varepsilon > 0$. Since $\lim_{x\to c} h(x) = L$ and $\lim_{x\to c} g(x) = L$ there exist $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$0 < |x-c| < \delta_1 \implies |h(x) - L| < \varepsilon, \text{ i.e., } L - \varepsilon < h(x) < L + \varepsilon,$$

and

$$0 < |x-c| < \delta_2 \implies |g(x) - L| < \varepsilon, \text{ i.e., } L - \varepsilon < g(x) < L + \varepsilon\,.$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Then for $0 < |x-c| < \delta$ we have

$$L - \varepsilon < h(x) \le f(x) \le g(x) < L + \varepsilon\,,$$

that is $|f(x) - L| < \varepsilon$. Thus, $\lim_{x\to c} f(x) = L$.

□

2.3 Limit laws

Our goal is to find laws which simplify finding a limit.

We have already seen the pinching theorem allows us to prove results that might look hard. We also have basic results such as:

Theorem 2.10. *If* $\lim_{x\to c} f(x) = L$ *and* $\lim_{x\to c} g(x) = 0$ *then* $\lim_{x\to c} f(x)g(x) = 0$.

The following theorem will allow us to compute lots of limits by using the examples we have seen before.

Theorem 2.11 (Limit Laws)**.** *If* $\lim_{x\to c} f(x) = L$ *and* $\lim_{x\to c} g(x) = M$ *then*

(i) $\lim_{x\to c}[f(x) + g(x)] = L + M$

(ii) $\lim_{x\to c}[\alpha f(x)] = \alpha L$ *for* $\alpha \in \mathbb{R}$

(iii) $\lim_{x\to c}[f(x)g(x)] = LM$

Remark 2.12. *(i) For the limit laws it is important that the individual limits exist. For example, for* $f(x) = \frac{1}{x}, g(x) = x$

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

whereas $\lim_{x\to 0} g(x) = 0$ *and* $\lim_{x\to 0} f(x)$ *does not exist; thus, in this case*

$$\lim_{x\to 0}[f(x)g(x)] \neq [\lim_{x\to 0} f(x)] \cdot [\lim_{x\to 0} g(x)] .$$

Thus, if you use the limit laws you have to state why you are allowed to so (mark reduction otherwise).

(ii) Note that the same limit laws apply for one–sided limits.

By induction we conclude from the Limit Laws:

Corollary 2.13. *Let* f_k *be functions and* $a_k \in \mathbb{R}$*, for* $k = 1, \ldots, n$*. If each* $\lim_{x\to c} f_k(x)$ *exists then*

$$\lim_{x\to c}\left(\sum_{k=1}^{n} a_k f_k(x)\right) = \sum_{k=1}^{n} a_k (\lim_{x\to c} f_k(x))$$

and

$$\lim_{x\to c}\left(\prod_{k=1}^{n} f_k(x)\right) = \prod_{k=1}^{n} \lim_{x\to c} f_k(x)$$

Applying the second result to the functions $f_k(x) = x$ we see that $\lim_{x\to c} x^n = c^n$, and applying the first result to $f_k(x) = x^k$ we see:

Corollary 2.14. *If* $P(x) = a_n x^n + \ldots + a_0$ *is a real polynomial and* $c \in \mathbb{R}$ *then*

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

Functions with $\lim_{x\to c} f(x) = f(c)$ are called *continuous* at c. We will discuss this very soon!

Theorem 2.15. *If* $\lim_{x\to c} g(x) = M$*,* $M \neq 0$*,* $\lim_{x\to c} f(x) = L$*, then*

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{L}{M} .$$

Applying this to the quotient of polynomials we have:

Corollary 2.16. *For rational functions* $R = \frac{P}{Q}$ *with polynomials* P *and* Q *we have*

$$\lim_{x\to c} R(x) = R(c)$$

provided that $Q(c) \neq 0$.

Another consequence is that we obtain a criterion for a limit not to exist:

Theorem 2.17. *If* $\lim_{x\to c} f(x) = L$, $L \neq 0$, *and* $\lim_{x\to c} g(x) = 0$ *then* $\lim_{x\to c} \dfrac{f(x)}{g(x)}$ *does not exist.*

The assumption $L \neq 0$ is crucial; for example, for $f(x) = x^2$ and $g(x) = x$ the limit exists and is $\lim_{x\to 0} \frac{f(x)}{g(x)} = 0$.

Proof. Proof by contradiction: Assume that $\lim_{x\to c} \dfrac{f(x)}{g(x)} = K$ for some $K \in \mathbb{R}$. Then by the limit laws

$$L = \lim_{x\to c} f(x) = \lim_{x\to c} g(x)\frac{f(x)}{g(x)} = \lim_{x\to c} g(x) \lim_{x\to c} \frac{f(x)}{g(x)} = 0 \cdot K = 0\,.$$

Since $L \neq 0$ this is a contradiction. □

We can only apply the limit laws if the limits exist. So there is no contradiction in the above argument if $\lim_{x\to c} \dfrac{f(x)}{g(x)}$ does not exist!

2.4 Trigonometric examples

Example 2.18.

$$\lim_{x\to 0} \sin x = 0, \quad \lim_{x\to 0} \cos x = 1\,.$$

Use $0 < |\sin x| < x$ for x near 0 and the Pinching Theorem. For the second statement use $\cos x = \sqrt{1 - \sin^2 x}$ for x near 0.

Another important consequence is:

Corollary 2.19.

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \quad \lim_{x\to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}\,.$$

2.5 Continuity

In ordinary language, “continuous” means that a process goes on without interruptions and abrupt changes. The mathematical notion is more precise.

Definition 2.20. Let $f : D \to \mathbb{R}$ a function and $c \in \mathbb{R}$ such that there exists $p > 0$ such that $(c - p, c + p) \in D$. Then f is called *continuous* at c if

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

In contrast to the definition of a limit, for continuity the function has to be defined at c. Thus, a function can only fail to be continuous at c if

(i) $\lim_{x \to c} f(x)$ exists but is not $f(c)$, or

(ii) $\lim_{x \to c} f(x)$ does not exist (but $f(c)$ does).

In the first case, we can "remove" the discontinuity by changing the value of f at c to be L, and we call it a *removable discontinuity*. The second case is called an *essential discontinuity*: however one changes finitely many values of f, the discontinuity will not go away.

Example 2.21. The *Dirichlet function*

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ -1, & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

has an essential discontinuity at every point.

We already have a small list of continuous functions: polynomials, rational functions $\frac{P}{Q}$ provided $Q(c) \neq 0$, absolute value function, square root function.

As before, we can construct new continuous functions by algebraic operations

Theorem 2.22. *If $f, g : D \to \mathbb{R}$ are continuous at $c \in D$, $\alpha \in \mathbb{R}$, then*

$$f + g, \quad f - g, \quad f \cdot g, \quad \alpha f, \quad \frac{f}{g} \quad \textit{(provided } g(c) \neq 0\textit{)}$$

are continuous at c.

Proof. Apply the Limit Laws! □

We now apply the ε, δ–criterion to get:

Theorem 2.23. *f is continuous at c if and only if $\forall \varepsilon > 0 \exists \delta > 0$:*

$$|x - c| < \delta \implies |f(x) - f(c)| < \varepsilon\,.$$

Proof. In difference to the definition of a limit, we can now allow $x = c$ because then $f(x) - f(c) = 0$, thus, $|f(x) - f(c)| < \varepsilon$ for all $\varepsilon > 0$! With this observation, the theorem follows immediately from the ε, δ definition of a limit. □

A form of this ε, δ definition of continuity was first given by Bernard Bolzano in 1817. Preliminary forms of a related definition of the limit were given by Cauchy. The formal definition and the distinction between pointwise continuity and uniform continuity were first given by Bolzano in the 1830s but the work wasn't published until the 1930s. Heine provided the first published definition of uniform continuity in 1872, but based these ideas on lectures given by Dirichlet in 1854.

Theorem 2.24. *If g is continuous at c, and f is continuous at $g(c)$ then $f \circ g$ is continuous at c.*

Proof. Let $\varepsilon > 0$. Since f is continous at $g(c)$ there exists $\delta_1 > 0$ such that

$$|t - g(x)| < \delta_1 \implies |f(t) - f(g(c))| < \varepsilon\,.$$

Since g is continuous at c, there exists $\delta > 0$ such that

$$|x - c| < \delta \implies |g(x) - g(c)| < \delta_1\,.$$

But then the previous statement shows (with $t = g(x)$) that

$$|x - c| < \delta \implies |g(x) - g(c)| < \delta_1 \implies |f(g(x)) - f(g(c))| < \varepsilon\,.$$

This shows that $f \circ g$ is continuous at c. $\square$

Example 2.25. $F(x) = \sqrt{\frac{x^2+1}{x-3}}$ is continuous for all $x \in (3, \infty)$: use $F = f \circ g$ with $f(x) = \sqrt{x}, g(x) = \frac{x^2+1}{x-3}$.

Definition 2.26. A function f is said to be *continuous on an interval* if it is continuous for all interior points of the interval, and one–sidedely continuous at the endpoints.

2.6 The Intermediate–Value Theorem

From our intuition, we think a continuous function on an interval shouldn't have holes or jumps. This is the content of the very important Intermediate Value Theorem. We will not prove this result here but it can be shown using the completeness of $\mathbb{R}$, that is, the Least Upper Bound Axiom.

Theorem 2.27 (Intermediate–Value Theorem)**.** *Let f be continuous on $[a, b]$ and K any number between $f(a)$ and $f(b)$. Then there exists at least one $c \in (a, b)$ with*

$$f(c) = K\,.$$

For $K = 0$, the statement is also known as Bolzano's theorem.

Note: c is in the open interval (a, b): it cannot take the values $c = a$ or $c = b$ since $f(a) \neq K$ and $f(b) \neq K$. Stating the Intermediate–Value Theorem as "... at least one $c \in [a, b]$..." gives mark reduction: although the statement is not false, the implication does not contain all information.

2.7 The Extreme Value Theorem

We first note that, if a function is continuous at a point, then in some small interval around that point it must be bounded.

Lemma 2.28. *If $f : D \to \mathbb{R}$ is continuous at $x = c$ there exist $\delta, B > 0$ such that, for all $x \in (c - \delta, c + \delta)$, $|f(x)| < B$.*

The extreme value theorem extends the previous result, to say a function is bounded on all closed domains. In fact it says more: every continuous function attains its maximum and minimum on closed intervals. Recall

Definition 2.29. f takes on a *maximum* on I if there exists $x \in I$ with $f(x) = \max f(I)$, that is, $f(t) \leq f(x)$ for all $t \in I$.

f takes on a *minimum* on I if there exists $x \in I$ with $f(x) = \min f(I)$, that is, $f(t) \geq f(x)$ for all $t \in I$.

(Note: later we will call these global maximum and global minimum to distinguish from the local extrema).

There are examples of functions on $[a, b]$ which are bounded but do not attain their maximum or minimum. However, if the function is continuous, this cannot happen. We will not prove this theorem since the proof, as in the case of the Intermediate-Value Theorem needs the Least Upper Bound Axiom.

Theorem 2.30 (Extreme–Value Theorem)**.** *Let f be continuous on a bounded closed interval $[a, b]$. Then f takes on both its maximum and its minimum on $[a, b]$.*

The extreme value theorem was originally proved by Bernard Bolzano in the 1830s in a work Function Theory but the work remained unpublished until 1930. Bolzano's proof consisted of showing that a continuous function on a closed interval was bounded, and then showing that the function attained its maximum and minimum value. Both proofs involved what is known today as the Bolzano–Weierstrass theorem, see Chapter 6.

Thus, continuous functions map bounded closed intervals $[a, b]$ onto bounded closed intervals $[m, M]$.

All assumptions of the Extreme–Value Theorem are needed!

- If the interval is not bounded, then a continuous function may not attain its extreme values: $f(x) = x^3$ on $[0, \infty)$.
- If the interval is not closed, then a continuous function may not attain its extreme values: $f(x) = x^3$ on $[0, 1)$.
- If the function is not continuous on $[a, b]$ then f may not attain its extreme values:
  $$f(x) = \begin{cases} 3, & x = 1 \\ x, & 1 < x < 5 \\ 3, & x = 5 \end{cases}.$$

2.8 Continuity and inverse functions

Definition 2.31. A function f is said to be

(strictly) ***increasing*** on an interval I if for all $x_1, x_2 \in I$, $x_1 < x_2$: $f(x_1) < f(x_2)$.

(strictly) ***decreasing*** on an interval I if for all $x_1, x_2 \in I$, $x_1 < x_2$: $f(x_1) > f(x_2)$.

Example 2.32. If f is increasing then f is one–to–one.

Lemma 2.33. *If f is continuous on (a, b) and one–to–one, then f is increasing or decreasing.*

Proof. Workbook. □

Theorem 2.34. *If f is one–to–one and continuous on (a, b) so is f^{-1}.*

2.9 Study guide

Go through the material of chapter 2, and identify the main topics and statements. Have a look some more examples: could you do reproduce them without looking at your notes? Do you understand what they are examples of? Are they counterexamples? Check the red text: be aware of the most commonly made errors!

To deepen your understanding, think about the following:

- What is the difference between the following statements:
  - for all $\varepsilon > 0$ there exists $\delta > 0$: $0 < |x - c| < \delta \implies |f(x) - f(c)| < \varepsilon$.
  - there exists $\alpha > 0$ so that for all $\beta > 0$: $0 < |x-c| < \alpha \implies |f(x)-f(c)| < \beta$.
  - for all $\alpha > 0$ there exists $\beta > 0$: $0 < |x - c| < \alpha \implies |f(x) - f(c)| < \beta$.

  Give in each case an example of a function satisfying the condition.

- Why are the following statements not characterising a limit?
  (i) If x gets closer to c then $f(x)$ gets closer to $f(c)$.
  (ii) By choosing x close to c we can have $f(x)$ arbitrarily close to $f(c)$.

  Give in both cases an example which satisfies the condition but doesn't have a limit at c.

- What is the difference between a function which has a limit at c and a function which is continuous at c?

- Find a continuous function without zeros. Is your example contradicting the Intermediate–Value Theorem? Does $f(x) = \sin(x^3) - 1, x \in \mathbb{R}$, attain the value -1.5? If so, find c with $f(c) = -1.5$ (you may want to use your Excel-skills).

Chapter 3

Derivatives

So far, continuity has given the existence of extreme values on a bounded, closed interval. But how do we actually find these extremas? Obviously, this question occurs in plenty of important problems in “real life”: can you maximise a profit? can you minimise material used? can you minimise energy? The answer involves derivatives.

3.1 Definition of the derivative

If we have an extreme value in the interior of an interval then its tangent line has to be horizontal. How do we find such points? How do we find the tangent line? Do all functions have tangent lines at all points? These are the mathematical questions we will have to answer.

We first observe that the *slope of the secant* through $f(x)$ and $f(x+h)$, $h \neq 0$, is given by the *difference quotient*

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

Now, for h approaching 0, ideally $f(x+h)$ slides towards $f(x)$, and the slope of the secant approaches the slope of the tangent line.

Definition 3.1. A function $f : (a,b) \to \mathbb{R}$ is said to be *differentiable* at $x \in (a,b)$ if

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

exists. In this case, the limit is called the *derivative* of f at x, and is denoted by

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

Moreover, the line through $f(x)$ with slope $f'(x)$ is called the *tangent line* of f at x, and the line through $f(x)$ with slope $-\frac{1}{f'(x)}$ is called the *normal line* of f at x if $f'(x) \neq 0$.

Example 3.2. (i) $f(x) = mx + b$, $m, b \in \mathbb{R}$ has $f'(x) = m$:

$$\frac{f(x+h) - f(x)}{h} = \frac{m(x+h) + b - (mx+b)}{h} = \frac{mh}{h} \overset{h \neq 0}{=} m$$

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

(ii) $f(x) = x^2$ has $f'(x) = 2x$:

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{2xh + h^2}{h} \overset{h \neq 0}{=} 2x + h$$

so that $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} 2x + h = 2x$.

(iii) $f(x) = \sqrt{x}$, $x \in [0, \infty)$, is differentiable for $x \neq 0$ with $f'(x) = \frac{1}{2\sqrt{x}}$:

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{\sqrt{x+h} - \sqrt{x}}{h} = \frac{\sqrt{x+h} - \sqrt{x}}{h} \cdot \frac{\sqrt{x+h} + \sqrt{x}}{\sqrt{x+h} + \sqrt{x}} \\ &= \frac{x + h - x}{h(\sqrt{x+h} + \sqrt{x})} \overset{h \neq 0}{=} \frac{1}{\sqrt{x+h} + \sqrt{x}} \end{aligned}$$

so that $f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{1}{\sqrt{x+h} + \sqrt{x}} = \frac{1}{2\sqrt{x}}$, provided $x \neq 0$.

(iv) $f(x) = \sin x$ has $f'(x) = \cos x$:

$$\begin{aligned} \frac{f(x+h) - f(x)}{h} &= \frac{\sin(x+h) - \sin x}{h} = \frac{\sin x \cos h + \sin h \cos x - \sin x}{h} \\ &= \sin x \left(\frac{\cos h - 1}{h} \right) + \left(\frac{\sin h}{h} \right) \cos x \end{aligned}$$

so that by limit laws $f'(x) = \lim_{h \to 0} \sin x \left(\frac{\cos h - 1}{h} \right) + \lim_{h \to 0} \left(\frac{\sin h}{h} \right) \cos x = \cos x$.

(v) $f(x) = \cos x$ has $f'(x) = -\sin x$.

(vi) $f(x) = \begin{cases} x + 1 & x \geq 0 \\ x - 1 & x < 0 \end{cases}$ is not differentiable at $x = 0$:

For $x = 0$ we have $f(0) = 1$ so that

$$\lim_{h \to 0+} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0+} \frac{(0+h) + 1 - 1}{h} = 1$$

whereas the limit

$$\lim_{h \to 0-} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0-} \frac{(0+h) - 1 - 1}{h} = \lim_{h \to 0-} \frac{h - 2}{h}$$

does not exist. Thus, $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$ does not exist either! (The function jumps at $x=0$: the secants from $f(h)=h-1$ to $f(0)=1$ have for $h<0$, and near 0, large slope).

You cannot compute the derivative of a function which is defined by cases by computing $f'(x)$ for $x<c$, $f'(x)$ for $x>c$ and showing $\lim_{x\to c-} f'(x) = \lim_{x\to c+} f'(x)$. The above example shows that in this case $\lim_{x\to c-} f'(x) = \lim_{x\to c+} f'(x) = 1$ but f is still not differentiable! Zero marks if you use this wrong argument.

If f is differentiable at c, then the equation of the tangent line at c is given by

$$y = f(c) + f'(c)(x-c)$$

and the equation of the normal line, if $f'(c) \neq 0$, by

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

(If $f'(c)=0$ then the normal line is the vertical line $x=f(c)$).

The tangent line approximates the graph of f near c. In other words, the function

$$g(x) = f(c) + f'(c)(x-c)$$

approximates the function f near c. We say that g *approximates* f *linearly*.

Note that a function can be continuous at some x but not differentiable: this could happen if the difference quotient diverges to ∞, or if this limit does not exist.

Example 3.3. $f(x)=|x|$ is continuous at $x=0$ but not differentiable at $x=0$.

However, if a function is differentiable then it must be continuous:

Theorem 3.4. *If f is differentiable at x then f is continuous at x.*

Proof. If $h \neq 0$ then

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

By Limit Laws:

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

which shows that $\lim_{h\to 0} f(x+h) = f(x)$ and f is continuous at x. □

Example 3.5. If a point moves on a straight line with constant speed v, then after time Δt the travelled distance is $\Delta s = v\Delta t$. Conversely, the average speed of a point is computed by $v = \frac{\Delta s}{\Delta t}$. If the point moves with variable speed then by considering very small time intervals, one can still assume that the speed in the small interval is roughly constant, and $v(t) = \frac{s(t+\Delta t)-s(t)}{\Delta t}$ in the small time period. Taking the limit, we indeed see that *speed* at time t is given by the derivative $v(t) = s'(t)$.

3.2 Rules of differentiation

As before, we want to use known derivatives to compute new ones.

Theorem 3.6. *Let $\alpha \in \mathbb{R}$, and let $f, g : (a, b) \to \mathbb{R}$ be differentiable at $x \in (a, b)$. Then $f \pm g, \alpha f$ are differentiable at x and*

$$(f \pm g)'(x) = f'(x) \pm g'(x), \quad (\alpha f)'(x) = \alpha f'(x)\,.$$

Proof. Left for you to do: use the definition of the derivative using limits of difference quotients, and use the limit laws. □

By induction we also have

$$\left(\sum_{k=1}^{n} \alpha_l f_k\right)' = \sum_{k=1}^{n} \alpha_k f_k'\,.$$

This says that differentiation is a "linear map". However, things become more complicated when considering products of two functions:

Theorem 3.7 (Product rule)**.** *If f, g are differentiable at x then so is $f \cdot g$ and*

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

Proof. We have

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

Now, g is continuous (since g is differentiable) and we can apply the Limit Laws:

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

□

Example 3.8. For all $n \in \mathbb{N}$: $f(x) = x^n$ has derivative $f'(x) = nx^{n-1}$.

Corollary 3.9. *A polynomial* $P(x) = a_n x^n + \ldots + a_0$ *is differentiable for all* $x \in \mathbb{R}$ *and*

$$P'(x) = n a_n x^{n-1} + \ldots + a_1 \,.$$

Example 3.10. Assume that g is a differentiable function and

$$F(x) = (x^3 - 5x) g(x) \,.$$

Compute $F'(2)$ provided that $g(2) = 3, g'(2) = -1$:

$$F'(x) = (3x^2 - 5) g(x) + (x^3 - 5x) g'(x)$$

so that $F'(2) = 23$.

Next, we investigate the reciprocal of a differentiable function:

Theorem 3.11 (Reciprocal rule)**.** *Let g be differentiable at x and $g(x) \neq 0$. Then $\frac{1}{g}$ is differentiable at x and*

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)} \,.$$

Proof. Since g is differentiable, it is continuous. Since $g(x) \neq 0$ this shows that $g(x+h) \neq 0$ for small h. Thus,

$$\lim_{h \to 0} \frac{1}{g(x+h)} = \frac{1}{g(x)} \,.$$

Now

$$\frac{\frac{1}{g(x+h)} - \frac{1}{g(x)}}{h} = \frac{g(x) - g(x+h)}{h g(x) g(x+h)} = \frac{g(x) - g(x+h)}{h} \cdot \frac{1}{g(x) g(x+h)}$$

so that by Limit Laws

$$\lim_{h \to 0} \frac{\frac{1}{g(x+h)} - \frac{1}{g(x)}}{h} = -\lim_{h \to 0} \frac{g(x+h) - g(x)}{h} \cdot \lim_{h \to 0} \frac{1}{g(x) g(x+h)} = -\frac{g'(x)}{g^2(x)}$$

□

Example 3.12. For all $n \in \mathbb{Z}$: $(x^n)' = nx^{n-1}$.

Combining the product and the reciprocal rule we obtain:

Theorem 3.13 (Quotient rule)**.** *If f, g are differentiable at x and $g(x) \neq 0$ then $\frac{f}{g}$ is differentiable at x and*

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

Proof.
$$\left(\frac{f}{g}\right)' = (f \cdot \frac{1}{g})' = f'\frac{1}{g} + f(-\frac{g'}{g^2}) = \frac{f'g - fg'}{g^2}\,.$$
□

Example 3.14. $(\tan x)' = \frac{1}{\cos^2 x} = 1 + \tan^2 x$ for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

In various source, the "Leibniz notation" is used: if y is a function of x then the derivative of y is denoted by $\frac{dy}{dx}$. If y is a function of t then the derivative of y is denoted by $\frac{dy}{dt}$. If we want to evaluate the derivative at a particular point x_0 then we write

$$y'(x_0) = \frac{dy}{dx}|_{x=x_0}\,.$$

With this notation, we can look at higher derivatives in the following way: if $f:(a,b)\to\mathbb{R}$ is differentiable on (a,b) then the derivative is a function $f':(a,b)\to\mathbb{R}$. If this function is differentiable, we denote by

$$f''(x) = \frac{d}{dx}(f'(x)) = \frac{d^2}{dx^2}(f(x))$$

the *second derivative* of f and by $f'''(x) = \frac{d}{dx}(f''(x))$ the *third derivative* of f. We denote by $f^{(n)}(x)$ the n^{th}–derivative of f.

Example 3.15. The *acceleration* a is the rate of change of the speed v. Since the speed is the rate of change of location s, we have

$$a = \frac{dv}{dt} = \frac{d^2s}{dt^2}\,.$$

We can now look at the composition of functions. That is, if y is a function of u and if u is a function of x then the chain rule states in Leibniz notation that

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

Written in function notation this reads as

Theorem 3.16 (Chain rule)**.** *If $F = f\circ g$ and g is differentiable at x and f differentiable at $g(x)$ then F is differentiable at x and*

$$F'(x) = (f\circ g)'(x) = f'(g(x))g'(x)\,.$$

"Proof:" Cancel du in the Leibniz notation. Done.

Of course, this is **not** a proof. Think about this: what would you want to cancel in $(f\circ g)' = (f'\circ g)g'$?

In this sense, the Leibniz notation is dangerous: it is tempting to consider dy and dx as independent objects (and indeed, there is a mathematical way to make this precise); however, we are not allowed to separate dy and dx (yet) but we read $\frac{dy}{dx}$ as "differentiate y with respect to the variable x".

Proof. We first recall

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

Thus, we consider the difference quotient

$$\frac{f(g(t))-f(g(x))}{t-x}=\underbrace{\frac{f(g(t))-f(g(x))}{g(t)-g(x)}}_{\to f'(g(x))}\underbrace{\frac{g(t)-g(x)}{t-x}}_{\to g'(x)}$$

What is wrong at this point? We cannot guarantee that $g(t)\neq g(x)$ near x, so on the right hand side we might divide by 0 — which is never a good idea!

To avoid this, we try to do something similar: we define a function which is the desired quotient as long as the denominator is not zero, and the predicted derivative otherwise:

$$H(y)=\begin{cases}\frac{f(y)-f(g(x))}{y-g(x)}, & y\neq g(x)\\ f'(g(x)), & y=g(x)\end{cases}\,.$$

This function H is continuous: the only critical point is $y=g(x)$ but there we have

$$\lim_{y\to g(x)}H(y)=\lim_{y\to g(x)}\frac{f(y)-f(g(x))}{y-g(x)}=f'(g(x))=H(g(x))\,.$$

For $t\neq x$ we have

$$\frac{f(g(t))-f(g(x))}{t-x}=H(g(t))\frac{g(t)-g(x)}{t-x}\,;$$

this holds since for $g(t)=g(x)$ both sides are 0, and otherwise by the definition of H.

Now, H and g are continuous (g is differentiable) so that the composition $F\circ g$ is continuous. Therefore, $\lim_{t\to x}H(g(t))=H(g(x))=f'(g(x))$ and the result follows from the limit laws.

□

The chain rule seems to have first been used by Leibniz. He used it to calculate the derivative of $\sqrt{a+bz+cz^2}$ as the composite of the square root function and the function $a+bz+cz^2$. He first mentioned it in a memoir with various mistakes in it. L'Hôpital uses the chain rule implicitly but also does not state it explicitly. The chain rule does not appear in any of Leonhard Euler's analysis books, even though they were written over a hundred years after Leibniz's discovery!

Example 3.17. Gravel is being poured by a conveyor onto a conical pile at the constant rate of 60π cubic feet per minute. Frictional forces within the pile are such that the height is always two–thirds of the radius. How fast is the radius of the pile changing at the instant the radius is 5 feet?

3.3 Derivatives of Inverse Functions

Contemplating a geometric picture, we observe that the tangent line of the inverse function f^{-1} should be well–defined wherever the tangent line of f is not horizontal.

Theorem 3.18. *Let $f : I \to \mathbb{R}$ be one–to–one and differentiable on an open interval I. Let $a \in I$ and $f(a) = b$. If $f'(a) \neq 0$ then f^{-1} is differentiable at b and*

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

What is wrong with the following "proof": Since $f \circ f^{-1}(x) = x$ we have by chain rule that $f'(f^{-1}(x)) \cdot (f^{-1}(x))' = 1$. Thus, if $f'(f^{-1}(x)) \neq 0$ then $(f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}$.

Of course, this argument doesn't work since we can only apply the chain rule, if we already know that f^{-1} is differentiable (but we don't!).

Even though the above argument NOT a proof, it is still a good way how to derive the formula for the derivative of the inverse in case you have forgotten it!

Obviously, the trigonometric functions are not one–to–one on their whole domain, so that we have to restrict the domain and consider

$$\begin{aligned} f(x) &= \sin(x), & x &\in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]\,, \\ g(x) &= \cos(x), & x &\in [0, \pi]\,, \\ h(x) &= \tan(x), & x &\in (-\tfrac{\pi}{2}, \tfrac{\pi}{2})\,. \end{aligned}$$

Restricting to the above domains the trigonometric functions are one–to–one, and have inverse functions

$$\begin{aligned} f^{-1}(x) &= \arcsin(x), & x &\in [-1, 1]\,, \\ g^{-1}(x) &= \arccos(x), & x &\in [-1, 1]\,, \\ h^{-1}(x) &= \arctan(x), & x &\in \mathbb{R}\,, \end{aligned}$$

with derivatives

$$\begin{aligned} \frac{d}{dx}\arcsin(x) &= \frac{1}{\sqrt{1-x^2}}, & x &\in (-1, 1)\,, \\ \frac{d}{dx}\arccos(x) &= -\frac{1}{\sqrt{1-x^2}}, & x &\in (-1, 1)\,, \\ \frac{d}{dx}\arctan(x) &= \frac{1}{1+x^2}, & x &\in \mathbb{R}\,. \end{aligned}$$

3.4 Newton's method

Here we just give a very short introduction.

The Intermediate–Value Theorem can be used to determine whether a continuous function has a zero. But how do we find it? The bisection method is one way to compute a zero.

A more efficient way is to use Newton's method: Given a differentiable function f, our goal is to find zeros of f on an interval $[a, b]$. Start with some $x_1 \in [a, b]$. If $f(x_1) = 0$ you were lucky and you are done. If not, since the tangent line describes the local shape of the function, we may try to find a better approximation of the zero by using the intersection of the tangent line with the x–axis instead of x_1: call this point x_2. Continue this process.

Let $x_n \in [a, b]$. The tangent line at x_n is given by $y = f(x_n) + f'(x_n)(x - x_n)$. The intersection x_{n+1} with the x–axis is given by $0 = f(x_n) + f'(x_n)(x_{n+1} - x_n)$. Solving for x_{n+1} we obtain

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}\,.$$

Of course, this recipe does not work in all cases. For example, the tangent line could be horizontal at some point x_n (and then we cannot divide by its slope!). Or in the process you may just jump between two different values, or the successive values may not converge to a zero.

However, the method will usually converge, provided the initial guess of x_1 is close enough to the unknown zero, and that $f'(x_1) \neq 0$.

Newton's method was described by Isaac Newton (1669), however, his description differs substantially from the modern description given above: Newton applies the method only to polynomials. Isaac Newton probably derived his method from a similar but less precise method by Vieta. The essence of Vieta's method can be found in the work of the Persian mathematician, Sharaf al-Din al-Tusi. A special case of Newton's method for calculating square roots was known much earlier and is often called the Babylonian method.

3.5 The Mean–Value Theorem

For a differentiable function we see geometrically that for any secant there is a line parallel to the secant which is tangent to the graph of the function f.

A special case of this theorem was first described by Parameshvara (1370–1460) from the Kerala school of astronomy and mathematics. The mean value theorem in its modern form was later stated by Augustin Louis Cauchy (1789–1857).

The mathematical precise statement is:

Theorem 3.19 (Mean–Value Theorem)**.** *Let f be differentiable on (a, b), continuous on $[a, b]$. Then there exists at least one number $c \in (a, b)$ for which*

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

As before, stating that $c \in [a, b]$ will give mark reduction: not all relevant information is given. In the use of the Theorem, many students forget to check that f is continuous on $[a, b]$.

Although the statement of the Mean–Value Theorem may seem unimpressive at first sight, the theorem has a myriad of application in studying the shape of functions and thereby give solutions to various type of extremal questions: how large can a profit be? How much energy can we save by using a particular parameter?

The proof is done in various steps, each interesting in themselves. First we show that the sign of the derivative at one point x_0 prescribes the shape of the graph near x_0.

Theorem 3.20. *Let f be differentiable at x_0.*

(i) If $f'(x_0) > 0$ then $f(x_0 - h) < f(x_0) < f(x_0 + h)$ for small enough $h > 0$.

(ii) If $f'(x_0) < 0$ then $f(x_0 + h) < f(x_0) < f(x_0 - h)$ for small enough $h > 0$.

Proof. Assume $f'(x_0) > 0$. Put $\varepsilon = f'(x_0)$. Since f is differentiable at x_0 there exists $\delta > 0$ such that

$$0 < |x - x_0| < \delta \implies \left|\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0)\right| < \varepsilon = f'(x_0)\,.$$

This implies $0 < \frac{f(x)-f(x_0)}{x-x_0} < 2f'(x_0)$.
Now, if $0 < h < \delta$ then $x = x_0 \pm h$ satisfies

$$0 < |x - x_0| = |\pm h| = h < \delta\,.$$

Thus,

$$0 < \frac{f(x_0 \pm h) - f(x_0)}{x_0 \pm h - x_0}\,.$$

The $+$ inequality implies $0 < f(x_0 + h) - f(x_0)$, the other one $0 > f(x_0 - h) - f(x_0)$. This shows the claim in the case $f'(x_0) > 0$. The other case can be done similarly.

□

Further important ingredients for the Mean-Value Theorem are the Extreme Value Theorem and Rolle's theorem. The first known formal proof was offered by Michel Rolle in 1691, which used the methods of differential calculus. The name "Rolle's theorem" was first used by Moritz Wilhelm Drobisch of Germany in 1834 and by Giusto Bellavitis of Italy in 1846.

Theorem 3.21 (Rolle's theorem)**.** *Suppose f is differentiable on (a, b) and continuous on $[a, b]$. If $f(a) = f(b) = 0$ then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. By the Extreme-Value Theorem there must be a maximum and a minimum of the function on $[a, b]$. The aim is to show that this extremum is attained in the interval (a, b): then we show that f has to have horizontal tangent line at this point.

Of course, we can assume that f is not vanishing identically on $[a, b]$: if that was the case then $f(x) = 0$ for all x implies $f'(x) = 0$ for all x and we are done.

Thus, we can assume that there exists at least one point with $f(x) > 0$ (or $f(x) < 0$) with $x \in (a, b)$. Take the first case (the other one works similarly). By the Extreme-Value Theorem, f attains its maximum on $[a, b]$ since it is a continous function. Since $f(x) > 0$, the maximum will be attained at some c with $f(c) > 0$ and thus $c \neq a, b$. Thus, $c \in (a, b)$ and f is differentiable at c.

If $f'(c) > 0$ the previous theorem gives points nearby with $f(c-h) < f(c) < f(c+h)$ with $h > 0$ small, which contradicts the fact that f attains a maximum at c. Similarly, if $f'(c) < 0$ then $f(c-h) > f(c) > f(c+h)$ with $h > 0$ small, which again contradicts the fact that f attains a maximum at c. Thus, $f'(c) = 0$.

□

Example 3.22. The polynomial $p(x) = 2x^3 + 5x - 1$ has exactly one root: Since p is a polynomial of degree 3 it has at least one zero (use the Intermediate-Value Theorem: $p(0) = -1 < 0$, $p(1) = 6 > 0$). Assume that there exist $x_1 < x_2$ with $p(x_1) = p(x_2) = 0$. Since p is a polynomial, it is continous on $[x_1, x_2]$ and differentiable on (x_1, x_2). We can now apply Rolle's Theorem: there exists $c \in (x_1, x_2)$ with $p'(c) = 0$.

But $p'(x) = 6x^2 + 5 > 0$ for all $x \in \mathbb{R}$! Contradiction.

Proof of the Mean–Value Theorem. We will use Rolle's theorem. To apply Rolle's theorem, we want to find a function g which is zero at a and b, such that its derivative at c is the difference between $f'(c)$ and the difference quotient $\frac{f(b)-f(a)}{b-a}$. Our first attempt is to write

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

Then $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}x$. However, at a and b this function is not zero. Can we repair this? Second attempt

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

We haven't changed the derivative (we only changed the function by constants) and have now also $g(a) = g(b) = 0$. Thus, Rolle's theorem gives $c \in (a, b)$ with $g'(c) = 0$. Thus $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

Example 3.23. Suppose f is differentiable on $(1, 4)$, continuous on $[1, 4]$ and $f(1) = 2$. Given that $2 \leq f'(x) \leq 3$ for all $x \in (1, 4)$ what is the least/greatest value f can take at $x = 4$?

We can apply the Mean-Value Theorem since f is differentiable on $(1, 4)$ and continuous on $[1, 4]$. Thus, there exists $c \in (1, 4)$ with

$$f'(c) = \frac{f(4) - f(1)}{4 - 1} = \frac{f(4) - 2}{3}\,.$$

Since $2 \leq f'(x) \leq 3$ we thus have

$$2 \leq \frac{f(4) - 2}{3} \leq 3$$

that is $8 \leq f(40) \leq 11$.

Since the derivative gives the slope of tangent lines, we can connect information given by the derivative to the shape of graph of the function.

Theorem 3.24 (Increasing/Decreasing Theorem)**.** *Suppose $f : I \to \mathbb{R}$ is differentiable on (a, b) and continuous on $[a, b]$. Then*

- *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. Let $x_1, x_2 \in [a, b]$, $x_1 < x_2$. Since f is differentiable on $(x_1, x_2) \subset (a, b)$ and f is continuous on $[x_1, x_2] \subset [a, b]$, the Mean–Value Theorem gives a $c \in (x_1, x_2)$ with

$$0 < f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Since $x_1 < x_2$ this implies $0 < f(x_2) - f(x_1)$ and f is increasing.
Similar arguments work for $f'(c) = 0$ or $f'(c) < 0$. □

Corollary 3.25. *If f is differentiable on an open interval (a, b) then: $f'(x) = 0$ for all $x \in (a, b)$ if and only if f is constant.*

From this we easily see a theorem which we will use frequently (in particular, for integrals):

Theorem 3.26. *Let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions on an interval (a, b). If $f'(x) = g'(x)$ for all $x \in (a, b)$, then $f(x) = g(x) + C$ for all $x \in (a, b)$ where $C \in \mathbb{R}$ is a constant.*

Proof. Consider $h = f - g$. Then $h'(x) = 0$ for all x, thus h is constant. □

3.6 Local extrema

Definition 3.27. Suppose f is a function on an interval $[a, b]$ and c an interior point of (a, b). Then f is said to have a

local maximum at c if $f(c) \geq f(x)$ for all x sufficiently close to c, that is, if there exists $\delta > 0$ such that $0 < |x - c| < \delta \implies f(c) \geq f(x)$.

local minimum at c if $f(c) \leq f(x)$ for all x sufficiently close to c, that is, if there exists $\delta > 0$ such that $0 < |x - c| < \delta \implies f(c) \leq f(x)$.

Thus, at a local extremum of a differentiable function we expect to have a horizontal tangent line.

Definition 3.28. Let $f : [a, b] \to \mathbb{R}$. If $c \in (a, b)$ and

$$f'(c) = 0 \quad \text{or} \quad f'(c) \text{ does not exist}$$

then c is called a *critical point* of f.

Thus, if a function has a local extremum at a point, then either it is not differentiable at the point or has horizontal tangent line.

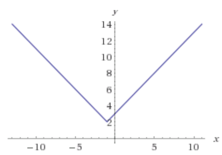
Theorem 3.29. *Let $f : [a, b] \to \mathbb{R}$. If f has a local extremum at $c \in (a, b)$ then c is a critical point of f.*

Proof. If f is not differentiable at c we are done. Consider now the case that $f'(c)$ exists but is not equal to 0. But then $f'(c) > 0$ or $f'(c) < 0$ and in either case there are x_1, x_2 arbitrarily close to c with $f(x_1) < f(c) < f(x_2)$. Contradiction.

$\square$

A function f may not have a local extremum at c in the interior of its domain but still c may be a critical point of f, e.g. $f(x) = x^3$ has a critical point at 0 but not a local maximum or maximum.

Example 3.30. Find the critical points of $f(x) = |x + 1| + 2$.



We can write

$$f(x) = \begin{cases} -x + 1 & x < -1 \\ x + 3 & x \geq -1 \end{cases}.$$

Since f is given by the absolute value function, we know that f is not differentiable at $x = -1$. Thus

$$f(x) = \begin{cases} -1 & x < -1 \\ DNE & x = -1 \\ 1 & x > -1 \end{cases},$$

and $x = -1$ is the only critical point. Indeed, f has a local minimum at $x = -1$.

So far, we found a necessary condition for c to be a local extremum: c has to be a critical point. Can we give sufficient conditions which imply that c is indeed a local extremum?

Theorem 3.31 (First–derivative test)**.** *Suppose c is a critical point of f and f is continuous at c. If there exists $\delta > 0$ such that*

(i) $f'(x) > 0$ for all $x \in (c-\delta, c)$ and $f'(x) < 0$ for all $x \in (c, c+\delta)$ then f has a local maximum at c.

(ii) $f'(x) < 0$ for all $x \in (c-\delta, c)$ and $f'(x) > 0$ for all $x \in (c, c+\delta)$ then f has a local minimum at c.

(iii) If f' keeps constant sign on $(c-\delta, c) \cup (c, c+\delta)$ then $f(c)$ is not a local extreme value.

Proof. Follows from the Increasing/Decreasing Theorem 3.24. ☐

The assumption that f is continuous at c is important. E.g., the function

$$f(x) = \begin{cases} 1 + 2x, & x \leq 1 \\ 5 - x, & x > 1 \end{cases}$$

does not attain at local extremum at $c = 1$ even though 1 is a critical point and the derivative of f changes sign.

The following second test is easier to apply, however, it works in even fewer cases than the First–Derivative test: we need to assume that the second derivative exists.

Theorem 3.32 (Second–Derivative test)**.** *Suppose $f'(c) = 0$ and $f''(c)$ exists.*

(i) If $f''(c) < 0$ then f has a local maximum at c.

(ii) If $f''(c) > 0$ then f has a local minimum at c.

Proof. Assume $f''(c) > 0$ (the other case is similar). Since f'' is the derivative of f' there exists $\delta > 0$ such that $f'(x) < f'(c) = 0$ for all $x \in (c-\delta, c)$ and $f'(x) > f'(c) = 0$ for all $x \in (c, c+\delta)$. Thus, the First-Derivative Test shows that f attains a local minimum at c.

☐

Note that the Second–Derivative test does not give any information if $f'(c) = 0$ and $f''(c) = 0$. Consider $f(x) = x^3$ and $f(x) = x^4$ at $x = 0$.

Example 3.33. The light reflection at a mirror satisfies the Fermat principle of shoterst distance. We consider two points a and b above a mirror whose horizontal distance is l. For which l is

$$L(x) = \sqrt{a^2 + x^2} + \sqrt{b^2 + (l - x)^2}$$

minimal?

If $x \to \pm\infty$ then $L(x) \to \infty$. Hence, there is a local minimum in $(-\infty, \infty)$ since L is differentiable (and thus continuous). We find the critical points by computing

$$L'(x) = \frac{x}{\sqrt{a^2 + x^2}} + \frac{-(l - x)}{\sqrt{b^2 + (l - x)^2}}.$$

Therefore, if $L'(c) = 0$ then c and $l - c$ have the same sign (thus are positive), and

$$\frac{c}{\sqrt{a^2 + c^2}} = \frac{(l - c)}{\sqrt{b^2 + (l - c)^2}}$$

is equivalent to

$$\frac{c^2}{a^2 + c^2} = \frac{(l - c)^2}{b^2 + (l - c)^2}.$$

This shows that $c^2 b^2 + c^2(l-c)^2 = a^2(l-c)^2 + c^2(l-c)^2$ and $bc = a(l-c)$ since $a, b, c, l-c > 0$. Therefore, the only critical point is $c = \frac{al}{b+a}$, and thus L has an absolute minimum at $c = \frac{al}{b+a}$.

In particular, $\frac{l-c}{b} = \frac{l(1-\frac{al}{b+a})}{b} = \frac{l}{b+a} = \frac{c}{a}$. In other words, denoting by α the angle of incidence and by β the angle of reflection, then

$$\cot \alpha = \frac{c}{a} = \frac{l - c}{b} = \cot \beta,$$

so we have shown the *law of reflection*.

3.7 Global extreme values

So far, we have only looked at local extreme values in the interior of the domain of the function. Can we now find the global extreme values? (We insert the word "global" to emphasise that we are looking for the overall maximal/minimal values, and not local extrema).

Recall our previous definition:

Definition 3.34. The function f is said to have an

(global) maximum at $d \in \mathrm{dom}(f)$ if $f(d) \geq f(x)$ for all $x \in \mathrm{dom}(f)$;

(global) minimum at $d \in \mathrm{dom}(f)$ if $f(d) \leq f(x)$ for all $x \in \mathrm{dom}(f)$.

Of course, a function has to be at least bounded to attain both global extreme values but could have a minimum or maximum when it's unbounded (e.g., $f(x) = x^2$ has a minimum even though it is not a bounded function). However, for a continuous function on a bounded closed interval we know that the function is bounded. Thus, we can apply the following strategy:

Strategy to find global extremas of a continuous function f on a closed bounded interval $[a, b]$:

(i) Find critical points of f.

(ii) Compare $f(a), f(b)$ and $f(c)$ for all critical points c of f.

Example 3.35. The function $f(x) = 1 + 4x^2 - \frac{1}{2}x^4$, $x \in [-1, 3]$, has critical points $x = 0, x = \pm 2$. Comparing the values of f at the critical points and the end points $x = -1, x = 3$ we see that f attains a (global) maximum at $x = 2$ and a (global) minimum at $x = 3$.

Common errors in finding the global values are the obvious ones:

- Using the above strategy even though the interval is not closed or not bounded.
- Forgetting to consider the critical points (including points where the derivatives does not exist).
- Forgetting to consider the end points of the interval.

3.8 Concavity and points of inflection

A further question about the general shape of a function is whether it is curving in a particular direction.

Definition 3.36. Let f be a function differentiable on an open interval I. The graph of f is said to be

concave up if f' increases on I, and

concave down if f' decreases on I.

Of course, the points where we change the type of concavity are important!

Definition 3.37. Let f be continuous at c, differentiable near c. Then $(c, f(c))$ is called a *point of inflection* if the graph of f changes its type of concavity at c, that is, if there exists $\delta > 0$ such that graph(f) is concave up (or concave down) on $(c - \delta, c)$ and the opposite concave down (or concave up) on $(c, c + \delta)$.

Since the second derivative describes the local shape of the graph of the first derivative, we can find the concavity type by examining the second derivative:

Theorem 3.38. *Let f be twice differentiable on an open interval I.*

(i) If $f''(x) > 0$ for all $x \in I$ then the graph of f is concave up.

(ii) If $f''(x) < 0$ for all $x \in I$ then the graph of f is concave down.

Proof. Increasing/Decreasing Theorem 3.24. □

Theorem 3.39. *If $(c, f(c))$ is a point of inflection then $f''(c) = 0$ or $f''(c)$ does not exist.*

Proof. First Derivative Test 3.31. □

3.9 Indeterminate forms and the Cauchy–Mean–Value Theorem

A modified version of the Mean–Value theorem can be used to compute indeterminate forms of type "$\frac{0}{0}$", that is, limits of fractions $\frac{f(x)}{g(x)}$ for $x \to c$ where $\lim\limits_{x \to c} f(x) = \lim\limits_{x \to c} g(x) = 0$.

Theorem 3.40 (Cauchy–Mean–Value Theorem)**.** *Let f, g be differentiable on (a, b), continuous on $[a, b]$. If $g'(x) \neq 0$ for all $x \in (a, b)$ then there exists $r \in (a, b)$ with*

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

Why is this a generalisation of the Mean–Value theorem? Consider $g(x) = x$.

We now apply this to functions with zeros at a:

Theorem 3.41 (L'Hôpital's Rule). *Suppose that* $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$ *with* $g(x) \neq 0$ *and* $g'(x) \neq 0$ *for all* x *near* c. *Then if*

$$\lim_{x\to c} \frac{f'(x)}{g'(x)} = L$$

exists, then $\lim_{x\to c} \frac{f(x)}{g(x)}$ *exists and equals* L.

A common error is to apply L'Hôpital's rule even though not both limits of f and g are zero. Moreover, if the limit of the quotient of the derivatives does not exist, the rule does not give any information about the original limit.

3.10 Taylor polynomials

Another important application of the Mean–Value Theorem are the Taylor polynomials. They are used to approximate a function.

Consider first the function $f(x) = (x-c)^n$, $c \in \mathbb{R}, n \in \mathbb{N}$.

Then we can approximate f linearly at c by its tangent

$$f(x) \sim f(c) + f'(c)(x-c) = P_1(x)\,.$$

Indeed, we have $f(c) = P_1(c)$ and $f'(c) = P_1'(c)$. However, the n^{th} derivative of f and P_1 do not coincide at c if $n > 1$, and P_1 is not a very good approximation in this case: $P_1(x) = 0$ is the constant zero function, whereas $f(x)$ is not constant.

Therefore, to approximate a function f we want to use a polynomial P_n of degree n instead of a polynomial of degree 1, and we require that the derivatives up to the n^{th} order of f and P_n agree.

Is there such a polynomial, and how do we find it?

Our example indicates that we want to write our polynomial in the form

$$P_n(x) = \sum_{l=0}^{n} a_l (x-c)^l$$

for some $a_l \in \mathbb{R}$, $l = 0, \ldots, n$. Then the requirement $P_n(c) = f(c)$ gives $a_0 = f(c)$. Furthermore, $P_n'(x) = \sum_{l=1}^{n} l a_l (x-c)^{l-1}$ gives with $P_n'(c) = f(c)$ that $a_1 = f'(c)$.

Using $P_n^{(k)}(x) = \sum_{l=k}^{n} l(l-1)\ldots(l-k+1) a_l (x-c)^{l-k}$ we obtain with $P_n^{(k)}(c) = f^{(k)}(c)$:

$$f^{(k)}(c) = k! a_k\,.$$

Definition 3.42. Let f be a function with n derivatives on an open interval I. Then for $c \in I$ the polynomial

$$P_n(x) = \sum_{l=0}^{n} \frac{f^{(l)}(c)}{l!}(x-c)^l = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

is called the n^{th} *Taylor polynomial* of f in powers of $(x-c)$.

For some reason, when computing the coefficients in the Taylor polynomial, it seems to be tempting to differentiate $f^{(n-1)}(c)$ to obtain $f^{(n)}(c)$: so the coefficients become zero quite quickly but unfortunately not rightfully. Instead you should take all needed higher derivatives first, before evaluating at c.

The concept of a Taylor polynomials was formally introduced by the English mathematician Brook Taylor in 1715. If the Taylor series is centred at $c = 0$, then that series is also called a Maclaurin polynomial, named after the Scottish mathematician Colin Maclaurin, who made extensive use of this special case of Taylor polynomial in the 18th century.

For every function with n derivatives, we can find the corresponding Taylor polynomial of order n. But how good is this approximation of the function f? A (partial) answer is the following Taylor formula. More results about Taylor approximation will be discussed in the second semester of Calculus and Analysis.

Theorem 3.43 (Taylor formula)**.** *If f has $n+1$ derivatives on an open interval I with $c \in I$ then*

$$f(x) = \sum_{l=0}^{n} \frac{f^{(l)}(c)}{l!}(x-c)^l + R_n(x)\,,$$

where the Lagrangian remainder R_n *satisfies*

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}$$

for some ξ between x and c. In particular,

$$\lim_{x \to c} \frac{R_n(x)}{(x-c)^n} = 0\,.$$

Thus, the remainder goes faster to zero then $(x-c)^n$ when x approaches c. However, this does not mean that f is well–approximated for all $x \in \operatorname{dom}(f)$.

Proof. Fix $x \in I$. Put

$$F(t) = \sum_{l=0}^{n} \frac{f^{(l)}(t)}{l!}(x-t)^l \quad \text{and} \quad G(t) = (x-t)^n\,.$$

Note first, that F is not the Taylor polynomial P_n of f: instead of considering the function depending on x, we consider the function for fixed x but varying c.

However, for $t = c$ we obtain $F(c) = P_n(x)$ and

$$F(x) - F(c) = f(x) - P_n(x) = R_n(x)\,.$$

The Cauchy Mean–Value theorem gives ξ between c and x with

$$\frac{F'(\xi)}{G'(\xi)} = \frac{F(x) - F(c)}{G(x) - G(c)} = \frac{R_n(x)}{-(x-c)^{n+1}}\,,$$

where we used that $G(x) = 0$ and $G(c) = (x-c)^{n+1}$. Now the product rule and chain rule give

$$\begin{aligned}
\frac{d}{dt}F(t) &= \sum_{l=0}^{n} \frac{f^{(l+1)}(t)}{l!}(x-t)^l - \sum_{l=0}^{n} \frac{f^{(l)}(t)}{l!} l (x-t)^{l-1} \\
&= \sum_{l=0}^{n} \frac{f^{(l+1)}(t)}{l!}(x-t)^l - \sum_{l=1}^{n} \frac{f^{(l)}(t)}{(l-1)!}(x-t)^{l-1} \\
&= \sum_{l=0}^{n} \frac{f^{(l+1)}(t)}{l!}(x-t)^l - \sum_{l=0}^{n-1} \frac{f^{(l+1)}(t)}{l!}(x-t)^l \\
&= \frac{f^{(n+1)}(t)}{n!}(x-t)^n\,,
\end{aligned}$$

and $\frac{d}{dt}G(t) = -(n+1)(x-t)^n$. Plugging this in, we see

$$R_n(x) = -\frac{F'(\xi)}{G'(\xi)}(x-c)^{n+1} = -\frac{\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n}{-(n+1)(x-\xi)^n}(x-c)^{n+1} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}\,.$$

□

3.11 Study guide

Go through the material of chapter 3, and identify the main topics and statements.

To deepen your understanding, think about the following:

- Draw continuous functions whose tangent line at the point 2 is
  - horizontal,
  - vertical,
  - does not exist, or
  - has slope -10.

  Try in each case to give a formula for the function. Verify that it is continuous, and satisfies the condition.
- Give examples of functions where $(fg)' \neq f'g'$.

- Give examples of functions where $\left(\frac{f}{g}\right)' \neq \frac{f'}{g'}$.
- Find two differentiable functions such that $(f \circ g)'(x) \neq f'(x)g'(x)$. Convince yourself that the correct formula is instead given by the chain rule.
- Discuss the difference between f^{-1} and $\frac{1}{f}$. What is the meaning of "inverse function"?
- Find situations in which Newton's method works/does not work.
- Does $f(x) = \sin(x^3) - 1, x \in \mathbb{R}$, attain the value -1.5? If so, find c with $f(c) = -1.5$ (use Newton's method). Compare your efficiency in solving with your attempt in Chapter 2.
- Find an example of a function which is differentiable on an open interval (a, b) but not continuous on $[a, b]$. Find such a function such that there is no tangent to the graph which is parallel to the secant between $f(a)$ and $f(b)$.
- Is the c in the Mean–Value Theorem unique? Prove your conjecture, or give an example in which it is not unique.
- What is the difference between Rolle's theorem and the Mean–Value Theorem?
- Why do we need the continuity assumption in the Increasing/Decreasing Theorem?
- Draw the graph of a function which is differentiable, has a local maximum at 1 and no minimum. Is the function concave up or down near 1 (or neither nor)? Is it possible that such a function has more than 2 inflection points?
- Find functions for which one cannot use the Second–Derivative test but the First Derivative test shows the existence of a local extremum.
- What is the difference between a local and an global extremum? Are there examples which are both local and global extremums (if so, give one!)
- Find a function which has no global maximum on $(0, 1)$.
- Why did we assume our function to be continuous in the "Finding global values"–strategy? What would you do if the function is not continuous?

Chapter 4

Sequences

We have seen that limits are at the core of calculus and analysis: every single notion is a limit in one sense or another! Examples where limits are used include the slope of the tangent line, the length of a curve, the area under a graph. Moreover, Calculus is at the basis of most modern mathematics. In other words, limits are the building blocks for all your future endeavours in mathematics.

You have seen the notion of a limit of a function, but now we are looking at the simplest example of a function: a sequence. Sequences often appear in mathematics and scientific experiments: when trying to approximate a value, we usually construct a sequence of numbers approaching the value up to a given error. An example for this is Newton's method. In particular, we are interested in infinite sequences and in the questions whether the elements of a sequence approach a limit.

4.1 Definition of a sequence

Here we discuss the proper definition and the notion of a limit.

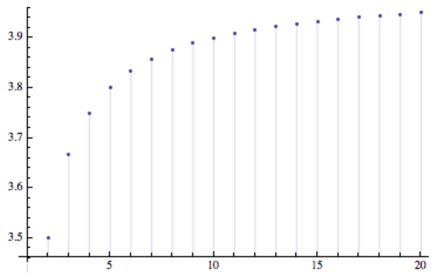
Definition 4.1. A *sequence* of real numbers is a real valued function defined on the natural numbers:

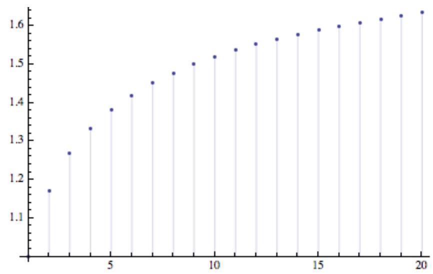
$$a : \mathbb{N} \to \mathbb{R}\,.$$

We write for short $a_n = a(n)$ and denote a sequence by $(a_n)_{n\in\mathbb{N}}$.

We often give sequences as follows:

$$(a_n)_{n\in\mathbb{N}},\; a_n = \frac{4n-1}{n}, \qquad \text{or} \qquad \left(\frac{2\sqrt{n}}{\sqrt{n}+1}\right)_{n\in\mathbb{N}}.$$





From our definition it is obvious that we can obtain new sequences by adding, subtracting, multiplying, dividing the n–th components of two sequences for all $n \in \mathbb{N}$ (after all, sequences are special cases of functions!):

Given two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ we obtain a new sequence $(c_n)_{n\in\mathbb{N}}$ by setting

- $c_n = a_n + b_n$
- $c_n = a_n - b_n$.
- $c_n = a_n \cdot b_n$.
- $c_n = \frac{a_n}{b_n}$, provided $b_n \neq 0$ for all $n \in \mathbb{N}$.

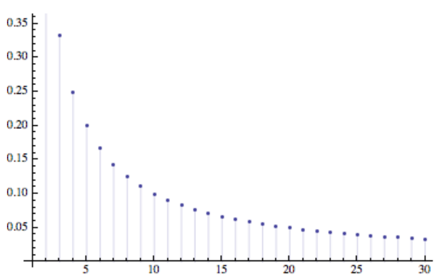
Our interest is to understand the behaviour of the sequence; what can we say about the shape of the graph of the function? Here are some useful notions:

Definition 4.2. A sequence $(a_n)_{n\in\mathbb{N}}$ is called

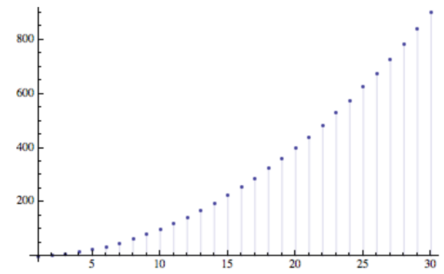
(i) *monotonic sequence* if
  - either $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$ (increasing),
  - or $a_n \geq a_{n+1}$ for all $n \in \mathbb{N}$ (decreasing).

(ii) *bounded sequence* if there exist $m, M \in \mathbb{R}$: $m \leq a_n \leq M$ for all $n \in \mathbb{N}$.

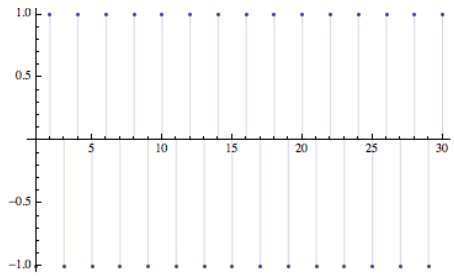
Example 4.3. The sequence $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ is (strictly) decreasing and bounded.



The sequence $(n^2)_{n\in\mathbb{N}}$ is (strictly) increasing but not bounded.



The sequence $((-1)^n)_{n\in\mathbb{N}}$ is bounded but not monotonic.



4.2 Limits of sequences

Our main question is however to investigate how a sequence behaves when n gets very large: does it approach a particular value, does it disappear to infinity or is behaved very irregularly?

For example the sequence

$$(a_n)_{n\in\mathbb{N}}, a_n = \frac{1}{n}$$

seems to be nicely behaved; the larger n gets, the closer we get to 0. So should we say that $(a_n)_{n\in\mathbb{N}}$ converges to a limit L if a_n gets close to L?

This definition is too vague: consider the sequence $(a_n)_{n\in\mathbb{N}}$,

$$a_n = \begin{cases} n, & n \neq 20 \\ 0, & n = 20 \end{cases}.$$

It gets close to 0 since the twentieth element is actually equal to zero – but that's clearly not what we understand as the limit when n growth really large.

So we need to incorporate that the a_n's should get close not once but when we go out to infinity.

How about the sequence $(a_n)_{n\in\mathbb{N}}$

$$a_n = \begin{cases} \frac{1}{n} & n \text{ odd} \\ n & n \text{ even} \end{cases}.$$

When n growth large, the odd elements of the sequence get closer and closer to 0 – but the even ones disappear to infinity! Again, this should not be meant when saying that a number is the limit of a sequence.

What we want is that L is the limit of the sequence $(a_n)_{n\in\mathbb{N}}$ if " a_n is arbitrary close to L for all n large enough".

What we need now is a mathematical way to formulate the above statement. When a_n is close to L then the distance between a_n and L is small. Recall that the *absolute value* measures the distance to 0, so we want that $|a_n - L|$ can be made arbitrarily small by taking n large enough. So given a small number ε we can find a number $K \in \mathbb{N}$ so that whenever n is bigger than K, the distance between a_n and L is smaller than ε, that is, $|a_n - L| < \varepsilon$.

Voilà, that's our definition:

Definition 4.4. Let $(a_n)_{n\in\mathbb{N}}$ be a sequence. We say that a_n has limit L,

$$\lim_{n\to\infty} a_n = L\,,$$

if for all $\varepsilon > 0$ there exists $K \in \mathbb{N}$ such that

$$|a_n - L| < \varepsilon \qquad \text{for all } n \geq K\,.$$

If a sequence has a limit it is said to be a *convergent sequence*, otherwise it is called a *divergent sequence*.

Note that the order of ε and K is crucial: we first choose how close we want to come to L and then have to find how large we have to be with n for this to be true.

Convince yourself that the "beginning" of a sequence has no impact on the convergence of the sequence. What is important is what happens for large n: "All's Well That Ends Well". In particular, one cannot investigate the convergence of a sequence by computing finitely many elements with a computer. In other words, if you compute a few values of a sequence and then only guess the limit you will get zero marks; a proper argument is needed.

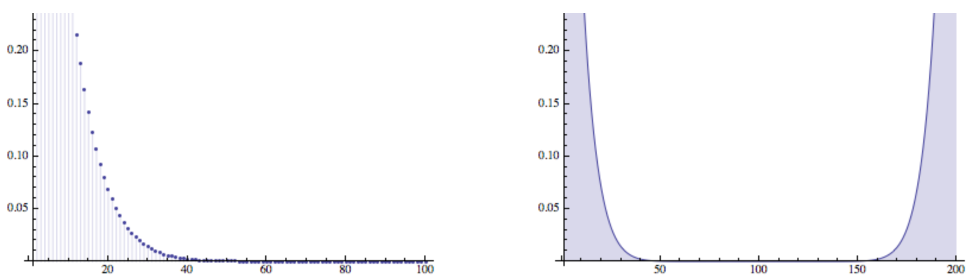


Figure 4.1: looks like it converges to 0 after 100 evaluations. After 200 the picture tells a different story!

Example 4.5. The sequence $(a_n)_{n\in\mathbb{N}}$, $a_n = \frac{4n-1}{n}$ converges to $L = 4$.

"Find K" Let us do some scratch work first to understand what we need to find.

We want to find $K \in \mathbb{N}$ such that $|\frac{4n-1}{n} - 4| < \varepsilon$ whenever $n \geq K$. Simplifying we get $|\frac{4n-1}{n} - 4| = \frac{|4n-1-4n|}{n} = \frac{1}{n}$. Thus, we take any $K \in \mathbb{N}$ with $\frac{1}{K} < \varepsilon$. (Note: we cannot say $\frac{1}{K} = \varepsilon$ because ε might not be rational).

"K works" Verify that our candidate for K indeed satisfies the conditions. This is the actual argument.

Let $\varepsilon > 0$ and let $K \in \mathbb{N}$ such that $\frac{1}{K} < \varepsilon$. Then for all $n \geq K$ we have

$$|\frac{4n-1}{n} - 4| = \frac{|4n-1-4n|}{n} = \frac{1}{n} \leq \frac{1}{K} < \varepsilon$$

since $n \geq K$ implies $\frac{1}{n} \leq \frac{1}{K}$. This shows that $\lim_{n\to\infty} \dfrac{4n-1}{n} = 4$.

If you are unfamiliar with the manipulations of inequalities, see the appendix in this Chapter.

Similarly, we show that the sequence $\left(\frac{2\sqrt{n}}{\sqrt{n}+1}\right)_{n\in\mathbb{N}}$ converges to 2.

The standard method to use has two parts:

" Find K" scratch work to find K.

"K works" verify that your candidate for K indeed satisfies the conditions.

Common errors:

(i) Argument only contains the "find" part. (This will result in major reduction in marks).

(ii) Argument does not contain the "find" part. (Usually, this will result in major reduction in marks unless it is kind of obvious how to find the K. If in doubt, show the work).

(iii) ε is not introduced or appears only after the K. (This will result in a zero mark: you can not find K if the ε is not given in advance).

Just to clarify this rather abstract question: can there be a sequence converging to two different limits? The answer is no:

Theorem 4.6. *If $(a_n)_{n\in\mathbb{N}}$ is a sequence with $\lim_{n\to\infty} a_n = L$ and $\lim_{n\to\infty} a_n = M$ then $L = M$.*

Proof. Theorem 11.3.2 in the supplement to section 11.3. $\square$

Theorem 4.7. *Every convergent sequence is bounded.*

The converse is false, there are bounded sequences which are divergent. For example $a_n = (-1)^n$.

However, the theorem implies that if a sequence is unbounded, then it is divergent.

Geometrically, we would expect a monotonic sequence to be convergent if it is bounded: it cannot disappear to ∞ but also cannot jump around between different values. More precisely:

Theorem 4.8. *If a sequence is bounded and monotonic then it is convergent.*

It is not so easy to *prove* a limit exists: it we will need the Least Upper Bound axiom for the real numbers that we saw in chapter 1.

4.3 Theorems for limits of sequences

Theorem 4.9. *If a sequence is bounded and monotonic then it is convergent.*

Of course, we don't want to use the ε–K criterion every single time. So, we try to find some general rules which allow to derive a limit, once we know simpler ones.

Limit laws If $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are convergent sequences then $(a_n \pm b_n)_{n\in\mathbb{N}}$, $(a_n b_n)_{n\in\mathbb{N}}$ and, provided $b_n \neq 0 \forall n$, $(\frac{a_n}{b_n})_{n\in\mathbb{N}}$ are convergent, and

$$\lim_{n\to\infty}(a_n \pm b_n) = \lim_{n\to\infty} a_n \pm \lim_{n\to\infty} b_n, \quad \lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n), \quad \lim_{n\to\infty}\frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}.$$

Note that we also assume $\lim_{n\to\infty} b_n \neq 0$.

For the limit laws it is important that the individual limits exist. For example, for $a_n = \frac{1}{n}, b_n = n$

$$\lim_{n\to\infty} a_n b_n = \lim_{n\to\infty} \frac{1}{n} n = \lim_{n\to\infty} 1 = 1$$

whereas $\lim_{n\to\infty} a_n = 0$ and you might have thought that $\lim_{n\to\infty} a_n b_n = 0$. But $\lim_{n\to\infty} b_n$ does not exist so therefore, you cannot use the limit laws to compute the limit. In fact, in this case

$$\lim_{n\to\infty} a_n b_n \neq (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n)\,.$$

A similar example is given by $a_n = \frac{1}{n}, b_n = (-1)^n$. You might be tempted to write

$$\lim_{n\to\infty} a_n b_n = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n) = 0$$

since $\lim_{n\to\infty} a_n = 0$. However, you are not allowed to use the limit laws, since $(b_n)_{n\in\mathbb{N}}$ is not convergent. You need to use a more refined argument to find whether the limit of $(a_n b_n)_{n\in\mathbb{N}}$ exists.

Thus, if you use the limit laws you have to state why you are allowed to so (mark reduction otherwise).

By induction we conclude from the Limit Laws:

Corollary 4.10. *Let $(a_n^{(k)})_{n\in\mathbb{N}}$ be convergent sequences of real numbers for $k = 1, \ldots, r$. Then*

$$\lim_{n\to\infty}\left(\sum_{l=1}^{r} \alpha_l a_n^{(l)}\right) = \sum_{l=1}^{r} \alpha_l (\lim_{n\to\infty} a_n^{(l)})$$

and

$$\lim_{n\to\infty}\left(\prod_{l=1}^{r} a_n^{(l)}\right) = \prod_{l=1}^{r} (\lim_{n\to\infty} a_n^{(l)})$$

Assume we have a situation where we know the behaviour of two sequences and have a third sequence lying in between like in the following picture:

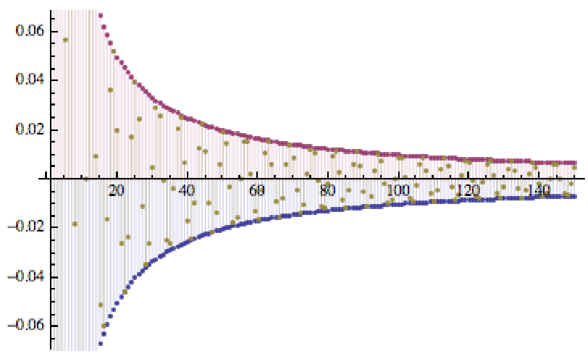


Figure 4.2: If the limits of the blue and red sequence are known, can we conclude the convergence of the green sequence?

Theorem 4.11 (Pinching theorem for sequences). *Let $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}, (c_n)_{n\in\mathbb{N}}$ be sequences such that there exists $N \in \mathbb{N}$ with*

$$a_n \leq b_n \leq c_n \qquad \textit{for all } n \geq N\,.$$

If $(a_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ are convergent with $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = L$ then $(b_n)_{n\in\mathbb{N}}$ is convergent with $\lim\limits_{n\to\infty} b_n = L$.

Proof. Let $\varepsilon > 0$. Since $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = L$ there exists $K_1 \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for all $n \geq K_1$ and $K_2 \in \mathbb{N}$ such that $|c_n - L| < \varepsilon$ for all $n \geq K_2$. Let $K = \max\{N, K_1, K_2\}$ then we have

$$L - \varepsilon < a_n < L + \varepsilon, \quad L - \varepsilon < c_n < L + \varepsilon, \quad \text{ and } a_n \leq b_n \leq c_n$$

for all $n \geq K$. Thus, combining all these inequalities we have

$$L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon$$

which shows that $L - \varepsilon < b_n < L + \varepsilon$ and thus $|b_n - L| < \varepsilon$ for all $n \geq K$. This shows that $\lim\limits_{n\to\infty} b_n = L$.

□

In particular, if

$|b_n| \leq c_n$ and $c_n \to 0$ then $b_n \to 0$.

4.4 Cauchy sequences and the Bolzano–Weierstrass theorem

We have seen that every convergent sequence is bounded, but the converse is not true. However, if we only consider selected (but infinitely many) elements one can always obtain a convergent subsequence. For example, $(-1)^n$ not convergent, but it is bounded, and contains two obvious convergent (constant!) subsequences.

Definition 4.12. Let $(a_n)_{n\in\mathbb{N}}$ a sequence. A subsequence of $(a_n)_{n\in\mathbb{N}}$ is a sequence $(b_n)_{n\in\mathbb{N}}$ such that for all $k \in \mathbb{N}$ there exists a $n_k \in \mathbb{N}$, $n_1 < n_2 < \ldots$, with $b_k = a_{n_k}$.

To understand better the relation between convergent subsequences and convergent sequences, we consider the notion of Cauchy sequences. Notice that if a sequence $(a_n)_{n\in\mathbb{N}}$ is convergent, then all elements a_n are close to the limit for large enough n. But then the distance between two elements also gets small. Sequences with this property are important in mathematics.

Definition 4.13. A sequence $(a_n)_{n\in\mathbb{N}}$ is called a *Cauchy sequence* if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|a_n - a_m| < \varepsilon \quad \forall m, n \geq N\,.$$

This definition does not mention any limits, so we are not just repeating the definition of convergent sequence.

However, We can prove all usual properties of convergent sequences for Cauchy sequences too. For example:

Theorem 4.14. *Every Cauchy sequence is bounded.*

Proof. If $(a_n)_{n\in\mathbb{N}}$ is Cauchy then (for $\varepsilon = 1$) there exists some $N \in \mathbb{N}$ such that

$$|a_n - a_m| < 1 \quad \forall n, m \geq N\,.$$

In particular

$$|a_n - a_N| < 1, \text{ and so } |a_n| < 1 + |a_N|, \quad \forall n \geq N\,.$$

So $M = \max\{|a_1|, \ldots, |a_{N-1}|, |a_N| + 1\}$ is a bound for $(a_n)_{n\in\mathbb{N}}$. □

Indeed:

Theorem 4.15. *Every convergent sequence is a Cauchy–sequence.*

Proof. Let $\varepsilon > 0$ and $L = \lim_{n\to\infty} a_n$. Then there exists $N \in \mathbb{N}$ such that

$$|a_n - L| < \frac{\varepsilon}{2}, \qquad \text{for all } n \geq N\,.$$

Therefore,

$$|a_n - a_m| = |a_n - L + L - a_m| \leq |a_n - L| + |L - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \text{for all } n, m \geq N\,.$$

This shows that $(a_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. □

In fact converse is true as well! We will see this in a little while.

Theorem 4.16 (Bolzano–Weierstrass theorem)**.** *Let $(a_n)_{n\in\mathbb{N}}$ be a bounded sequence. Then there exists a convergent subsequence $(a_{n_k})_{k\in\mathbb{N}}$ of $(a_n)_{n\in\mathbb{N}}$.*

Proof. The proof is a constructive one, that is, it not only gives the existence of this subsequence but also shows how a bisection method can be used to find it.

First note that, since $(a_n)_{n\in\mathbb{N}}$ is bounded there exists $M > 0$ with $|a_n| \leq M$ for all $n \in \mathbb{N}$. Next we show that there exists, for all $n \in \mathbb{N}$, a closed interval $I_n = [\alpha_n, \beta_n]$ such that

- the length $|I_n| = \beta_n - \alpha_n = \frac{1}{2^{n-1}}M$
- $I_n = [\alpha_n, \beta_n]$ contains infinitely many terms of the sequence $(a_n)_{n\in\mathbb{N}}$
- α_n is an increasing sequence
- β_n is a decreasing sequence

We prove this by induction:

Base case: Let $I_0 = [\alpha_0, \beta_0] = [-M, M]$, of length $2M = \frac{1}{2^{-1}}M$.

Inductive step: If we know the result is true for some $n = k$, we can divide I_k into two equal parts

$$[\alpha_k, \tfrac{1}{2}(\alpha_k + \beta_k] \text{ and } [\tfrac{1}{2}(\alpha_k + \beta_k, \beta_k]$$

One of these will contain infinitely terms of the sequence, call it $I_{k+1} = [\alpha_{k+1}, \beta_{k+1}]$. In either case, as $[\alpha_{k+1}, \beta_{k+1}] \subset [\alpha_k, \beta_k]$, we have $\alpha_k \leq \alpha_{k+1} < \beta_{k+1} \leq \beta_k$. The length is $|I_{k+1}| = \frac{1}{2}|I_k| = \frac{1}{2}\frac{1}{2^{k-1}}M = \frac{1}{2^k}M$.

Thus, the result for $n = k$ implies the result for $n = k + 1$.

From these bisections of intervals we can construct a subsequence. Since each I_k has infinitely terms of the sequence, just pick one from each, $a_{n_k} \in I_k$, making sure that $n_0 < n_1 < n_2 < \ldots$.

To see this subseqence converges, we note that

- the sequences (α_k) and (β_k) converge, as they are bounded and monotonic.
- they converge to the same limit as $\beta_k - \alpha_k \to 0$.
- a_{n_k} also converges to the same limit, by the Pinching theorem, as $\alpha_k \leq a_{n_k} \leq \beta_k$.

□

Using the Bolzano–Weierstrass theorem we can prove that indeed all Cauchy–sequences are convergent.

Theorem 4.17. *Every Cauchy sequence is convergent.*

Proof. Let $(a_n)_{n\in\mathbb{N}}$ be a Cauchy sequence. We have seen that that $(a_n)_{n\in\mathbb{N}}$ is bounded. By Bolzano–Weierstrass we know also that it has a subsequence $(a_{n_k})_{k\in\mathbb{N}}$ converging to some limit $a \in \mathbb{R}$. Using the triangle inequality, we will prove that the Cauchy sequence $(a_n)_{n\in\mathbb{N}}$ also converges, to the same limit a.

Let $\varepsilon > 0$. Since $(a_n)_{n\in\mathbb{N}}$ is a Cauchy sequence there is $N \in \mathbb{N}$ with

$$|a_n - a_m| < \frac{\varepsilon}{2} \quad \forall n, m \geq N\,.$$

Since $(a_{n_k})_{k\in\mathbb{N}}$ converges to a there exists $k \in \mathbb{N}, k \geq N$ with

$$|a_{n_k} - a| < \frac{\varepsilon}{2}\,.$$

Since $n_k \geq k \geq N$ we have for $n \geq N$

$$|a_n - a_{n_k}| < \frac{\varepsilon}{2}\,.$$

But then

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \varepsilon \quad \forall n \geq N\,.$$

□

Chapter 5

Integration

Integration is based on a very old problem: can we find the area under a graph? The first documented systematic technique capable of determining areas is the method of exhaustion of the ancient Greek astronomer Eudoxus (ca. 370 BC), which sought to find areas by breaking them up into an infinite number of shapes for which the area or volume was known. This method was further developed and employed by Archimedes in the 3rd century BC and used to calculate areas for parabolas and an approximation to the area of a circle. We follow a similar approach.

5.1 The definite integral

First, we formulate our question more precisely: if f is a continuous, positive function on $[a, b]$ what is the area A under the graph of f?

It is also possible to find the area defined by the graph of a function that is not positive and not continuous. Even if the function is not continuous we will usually assume that the function is bounded, but even this is not really necessary as we will see when we consider *improper* integrals.

The obvious approach is to split the considered interval into subintervals: Let $x_0 = a < x_1 < \ldots < x_n = b$ and consider the intervals $[x_0, x_1]$, $[x_1, x_2]$, $\ldots$, $[x_{n-1}, x_n]$. The set

$$P = \{x_0, x_1, \ldots, x_n\}$$

is then called a *partition* of $[a, b]$. We will write Δx_i for the *width* of the subinterval $[x_{i-1}, x_i]$,

$$\Delta x_i \quad = \quad x_i - x_{i-1}$$

On each of these intervals, the area under the graph is approximated by the area of a rectangle whose height is given by the value of the function at some point. Adding these rectangle areas, the sum should be an approximation of the area under the graph.

To show that this approximation works, we choose special values for the height:

$$m_i = \text{glb}\{f(x) : x \in [x_{i-1}, x_i]\}$$

and

$$M_i = \text{lub}\{f(x) : x \in [x_{i-1}, x_i]\}$$

respectively. The union of the rectangles r_i and R_i given by the heights m_i and M_i have area smaller respectively bigger than the area A under the graph.

Thus, denoting by

$$L_f(P) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = \sum_{i=1}^{n} m_i \, \Delta x_i$$

the *lower sum* of f and by

$$U_f(P) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = \sum_{i=1}^{n} M_i \, \Delta x_i$$

the *upper sum* of f we have

$$L_f(P) \leq A \leq U_f(P) \, .$$

Example 5.1. If a mass point is in free fall, the speed is proportional to the elapsed time, that is, $v(t) = gt$ for some $g > 0$. To compute the travelled distance, we use upper and lower sums for the partition $P = \{0, \frac{T}{n}, \dots, \frac{T(n-1)}{n}, T\}$: with $m_i = \frac{g(i-1)T}{n}$ and $M_i = \frac{giT}{n}$ we obtain

$$L_P(v) = \sum_{i=1}^{n} \frac{g(i-1)T}{n} \left(\frac{T}{n}\right), \quad L_P(v) = \sum_{i=1}^{n} \frac{giT}{n} \left(\frac{T}{n}\right) \, .$$

We can compute these sums further

$$L_P(v) = \frac{gT^2}{n^2} \sum_{i=1}^{n} (i-1) = \frac{gT^2}{n^2} \frac{n(n-1)}{2} = \frac{gT^2}{2}(1 - \frac{1}{n})$$

and

$$U_P(v) = \frac{gT^2}{2}(1 + \frac{1}{n}) \, .$$

In particular, we see that

$$L_p(v) \leq \frac{gT^2}{2} \leq U_p(v) \, .$$

For large n we see that $\frac{1}{n}$ gets close to zero, and we expect the fall distance in time T to be $\frac{gT^2}{2}$.

Considering the above example: do we know that $L_f(P), U_f(P)$ converge to the same value?

Note that if we replace a partition P by a *finer* partition, the upper sum will get smaller, and the lower sum will get larger. Given any bounded function $f : [a,b] \to \mathbb{R}$ and two partitions P and Q we can use the partition $P \cup Q$ to see that

$$L_f(P) \leq L_f(P \cup Q) \leq U_f(P \cup Q) \leq U_f(Q) \, .$$

Therefore: the set of all possible lower sums

$$\{L_f(P) : P \text{ is partition of } [a,b]\}$$

is bounded above, and so it has a least upper bound L_f, called the *lower integral* of f on $[a,b]$.

Similarly: the set of all possible upper sums

$$\{U_f(P) : P \text{ is partition of } [a,b]\}$$

is bounded below, and so it has a greatest lower bound U_f, called the *upper integral* of f on $[a,b]$.

To summarise: for any bounded function $f : [a,b] \to \mathbb{R}$ we have

$$L_f(P) \le L_f \le U_f \le U_f(P)$$

for any partition P.

A bounded function f on a bounded closed interval is called *Riemann integrable* or just *integrable* if its upper and lower integrals agree.

Theorem 5.2. *Any a continuous function f on $[a,b]$ is integrable. There exists a unique $I \in \mathbb{R}$ with*

$$L_f(P) \le I \le U_f(P)$$

for all partitions P of $[a,b]$.

Proof. We claim that $U_f = L_f$ and thus $I = U_f$.

Since f is continuous on $[a,b]$ it is uniformly continuous, and for $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$ whenever $|x - y| < \delta$. Let P be a partition of $[a,b]$ with $\max\{|x_i - x_{i-1}|\} < \delta$. Then

$$U_f(P) - L_f(P) = \sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) < \frac{\varepsilon}{b-a}\sum_{i=1}^{n}(x_i - x_{i-1}) = \varepsilon\,.$$

But then $0 \le U - L \le U_f(P) - L_f(P) < \varepsilon$. Since ε was arbitrary, this shows that $U = L$. □

Definition 5.3. Let f be an integrable function on $[a,b]$. The unique number I with $L_f(P) \le I \le U_f(P)$ for all partitions P of $[a,b]$ is called the *definite integral* of f from a to b and is denoted by

$$I = \int_a^b f(x)dx\,.$$

In other words, we can use upper and lower sums to approximate the area under the graph of a function. We return to the question how to compute the integral numerically at the end of the section.

The symbol "$\int$" goes back to Leibniz: it is an elongated S for the Latin word "summa" (=sum).

Remark 5.4. *Note that the letter x in the integrand is a "dummy variable"; we could also write $I = \int_a^b f(t)dt$.*

Corollary 5.5. *If f is non–negative and integrable on $[a, b]$, the the area A below the graph of f from a to b is given by*

$$A = \int_a^b f(x)dx\,.$$

Theorem 5.6 (Riemann's integrability criterion)**.** *A bounded function $f : [a, b] \to \mathbb{R}$ is integrable if and only if for all $\varepsilon > 0$ there exists a partition P of $[a, b]$ such that $U_f(P) - L_f(P) < \varepsilon$.*

Remark 5.7. *The integrability theorem says we need to find partitions to make the difference between the upper and lower integrals as small as we like,*

$$U_f(P) - L_f(P) \;=\; \sum_{i=0}^{n} M_i\Delta x_i - \sum_{i=0}^{n} m_i\Delta x_i \;=\; \sum_{i=0}^{n}(M_i - m_i)\Delta x_i \;<\; \varepsilon\,.$$

One useful technique for proving integrability is to consider partitions P_n of the interval $[a, b]$ into n equal parts, so that we have two bounded monotonic sequences $(L_f(P_n))_{n\geq 1}$ and $(U_f(P_n))_{n\geq 1}$. Integrability means these sequences have the same limit $L = U$. Since all the widths $\Delta x_i = \Delta x = (b - a)/n$, we have

$$U_f(P_n) - L_f(P_n) \;=\; \frac{b-a}{n}\sum_{i=0}^{n}(M_i - m_i).$$

A good application of taking partitions with all widths equal is the proof that any monotonic bounded function $f : [a, b] \to \mathbb{R}$ is integrable. *For example, if f is monotonic decreasing then the maximum value on each subinterval is the value at the beginning of the subinterval, $M_i = f(x_i)$, and similarly $m_i = f(x_{i+1})$. Therefore in the difference between the upper and lower sums we get*

$$U_f(P_n) - L_f(P_n) \;=\; \frac{b-a}{n}\sum_{i=0}^{n}(M_i - m_i) \;=\; \frac{b-a}{n}\sum_{i=0}^{n}(f(x_i) - f(x_{i+1}))$$

Almost all of the terms here cancel, leaving

$$\frac{b-a}{n}\sum_{i=0}^{n}(f(x_i) - f(x_{i+1})) \;=\; \frac{b-a}{n}\left(f(x_0) - f(x_n)\right) \;=\; \frac{(b-a)(f(a) - f(b)}{n}\,.$$

The limit of this sequence as $n \to \infty$ is zero, so f is integrable.

While Newton and Leibniz provided in the 17th century a systematic approach to integration, their work lacked a degree of rigour. Integration was first rigorously formalised, using limits, by Riemann.

Theorem 5.8 (Additivity). *If f is integrable on $[a,b]$, $c \in (a,b)$, then*

$$\int_a^c f(t)dt + \int_c^b f(t)dt = \int_a^b f(t)dt\,.$$

Proof. Here is a sketch of the proof: Take any partition P, without loss of generality we can assume that $c \in P$ (otherwise consider the partition $P \cup \{c\}$). Divide P into two parts P_1, P_2 to obtain partitions of $[a,c]$, $[c,b]$. Then the upper and lower sums of the partitions squeeze in $\int_a^c f(x)dx$ and $\int_c^b f(x)dx$ respectively; and since $U_f(P) = U_f(P_1) + U_f(P_2)$ and $L_f(P) = L_f(P_1) + L_f(P_2)$ we obtain the result. □

Remark 5.9. *We define for an integrable function f on $[a,b]$*

$$\int_b^a f(x)dx = -\int_a^b f(x)dx\,.$$

Then $\int_a^a f(x)dx = 0$, and the above additivity property holds for every a, b, c.

5.2 The Fundamental Theorem of Calculus

The major advance in integration came in the 17th century with the independent discovery of the fundamental theorem of calculus by Newton and Leibniz. The theorem demonstrates a connection between integration and differentiation. This connection, combined with the comparative ease of differentiation, can be exploited to calculate integrals.

If $f : [a,b] \to \mathbb{R}$ is bounded, $|f(x)| \leq B$, and integrable then we can prove that the integral

$$F(x) = \int_{x_0}^x f(t)dt$$

for $x_0 \in [a,b]$, is a uniformly continuous function $F : [a,b] \to \mathbb{R}$.

Given $\varepsilon > 0$ then we see that if $|x_2 - x_1| < \delta = \varepsilon/B$ then, by additivity,

$$|F(x_2) - F(x_1)| \;=\; \left|\int_{x_1}^{x_2} f(t)dt\right| \;\leq\; \int_{x_1}^{x_2} |f(t)|dt \;\leq\; |x_2 - x_1|\,B \;<\; \varepsilon.$$

Now if f is also continuous, we can prove that its integral F is differentiable:

Theorem 5.10. *Let f be continuous on $[a,b]$ and $x_0 \in [a,b]$. Then the function*

$$F(x) = \int_{x_0}^x f(t)dt$$

is differentiable on (a,b) with derivative

$$F'(x) = f(x) \quad \forall x \in (a,b)\,.$$

Proof. Observe that to prove $F'(c) = f(c)$ for $c \in (a, b)$ we can look at the definition of $F'(x)$ as a limit of a difference quotient, and prove

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| \to 0 \qquad \text{as } x \to c.$$

We can rewrite this so we can use the fact that f is continuous at $x = c$:

$$\frac{F(x) - F(c)}{x - c} - f(c) = \frac{\int_c^x (f(t) - f(c))dt}{x - c}$$

Let $\varepsilon > 0$. By continuity of f there exists $\delta > 0$ such that $|t - c| < \delta$ implies $|f(t) - f(c)| < \varepsilon$. This completes the proof. We have the limit we want because if $|x - c| < \delta$ then

$$\left| \frac{\int_c^x (f(t) - f(c))dt}{x - c} \right| \le \frac{\int_c^x |f(t) - f(c)|dt}{|x - c|} \le \frac{\varepsilon \cdot |x - c|}{|x - c|} = \varepsilon.$$

□

This gives rise to the following definition:

Definition 5.11. Let f be continuous on $[a, b]$. A function F is called an *antiderivative* for f on $[a, b]$ if

(i) F is continuous on $[a, b]$, and

(ii) $F'(x) = f(x)$ for all $x \in (a, b)$.

Antiderivatives are not quite unique. But if two functions have the same derivative then their difference has derivative zero, which means the difference is a constant! The previous theorem gave one possible antiderivative for any continuous function f on $[a, b]$, so we know that any antiderivative F will satisfy $F(x) - \int_c^x f(t)dt = \text{constant}$. Taking $x = c$ tells us the constant must be $F(c)$, so

$$\int_c^x f(x)dx = F(x) - F(c).$$

In fact, this is true for all integrable functions, not just continuous functions.

Theorem 5.12 (Fundamental Theorem of Calculus)**.** *If f is integrable on $[a, b]$ then If F is any antiderivative for f on $[a, b]$ then*

$$\int_a^b f(t)dt = F(b) - F(a)\,.$$

As before, we collect tools to compute integrals from known ones.

Lemma 5.13 (Linearity). *For f, g continuous on $[a, b]$, $\alpha \in \mathbb{R}$:*

$$\int_a^b \alpha f(t)dt = \alpha \int_a^b f(t)dt$$

and

$$\int_a^b (f(t) + g(t))dt = \int_a^b f(t)dt + \int_a^b g(t)dt\,.$$

Proof. Homework, see (5.4.3)–(5.4.4). □

We often write

$$\int f(x)dx = F(x) + C$$

to indicate that F is an antiderivative of f. This is meant by the notion *indefinite integral*. In this case, C is called the *constant of integration*.

An object moves along a coordinate line with velocity $v(t) = 2 - 3t + t^2$ units per second. Its initial position (position at time $t = 0$) is 2 units to the right of the origin. Find the position of the object 4 seconds later.

Common errors:

- A tempting mistake is to forget the "dx" under the integration symbol (slight mark reduction). In most cases, you might reasonable argue that it is obvious which variable you are integrating; however, occasionally, the integrand will depend on a constant, and therefore it is important to know which variable is used. So, better get used to using the proper notation.
- Another frequent error is to write $F(x) = \int_0^x f(x)dx$, that is, to use the same symbol for the limit and the integration variable. This may lead to serious confusion in complicated situations so try to avoid it.
- A very common mistake is to forget the constant of integration for the indefinite integral. Depending on the problem, this might lead to heavy mark reduction.

Example 5.14. Let $c(t) = (x(t), y(t))$ be a planar curve where $x, y : I \to \mathbb{R}$ are differentiable functions.
The arclength of c from a to b ($a, b \in I$) is given by

$$L_a^b(c) = \int_a^b \sqrt{(x'(t))^2 + (y'(t))^2}dt\,.$$

This can be seen by approximating the length of the arc by the length of inscribed polygons.

In particular, the circumference of an ellipse $c(t) = (a \sin t, b \cos t)$ with semi–axes $a > b > 0$ is given by $4aE(e)$ where $e = \sqrt{\frac{a^2-b^2}{a^2}}$ is the eccentricity of the ellipse and $E(k)$ is the (complete) *elliptic integral of second kind*

$$E(k) = \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \sin^2 x}\, dx\,.$$

For $k \neq 1$ this integral cannot be expressed in elementary functions.

5.3 The Substitution Law

We discuss a method to compute integrals. Note that integration is less algorithmic then differentiation: not all integrals can be computed explicitly as we have seen in the example of the circumference of an ellipse.

Consider $F(x) = (x^2 - 1)^5$. Then $F'(x) = 5(x^2 - 1)^4(2x)$ by the chain rule: writing $F(x) = G(u(x))$ with $u(x) = x^2 - 1$ and $G(u) = u^5$. Thus,

$$\int (x^2-1)^4 x dx = \frac{1}{10}\int G'(u(x))u'(x)dx = \frac{1}{10}\int \frac{d}{dx}(G \circ u)(x)dx = \frac{1}{10}(x^2-1)^5 + C\,.$$

This example can be generalised:

Theorem 5.15 (Substitution Law)**.** *If f is continuous with antiderivative F, then*

$$\int f(u(x))u'(x)dx = F(u(x)) + C$$

for all functions u with $\operatorname{range}(u) \subset \operatorname{dom}(f)$ *and continuous derivative u'.*

To compute the definite integral we have to take care of the limits of the integral:

Theorem 5.16 (Change of Variable Formula)**.** *If f is continuous on $[a, b]$ with antiderivative F, then*

$$\int_a^b f(u(x))u'(x)dx = \int_{u(a)}^{u(b)} f(u)du\,.$$

for all functions u with $\operatorname{range}(u) \subset [a, b]$ *and continuous derivative u'.*

Example 5.17. The energy of a harmonic oscillation $f(t) = A \cos t$ is given by the integral

$$E(f) = \int_0^{2\pi} A^2 \cos^2(t)dt$$

We compute by change of variable $x(t) = \frac{\pi}{2} - t, x'(t) = -1$

$$\int_0^{\frac{\pi}{2}} \cos^2 x dx = \int_{\frac{\pi}{2}}^0 \cos^2(\frac{\pi}{2} - t)(-1)dt = -\int_{\frac{\pi}{2}}^0 \sin^2(t)dt = -\int_{\frac{\pi}{2}}^0 1 - \cos^2(t)dt$$

where we used that $\cos(\frac{\pi}{2} - t) = \sin(t)$. Therefore,

$$2\int_0^{\frac{\pi}{2}} \cos^2 x dx = \int_0^{\frac{\pi}{2}} 1 dt = \frac{\pi}{2}$$

and the energy is given by $E(f) = A^2\pi$.

Example 5.18. We can compute $\int_0^\pi \sin x dx = -\cos x|_0^\pi = 2$. But also by substituting $\sin x = u(x)$:

$$\int_0^\pi \sin x dx = \int_{u(0)=0}^{u(\pi)=0} \frac{1}{\sqrt{1-u^2}} du = 0\,.$$

What is wrong with our computation? Have a look at what happens at $x = \frac{\pi}{2}$.

Common errors:

- Forgetting the derivative of u.
- Forgetting to change the limits in the definite integral.
- Forgetting to check if the integrands are continuous.

5.4 Properties of the definite integral

We will list a few very useful facts on the definite integral.

Let f, g be continuous on $[a, b]$. Then

(i) If $f(x) \geq 0$ for all $x \in [a, b]$ then $\int_a^b f(x)dx \geq 0$.

(ii) If $f(x) > 0$ for all $x \in [a, b]$ then $\int_a^b f(x)dx > 0$.

(iii) If $f(x) \leq g(x)$ for all $x \in [a, b]$ then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$ (*monotony of integral*).

(iv) $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$.

(v) If $m = min\{f(x) : x \in [a, b]\}, M = max\{f(x) : x \in [a, b]\}$ then $m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$.

(vi) $\frac{d}{dx}\int_a^{u(x)} f(t)dt = f(u(x))u'(x)$ if u is differentiable function.

(vii) If f is odd then $\int_{-a}^a f(x)dx = 0$.

(viii) If f is even then $\int_{-a}^a f(x)dx = 2\int_0^a f(x)dx$.

A nice property is also given by the following theorem:

Theorem 5.19 (Mean–Value Theorem for Integrals)**.** *If f is continuous on $[a, b]$ then there exists $c \in (a, b)$ such that*

$$\int_a^b f(x)dx = f(c)(b-a)\,.$$

Proof. Homework: use the Mean–Value Theorem . ☐

5.5 Aside: transcendental functions

Algebraic numbers are real numbers which satisfy polynomial equations with integer coefficients, e.g. $\sqrt{2}$ satisfies $x^2 - 2 = 0$ and is algebraic. Real numbers which are not algebraic, e.g., π, are called *transcendental numbers*.

Similarly, functions can satisfy polynomial equations with polynomial coefficients, e.g. $f(x) = 2\sqrt{x} - 3x^2$ satisfies $f^2(x) + 6x^2 f(x) + (9x^4 - 4x) = 0$. Such a functions is called an *algebraic function*. A function which is not algebraic is called *transcendental function*. In this section we will discuss some examples (however, we will not show why they are transcendental).

The Logarithm

We have $x^n = \frac{d}{dx}\left(\frac{x^{n+1}}{n+1}\right)$ as long as $n \neq -1$. For $n = -1$ we can define

$$L(x) = \int_1^x \frac{1}{t}dt, \quad x > 0\,,$$

($\frac{1}{t}$ is continuous for $x > 0$). What are the properties of this function?

Since $L'(x) = \frac{1}{x} > 0$ for $x > 0$ we know that L increases. Moreover $L(1) = 0$ so that $L(x) < 0$ for $x < 1$ and $L(x) > 0$ for $x > 1$.

Theorem 5.20. *For all $a, b \in (0, \infty)$: $L(ab) = L(a) + L(b)$.*

As an immediate consequence we have:

Corollary 5.21. $\forall a, b > 0$: $L\left(\frac{a}{b}\right) = L(a) - L(b)$.

Proof. Homework. ☐

Corollary 5.22. *$\forall a > 0, \forall r \in \mathbb{Q}$: $L(a^r) = rL(a)$.*

Corollary 5.23. range$(L) = \mathbb{R}$.

Since L is continuous and increasing, and has range$(L) = \mathbb{R}$, we know that there is a unique number which is mapped to 1.

Definition 5.24. The unique number $e \in (0, \infty)$ with $L(e) = 1$ is called the *Euler number*.

The value of e is roughly $e \sim 2.7182\ldots$. An approximation of e can be found for example by Newton's method. In fact, the Euler number e is a transcendental number.

Definition 5.25. The function L is called the *natural logarithm* and is written as

$$L(x) = \ln x\,.$$

Hence, summarising the previous results:

$$\begin{gathered} \ln 1 = 0, \quad \ln e = 1 \\ \ln(ab) = \ln(a) + \ln(b), \quad \ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b) \\ \ln(a^r) = r\ln(a), \quad \forall r \in \mathbb{Q}\,. \end{gathered}$$

Note that for $f(x) = \ln|x|$, $x \neq 0$, we have $\frac{d}{dx}\ln|x| = \frac{1}{x}$ so that

$$\int \frac{1}{x}dx = \ln|x| + C$$

on every interval which does not include 0. A common error is to write $\int \frac{1}{x}dx = \ln x + C$: but this is false on domains containing negative numbers.

We can use the logarithm to differentiate products of functions; using *logarithmic differentiation*: for $g = \prod_{i=1}^n g_i$ with differentiable functions g_i we have

$$\ln|g| = \ln\left(\prod_{i=1}^n |g_i|\right) = \sum_{i=1}^n \ln|g_i| \,.$$

Therefore, by differentiating both sides using the chain rule, we have (away from the zeros of g):

$$\frac{g'}{g} = \sum_{i=1}^n \frac{g_i'}{g_i}$$

so that

$$g' = g\left(\sum_{i=1}^n \frac{g_i'}{g_i}\right) .$$

The Exponential Function

Since the logarithm is an increasing function, it is also one–to–one, that is, there exists a function $E : \mathbb{R} \to (0, \infty)$ with

$$\ln(E(x)) = x \quad \forall x \in \mathbb{R} \,.$$

Since $\ln(e^r) = r\ln(e) = r$ for $r \in \mathbb{Q}$ we know that $E(r) = e^r$. Thus, the function E extends the *exponential function* from rational numbers to all real numbers. We denote this extension by

$$e^x = E(x) \,.$$

In particular, by definition $\ln(e^x) = x$ for all $x \in \mathbb{R}$. Using the properties of the logarithm function we can find the corresponding properties of the exponential function.

Properties:

(i) $e^x > 0$ for all $x \in \mathbb{R}$.

(ii) $e^{\ln x} = x$ for all $x > 0$.

(iii) $e^{a+b} = e^a e^b$ for all $a, b \in \mathbb{R}$.

(iv) $e^{-b} = \frac{1}{e^b}$ for all $b \in \mathbb{R}$.

(v) $\frac{d}{dx}e^x = e^x$ for all $x \in \mathbb{R}$.

Using the logarithm function we can also define exponentials to any base.

Definition 5.26. For $x > 0$ put

$$x^r = e^{r\ln x}, \quad \forall r \in \mathbb{R} \,.$$

Using the laws of logarithm we immediately see that

$$x^{r+s} = x^r x^s, \quad x^{r-s} = \frac{x^r}{x^s}, \quad \forall r, s \in \mathbb{R}.$$

The derivative of the function $f(x) = x^r$, $x \in (0, \infty)$, $r \in \mathbb{R}$, is given by

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

For the function $f(x) = p^x$, $x \in \mathbb{R}$, $p > 0$ we have

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

An important application of the exponential function is the exponential growth/decay.

Theorem 5.27 (Exponential growth/decay)**.** *If $f'(t) = kf(t)$ for some $k \in \mathbb{R}$ then there exists a constant $c \in \mathbb{R}$ such that*

$$f(t) = ce^{kt}, \quad t \in \mathbb{R}.$$

Exponential growth/decay can be seen in various topics such as biology and economy.

Example 5.28. The world population in 1980 P was roughly 4.5 billion, in 2000 it was at 6 billion. What do you expect the population to be in 2002 and in 2012?

We use a model that says that the rate of increase of a population at the time t is proportional to the size of the population at time t.

Writing $P(t)$ = population in billion at t years after 1980 we have

$$P(0) = 4.5, \quad P(20) = 6.$$

Assuming exponential growth we have $P(t) = \frac{9}{2}e^{kt}$. Evaluating at $t = 20$ we have $6 = \frac{9}{2}e^{20k}$ and $\ln \frac{4}{3} = 20k$. Thus,

$$k = \frac{1}{20} \ln(\frac{4}{3}) \approx 0.0143$$

and

$$P(t) = \frac{9}{2} e^{\frac{t}{20} \ln \frac{4}{3}}.$$

Thus, our estimate for the world population in 2002 is $P(22) \approx 6.164$. The reported figure was 6.2 billion. Moreover, our model gives $P(32) \approx 7.11$ for the population in 2012, with reported figures around 7.04 billion.

Hyperbolic functions

Define the following hyperbolic sine, hyperbolic cosine and hyperbolic tangent by the following functions

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2}, \quad \tanh x = \frac{\sinh x}{\cosh x}, \quad x \in \mathbb{R}.$$

Note that the hyperbolic tangent is indeed defined for all $x \in \mathbb{R}$ since $\cosh x > 0$ for all $x \in \mathbb{R}$.

(You will in later semesters see that $\sin(x) = \frac{e^{ix}-e^{-ix}}{2}$ where i is the complex number with $i^2 = -1$. Thus, the hyperbolic sine is indeed in its definition similar to the sine function).

The hyperbolic functions satisfy similar addition theorems as the trigonometric functions, e.g.

$$\begin{aligned} 1 &= \cosh^2(x) - \sinh^2(x), \\ \sinh(t+s) &= \sinh(t)\cosh(s) + \cosh(t)\sinh(s), \\ \cosh(t+s) &= \sinh(t)\sinh(s) + \cosh(t)\cosh(s)\,. \end{aligned}$$

The derivatives of these functions are given by

$$\frac{d}{dx}\sinh x = \cosh x, \quad \frac{d}{dx}\cosh x = \sinh x, \quad \frac{d}{dx}\tanh x = \frac{1}{\cosh^2 x} = 1 - \tanh^2 x\,.$$

5.6 Further integration methods

We now give further methods to compute integrals. As there is no trick which works for every integral, it requires a lot of practise to find which method works. Note however, that not every integral can be computed explicitly in terms of known elementary functions.

Recall first the product rule $(uv)' = u'v + uv'$ for differentiable functions u, v. We can use this to calculate integrals

Theorem 5.29 (Integration by parts)**.** *If u, v are differentiable with continuous derivatives then*

$$\int u'(x)v(x)dx = u(x)v(x) - \int u(x)v'(x)dx\,.$$

For the definite integral this gives

$$\int_a^b u'(x)v(x)dx = u(x)v(x)\big|_a^b - \int_a^b u(x)v'(x)dx\,.$$

One of many tricks of integration is to remember the various equalities for trigonometric functions, e.g., in the following example we use $\sin^2 t + \cos^2 t = 1$ and substitution:

Example 5.30.

$$\int \sin^2 x \cos^5 x dx = \frac{1}{3}\sin^3 x - \frac{2}{5}\sin^5 x + \frac{1}{7}\sin^7 x + C\,.$$

Another method is to write a rational function as sum of more elementary rational functions. We demonstrate *partial fractions* by example:

Example 5.31. Let $f(x) = \frac{2x}{x^2-x-2}$. Since $x^2 - x - 2 = (x-2)(x+1)$ we try to write

$$f(x) = \frac{A}{x-2} + \frac{B}{x+1}$$

with $A, B \in \mathbb{R}$. Finding the common denominator we get

$$\frac{2x}{x^2-x-2} = \frac{A}{x-2} + \frac{B}{x+1} = \frac{A(x+1)+B(x-2)}{(x-2)(x+1)} = \frac{(A+B)x + A - 2B}{(x-2)(x+1)}$$

Comparing we see that $(A+B)x + A - 2B = 2x$. Since two polynomials are equal only if the coefficients coincide this implies that

$$2 = (A+B) \quad \text{and} \quad A - 2B = 0\,.$$

Solving these linear equations (compare with Linear Algebra 1!) we have $A = \frac{4}{3}, B = \frac{2}{3}$. Thus,

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

This helps to compute the indefinite integral

$$\int \frac{2x}{x^2-x-2}dx = \int \frac{4}{3(x-2)}dx + \int \frac{2}{3(x+1)}dx = \frac{4}{3}\ln|x-2| + \frac{2}{3}\ln|x+1| + C\,.$$

Example 5.32. If $f(x) = \frac{2x^2+3}{x(x-1)^2}$ has a multiple zero, the ansatz is

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

As before, we compute $A = 3, B = -1, C = 5$ so that we can integrate

$$\begin{aligned}\int \frac{2x^2+3}{x(x-1)^2}dx &= \int \frac{3}{x}dx - \int \frac{1}{x-1}dx + \int \frac{5}{(x-1)^2 dx} \\ &= 3\ln|x| - \ln|x-1| - \frac{5}{x-1} + \tilde{C}\,.\end{aligned}$$

Example 5.33. If $f(x) = \frac{x^2+5x+2}{(x+1)(x^2+1)}$ has a factor which never vanishes, the ansatz is

$$f(x) = \frac{A}{x+1} + \frac{Bx+C}{x^2+1}\,.$$

As before, we compute $A = -1, B = 2, C = 3$ so that we can integrate, using substitution,

$$\int \frac{x^2+5x+2}{(x+1)(x^2+1)}dx = -\ln|x+1| + \ln|x^2+1| + 3\arctan x + \tilde{C}\,.$$

Numerical Integration

You have already worked with various numerical approximations of integrals. Here is a quick list of methods:

(i) Left, right and midpoint estimate:

$$\sum_{i=0}^{n-1} \frac{f(x_i)}{n}(b-a), \quad \sum_{i=1}^{n} \frac{f(x_i)}{n}(b-a), \quad \sum_{i=0}^{n-1} \frac{f(\frac{x_i+x_{i+1}}{2})}{n}(b-a)\,,$$

(ii) trapezoidal rule: $\frac{b-a}{n}\sum_{i=0}^{n-1}\frac{f(x_i)+f(x_{i+1})}{2}$

(iii) Simpson rule: $\frac{b-a}{6n}\big(f(x_0)+f(x_n)+2(\sum_{i=1}^{n-1} f(x_i))+4(\sum_{i=0}^{n-1} f(\frac{x_i+x_{i+1}}{2}))\big)$.

All these methods are in particular important if the integral cannot be computed in terms of elementary functions.

5.7 Improper integrals

In the definition of the definite integral we assume that f is continuous on the bounded interval $[a,b]$. However, in some cases we can define an integral even if f is continuous on a half–open interval.

Definition 5.34. (i) If f is continuous on $[a,b)$, $b \in \mathbb{R}$ or $b = \infty$, and if $L = \lim_{R\to b^-}\int_a^R f(x)dx$ exists, we say that the *improper integral* of f *converges* to L, and write

$$\int_a^b f(x)dx = L\,.$$

Otherwise, we say that the improper integral *diverges*.

(ii) If f is continuous on $(a,b]$, $a \in \mathbb{R}$ or $a = -\infty$, and if $L = \lim_{R\to a^+}\int_R^b f(x)dx$ exists, we say that the *improper integral* of f *converges* to L, and write

$$\int_a^b f(x)dx = L\,.$$

Otherwise, we say that the improper integral *diverges*.

Example 5.35. A famous improper integral is

$$\int_1^\infty \frac{1}{x^p} \quad \text{which} \begin{cases} \text{converges for } p>1 \\ \text{diverges for } 0<p\leq 1 \end{cases}$$

Example 5.36. In probability the *Gaussian distribution* plays an important role. The integrand is the bell curve

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)}$$

where μ is the expectation and σ^2 is the variance. The case $\mu = 0$ and $\sigma^2 = 1$ is called the *standard normal distribution*.

There is no elementary antiderivative of f. However, with methods of Calculus and Analysis II one can compute the Gaussian integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}\,.$$

Here, the improper integral is defined by

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \int_{0}^{\infty} e^{-x^2} dx + \int_{-\infty}^{0} e^{-x^2} dx\,.$$

Chapter 6

Differential Equations

Differential equations are on the core of many (real life and mathematical) problems.

We already met some differential equations: we know that solutions of the *first order differential equation*

$$\frac{d}{dt}y = ky$$

are given by $y = Ce^{kt}$ (exponential growth/decay), or that the

$$y'' = -\omega^2 y$$

has solutions $y(x) = C_1 \sin(\omega x) + C_2 \cos(\omega x)$ (harmonic oscillations).

6.1 What is a differential equation

Differential equations are equations involving one or more derivatives of an unknown function. We distinguish among two main types of differential equations:

- Ordinary Differential Equations (ODEs): these are differential equations in which the unknown is a function (the dependent variable) of a single variable (the independent variable).

  For example Newton's law of motion[1]

  $$\frac{d}{dt}(mv(t)) = f(t),$$

  expressing the fact that the rate of changeof momentum is equal to the applied externalforce, is an ODE for the unknown velocity $v = v(t)$ in function of time.

- Partial Differential Equations (PDEs): these are differential equations in which the dependent variable is a function of more than one variable. The derivatives appearing in the differential equations are *partial derivatives*[2] with respect to the few independent variables, hence the name *partial* differential equations.

[1]Try to get use to the alternative notation for derivatives $\frac{d}{dt}f(t) = f'(t)$.

[2]Functions of several variables and partial derivatives are the topic of a later chapter.

For example, the (one-dimensional) heat equation

$$\frac{\partial}{\partial t}(T(t,x)) = \frac{\partial^2}{\partial x^2}(T(t,x))$$

which is derived from Fourier's law and conservation of energy and expresses the evolution of temperature T along a bar in function of two independent variables: the position in the bar x and the time t.

As you can see from the examples, differential equations are very useful in modelling real problems. This is because often a phenomenon (being it natural, physical, or anything else) can be described by a law which involves the rate of variation of a given quantity (like the momentum in Newton's law of motion).

Example 6.1. A simple law for describing the growth of a population of bacteria is the following. Bacteria reproduce by division. Every bacteria divides after some period of time (say, one hour). So, clearly the rate at which the number of bacteria $N = N(t)$ grows with time is proportional to its number. So we have

$$\frac{d}{dt}(N(t)) = rN(t),$$

where r is the *rate constant* at which bacteria reproduce.

Here we will only consider ODEs. A formal definition can be given as follows.

Definition 6.2. An ordinary differential equation (ODE) is an equation

$$\Phi(x, y, y', \ldots, y^{(n)}) = 0,$$

involving one or more derivatives of the independent variable y with respect to the dependent variable x. If the maximum order of derivatives appearing in the equation is equal to$n$, the ODE is an n^{th}-order ODE. A function $y = y(x)$ satisfying the differential equation identically on an interval is called a solution on that interval.

Example 6.3. We all know that $\sin'(x) = \cos(x)$. Hence the function $\sin(x)$ is a solution on the whole real line $\mathbb{R}$ of the 1^{st}-order differential equation

$$y'(x) = \cos(x),$$

as you can immediately verify by substitution. We can also consider systems of differential equations. For instance we may look for a couple of functions $y(x)$ and $z(x)$ such that

$$\begin{aligned} y'(x) &= z(x) \\ x'(x) &= -y(x). \end{aligned} \tag{6.1}$$

This is a system of 1^{st} order differential equations. Can you guess a solution of this system of differential equations? It's easy: $y(x) = \sin(x)$ and $z(x) = \cos(x)$ will work!

Notice that if you differentiate the first equation with respect to x you get:

$$y''(x) = z'(x) = -y(x),$$

having used the second equation. Similarly, by differentiating the second equation you get:

$$z''(x) = -y'(x) = -x(x).$$

Hence if $y(x)$ and $z(x)$ solve the system of differential equation (6.1) then they both solve the differential equation

$$y''(x) = -y(x).$$

This is an example of a 2^{nd}-order differential equation! It is often possible to reduce systems of 1^{st} order differential equations into a single equation of higher order.

Example 6.4. Solve the differential equation

$$y'(x) = y(x).$$

Solution We all know which one is the function whose derivative it itself: it's the exponential function. Indeed $y(x) = e^x$ solve the given differential equation on the whole real line. In facts, for any $C \in \mathbb{R}$ the function

$$y(x) = Ce^x,$$

is a solution! So the given equation has infinitely many solutions. □

The situation encountered in the previous example – the presence of *many* solutions – is typical.

Definition 6.5. We define general solution of a differential equation the set of all solutions of the equation, each taken with its maximal interval of definition.

The general solution of equation of Example 6.4 on the whole real line is the set of functions $S = \{Ce^x : C \in \mathbb{R}\}$. Indeed we know already that any such function is a solution. Also, as you have seen in Calculus & Analysis I, the exponential function is the *only* function whose derivative is itself. So S contains all solutions.

Example 6.6. Determine the evolution of a population of bacteria $N = N(t)$ in function of time given that they reproduce with rate r and that the population at time $t = 0$ is given by N_0.

Solution We have seen in Example 6.1 that the population evolves according to the 1^{st}-order ordinary differential equation $\frac{d}{dt}(N(t)) = rN(t)$. This is essencialy the equation of the previous example. The solutions are given by

$$N(t) = Ce^{rt}, \qquad C \in \mathbb{R},$$

as you can easily verify. Imposing that we start with N_0 bacteria at time $t = 0$, we get $N_0 = N(0) = C$. So the *initial condition* $N(0) = N_0$ selects a single solution. In other words, the combination of the evolution law *and* the starting population permits to uniquely determine the evolution of the population. □

The previous example is a typical example of an initial value problem.

An initial value problem (IVP) for an ODE is a problem consisting of

- a differential equation

- a set of prescribed values for the solutions and its derivatives at the *initial point* that permit to identify a unique solution of the differential equation.

The solution of an IVP is valid on the largest interval containing the initial point and in which the solution function satisfies the differential equation. It is a particular solution of the differential equation.

To get started, we shall restrict ourselves to so-called "quasilinear" 1^{st}-order ordinary differential equations, that can be written in the form

$$y'(x) = g(x, y), \tag{6.2}$$

for some continuous function g of both the independent variable x and dependent variable y. In terms of the original Definition 6.2, this means that $\Phi(x, y, y') = y'(x) - g(x, y)$.

An IVP for (6.2) reads:

$$\begin{cases} y'(x) = g(x, y), \\ y(x_0) = y_0. \end{cases}$$

Example 6.7. Solve the IVP

$$\begin{cases} y'(x) = xe^{-x^2/2}, \\ y(0) = 1. \end{cases}$$

Solution First, we need to try find all solution of the differential equations. Second, we selection the particular solution that satisfies the initial value condition.

The equation can simply be solved by integration with respect to x. We have

$$\int y'(x)dx = \int xe^{-x^2/2}dx + C,$$

where $C \in \mathbb{R}$ is the constant of integration. And, since $y = y(x)$ we have $dy = y'(x)dx$, and so

$$\int dy = \int e^{-x^2/2}d(x^2/2) + C.$$

Now integrating both sides (with respect to y and x, respectively), we get:

$$y(x) = -e^{-x^2/2} + C.$$

Now we need to select *the* solution of the IVP. Imposing the initial condition yields:

$$1 = y(0) = -1 + C,$$

hence $C = 2$ and the solution of the IVP is $y(x) = -e^{-x^2/2} + 2$ $\square$

From the above example we learn two facts:

1. Equations of the form $y'(x) = f(x)$ can be solved by integration.

2. The constant of integration is fixed by the initial condition.

6.2 First-order separable ODEs

The integrable equations seen at the end of the previous section are a special case of a larger class of equation that can be solved by integration: separable equations.

Definition 6.8. A 1^{st}-order separable (quasilinear) ordinary differential equation is an equation of the form

$$y'(x) = f(x)g(y), \tag{6.3}$$

with f, g continuous functions.

Alternatively, we could define separable ODEs as those which can be put in the form $q(y)y' + p(x) = 0$. The two definitions are essentially equivalent.

Given a separable equation in the form (6.3) and assuming that I is an interval where $g \neq 0$, we can divide through by $g(y)$ and integrate with respect to x:

$$\int \frac{y'(x)}{g(y)}dx = \int f(x)dx + C \qquad \Rightarrow \qquad \int \frac{dy}{g(y)} = \int f(x)dx + C.$$

If the functions $\frac{1}{g(y)}$ and $f(x)$ can be integrated, the solution can be found by integration for all $x \in I$.

Example 6.9. Find the general solution of the separable differential equation

$$y' = x^3 e^{-x^2/2}. \tag{6.4}$$

Solution The equation is separable with $f(x) = x^3 e^{-x^2/2}$ and $g(y) = 1$. Now $g(y) \neq 0$ always, so we can freely divide through by $g(y)$ and solve by separating variables:

$$\int dy = \int x^3 e^{-x^2/2}dx + C \quad \Rightarrow \quad y = \frac{1}{2}\int x^2 e^{-x^2/2}dx^2 + C \quad \Rightarrow \quad y = -(2+x^2)e^{-x^2/2} + C \qquad C \in \mathbb{R}.$$

The solutions found are valid for $x \in \mathbb{R}$. □

Example 6.10. Find the general solution of the separable differential equation

$$y' = y(1-y). \tag{6.5}$$

Solution The equation is separable with $f(x) = 1$ and $g(y) = y(1-y)$. Now $g(y) = 0$ if and only if $y = 0$ or $y = 1$. So the equation can be solved by separating variables in the three intervals $y < 0$, $0 < y < 1$), and $y > 1$. On any such interval, we have:

$$\int \frac{dy}{y(1-y)} = \int dx + C_1 \qquad \Rightarrow \qquad \int \left(\frac{1}{y} + \frac{1}{1-y}\right) dy = \int dx + C_1 \qquad C_1 \in \mathbb{R},$$

and thus

$$\ln|y| - \ln|1-y| = x + C \qquad \Rightarrow \qquad \ln\left|\tfrac{y}{1-y}\right| = x + C_1$$

$$\Rightarrow \qquad \left|\tfrac{y}{1-y}\right| = e^{x+C_1} = e^{C_1}e^x \qquad \Rightarrow \qquad \left|\tfrac{y}{1-y}\right| = C_2 e^x \qquad C_2 = e^{C_1} > 0.$$

Now $y/(1-y)$ is negative if $y<0$ or $y>1$ and positive if $0<y<1$. We have:

$$\left|\frac{y}{1-y}\right| = C_2 e^x \quad \Rightarrow \quad -\frac{y}{1-y} = C_2 e^x \quad \Rightarrow \quad y = \frac{C_2 e^x}{C_2 e^x - 1} \quad C_2 > 0 \qquad \text{for } y < 0 \text{ or } y > 1.$$

For $x = \ln(1/C_2)$ we have $C_2 e^x - 1 = 0$. So the obtained solution is valid only if $x < \ln(1/C_2)$ if $y < 0$ (as $C_2 > 0$) and if $x > \ln(1/C_2)$ if $y > 1$. Similarly,

$$\left|\frac{y}{1-y}\right| = C_2 e^x \quad \Rightarrow \quad \frac{y}{1-y} = C_2 e^x \quad \Rightarrow \quad y = \frac{C_2 e^x}{C_2 e^x + 1} \quad C_2 > 0 \qquad \text{for } 0 < y < 1.$$

In this last case, $y(x) \in (0,1)$ for all $x \in \mathbb{R}$ (because $C_2 > 0$), so the solution is valid for all $x \in \mathbb{R}$.

Finally we notice that $y = 0$ and $y = 1$ are also solutions of the differential equation valid for all $x \in \mathbb{R}$. $\square$

The situation seen with the previous example is typical of separable equations. You always need to examine the case $g(y) = 0$ separately. Notice that if $g(y) = 0$ then $y' = 0$ thanks to the differential equation (cf. (6.3)). So the values y for which $g(y) = 0$ are constant stationary solutions of the equation.

Definition 6.11. A stationary or equilibrium solution for a differential equation $y' = f(x)g(y)$ is any solution $y(x) = \mathit{Constant}$. The stationary solutions can be found by solving for y the equation $g(y) = 0$.

Exercises

Example 6.12. Find the general solution of the differential equation

$$\frac{dy}{dx} = x^2 e^{x-y}.$$

Solve the initial value problem for the above equation with initial value $y(0) = \ln 2$ and discuss its validity.

Example 6.13. Show that the differential equation

$$y\frac{dy}{dx} = x^2y^2 + x^2 + y^2 + 1.$$

is separable and find the general solution. Find the solution of the initial value problem for the above equation with initial value $y(0) = 1$ and discuss its validity.

Example 6.14. Consider the differential equation

$$y' = x(k - y),$$

where $k \in \mathbb{R}$ is a fixed constant. Find the general solution and check that solutions behaves as expected as $x \to \pm\infty$.

Example 6.15. Consider *homogeneous first-order* differential equations (a definition you may find easy to work with is: $\frac{dy}{dx} = f(x,y)$ is homogeneous if $f(x,y) = f(1, y/x)$); give their definition and explain how they can be solved in general. Show that the differential equation

$$\frac{dy}{dx} = \frac{x+y}{x},$$

is homogeneous and solve it.

6.3 First-order linear ODEs

According to our classification of differential equations, a first-order linear ODE is a differential equation in the form.

$$y' + p(x)y = q(x). \tag{6.6}$$

Here $y = y(x)$ is an unknown function and p, q are given **continuous** functions of the independent variable x only.

We shall consider solving Equation (6.6) on a interval $I \subset \mathbb{R}$. To this end, we distinguish three cases:

• $p(x) = 0$ for all $x \in I$. In this case Equation (6.6) reduces to

$$y' = q(x) \quad \Rightarrow \quad y(x) = \int q(x)dx + C \qquad C \in \mathbb{R}.$$

The general solution is given by all antiderivatives of q and they can simply be computed by integration.

• $q(x) = 0$ for all $x \in I$. In this case Equation (6.6) reduces to

$$y' + p(x)y = 0.$$

We call this a first-order linear homogeneous equation. It clearly is separable:

$$\frac{dy}{y} = -\frac{dx}{p(x)}$$

hence, as usual, it can be solved as by integrating both sides.

• $p(x), q(x) \neq 0$ for at least some $x \in I$. In this case the equation can be solved by the integrating factor method:

1. Find an integral of $p(x)$:

$$H(x) = \int p(x)dx. \tag{6.7}$$

The function $e^{H(x)}$ is called an integrator factor and it is defined up to a constant (note that indeed in the equation defining $H(x)$ we omitted the constant of integration as *any* integral will do). Such name comes from the fact that if we multiply the equation by the integrating factor we get:

$$\begin{aligned} & e^{H(x)}y' + e^{H(x)}p(x)y = e^{H(x)}q(x) \\ \Leftrightarrow \quad & \frac{d}{dx}\left[e^{H(x)}y\right] = e^{H(x)}q(x) \end{aligned}$$

and now, being a derivative, the right-hand side is trivially integrated:

$$e^{H(x)}y = \int e^{H(x)}q(x)dx + C \qquad C \in \mathbb{R},$$

2. and the general solution is given as

$$y(x) = e^{-H(x)} \left(\int e^{H(x)} q(x) dx + C \right) \qquad C \in \mathbb{R}. \tag{6.8}$$

Follow my advice: do not remember the formula by heart. Just remember how we got it. Let us see how this works in practice with a couple of examples.

Example 6.16. Solve the general first-order linear ODE with constant coefficients, that is the equation

$$y' + ay = b \qquad a, b \in \mathbb{R}, \text{with } a \neq 0.$$

Solution We apply the integrating factor method. As $p(x) = a$ and $\int a\,dx = ax + C$, we can take $H(x) = ac$ and an integrating factor is given by

$$e^{H(x)} = e^{ax}.$$

Multiplying the equation by the integrating factor we get

$$\frac{d}{dx}\left[e^{ax} y\right] = e^{ax} b \quad \Rightarrow \quad e^{ax} y = \frac{b}{a} e^{ax} + C \quad \Rightarrow \quad y = \frac{b}{a} + Ce^{-ax},$$

for $C \in \mathbb{R}$. □

Example 6.17. Find the general solution of the first-order linear equation

$$y' + \frac{y}{x} = 1 \qquad x > 0.$$

Solution Here $p(x) = \frac{1}{x}$ and $q(x) = 1$, hence $H(x) = \int \frac{1}{x} dx = \ln x$ for $x > 0$. Hence an integrating factor is given by

$$e^{H(x)} = e^{\ln x} = x.$$

Multiplying the equation by the integrating factor we get

$$\frac{d}{dx}(xy) = x \quad \Rightarrow \quad xy = \int x\,dx = \frac{1}{2}x^2 + C$$

and thus the general solution of the equation is given by

$$y = \frac{x}{2} + \frac{C}{x}.$$

□

As for the other types of differential equations considered so far, a particular solution will be fixed by imposing an initial condition. In general the

$$\text{initial-value problem} \qquad \begin{cases} y' + p(x)y = q(x) & x \in I = (x_0, x_1) \\ y(x_0) = y_0, \end{cases}$$

will have a unique solution.

Example 6.18. Find the solution of the initial-value problem

$$\begin{cases} y' + ay = b & x > 0 \\ y(0) = 0, \end{cases}$$

and discuss the behaviour of the solution as $x \to \infty$ in function of the coefficients a and b.

Solution We already found the general solution of the equation in Example 6.16. This is given by: $y = \frac{b}{a} + Ce^{-ax}$. We fix the constant by imposing the initial condition:

$$0 = y(0) = \frac{b}{a} + C \quad \Rightarrow \quad C = -\frac{b}{a},$$

and hence the solution of the initial-value problem is given by

$$y(x) = \frac{b}{a}(1 - e^{-ax}).$$

Now if $a > 0$ then $y(x) \overset{x\to\infty}{\longrightarrow} \frac{b}{a}$, while if $a < 0$ the solution tends to $+\infty$ if $b > 0$ and $-\infty$ if $b < 0$. □

Exercises

Example 6.19. Solve the differential equation

$$y' - x^n y = x^n.$$

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

Example 6.20. Solve initial value problem

$$\begin{cases} y' = \dfrac{x^2 + y}{x} \\ y(1) = 0. \end{cases}$$

6.4 Second-order linear ODEs

A second order differential equation is said to be *linear* if it can be written as

$$y'' + p(x)y' + q(x)y = f(x). \tag{6.9}$$

The differential equation (6.9) is *homogeneous* if $f \equiv 0$ or *inhomogeneous* if $f \not\equiv 0$.

It is necessary to solve the homogeneous equation

$$y'' + p(x)y' + q(x)y = 0 \tag{6.10}$$

in order to solve the inhomogeneous equation (6.9).

The next theorem gives sufficient conditions for existence and uniqueness of solutions of initial value problems for (6.10). We omit the proof.

Theorem 6.21. *Suppose p, q and f are continuous on an open interval (a,b), let x_0 be any point in (a,b), and let k_0 and k_1 be arbitrary real numbers. Then the initial value problem*

$$y'' + p(x)y' + q(x)y = f(x),\ y(x_0) = k_0,\ y'(x_0) = k_1$$

has a unique solution on (a,b).

Let us see 3 worked examples of homogeneous second order IVPs.

Example 6.22. The coefficients of y' and y in

$$y'' - y = 0 \tag{6.11}$$

are the constant functions $p \equiv 0$ and $q \equiv -1$, which are continuous on $(-\infty,\infty)$. Therefore Theorem 6.21 implies that every initial value problem for (6.11) has a unique solution on $(-\infty,\infty)$.

a) Verify that $y_1 = e^x$ and $y_2 = e^{-x}$ are solutions of (6.11) on $(-\infty,\infty)$.

b) Verify that if c_1 and c_2 are arbitrary constants, $y = c_1e^x + c_2e^{-x}$ is a solution of (6.11) on $(-\infty,\infty)$.

c) Solve the initial value problem

$$y'' - y = 0, \quad y(0) = 1, \quad y'(0) = 3. \tag{6.12}$$

Example 6.23. Let ω be a positive constant. The coefficients of y' and y in

$$y'' + \omega^2 y = 0 \tag{6.13}$$

are the constant functions $p \equiv 0$ and $q \equiv \omega^2$, which are continuous on $(-\infty,\infty)$. Therefore Theorem 6.21 implies that every initial value problem for (6.13) has a unique solution on $(-\infty,\infty)$.

a) Verify that $y_1 = \cos\omega x$ and $y_2 = \sin\omega x$ are solutions of (6.13) on $(-\infty,\infty)$.

b) Verify that if c_1 and c_2 are arbitrary constants then $y = c_1\cos\omega x + c_2\sin\omega x$ is a solution of (6.13) on $(-\infty,\infty)$.

c) Solve the initial value problem

$$y'' + \omega^2 y = 0, \quad y(0) = 1, \quad y'(0) = 3. \tag{6.14}$$

Theorem 6.24. *If y_1 and y_2 are solutions of the homogeneous equation*

$$y'' + p(x)y' + q(x)y = 0 \tag{6.15}$$

on (a,b), then any linear combination

$$y = c_1y_1 + c_2y_2 \tag{6.16}$$

of y_1 and y_2 is also a solution of (6.15) on (a,b).

If a, b, and c are real constants and $a \neq 0$, then

$$ay'' + by' + cy = F(x)$$

is said to be a constant coefficient equation. First, we consider the homogeneous constant coefficient equation

$$ay'' + by' + cy = 0. \tag{6.17}$$

To do so, we consider the characteristic polynomial of (6.17),

$$p(r) = ar^2 + br + c$$

We can see that

$y = e^{rx}$ is a solution of (6.17) if and only if $p(r) = 0$.

The roots of the characteristic equation are given by the quadratic formula

$$r = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{6.18}$$

We consider three cases:

CASE 1. $b^2 - 4ac > 0$, so the characteristic equation has two distinct real roots.

Example: Solve the initial value problem

$$y'' + 6y' + 5y = 0, \quad y(0) = 3,\ y'(0) = -1.$$

Solution: The characteristic polynomial is

$$p(r) = r^2 + 6r + 5 = (r + 1)(r + 5).$$

Therefore the general solution is

$$y = c_1 e^{-x} + c_2 e^{-5x}.$$

Imposing the initial conditions $y(0) = 3$, $y'(0) = -1$ in yields

$$\begin{array}{rcr} c_1 + \ \ c_2 & = & 3 \\ -c_1 - 5c_2 & = & -1. \end{array}$$

The solution of this system is $c_1 = 7/2, c_2 = -1/2$. Therefore the solution of the IVP is

$$y = \frac{7}{2}e^{-x} - \frac{1}{2}e^{-5x}.$$

CASE 2. $b^2 - 4ac = 0$, so the characteristic equation has a repeated real root.

Example: Solve the initial value problem

$$y'' + 6y' + 9y = 0, \quad y(0) = 3,\ y'(0) = -1. \tag{6.19}$$

Solution: The characteristic polynomial is

$$p(r) = r^2 + 6r + 9 = (r + 3)^2,$$

and the general solution is

$$y = e^{-3x}(c_1 + c_2 x)$$

Imposing the initial conditions $y(0) = 3$, $y'(0) = -1$ yields $c_1 = 3$ and $-3c_1 + c_2 = -1$, so $c_2 = 8$. Therefore the solution of (6.19) is

$$y = e^{-3x}(3 + 8x).$$

CASE 3. $b^2 - 4ac < 0$, so the characteristic equation has complex roots.

Example: Solve the initial value problem

$$y'' + 4y' + 13y = 0, \quad y(0) = 2, \; y'(0) = -3. \tag{6.20}$$

Solution: The characteristic polynomial of (6.20) is

$$p(r) = r^2 + 4r + 13 = r^2 + 4r + 4 + 9 = (r + 2)^2 + 9.$$

whose roots are $-2 \pm 3i$. Therefore the general solution is

$$y = e^{-2x}(c_1 \cos 3x + c_2 \sin 3x) \tag{6.21}$$

Imposing the condition $y(0) = 2$ in (6.21) shows that $c_1 = 2$. Differentiating (6.21) yields

$$y' = -2e^{-2x}(c_1 \cos 3x + c_2 \sin 3x) + 3e^{-2x}(-c_1 \sin 3x + c_2 \cos 3x),$$

and imposing the initial condition $y'(0) = -3$ here yields $-3 = -2c_1 + 3c_2 = -4 + 3c_2$, so $c_2 = 1/3$. Therefore the solution of (6.20) is

$$y = e^{-2x}(2 \cos 3x + \frac{1}{3} \sin 3x).$$

To summarise:

Theorem 6.25. *Let $p(r) = ar^2 + br + c$ be the characteristic polynomial of*

$$ay'' + by' + cy = 0. \tag{6.22}$$

Then

a) *If $p(r) = 0$ has distinct real roots r_1 and r_2, then the general solution of (6.22) is*

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x}.$$

b) *If $p(r) = 0$ has a repeated root r_1, then the general solution of (6.22) is*

$$y = e^{r_1 x}(c_1 + c_2 x).$$

c) *If $p(r) = 0$ has complex conjugate roots $r_1 = \lambda + i\omega$ and $r_2 = \lambda - i\omega$ (where $\omega > 0$), then the general solution of (6.22) is*

$$y = e^{\lambda x}(c_1 \cos \omega x + c_2 \sin \omega x).$$

Now we consider the inhomogeneous constant coefficient equation

$$ay'' + by' + cy = 0. \tag{6.23}$$

whose complementary equation is the associated homogeneous part that we say in (6.17). The next theorem shows how to find the general solution of (6.23) if (somehow) we know one particular solution y_p and we also know the general solution of the complementary equation.

Theorem 6.26. *Suppose p, q, and f are continuous on $(a, b.)$ Let y_p be a particular solution of*

$$y'' + p(x)y' + q(x)y = f(x) \tag{6.24}$$

on (a, b), and let $y = c_1y_1 + c_2y_2$, where c_1 and c_2 are constants, be the general solution of the complementary equation

$$y'' + p(x)y' + q(x)y = 0$$

on (a, b). Then y is a solution of (6.24) on (a, b) if and only if

$$y = y_p + c_1y_1 + c_2y_2.$$

Sometimes it is not hard to guess the form of a particular solution.

Example 6.27. Solve the initial value problem

$$y'' - 2y' + y = -3 - x + x^2, \quad y(0) = -2, \quad y'(0) = 1.$$

Solution: The characteristic polynomial of the complementary equation

$$y'' - 2y' + y = 0$$

is $r^2 - 2r + 1 = (r - 1)^2$, so $y = (c_1 + c_2x)e^x$ is the general solution of the complementary equation. To guess a form for a particular solution, we note that substituting a second degree polynomial $y_p = A + Bx + Cx^2$ into the left side of the differential equation will produce another second degree polynomial with coefficients that depend upon A, B, and C. The trick is to choose A, B, and C correctly. If

$$y_p = A + Bx + Cx^2 \quad \text{then} \quad y_p' = B + 2Cx \quad \text{and} \quad y_p'' = 2C,$$

so

$$\begin{aligned} y_p'' - 2y_p' + y_p &= 2C - 2(B + 2Cx) + (A + Bx + Cx^2) \\ &= (2C - 2B + A) + (-4C + B)x + Cx^2 = -3 - x + x^2. \end{aligned}$$

Equating coefficients of like powers of x on the two sides of the last equality yields

$$\begin{aligned} C &= 1 \\ B - 4C &= -1 \\ A - 2B + 2C &= -3, \end{aligned}$$

so $C = 1$, $B = -1 + 4C = 3$, and $A = -3 - 2C + 2B = 1$. Therefore $y_p = 1 + 3x + x^2$ is a particular solution and Theorem 6.26 implies that

$$y = 1 + 3x + x^2 + e^x(c_1 + c_2 x) \tag{6.25}$$

is the general solution. Imposing the initial condition $y(0) = -2$ in (6.25) yields $-2 = 1 + c_1$, so $c_1 = -3$. Differentiating (6.25) yields

$$y' = 3 + 2x + e^x(c_1 + c_2 x) + c_2 e^x,$$

and imposing the initial condition $y'(0) = 1$ here yields $1 = 3 + c_1 + c_2$, so $c_2 = 1$. Therefore the solution of the IVP is

$$y = 1 + 3x + x^2 - e^x(3 - x).$$

This guesswork method might not always work: the easy differential equation

$$y'' + y' = 1 + 2x$$

does not have any particular solution of the form $y_p = A + Bx$ where A, B are constants! But it can be integrated to give a differential equation

$$y' + y = C + x + x^2$$

which can be solved using an integrating factor.

The differential equation

$$ay'' + by' + cy = ke^{\alpha x},$$

where a, b, c, k, has a particular solution y_p:

a) If $e^{\alpha x}$ isn't a solution of the complementary equation

$$ay'' + by' + cy = 0,$$

then $y_p = Ae^{\alpha x}$.

b) If $e^{\alpha x}$ is a solution of the complementary equation but $xe^{\alpha x}$ is not, then $y_p = Axe^{\alpha x}$.

c) If both $e^{\alpha x}$ and $xe^{\alpha x}$ are solutions of the complementary equation then $y_p = Ax^2e^{\alpha x}$.

In each case the constant A is an unknown that still has to be found by substituting y_p into the differential equation.

We look at just one more type of inhomogeneous equation:

Theorem 6.28. *a) Suppose ω is a positive number and P and Q are polynomials. Let k be the larger of the degrees of P and Q. Then the equation*

$$ay'' + by' + cy = P(x)\cos\omega x + Q(x)\sin\omega x$$

has a particular solution

$$y_p = A(x)\cos\omega x + B(x)\sin\omega x,$$

where

$$A(x) = A_0 + A_1 x + \cdots + A_k x^k \quad \text{and} \quad B(x) = B_0 + B_1 x + \cdots + B_k x^k,$$

provided that $\cos \omega x$ *and* $\sin \omega x$ *are not solutions of the complementary equation.*
b) The equation

$$y'' + \omega^2 y = P(x) \cos \omega x + Q(x) \sin \omega x$$

has a particular solution

$$y_p = A(x) \cos \omega x + B(x) \sin \omega x,$$

where

$$A(x) = A_0 x + A_1 x^2 + \cdots + A_k x^{k+1} \quad \text{and} \quad B(x) = B_0 x + B_1 x^2 + \cdots + B_k x^{k+1}.$$

In both cases, the values of the constants $A_0, B_0, A_1, B_1 \ldots, A_k, B_k$ are unknowns that still have to be found by substituting y_p into the differential equation.

Chapter 7

Infinite series

Using what we learnt about convergence of sequences, we now start the study of infinite sums, the so-called infinite series, or series for short.

The question is very simple: can we evaluate sums made up of infinite terms? To start with, we need an easy way to represent the terms of the sum. The notion of sequences developed in the previous section is just what we need. Indeed, we rephrase our question as follows: given a sequence $(a_n)_{n\geq 1}$ can we add up its terms, namely consider the *infinite sum*

$$a_1 + a_2 + a_3 + \cdots + a_n + \cdots$$

or, using the sigma notation,

$$\sum_{n=1}^{\infty} a_n.$$

As we shall see, it turns out that the concept of convergence of sequences is the right one for defining infinite sums. In other words, we shall give a meaning to infinite sums by resorting to the concept of limit of sequences.

Definition 7.1. For a sequence $(a_n)_{n\geq 1}$ of real numbers, its n-th partial sum s_n is

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{k=1}^{n} a_k.$$

If this new sequence of partial sums $(s_n)_{n\geq 1}$ converges, we say that the sequence of terms $(a_n)_{n\geq 1}$ is summable and that the infinite series is convergent

$$s = \sum_{n=1}^{\infty} a_n = \lim_{n\to\infty} s_n.$$

Notice that the sum of a convergent series is unique by definition (*cf.* Theorem **??**). Also notice that the fact of starting the sum with $n = 1$ is inessential. In particular, in many cases we will consider series that start from 0-th index, i.e., for $(a_n)_{n\geq 0}$, we define the

n-th partial sum as $s_n = \sum_{k=0}^{n} a_k$, (the 0-th partial sum is included!) and the infinite series as $\sum_{n=0}^{\infty} a_n$.

Ok, all this is very nice, but: are there convergent series? The following are examples showing that convergent series *do* exists.

Example 7.2. Show that the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.

Solution We consider the sequence of partial sums $(s_n)_{n\geq 1}$, where

$$s_n = \sum_{k=1}^{n} \frac{1}{k^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2}.$$

Since $\frac{1}{k^2} > 0$ for all $k \in \mathbb{N}$, we can easily see that

$$s_{n+1} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} + \frac{1}{(n+1)^2} \geq 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} = s_n,$$

that is, $(s_n)_{n\geq 1}$ is increasing. Moreover, we have

$$\frac{1}{k^2} = \frac{1}{k \cdot k} \leq \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k},$$

for $k = 2, 3, \ldots$, and hence,

$$\begin{aligned} s_n &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2} \leq 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 1 + 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{n-1} - \frac{1}{n} = 2 - \frac{1}{n} \leq 2 \quad \text{for all } n = 2, 3, \ldots. \end{aligned}$$

But since $s_n > 0$ for all $n \in \mathbb{N}$ and $s_n \leq 2$ for $n = 2, 3, \ldots$, we conclude that $(s_n)_{n\geq 1}$ is also bounded. Thus, $(s_n)_{n\geq 1}$ is monotonic and bounded. Then, from Theorem **??** we conclude that $(s_n)_{n\geq 1}$ converges and, therefore, the series converges. □

So we do have an example of a series that converges. But are there cases were we can actually compute the infinite sum? Here is a (fundamental!) example.

Example 7.3 (Geometric series)**.** Show that the infinite series

$$\sum_{n=0}^{\infty} c^n = \frac{1}{1-c} \quad \text{if } |c| < 1,$$

and that it diverges for all other values of c. This is the so-called *geometric series*; every term in the sum is given by the product of the previous term and c.

Solution For $c \neq 1$, recall the identity (the proof of this is Exercise 7.9)

$$\sum_{k=0}^{n} c^k = \frac{1-c^{n+1}}{1-c}.$$

Thus,

$$\sum_{n=0}^{\infty} c^n = \lim_{n\to\infty} \sum_{k=0}^{n} c^k = \lim_{n\to\infty} \frac{1-c^{n+1}}{1-c} = \frac{1-\lim_{n\to\infty} c^{n+1}}{1-c} = \frac{1-c\lim_{n\to\infty} c^n}{1-c} = \frac{1-c\cdot 0}{1-c} = \frac{1}{1-c},$$

since, as shown in Exercise **??**, we have that $\lim_{k\to\infty} c^k = 0$ if $-1 < c < 1$.

Now assume c is *not* such that $-1 < c < 1$, i.e., $|c| \geq 1$. Then we can see that $\lim_{k\to\infty} c^k \neq 0$ (in particular, we have $\lim_{k\to\infty} |c|^k = +\infty$). The following Theorem 7.4 shows that, if this is the case, than the series cannot converge. $\square$

Theorem 7.4 (The vanishing condition)**.** *If* $\sum_{n=0}^{\infty} a_n$ *converges, then* $\lim_{n\to\infty} a_n = 0$.

Proof. Assume that $\sum_{n=0}^{\infty} a_n = s$ for some $s \in \mathbb{R}$. We have

$$s_{n-1} + a_n = (a_1 + a_2 + \cdots + a_{n-1}) + a_n = s_n \quad \text{and thus,} \quad a_n = s_n - s_{n-1}.$$

Now, taking limits, we obtain

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = s - s = 0.$$

$\square$

Example 7.5. The series $1+1+1+1+\cdots$ is divergent. Indeed this sum diverges to $+\infty$.

Solution The vanishing condition is not satisfied as the $\lim_{n\to\infty} 1 = 1 \neq 0$. Clearly, the n^{th} partial sum is given by $s_n = n$ and it diverges to $+\infty$. $\square$

Example 7.6. The series $1-2+1-2+\cdots$ is divergent. Indeed it oscillates: the odd partial sums add up to $+1$ and the even partial sums add up to -1.

Solution The sequence $1, -2, 1, -2, \ldots$ has no limit. If it had, than every subsequence would share the same limit (the proof of this fact is left as an exercise). But, the two subsequence given by the even and odd terms:

$$(a_{2n})_{n\in\mathbb{N}} := (1)_{n\in\mathbb{N}}, \qquad (a_{2n+1})_{n\in\mathbb{N}} := (-2)_{n\in\mathbb{N}}$$

are both constant sequences thus they respectively converge to their constant values $+1$ and -2. As the latter numbers differ, they cannot be simultaneously equal to the common limit of the original sequence.

Let us analyse how the partial sums $(s_n)_{n\geq 1}$ look like. Clearly,

$$s_n = \begin{cases} +1 & \text{if } n \text{ odd;} \\ -1 & \text{if } n \text{ even.} \end{cases}$$

The subsequence $(s_{2n+1})_{n\in\mathbb{N}}$ is the constant sequence of $+1$ and $(s_{2n})_{n\in\mathbb{N}}$ is the constant sequence of -1. Thus, both these subsequences are convergent, while the original sequence $(s_n)_{n\geq 1}$ oscillates between the values $+1$ and -1. □

The above theorem is quite intuitive. It is to be expected that, summing up infinite numbers, these must get smaller and smaller in absolute value, if we want to hope for the sum to be finite!

However, the converse is *not* true. We say that $\lim_{n\to\infty} a_n = 0$ is a necessary, but not sufficient, condition for the convergence of the associated series. This fact is proved by the following (fundamental!) example.

Example 7.7 (Harmonic series)**.**

Show that, although $\lim_{n\to\infty} \frac{1}{n} = 0$, the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

Solution The fact that $\lim_{n\to\infty} \frac{1}{n} = 0$ was shown in Exercise **??**. To show that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, we consider

$$s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \quad \text{and} \quad s_{2n} = 1 + \frac{1}{2} + \cdots + \frac{1}{n} + \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n},$$

and, thus

$$s_{2n} - s_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} > \underbrace{\frac{1}{2n} + \frac{1}{2n} + \cdots + \frac{1}{2n}}_{n \text{ times}} = n \cdot \frac{1}{2n} = \frac{1}{2}.$$

Therefore, for every $n \in \mathbb{N}$, we have that

$$s_{2n} - s_n > \frac{1}{2}.$$

Thus the sequence $(s_n)_{n\geq 1}$ of partial sums is *not* a Cauchy sequence (as we can easily verify by choosing $\varepsilon = \frac{1}{2}$ and $m = 2n$ in Definition **??**), and therefore, it does not converge (because if it did converge it should also be a Cauchy sequence according to Theorem **??**). Since the sequence of partial sums of the series does not converge, we conclude that the series diverges. □

As we did for sequence, we can define an *algebra* for series. This is most useful as it will allow us to study complicated series by considering simpler ones.

Theorem 7.8 (Linearity)**.** *Let $\lambda \in \mathbb{R}$, $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be convergent series.*

Then the series $\sum_{n=1}^{\infty}(a_n + b_n)$ and $\sum_{n=1}^{\infty} \lambda a_n$ converge, and we have

$$\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n, \quad \text{and} \quad \sum_{n=1}^{\infty} \lambda a_n = \lambda \sum_{n=1}^{\infty} a_n,$$

respectively.

Proof. The proof follows by considering the definition of a convergent series (Definition 7.1) and the properties of the limits of sequences described in Theorem **??**, and is left as an exercise. □

Exercises

Example 7.9. Prove by Induction that:

$$\sum_{k=0}^{n} c^k = \frac{1 - c^{n+1}}{1 - c}$$

Example 7.10. Determine if the following series converge or diverge. Justify your answer.

- ♣ $\sum_{n=0}^{\infty} 2^n$
- ♣ $\sum_{n=0}^{\infty} (-1)^n$
- ♣ $\sum_{n=0}^{\infty} \frac{1}{\sqrt{n}}$
- ♣ $\sum_{n=0}^{\infty} \frac{n^3}{n^5 + 2n^4 + n^2 + 5}$
- ♣ $\sum_{n=0}^{\infty} \frac{1}{n^3}$

7.1 Convergence tests for series with nonnegative terms

Here, we present various criteria that enable us to analyse the convergence of infinite series.

We start with the study of series whose terms a_n are *all* nonnegative. Clearly, everything we say here applies to series whose terms are all nonpositive (why?). Further, it is clear that, if a few terms are of a different sign, it does not matter much: all our criteria are still valid as long as *ultimately*, *i.e.* from a sufficiently large n onwards, the sign does not change[1] as the only thing that matters is what happens at the limit.

[1] For example, the sequence $a_1 = -1, a_2 = 1/2, a_3 = -1/3, a_4 = 1/4, a_5 = 1/5, \ldots, a_n = 1/n, \ldots$ would define a definitively nonnegative series

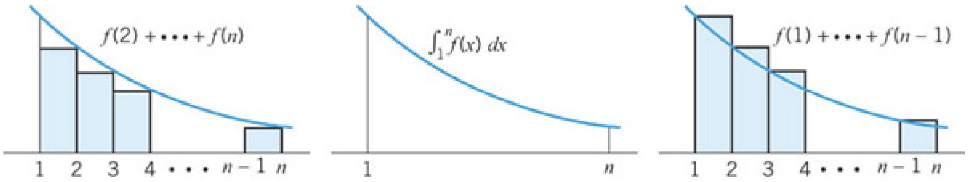


Figure 7.1: Depiction of the Integral Test.

Ok, so assume all a_n are nonnegative. Question: what can happen to the series $\sum_{n=1}^{\infty} a_n$? Every time we add one more term, the sum increases or at most stays the same if the new term was zero. In other words, the partial sums with nonnegative terms are monotonic nondecreasing, and for the associated series we have only two possibilities: a series of nonnegative terms either converges or diverges to $+\infty$. This clear cut situation is what makes these series easier to analyse. We can summarise these statements with the following theorem.

Theorem 7.11. *A series whose terms are ultimately nonnegative converges if and only if the sequence of partial sums is bounded above.*

Theorem 7.12 (Integral test)**.** *Suppose that f is a* **continuous**, **positive**, *and* **nonincreasing** *function on $[1, \infty)$, and that $f(n) = a_n$ for all n. Then*

$$\sum_{n=1}^{\infty} a_n \quad \textit{converges if and only if} \quad \int_1^{\infty} f(x)\,dx \quad \textit{converges.}$$

We give here a hint of the proof. The idea is very well depicted in Figure 7.1. The partial sums can be seen as lower and upper sums bounding the integral from below and from above (remember how the integral is defined in the first place?). Indeed, as shown in the figure, since for each $n \in \mathbb{N}$ and $x \in [n, n+1]$ we have $f(n+1) \leq f(x) \leq f(n)$, we must have

$$\sum_{k=2}^{n} f(k) \leq \int_1^{n} f(x)\,dx \leq \sum_{k=1}^{n-1} f(k) \qquad \forall n \in \mathbb{N}. \tag{7.1}$$

The first inequality tells us that if the improper integral is convergent, then so is the infinite series, and the second inequality tells us the converse.

Let us see now how the test can be put at use. The following is perhaps the most important example of family of series that can be analysed using the integral test. The improper integrals used here are also worth remembering.

Example 7.13 (p-series)**.** The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \begin{cases} \text{diverges for} & 0 < p \leq 1 \\ \text{converges for} & 1 < p. \end{cases} \tag{7.2}$$

By the integral test, the convergence of the p-series is equivalent to the existence of $\displaystyle\int_1^{\infty} \frac{1}{x^p}\,dx$, and this integral exists only if $p > 1$, *cf.* Exercise 7.21.

Notice that the p-series for $p = 1$ is nothing else than the harmonic series, hence the above is an alternative way to prove that the harmonic series is divergent. The harmonic series sits right at the turning point between convergent and divergent p-series: its terms are small but quite not small enough...

Let us go back to the *meaning* of the Integral Test by examining Figure 7.1 a bit more closely. It is evident from the figure that the integral of the function f, that is the area below the blue curve $y = f(x)$ is somehow close to the sum of few terms of the series, and vice-versa. This idea can be exploited as follows.

Using the integral test to estimate the sum of a series. Let $s = \sum_{n=1}^{\infty} f(n)$ be a (nonnegative) convergent series defined through the function $f : [1, \infty) \to \mathbb{R}$ such that $\int_1^{\infty} f(x)\,dx$ is convergent. Reasoning as we did to get (7.1), it easy to see that [2]

$$\int_{n+1}^{\infty} f(x)\,dx \le s - s_n = \sum_{k=n+1}^{\infty} f(k) \le \int_n^{\infty} f(x)\,dx, \qquad n = 2, 3, \dots,$$

or, equivalently, s belongs to the interval $s \in [s_n + \int_{n+1}^{\infty} f(x)\,dx, s_n + \int_n^{\infty} f(x)\,dx]$, which thus gives an estimation of the unknown value of the infinite sum. This idea is used in Exercise 7.22 to compute an estimation of the p-series' value.

The following is arguably the most useful test for infinite series.

Theorem 7.14 (Comparison test)**.** *Suppose that* $0 \le a_n \le b_n$ *for all* $n \in \mathbb{N}$. *Then:*

(i) if the series $\displaystyle\sum_{n=1}^{\infty} b_n$ *converges then the series* $\displaystyle\sum_{n=1}^{\infty} a_n$ *converges;*

(ii) (equivalently:) if the series $\displaystyle\sum_{n=1}^{\infty} a_n$ *diverges then the series* $\displaystyle\sum_{n=1}^{\infty} b_n$ *diverges.*

Proof. (i) Since $0 \le a_n \le b_n$ for all $n \in \mathbb{N}$, both sequences of partial sums $\sum_{k=1}^{n} a_k$ and $\sum_{k=1}^{n} b_k$ are monotonic nondecreasing. Moreover, for all $n \in \mathbb{N}$

$$0 \le \sum_{k=1}^{n} a_k \le \sum_{k=1}^{n} b_k \le \sum_{k=1}^{\infty} b_k \in \mathbb{R},$$

as $\displaystyle\sum_{n=1}^{\infty} b_n$ is convergent. Hence, the series $\displaystyle\sum_{n=1}^{\infty} a_n$ converges by the Monotonic Convergence Theorem.

[2] Try to draw the figure corresponding to these bounds as Figure 7.1 corresponds to the bounds in (7.1).

(ii) This statement is equivalent to the previous one (it is the *contrapositive* version of (i)). □

Example 7.15. Show that the series $\displaystyle\sum_{n=0}^{\infty} \frac{1}{\sqrt{n}}$ diverges.

Solution We have $\sqrt{n} \leq n$ for $n \in \mathbb{N}$ (since $n \geq 1$). Therefore,

$$\frac{1}{n} \leq \frac{1}{\sqrt{n}}.$$

But, we know from Example 7.7 that $\sum_{n=0}^{\infty} \frac{1}{n}$ diverges, and thus, from Theorem 7.14(ii), $\sum_{n=0}^{\infty} \frac{1}{\sqrt{n}}$ diverges. □

Theorem 7.16 (Root and Ratio test). *Let $(a_n)_{n\geq 1}$ a positive sequence and suppose that the following limit*

$$\rho = \lim_{n\to\infty} \frac{a_{n+1}}{a_n} \qquad \textit{(RATIO TEST)} \tag{7.3}$$

or

$$\rho = \lim_{n\to\infty} (a_n)^{1/n} \qquad \textit{(ROOT TEST)} \tag{7.4}$$

exists in $\mathbb{R}$ or is $+\infty$. If

(i) $0 \leq \rho < 1$, then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ converges;

(ii) $\rho > 1$, then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ diverges;

(iii) $\rho = 1$, then the test is inconclusive.

Remark 7.17. *If $\rho = 1$ the above tests are inconclusive in the sense that the series may converge or diverge. Indeed, there exist cases where $\rho = 1$ for a convergent series (e.g., for $\sum_{n=0}^{\infty} \frac{1}{n^2}$ we have $\rho = 1$ and the series converges), and cases where $\rho = 1$ for a divergent series (e.g., for the Harmonic series $\sum_{n=0}^{\infty} \frac{1}{n}$ we have $\rho = 1$ and the series diverges, see Example 7.7).*

Example 7.18. Show that $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!}$ converges.

Proof. We use the ratio test. We compute

$$\frac{a_{n+1}}{a_n} = \frac{\frac{1}{(n+1)!}}{\frac{1}{n!}} = \frac{n!}{(n+1)!} = \frac{n!}{n!(n+1)} = \frac{1}{n+1}, \text{ and, therefore } \lim_{k\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1}{n+1} = 0 < 1.$$

hence the series converges. □

Example 7.19. Show that $\sum_{n=1}^{\infty} \frac{1}{(\ln n)^n}$ converges.

Proof. We use the root test:

$$\lim_{n\to\infty} (a_n)^{1/n} = \lim_{n\to\infty} \frac{1}{\ln n} = 0 < 1,$$

hence the series converges. □

So, which test should we use? As with integration, there is no definitive answer, but the 'look' of the series at hand might give you a hint. In general:

- The ratio test is particularly effective with factorials and with combinations of powers and factorials;
- The root test is used only if powers are involved. It is used more rarely than the ratio test is;
- Use the comparison test if the terms are rational functions of n. The ratio and root tests are seldom effective in this case.
- The integral test is most useful when the terms have the form of a derivative.

Exercises

Example 7.20. Determine if the series $\sum_{n=1}^{\infty} \frac{n^4}{2^n}$ converges justifying your assertion.

Example 7.21. By evaluating the $\lim_{a\to\infty} \int_1^a \frac{1}{x^p}\, dx$ for $p \in (0, \infty)$ show that the p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges only if $p > 1$.

Example 7.22. Let $s = \sum_{n=1}^{\infty} \frac{1}{n^2}$ and s_n the associated partial sum. Use the integral estimation method to find $n \in \mathbb{N}$ such that the error given by using s_n to estimate s, that is the quantity $|s - s^n|$, is less than 0.01. Can you think of a way to get a better approximation using the same number of terms of the sum?

Example 7.23. Prove directly the second statement of Theorem 7.14.

7.2 Absolute and conditional convergence

Definition 7.24. We say that a series $\sum_{k=1}^{\infty} a_k$ converges absolutely if $\sum_{k=1}^{\infty} |a_k|$ converges. We say that a series is *conditionally convergent* if it converges but does not converge absolutely.

The notion of absolute convergence is in general stronger than the notion of convergence, as shown by the following result.

Theorem 7.25. *If a series converges absolutely, then it converges.*

Proof. Since $0 \leq a_n + |a_n| \leq 2|a_n|$ and $\sum 2|a_n|$ converges, the comparison test tells us that $\sum a_n + |a_n|$ converges. By linearity we can subtract $\sum |a_n|$ to see that $\sum a_n$ converges. □

Note that for nonnegative or nonpositive series, the notion of convergence and absolute convergence coincide: these series are either (absolutely) convergent or they diverge. Moreover, when it diverges, it's sum diverges to $+\infty$. This is *not* the case in general. Series with both positive and negative terms can

- converge absolutely (in which case, we also have convergence by Theorem 7.25); or
- converge conditionally; or
- diverge. In this case, it can either diverge to $+\infty$ or $-\infty$ or it can *oscillate*, *cf.* Exercise 7.6

Checking for absolute convergence is the easier bit: all tests for series with nonnegative terms can be used to test for absolute convergence[3]! As for testing for conditional convergence, we only have the following test.

Theorem 7.26 (alternating series test/Leibniz[4] criterion)**.** *Let $(a_n)_{n\geq 1}$ be a decreasing and positive sequence. Then, the series*

$$\sum_{n=1}^{\infty}(-1)^{n+1}a_n = a_1 - a_2 + a_3 - a_4 + \dots \quad \textit{converges iff} \quad \lim_{n\to\infty} a_n = 0$$

Proof. If the series converges, then by the vanishing condition, (Theorem 7.4), we know that $(-1)^{n+1}a_n \to 0$, which implies $a_n \to 0$.

Conversely, we separate the sequence of partial sums $(s_n)_{n\geq 0}$ into two sequences, the sequence of even partial sums $(s_{2n})_{n\geq 0}$ and the sequence of odd partial sums $(s_{2n+1})_{n\geq 0}$. As the terms a_n are decreasing, the even and odd partial sums of the alternating series are monotonic increasing and decreasing, respectively

$$s_{2n+2} - s_{2n} = a_{2n+1} - a_{2n+2} \geq 0, \qquad s_{2n+1} - s_{2n-1} = -a_{2n} + a_{2n+1} \leq 0$$

Also all the even partial sums are less than all the odd ones,

$$0 = s_0 \leq s_2 \leq s_4 \leq \dots \leq s_{2n} \leq s_{2n} + a_{2n+1} = s_{2n+1} \leq \dots \leq s_5 \leq s_3 \leq s_1$$

[3] The relevant test should be applied to the series of the terms taken in absolute value. For instance, suppose we want to check for absolute convergence of the series $\sum_{n=1}^{\infty} a_n$ using the ratio test. Then, we just apply the test to the series $\sum_{n=1}^{\infty} |a_n|$, that is, we evaluate the limit $\rho = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$. You can try this with the alternating harmonic series.

Since the even ones are monotonic increasing bounded above (by s_1)
and the odd ones are monotonic decreasing bounded below (by s_0)
we have convergence of bounded monotonic sequences:

$$s_{2n} \to L, \qquad s_{2n+1} \to L', \qquad \textbf{and} \;\; a_{2n+1} = s_{2n+1} - s_{2n} \to L' - L.$$

If $a_n \to 0$ then $a_{2n+1} \to 0$ so $L = L'$ and the sequence of partial sums (the whole sequence, not just the odd or even parts) converges. That is, the series converges to $s = L = L'$. □

Example 7.27 (The alternating harmonic series)**.** The series $\displaystyle\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ converges.

Solution We apply Theorem 7.26. We observe that $(1/n)_{n\in\mathbb{N}}$ is a decreasing sequence (as, simply, $\frac{1}{n} > \frac{1}{n+1}$ for all $n \in \mathbb{N}$), and that $\displaystyle\lim_{n\to\infty}\frac{1}{n} = 0$ (see Example **??**, for details). Thus, as all hypothesis of Theorem 7.26 are satisfied, the series converges. □

Note, that although $\displaystyle\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ converges, the Harmonic series $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n}$ diverges, as shown in Example 7.7. Hence, the alternating harmonic series is an example of a conditionally convergent series.

There is a profound difference between absolute and conditional convergence. A series $\sum_{n=1}^{\infty} a_n$ converges absolutely if its terms $(a_n)_{n\geq 1}$ decrease in size fast enough that their sum can be finite even if no cancellation occurs due to terms of opposite sign. On the contrary, we have conditional convergence when cancellation *is* required to make the series converge as the terms decrease too slowly, as in the case of the alternating harmonic series. As a consequence, if we rearrange the terms of a conditionally convergent series, we may get different results! This is not the case for absolutely convergent series: all rearrangements of an absolutely convergent series converge absolutely to the same sum. In sharp contrast, a series that is only conditionally convergent can be rearranged to converge to any number we please[5].

Theorem 7.28 (Rearrangments of series)**.** *Let $\sum_{n=1}^{\infty} a_n$ be a convergent series.*

- *If the series converges absolutely than every rearrangement[6] of its terms converges to the same limit.*
- *If the series is conditionally convergent, then its terms can be rearranged so as to make the series converge conditionally to any real number or diverge to $\pm\infty$.*

The next result addresses the question of multiplication of series. Notice that it only applies to absolutely convergent series.

Theorem 7.29 (product rule)**.** *Let $\displaystyle\sum_{n=1}^{\infty} a_n$ and $\displaystyle\sum_{n=1}^{\infty} b_n$ be absolutely convergent series. Define $c_n = \displaystyle\sum_{k=1}^{n} a_k b_{n-k}$. Then the series $\displaystyle\sum_{n=1}^{\infty} c_n$ converges, and we have*

[5] This fact was proven by the german mathematician B. Riemann in 1867.

[6] A rearrangment is obtained by applying a permutation to the set of terms $(a_n)_{n\geq 1}$. A permutation is a bijection from the set of positive integers to itself.

$$\sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} a_n \sum_{n=1}^{\infty} b_n.$$

Proof. The proof is beyond of the scope of the present notes. □

Let us now apply Theorem 7.29 to calculate the "square" of the Geometric series.

Example 7.30. Show that, for $-1 < \gamma < 1$, we have $\displaystyle\sum_{n=0}^{\infty}(n+1)\gamma^n = \left(\sum_{n=0}^{\infty}\gamma^n\right)^2 = \left(\frac{1}{1-\gamma}\right)^2$.

Proof. We know from Example 7.3 that $\displaystyle\sum_{n=0}^{\infty}\gamma^n = \frac{1}{1-\gamma}$, for $-1 < \gamma < 1$. We apply Theorem 7.29 with $a_n = \gamma^n$ and $b_n = \gamma^n$, for all $n \in \mathbb{N}$, and for a γ such that $-1 < \gamma < 1$. We calculate the c_n's from Theorem 7.29:

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n} \gamma^k \gamma^{n-k} = \sum_{k=0}^{n} \gamma^n = \gamma^n \sum_{i=0}^{n} 1 = \gamma^n(\underbrace{1+1+\cdots+1}_{n+1 \text{ times}}) = (n+1)\gamma^n.$$

Thus,

$$\sum_{n=0}^{\infty}(n+1)\gamma^n = \sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \left(\sum_{n=0}^{\infty}\gamma^n\right)\left(\sum_{n=0}^{\infty}\gamma^n\right) = \left(\sum_{n=0}^{\infty}\gamma^n\right)^2 = \left(\frac{1}{1-\gamma}\right)^2.$$

□

Exercises

Example 7.31. Determine the convergence properties of the series $\displaystyle\sum_{n=1}^{\infty}\frac{\cos(n\pi)}{n^2}$ justifying your assertion.

Example 7.32. Determine the convergence properties of the series $\displaystyle\sum_{n=1}^{\infty}(-1)^n\frac{1}{n^p}$ in function of the parameter $p > 0$ justifying your assertions.

Example 7.33 (Estimating the sum of an alternating series)**.** Let $(a_n)_{n\geq 1}$ a decreasing and positive sequence such that $\lim_{n\to\infty} a_n = 0$ (these are the hypothesis of Theorem 7.26) and consider the convergent alternating series $s = \sum_{n=1}^{\infty}(-1)^{n+1}a_n$. Show that:

$$|s - s_n| \leq a_{n+1} \qquad \forall n \in \mathbb{N}. \tag{7.5}$$

Example 7.34. Show that the series $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{1+2^n}$ is convergent. Use (7.5) to estimate the error committed by truncating its sum by considering the first 6 terms only.

Example 7.35. As we will see in the next section, $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln 2$, that is the alternating harmonic series is conditionally convergent to $\ln 2$. Find a rearrangement of the alternating harmonic series that makes it

- divergent to ∞;
- convergent to $\frac{\ln 2}{2}$. *[Hint: your rearrangment should be such that $\frac{1}{2}$ can be factorised. For instance, the three terms $1 - \frac{1}{2} - \frac{1}{4} = \frac{1}{2} - \frac{1}{4} = \frac{1}{2}\left(1 - \frac{1}{2}\right)$]*

7.3 Power series

Let us start with the definition.

Definition 7.36 (Power series)**.** A series of the form

$$\sum_{n=0}^{\infty} a_n (x-c)^n, \tag{7.6}$$

defines a power series of $x - c$ or a power series about $c \in \mathbb{R}$. The constants $(a_n)_{n\geq 1}$ are the coefficients of the power series and c is its centre[7].

The power series can be seen as a function of x at all $x \in \mathbb{R}$ for which the series converges. For each x we get a series of real numbers and the question if such series converges (conditionally or absolutely) arises. So, in principle, the study of a power series involves the study of the convergence of as many series as there are real numbers! Fortunately, the behaviour of power series is, in facts, very simple, as the following results show. Essentially only three possibilities can arise: a power series can either converge only at its centre[8] or in an interval centred in c, or on the whole real line.

Lemma 7.37 (Abel)**.** *Assume that the power series $\sum_{n=0}^{\infty} a_n (x-c)^n$ converges at some $x_0 \neq c$. Then it converges absolutely for every x that is closer to c than x_0 is, that is*

$$\textit{convergence at } x_0 \quad \Rightarrow \quad \textit{absolute convergence} \quad \forall x \in \mathbb{R} : |x-c| < |x_0 - c|. \tag{7.7}$$

Proof. Since, by hypothesis, the series $\sum_{n=0}^{\infty} a_n (x_0 - c)^n$ is convergent, we must have (by the vanishing condition) that $\lim_{n\to\infty} a_n (x_0 - c)^n = 0$ and thus (as convergent sequences are bounded, Theorem **??**) there exists $K \in \mathbb{R}$ such that $|a_n (x_0 - c)^n| \leq K$. Let

$$r := \frac{|x-c|}{|x_0 - c|} \quad \Rightarrow \quad r < 1 \quad \text{by (7.7)}.$$

[8]Note that all power series converge at their centre simply because when $x = c$ we get the sereis $a_0, 0, 0, \ldots$ which sums up to a_0 (here, we are using the convention that $0^0 = 1$).

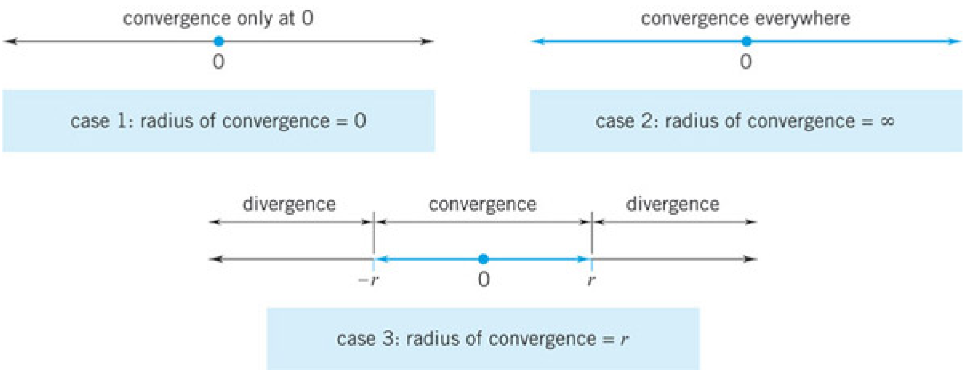


Figure 7.2: Interval of convergence for power series.

Then,

$$\sum_{n=0}^{\infty} |a_n(x-c)^n| = \sum_{n=0}^{\infty} |a_n(x_0-c)^n| \left|\frac{x-c}{x_0-c}\right|^n \leq K \sum_{n=0}^{\infty} r^n = K\frac{1}{1-r} < \infty,$$

as this is a convergent Geometric series. Hence we have proven that $\sum_{n=0}^{\infty} |a_n(x-c)^n|$ converges absolutely. □

The Lemma together with the fact that a power series clearly converges absolutely to its centre, yields the following crucial theorem summarising the convergence properties of power series.

Theorem 7.38. *For any power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ one of the following must hold:*

(i) *the series converges only at $x = c$, or*

(ii) *the series converges for all $x \in \mathbb{R}$, or*

(iii) *there exists $r > 0$ such that the series converges for every x such that $|x-c| < r$ and diverges for every x such that $|x-c| > r$. In this case the series may or may not converge at either of the two endpoints $x = c \pm r$.*

Convergence is always absolute, except possibly at the endpoints.

The three cases above can be put together by saying that the set of values for which a power series $\sum_{n=0}^{\infty} a_n(x_0-c)^n$ converges is an interval centred at c and of radius equal to (i) 0, (ii) ∞, or (iii) some $0 < r < \infty$; the three cases are depicted in Figure 7.2. The interval is named interval of convergence and the radius of the interval is named radius of convergence. If the radius is not zero or infinite, the interval of convergence can take any of the following forms:

$$[c-r, c+r], \text{ or } (c-r, c+r], \text{ or } [c-r, c+r), \text{ or } (c-r, c+r).$$

In conclusion, all we need to do in order to establish were a given power series converges is to determine its radius of convergence, and probably to study the convergence also at the end points of the interval of convergence.

Often, the radius of convergence can be found by applying the ratio test. Applied to the series $\sum_{n=0}^{\infty} |a_n(x-c)^n|$ the ratio test yields convergence if the following limit exists and is strictly less than one, so

$$\lim_{n\to\infty}\left|\frac{a_{n+1}(x-c)^{n+1}}{a_n(x-c)^n}\right| = |x-c|\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| < 1 \qquad \Leftrightarrow \qquad |x-c| < 1\Big/\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = r$$

Thus we have proved the following.

Theorem 7.39 (of D'Alembert/Ratio test for power series)**.** *The radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ is given by*

$$r = 1/\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|,$$

if such limit exists, even if the limit is equal to 0 *or* $+\infty$ *in which case the radius of convergence is equal to* $+\infty$ *and* 0*, respectively.*

As similar result based on applying the root test is left as an exercise, *cf.* Exercise 7.43

Example 7.40. Study the convergence of the following power series

$$\sum_{n=0}^{\infty}\frac{1}{n!}x^n, \qquad \sum_{n=1}^{\infty}\frac{n^2+1}{n^2+4}x^n, \qquad \sum_{n=0}^{\infty} n!x^n.$$

Solution For the first, we have shown in Example 7.18 that

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{\frac{1}{(n+1)!}}{\frac{1}{n!}} = \frac{n!}{(n+1)!} = \frac{1}{n+1}, \text{ and, therefore } \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{1}{n+1} = 0,$$

which gives $\frac{1}{r} = 0$, i.e., $r = \infty$ and, thus the interval of convergence is $(-\infty,\infty) = \mathbb{R}$. For the second, we have

$$\frac{1}{r} = \lim_{n\to\infty}\left|\frac{\frac{(n+1)^2+1}{(n+1)^2+4}}{\frac{n^2+1}{n^2+4}}\right| = \lim_{n\to\infty}\left|\frac{((n+1)^2+1)(n^2+4)}{((n+1)^2+4)(n^2+1)}\right| = \lim_{n\to\infty}\left|\frac{n^4+2n^3+6n^2+8n+8}{n^4+2n^3+6n^2+2n+5}\right| = 1.$$

Thus $r = 1$, and therefore the power series converges (at least) for $x \in (-1,1)$. Does it converges at the interval of convergence endpoints? For $x = 1$ we get the series $\sum_{n=1}^{\infty}\frac{n^2+1}{n^2+4}$. As the $\lim_{n\to\infty}\frac{n^2+1}{n^2+4} = 1 \neq 0$, this series does not converge. Similarly, the series obtained with $x = -1$ does not converge as its defining sequence has no limit. In conclusion, this series converges (absolutely) only for $x \in (-1,1)$.

Finally, for the third series, the ratio tests yields

$$\frac{1}{r} = \lim_{n\to\infty}|\frac{(n+1)!}{n!}| = \lim_{n\to\infty}(n+1) = \infty.$$

Hence, the radius of convergence of this power series is 0. The series converges only at its centre $c = 0$. □

Notice that the first series of Example 7.40 is the series considered in (**??**) as a *representation* of the exponential function. Now we know that such series does converge for every $x \in \mathbb{R}$.

Indeed a power series can be considered as a function defined on its interval of convergence. The following theorem establishes the rules of differentiation and integration of power series: essentially it shows that power series can be considered as infinite polynomials.

Theorem 7.41 (differentiation and integration of power series)**.** *Let $r > 0$ be the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(x-c)^n$, and define the function $f : (-r, r) \to \mathbb{R}$ as*

$$f(x) := \sum_{n=0}^{\infty} a_n(x-c)^n.$$

Then f is infinite times differentiable in $(-r, r)$ and

$$f'(x) = \left(\sum_{n=0}^{\infty} a_n x^n\right)' = \sum_{n=1}^{\infty} n a_n x^{n-1}, \tag{7.8}$$

and, in general,

$$f^{(m)}(x) = \left(\sum_{n=0}^{\infty} a_n x^n\right)^{(m)} = \sum_{n=m}^{\infty} n(n-1)\dots(n-m+1)a_n x^{n-m}. \tag{7.9}$$

The series in (7.8) *and* (7.9) *have radius of convergence r.*
Similarly, the function $F : (-r, r) \to \mathbb{R}$

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1},$$

is defined by a power series with radius of convergence r and is such that

$$\int f(x)dx = \int \left(\sum_{n=0}^{\infty} a_n(x-c)^n\right) dx = \sum_{n=0}^{\infty} \left(\int a_n(x-c)^n dx\right) = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-c)^{n+1} + C = F(x) + C. \tag{7.10}$$

Proof. The proof is beyond the scope of these notes. □

Example 7.42. Find the power series representation of

(i) $\frac{1}{(1-x)^2}$,

(ii) $\ln(1+x)$,

by differentiation and integration of the geometric series

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \qquad -1 < x < 1.$$

State the range of validity of the power series representation you have obtained.

Solution

(i) $\dfrac{1}{(1-x)^2} = \left(\dfrac{1}{(1-x)}\right)' = \left(\sum_{n=0}^{\infty} x^n\right)' = \sum_{n=0}^{\infty}(x^n)' = \sum_{n=1}^{\infty} nx^{n-1} = \sum_{n=0}^{\infty}(n+1)x^n$ for $-1 < x < 1$. Notice that this is the same result found in Example 7.30 using the product rule for absolutely convergent series, Theorem 7.29.

(ii) Here the idea is to use the fact that $\int \frac{1}{1+x} dx = \ln(1+x) + C$ and

$$\frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty}(-x)^n = \sum_{n=0}^{\infty}(-1)^n x^n \qquad -1 < x < 1.$$

Hence,

$$\ln(1+x) = \int \frac{1}{1+x} dx = \int \left(\sum_{n=0}^{\infty}(-1)^n x^n\right) dx = \sum_{n=0}^{\infty}\left(\int (-1)^n x^n dx\right) = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{n+1} + C,$$

having again radius of convergence 1. At $x = 0$ both $\ln(1+x)$ and the series on the right-hand side are 0, hence $C = 0$ giving

$$\ln(1+x) = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{n+1}.$$

Notice that for $x = -1$ the above series reduces to the harmonic series while for $x = 1$ it reduces to the alternating harmonic series. Hence, as an aside, we have proven[9] that

$$\ln 2 = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}.$$

In conclusion, the above representation of the function $\ln(1+x)$ is valid for $-1 < x \leq 1$.

□

Exercises

Example 7.43. (Cauchy's Theorem) By applying the root test to the generic power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ with real coefficients, deduce a criterion for determining the radius of convergence.

7.4 Taylor series

We start by recalling Taylor's theorem[10].

[9] In facts, we have only proven that the series converges for $x = 1$ and *not* that the sum is indeed equal to $\ln 2$. To prove the latter we need the theorem that says that a power series is a *continuous* function on its domain of definition.

[10] You have already seen this theorem in Calculus & Analysis I.

Theorem 7.44 (Taylor's Theorem)**.** *Let $a < b$ define the interval $I = [a, b]$, $c \in I$, and $f : I \to \mathbb{R}$ be a continuous function such that the derivatives up to order $n+1$ (inclusive) exist at every point in I. Then, for every $x \in I$, there exists ξ between c and x, such that*

$$f(x) = P_n(x) + R_n(x), \tag{7.11}$$

where

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k$$

is the Taylor polynomial of degree n of f at c (Maclaurin polynomial of degreen n of f if $c = 0$) and, for some ξ between c and x,

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}, \tag{7.12}$$

is the Taylor residual of n^{th} order of f at c and is such that

$$\lim_{x \to c} \frac{R_n(x)}{(x-c)^n} = 0. \tag{7.13}$$

The representation given by (7.11) with R_n given by (7.12) is called Taylor's formula with Lagrange remainder[11].

We already discussed some properties of the Taylor polynomial right at the beginning of the chapter. All derivatives of P_n up to order n coincide with those of f at $x = c$, and this fact led us to predict that P_n should give a good approximation of f, at least in the vicinity of c. The remainder in Taylor's formula quantifies the error committed by approximating f by its Taylor polynomial, showing that this is in facts smaller that $(x-c)^n$. In particular, if f is infinitely differentiable and for all x such that $\lim_{n\to\infty} R_n(x) = 0$, the residual can be made arbitrarily small by taking larger and larger n.

The quantification of the residual given by (7.13), is so important that we have a notation to express it.

Definition 7.45 (The little-o notation)**.** We say that a function $g(x)$ is little-$o(x^n)$ (read 'little oh of x to the power n') and we write

$$g(x) = o(x^n) \quad \text{if} \quad \lim_{x \to 0} \frac{g(x)}{x^n} = 0. \tag{7.14}$$

So, Theorem 7.44 is stating that $R_n = o((x-c)^n)$. Furthermore, Equation (7.12) can be used to bound the error made when approximating a function by its Taylor polynomial of order n, as:

$$|f(x) - P_n(x)| = |R_n(x)| \le \max_{t \in J} |f^{(n+1)}(t)| \frac{|(x-c)^{n+1}|}{(n+1)!}, \tag{7.15}$$

hence the right-hand side of the above inequality can be used to evaluate the distance between $f(x)$ and $p_n(x)$.

[11] Joseph–Louis Lagrange (1736–1813) is known as a French mathematician, but in fact he was Italian: his original name was Giuseppe Luigi Lagrangia! The reason why we give a name to the remainder formula is that various other formulas can be used.

Example 7.46. Show that the exponential function admits the power series expansion given in (**??**)

Solution Let $f(x) = e^x$. In Example **??** we saw that the Maclaurin polynomial (Taylor polynomial at $c = 0$) of degree n of f is given by $P_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$. As f is infinitely differentiable (i.e., $f^{(n)}$ exist for all $n \in \mathbb{N}$), we can consider the Taylor formula of f of any order, that is

$$e^x = P_n(x) + R_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^{\xi}}{(n+1)!} x^{n+1}, \tag{7.16}$$

is true for all $n \in \mathbb{N}$. It is not too difficult to prove that for all $x \in \mathbb{R}$ the $\lim_{n\to 0} R_n(x) = 0$ (this is Exercise 7.51). Taking limits in (7.16) it immediately follows that

$$e^x = \lim_{n\to\infty} e^x = \lim_{n\to\infty} P_n(x) + \lim_{n\to\infty} R_n(x) = \lim_{n\to\infty} P_n(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k,$$

for all $x \in \mathbb{R}$. This is the power series representation of the exponential function. $\square$

The above discussion and example shows that **if** $R_n \to 0$ as $n \to \infty$, then a function f can be approximated by its Taylor polynomial to arbitrary precision! Moreover, taking the approximation to the limit as $n \to \infty$ yields a power series representation of the given function! Thus, it is possible to use Taylor's Theorem in order to construct power series expansions of well known functions.

Definition 7.47. Let I be an open interval, $c \in I$, and $f : I \to \mathbb{R}$ an infinitely differentiable function. We define the Taylor series of f at the point x about the point c as

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!} (x-c)^n.$$

When $c = 0$, i.e., when we have

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

the Taylor series is sometimes also called MacLaurin series.

In conclusion, the Taylor series of a function f is a power series. As seen in Example 7.46, in order to express an (infinitely differentiable) function f into power series, we can find its Taylor series and then prove that the Taylor residual tends to 0 as $n \to \infty$. The obtained representation is then valid in the interval of convergence of the power series.

Now the following question arises: can there be power series representation of a function *other* than the function's Taylor series? The answer is **NO**, as the following theorem states.

Theorem 7.48. *Suppose*

$$f(x) = \sum_{n=0}^{\infty} a_n (x-c)^n \qquad \textit{for} \quad x \in (c-r, c+r), \tag{7.17}$$

where $r > 0$. *Then,*

$$a_n = \frac{f^{(n)}(c)}{n!} \qquad n = 0, 1, 2, \ldots \tag{7.18}$$

Proof. To prove this result, all we need to do is to differentiate term by term (7.17) and evaluate the result at $x = c$, exactly as was done for polynomials at the beginning of Section **??**. Theorem 7.41 ensures that each differentiation is allowed and yields a series having the same radius of convergence of the original series. $\square$

The notion of functions that coincides with their Taylor series is so important that we give it a name.

Definition 7.49. A function f is analytic at c if the Taylor series of f at c converges to f in an open interval containing c.

The following are few Taylor series expansion of elementary functions.

Example 7.50. These are Maclaurin (power) series you must remember together with their interval of validity:

$$\tfrac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 \cdots \qquad \forall x \in (-1,1),$$

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \cdots \qquad \forall x \in \mathbb{R},$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} \cdots \qquad \forall x \in \mathbb{R},$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \cdots \qquad \forall x \in \mathbb{R},$$

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} \cdots \qquad \forall x \in (-1,1],$$

$$\tan^{-1} x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} \cdots \qquad \forall x \in [-1,1],$$

$$\sinh x := \frac{e^x - e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1} = x + \frac{x^3}{3!} + \frac{x^5}{5!} \cdots \qquad \forall x \in \mathbb{R},$$

$$\cosh x := \frac{e^x + e^{-x}}{2} = \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} \cdots \qquad \forall x \in \mathbb{R}.$$

Exercises

Example 7.51. Show that the remainder of the Taylor formula at $c = 0$ of the exponential function $f(x) = e^x$ tends to zero for all $x \in \mathbb{R}$ as $n \to \infty$ by proving that

$$\lim_{n\to\infty} \left| \frac{e^\xi}{(n+1)!} x^{n+1} \right| = 0,$$

where ξ is fixed.

Example 7.52. Show that the Maclaurin series representation of $\sin x$ given in Example 7.50 is defined on the whole real line. Deduce from the Maclaurin representation of $\sin x$ that of $\cos x$ and its range of validity.

Example 7.53. Use Maclaurin (Taylor centred at 0) polynomials of $f(x) = 4\sqrt{(1+x)}$ to approximate $\sqrt{17}$ to 3 decimal places. [Hint: *you should first find an x such that* $f(x) = \sqrt{17}$.]

Chapter 8

Differential calculus of functions of more than one variable

8.1 Introduction

In this chapter we are going to study real functions of several variables. In Calculus & Analysis I you have studied functions of one real variable, herein denoted single-variable function. That is functions $f : D \to \mathbb{R}$ associating to each real number x in the function's domain D a real number $y = f(x)$. Now we are going to extend such concept and the various definitions of limit, continuity, and derivative (as for multi-dimensional integration, you'll have to wait until your 3^{rd} semester) to functions of several variables, that is objects of the kind $y = f(x_1, x_2, \dots, x_n)$ in which the function's value depends on several variables.

Before getting into the definitions, let us start with a couple motivating examples.

Example 8.1. Suppose we want to describe the elevation above see level of each point on the surface of a mountain. For simplicity, suppose that the mountain just looks like a cone, with the base at see level. The altitude can be represented by the function

$$f : D \longrightarrow \mathbb{R}, \qquad z = f(x, y),$$

associating to each point in the cone's base to the corresponding altitude. Here, the cone base is represented by a subset of the real plane $D \subset \mathbb{R}^2$: this is the map of the mountain. Each point in D can be uniquely represented by a pair of coordinates (x, y).

Example 8.2. Similarly, to represent the temperature in each point of your study room, we can use a function of three variables:

$$f : D \longrightarrow \mathbb{R}, \qquad T = f(x, y, z).$$

Here, the domain $D \subset \mathbb{R}^3$ describes the room and the the output value T the temperature as a function of position in space.

Example 8.3. What if we want also to keep track on how the temperature changes with time? Well, just add one more variable, t for time, and describe the temperature as a function of both position in space and time:

$$f : D \longrightarrow \mathbb{R}, \qquad T = f(t, x, y, z).$$

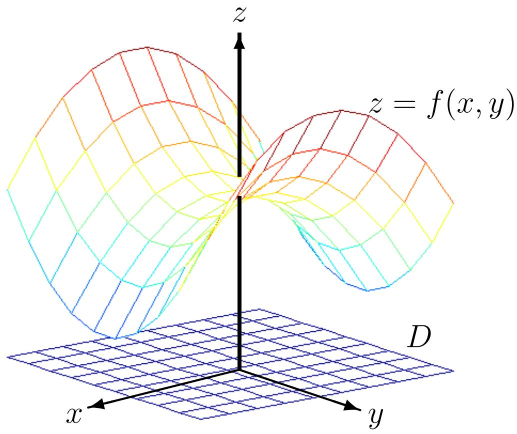


Figure 8.1: Graph of the function $f(x, y) = y^2 - x^2 + 5$.

Here, the domain $D \subset \mathbb{R}^4$ describes the time-space domain given by the room the time interval of interest.

You understand from the above examples that we can keep going forever imagining situations were more and more variables are needed in order to describe a given phenomenon.

Definition 8.4. A function f of n real variables is a rule assigning a unique real number $f(x_1, x_2, \ldots, x_n)$ to each point $(x_1, x_2, \ldots, x_n) \in D \subset \mathbb{R}^n$ called the domain of the function. The set of all real numbers $f(x_1, x_2, \ldots, x_n)$ for $(x_1, x_2, \ldots, x_n) \in D$ is called the range of f.

Often, the vector notation $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ will be used. Also, in the special cases of functions of two or three variables, we will use the more familiar notation $f(x, y)$ and $f(x, y, z)$, respectively.

How do we represent and draw a function of several variables? Well, as far as drawing is concerned, we are limited to just functions of two variables. The graph of a function of two variables (*bivariate*) $f(x, y)$ can be represented as the set of points in space having coordinates $(x, y, f(x, y))$ in function of the points $(x, y) \in D$ in the domain of the function. Hence, D can be represented as a subset of the xy-plain, and the graph of the function is the surface made of all points in space satisfying the equation $z = f(x, y)$. An example is given in Figure 8.1

Another way to represent functions is through level curves. To explain this, let us go back to our Example 8.1 of the representation of a conic-looking mountain. In a map, the altitude would be represented by lines bearing the same altitude, as shown in Figure 8.2. These are the level curves of the altitude function, which indeed we plot on the map: the xy-plane. Notice that they give a very good impression on how the mountain shape looks like.

Definition 8.5. Let $f : D \to \mathbb{R}$ with $D \subset \mathbb{R}^2$ a real function of two variables. For every real number c in the range of f, we define the c-contour or level curve of f as the set of points in the xy-plane such that

$$f(x, y) = c.$$

For example, the level curves of the function $f(x, y) = y^2 - x^2 + 5$ of Figure 8.1 are the hyperbolas $y^2 - x^2 + 5 = c$ shown in Figure 8.3

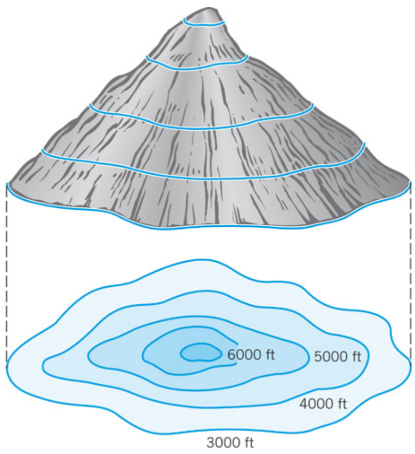


Figure 8.2: Mapping of a mountain by level curves.

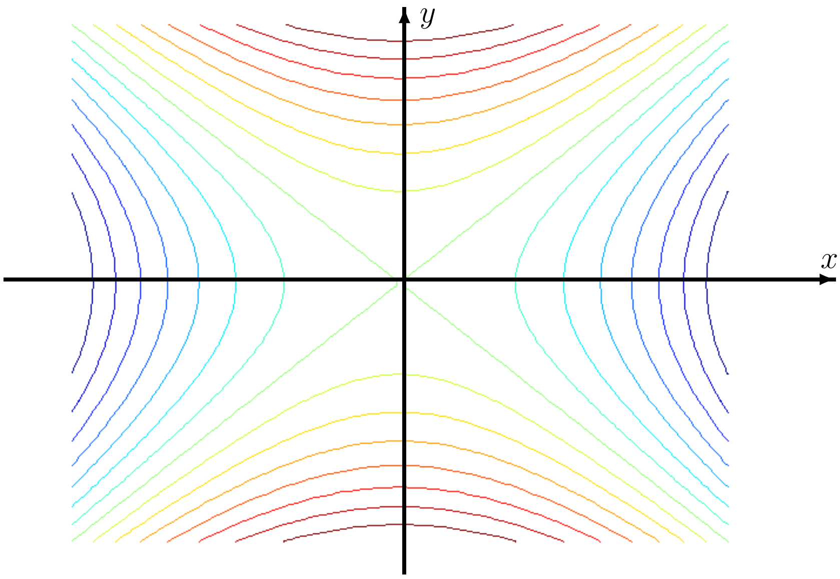


Figure 8.3: Level curves of the function $f(x, y) = y^2 - x^2 + 5$

As said above, representing functions on a sheet of paper becomes prohibitive as the number of variables grow. We are limited to level curves and graph of functions of two variables, and level surfaces (see definition below) of functions of three variables, as these still *live* in $\mathbb{R}^3$. Still we can define (and study!) them and generalise the concepts of surfaces and level curves.

The graph of a function $f : D \to \mathbb{R}$ of n variable, is the set of points of $\mathbb{R}^{n+1}$ given by

$$\text{graph}(f) = \{(\mathbf{x}, f(\mathbf{x})) \,:\, \mathbf{x} \in D\} \subset \mathbb{R}^{n+1}.$$

It is an n-dimensional surface, or hypersurface, in $\mathbb{R}^{n+1}$. A level set or level hypersurface is the set of points of $\mathbb{R}^n$ given by

$$L_c(f) = \{\mathbf{x} \in D \,:\, f(\mathbf{x}) = c\} \subset \mathbb{R}^n.$$

It is an hypersurface in $\mathbb{R}^n$.

The concepts of graphs and hypersurfaces is important, although plotting is out of question. Taking again the example of the mountain, we may want to get to the other side of the mountain without moving up and down, or instead move along it in a way to climb up as fast as possible. This can be achieved by moving parallel or somehow orthogonally to the level curve we are on.

For instance, the concept of *gradient* will answer the question on how to find the steepest direction along a surface. The gradient is a special combination of derivatives of the function: multi-variable differential calculus is what we need.

8.2 Limits and continuity

As you have seen in Calculus & Analysis I, the concept of *limit* defines the behaviour of a function as a given point is approached. Thus, in order to define limits, we need a concept of a distance. The Euclidean metric is all we need. In the real line, this was given by the absolute value of the difference. So, given $x, y \in \mathbb{R}$, their distance is $|x-y|$. This way, we could define the (open) *neighborhood* of a given real number x_0 as the set of points $B_\delta(x_0) = \{x \in \mathbb{R} : |x - x_0| < \delta\}$, for $\delta \in \mathbb{R}^+$. This is just an open interval centred at x_0, that is $(x_0 - \delta, x_0 + \delta)$.

These concepts are easily generalised to $\mathbb{R}^n$. For any $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$ in $\mathbb{R}^n$, we define their distance[1] as $\|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$. Now we are in a position to give the definition of neighbourhood.

Definition 8.6. A neighbourhood of a point $\mathbf{x}_0 \in \mathbb{R}^n$ is a set of the form

$$B_\delta(x_0) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| < \delta\},$$

where δ is some number greater than zero.

In the plane $\mathbb{R}^2$, a neighbourhood of a point $\mathbf{x}_0 = (x_0, y_0)$ is a disk centred at (x_0, y_0). In the three-space $\mathbb{R}^3$, a neighbourhood of a point $\mathbf{x}_0 = (x_0, y_0, z_0)$ is a ball centred at (x_0, y_0, z_0). In general, we say that a neighbourhood is a n-dimensional open ball centred on $\mathbf{x}_0$ and of radius δ, hence the notation $B_\delta(x_0)$.

Further *topological* definitions, such as of interior, boundary, and exterior points, open and closed sets also are obviously extended from the once given in the real line.

Definition 8.7 (Topology of $\mathbb{R}^n$)**.** Let $S \subset \mathbb{R}^n$.

- A point $x_0 \in S$ is said to be an interior point of the set S if the set S contains some neighbourhood of x_0,
- The set of all interior points of S is called the interior of S,
- A point $x_0 \in \mathbb{R}^n$ is said to be a boundary point of the set S if every neighbourhood of x_0 contains points that are in S as well as points that are not in S,
- The set of all boundary points of S is called the boundary of S,
- A point $x_0 \in S$ is said to be an isolated point of the set S if there exists neighbourhoods of x_0 that contains no other points of S (note that an isolated point is also a boundary point),

[1]Note that, for $n = 1$, that is back to the real line, $|x|$ and $\|x\|$ are just two ways to denote the very same quantity.

- A point $x_0 \in \mathbb{R}^n$ is said to be of accumulation for the set S if it is not an isolated point, that is if every neighbourhood of x_0 contains points of S other that x_0,
- The set S is said to be open if it coincides with its interior (that is, all its points are internal),
- The set S is said to be closed if it contains its boundary,

We are now ready to give the definition of limit and soon after of continuity of a function of several variables.

Definition 8.8 (Limit of a function of several variables)**.** Let $f : D \to \mathbb{R}$, with $D \subset \mathbb{R}^n$. We say that the function f approaches the limit L at $\mathbf{x}_0$ and write

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = L, \tag{8.1}$$

or, equivalently,

$$\lim_{\mathbf{h}\to\mathbf{0}} f(\mathbf{x}_0 + \mathbf{h}) = L,$$

if

(i) every neighbourhood of $\mathbf{x}_0$ contains points of the domain of f different from $\mathbf{x}_0$, that is,

$$\forall\delta > 0, \exists\mathbf{x} \in B_\delta(\mathbf{x}_0) : \mathbf{x} \in D,$$

(ii) for each positive number ε there exists a positive number $\delta = \delta(\varepsilon)$ such that $|f(\mathbf{x}) - L| < \varepsilon$ holds whenever $\mathbf{x}$ is in the domain of f and satisfies $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$, that is

$$\forall\varepsilon, \exists\delta = \delta(\varepsilon) : \forall\mathbf{x} \in D \cap B_\delta(\mathbf{x}_0) \setminus \{\mathbf{x}_0\}, |f(\mathbf{x}) - L| < \varepsilon.$$

A few comments about the limit definition:

- We do not define limits for isolated point; this is why condition (i) is there,
- If the limit exists it is unique (we refrain to do so here, but be aware that this statement should be proved),
- The algebra of limits seen for single-variable functions (sum of the limit equals limit of the sum, etc.) extend to the several variables case:
  If $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = L$ and $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = M$, then

  (i) $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} (f(\mathbf{x}) \pm g(\mathbf{x})) = L \pm M$

  (ii) $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})g(\mathbf{x}) = LM$

  (iii) $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x})}{g(\mathbf{x})} = \frac{L}{M}$ provided $M \neq 0$

  (iv) $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} F(f(\mathbf{x})) = F(L)$ provided $F(\mathbf{x})$ is continuous at L.

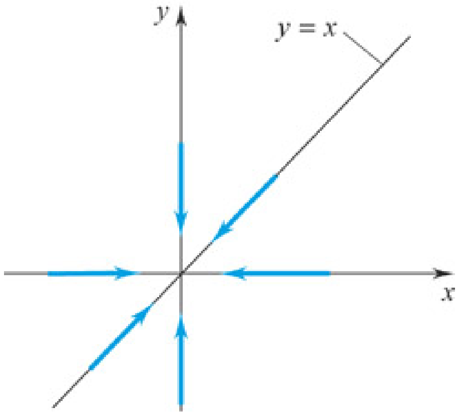


Figure 8.4: Approaching $(0,0)$ along different lines.

- For a single-variable function the existence of the limit depends on the behaviour of the function as x approaches x_0 from the left and the right, that is[2]

$$\text{for a single-variable function} \quad \lim_{x\to x_0} f(x) = L \Leftrightarrow \lim_{x\to x_0^+} f(x) = L = \lim_{x\to x_0^-} f(x).$$

Similarly, *eg.* for a function of two variables,

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L \Leftrightarrow f \text{approaches } L \text{ along any curve through } (x_0,y_0) \text{ lying in the domain of}$$

The last bullet above implies that checking limits of functions of several variables is a much more complex exercise than it is for single-variable functions.

Example 8.9. Show that the function

$$f(x,y) = \frac{2xy}{x^2+y^2},$$

does not have a limit at the origin.

Solution Notice that f is defined everywhere apart from the origin: its domain of existence is $D = \mathbb{R}^2 \setminus \{(0,0)\}$. Let us investigate the behaviour of f as we approach the origin along few special paths, namely the lines depicted in Figure 8.4. Along the x-axis, that is for $y = 0$, the function becomes $f(x,0) = 0$ for all $x \neq 0$. Thus, the limit of f along the axis $y = 0$ is equal to 0. The same is true along the y-axis. Now, f will have a limit at the origin only if we get the same limit, namely 0, as we approach from *any* other curve. This is not the case as, for instance, along the line $y = x$ we have $f(x, y = x) = f(x,x) = \frac{2x^2}{x^2+x^2} = 1$ for all $x \neq 0$. Thus the limit along $y = x$ is equal to $1 \neq 0$. □

Example 8.10. Show that the function

$$f(x,y) = \frac{x^2y}{x^2+y^2},$$

does have a limit at the origin.

[2]This is Theorem 2.5 in the Calculus & Analysis I Lecture Notes.

Solution Again, this function is defined in the whole of the real plane, excluding the origin. But, we are told that it does have a limit at the origin. Let us first find a candidate limit. We can do that, for instance, by studying the behaviour of f along the x-axis. As in the previous example, $f(x, 0) = 0$ for all $x \neq 0$, thus the limit along the x-axis is equal to 0. It remains to prove that 0 *is* the limit, that is we need to prove that

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

Since $x^2 \leq x^2 + y^2$, it holds

$$0 \leq |f(x, y) - 0| = \left|\frac{x^2y}{x^2 + y^2}\right| \leq |y| \leq \sqrt{x^2 + y^2}.$$

We are now ready to verify the formal definition of limit. For any $\varepsilon > 0$, fixing $\delta = \varepsilon$ guarantees that if $(x, y) \in B_\delta(0, 0) \setminus \{(0, 0)\}$, that is if $0 < \sqrt{x^2 + y^2} < \delta$, then $|f(x, y) - 0| < \delta = \varepsilon$.

Notice that we could have 'seen' this straight away by noticing that, as the quantity on the right-hand side tends to 0 as $(x, y) \to (0, 0)$, the limit of f at the origin must be 0 by the Pinching Theorem. $\square$

Definition 8.11. The function f is continuous at the point $\mathbf{x}_0$ if

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

We say that f is continuous on the set S if it is continuous at any point in S.

Note that the definition of continuity at one point x_0 asks for three facts to be true:

- f is defined at $\mathbf{x}_0$,
- f has limit as $\mathbf{x} \to \mathbf{x}_0$,
- the limit equals the value of f at $\mathbf{x}_0$.

Be aware that all these three facts must be checked when verifying continuity.

As the notion of continuity is based on that of limit, again we have that the sum of continuous functions is continuous, the composition of continuous functions is continuous, etc. Thus the following are obvious examples of continuous functions:

$$\begin{aligned} f(x, y) &= 1 + x + xy + xy^5 && \text{(polynomials in several variables)} \\ f(x, y, z) &= \cos(z)e^{\sin(x^2 - y)} && \text{(compositions of continuous elementary functions)}, \end{aligned}$$

while, for instance, quotients of polynomials, which are named rational functions, are continuous except where the denominator is zero. The function of Example 8.10, namely

$$f(x, y) = \frac{x^2y}{x^2 + y^2} \quad \text{is continuous except on the origin,}$$

although it does have limit at the origin. Notice that f is not defined at the origin, hence it cannot be continuous there although the limit exists. In such cases, we can extend by continuity: the following function is defined and continuous on the whole of $\mathbb{R}^2$:

$$f(x,y) = \begin{cases} \frac{x^2y}{x^2+y^2} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0). \end{cases}$$

Of course, it is not always possible to extend by continuity where the function is not defined: we need the function to have at least a limit there. For instance, although we can certainly modify the function of Example 8.9 so that it is defined at the origin:

$$f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0), \end{cases} \tag{8.2}$$

this function is defined but not continuous at the origin. Notice, though that it is continuous everywere in each of its variables separately, that is

$$\lim_{x\to x_0} f(x,y_0) = f(x_0,y_0), \qquad \lim_{y\to y_0} f(x_0,y) = f(x_0,y_0),$$

for all $x_0, y_0 \in \mathbb{R}$.

Indeed, given a function of several variables, we can always fix all but one variable and consider the single-variable function so obtained. The example above shows that, in general, continuity with respect to each variable does not imply continuity.

Exercises.

Example 8.12. Consider the function of two variables

$$f(x,y) = \frac{2x^2y}{x^4+y^2}.$$

- Discuss the domain of existence of f.
- Show that, when restricted to any line through the origin, f approaches 0 at the origin.
- Show that, f does not have a limit at the origin. *Hint: study the behaviour of f along the parabola $y = x^2$.*

8.3 Partial derivatives

As we just saw at the end of the previous section, we can always get a (*partial*) information about a function of several variables by fixing all but one variable. This technique is rather useful as we know already a whole lot about single-variable functions.

Definition 8.13 (Partial derivatives; two variables)**.** Let f be a function of two variables x, y. The partial derivatives of f with respect to x and y are, respectively, the functions denoted by f_x and f_y and defined by setting

$$f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}, \tag{8.3}$$

$$f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}. \tag{8.4}$$

The partial derivatives measures the rate of change of f with respect to one variable.

Notations for first partial derivatives The following are other notations for first partial derivatives you should be aware of:

$$f_x(x,y) = f_1(x,y) = \frac{\partial f}{\partial x}(x,y) = D_1 f(x,y),$$
$$f_y(x,y) = f_2(x,y) = \frac{\partial f}{\partial y}(x,y) = D_2 f(x,y).$$

Finding the partial derivatives of a function, is pretty straightforward if you know how to take derivatives of single-variable functions. Indeed, by definition, the partial derivative, say, with respect to x, is the derivative of the function when y is fixed. The procedure is illustrated with the following example.

Example 8.14. Find the first partial derivatives of the function $f(x,y) = x\arctan(xy) + e^{2y}$. Evaluate the partial derivatives at the point $(x,y) = (1,0)$.

Solution We have:

$$f_x(x,y) = \arctan(xy) + x\tfrac{y}{1+(xy)^2} = \arctan(xy) + \tfrac{xy}{1+(xy)^2} \quad \Rightarrow \quad f_x(1,0) = 0,$$
$$f_y(x,y) = x\tfrac{x}{1+(xy)^2} + 2e^{2y} = \tfrac{x^2}{1+(xy)^2} + 2e^{2y} \quad \Rightarrow \quad f_y(1,0) = 3.$$

□

Now look back to the Definition 8.13 of partial derivatives. Clearly, the existence of partial derivatives at one point only depends on the behaviour of the function along a line parallel to one axis through that point (*cf.* Figure 8.5 below). On the contrary, continuity depends on the behaviour of the function in all directions. Thus, it should not be surprising that for functions of several variables, the existence of partial derivatives does not guarantee continuity, as the following example shows.

Example 8.15. Show that the function in (8.2) admits partial derivatives on the whole of $\mathbb{R}^2$ and, in particular, at the origin.

Solution Away from the origin, that is far all $(x,y) \neq (0,0)$ we have:

$$f_x(x,y) = \frac{2y(x^2+y^2) - 4x^2y}{(x^2+y^2)^2} = 2\frac{y^3 - x^2y}{(x^2+y^2)^2},$$
$$f_y(x,y) = \frac{2x(x^2+y^2) - 4xy^2}{(x^2+y^2)^2} = 2\frac{x^3 - 2xy^2}{(x^2+y^2)^2}.$$

Let us now concentrate on the origin. Here, the ordinary rules for differentiation cannot be used, and we need to resort to the definition of partial derivatives as difference quotients:

$$f_x(0,0) = \lim_{x\to 0} \frac{f(x,0) - f(0,0)}{x} = \lim_{x\to 0} \frac{0-0}{x} = 0.$$

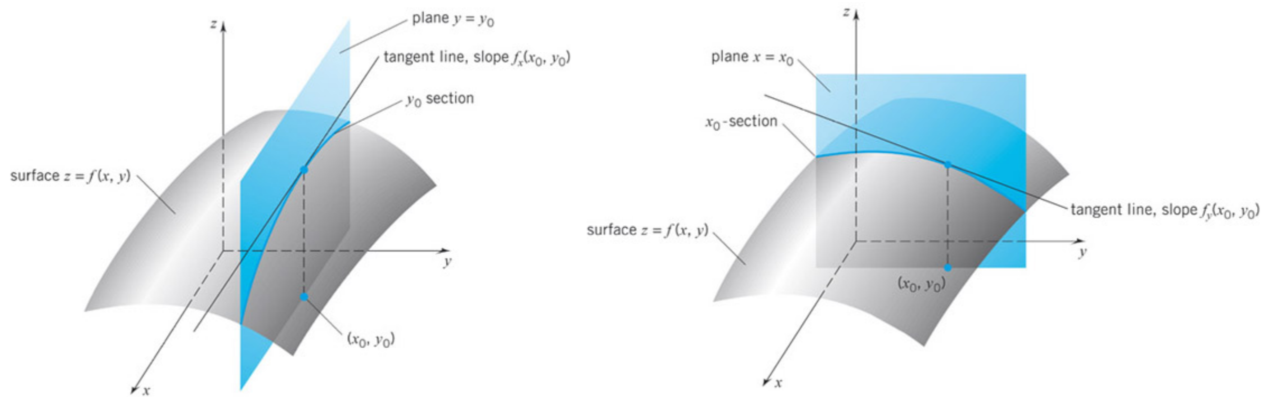


Figure 8.5: Depiction of the geometrical meaning of the first partial derivatives.

Similarly, $f_y(0,0) = 0$.

□

Thus we have verified that the existence of all partial derivatives does not imply continuity. This is in contrast with what we already know to be true for single-variable functions, namely that the existence of the derivative (that is, *differentiability*) guarantees continuity. Notice that, while single-variable functions are defined differentiable if their derivative exists, *we have not yet defined what differentiability means for function of several variables.* Such definition will be given in the next section: all will be clear then! In facts, the first step is to understand what a partial derivative is *geometrically.*

Geometric interpretation of first partial derivatives As the derivative of a single-variable function at a point gives the slope of the line tangent to the curve $y = f(x)$ at that point, the first partial derivatives give the slope of the lines tangent to the surface $z = f(x, y)$ in the direction of the coordinate axis, see Figure 8.5.

The concept of partial derivatives extends to function of three variables x, y, z. More generally, a function f of n variables $x_1, \ldots, x_n$ has n first partial derivatives given by

$$f_{x_k}(x_1, \ldots, x_n) = f_k(x_1, \ldots, x_n) = \lim_{h\to 0} \frac{f(x_1, \ldots, x_{k-1}, x_k + h, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_k, \ldots, x_n)}{h}, \tag{8.5}$$

for $k = 1, 2, \ldots, n$.

Derivatives of Higher order Again, we consider first just functions of two variables. Suppose that f is a function of the two variables x, y admitting first partial derivatives in its domain of definition. As the partial derivatives f_x and f_y are again functions of x and y, they may themselves possess partial derivatives $(f_x)_x$, $(f_x)_y$, $(f_y)_x$ $(f_y)_y$. These functions are the second-order partial derivatives of f. For these, we introduce the following notation. The two pure second partial derivatives

$$f_{xx} = f_{11} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) := (f_x)_x,$$
$$f_{yy} = f_{22} = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) := (f_y)_y,$$

and two mixed second partial derivatives

$$f_{xy} = f_{12} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) := (f_x)_y,$$
$$f_{yx} = f_{21} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) := (f_y)_x.$$

These are, by definition, calculated by taking partial derivatives of already calculated partial derivatives.

Example 8.16. Calculate the second partial derivative of the function $f(x, y) = ye^{x^2} + xy$.

Solution We start by calculating the first partial:

$$f_x(x, y) = 2xye^{x^2} + y \qquad f_y(x, y) = e^{x^2} + x,$$

and then get the second partial but taking the partial derivatives of f_x and f_y:

$$f_{xx} = (2xye^{x^2} + y)_x = (2y + 4x^2y)e^{x^2} \qquad f_{xy} = (2xye^{x^2} + y)_y = 2xe^{x^2} + 1$$
$$f_{yx} = (e^{x^2} + x)_x = 2xe^{x^2} + 1 \qquad f_{yy} = (e^{x^2} + x)_y = 0.$$

□

Notice that, in the example above, it happened that $f_{xy} = f_{yx}$. It turns out that this is not by chance: it is due to continuity, as the following theorem states.

Theorem 8.17 (Equality of mixed partials)**.** *On every open set on which f and its first and second partials f_x, f_y, f_{xy}, and f_{yx} are continuous, it holds*

$$f_{xy} = f_{yx}. \tag{8.6}$$

Proof. The proof is beyond the scope of these notes. □

The theorem generalises to function of any number n of variables.

Example 8.18. Calculate all first and second partials of $f(x, y, z) = \sin(x)yz^3$. Verify the equality of the mixed partials, namely:

$$f_{xy} = f_{yx}, \quad f_{xz} = f_{zx}, \quad f_{yz} = f_{zy}.$$

Solution The three first partial derivatives of f are:

$$\cos(x)yz^3, \quad \sin(x)z^3, \quad 3\sin(x)yz^2.$$

We get the second partials by computing, for each first partials, its three first partials:

$$f_{xx} = -\sin(x)yz^3, \quad f_{xy} = \cos(x)z^3, \quad f_{xz} = 3\cos(x)yz^2,$$
$$f_{xy} = \cos(x)z^3, \quad f_{yy} = 0, \quad f_{yz} = 3\sin(x)z^2,$$
$$f_{xz} = 3\cos(x)yz^2, \quad f_{yz} = 3\sin(x)z^2, \quad f_{zz} = 6\sin(x)yz.$$

□

Of course, the process of taking partial derivatives can be carried on, and partial derivatives of any order be calculated. For instance, Taylor polynomials and series of function of several variables can be defined. We will not pursue in this direction here[3], but you will find a hint in the next section, were linear approximation and Taylor polynomials of second order are discussed.

8.4 The differential, the gradient, and linear approximation

In this section we extend the notion of differentiability from single-variable functions to functions of several variables.

Let us start by recalling the definition of differential in the single-variable case. A function f of one variable is differentiable at $x_0 \in \mathbb{R}$ if the limit

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h},$$

exists. Moreover, if this is the case, the limit is the derivative of f at x_0, denoted by $f'(x_0)$, so

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} =: f'(x_0),$$

or, equivalently,

$$\lim_{h\to 0} \frac{f(x_0+h)-(f(x_0)+f'(x_0)h)}{h} = 0.$$

This last equation is telling us the following: as h gets smaller and smaller (that is, approaching the limit $h \to 0$), the difference between $f(x_0+h)$ and $f(x_0)+f'(x_0)h$ gets smaller than h. Using the *little-o* notation introduced in Definition 7.45, we can say that

$$f(x_0+h)-(f(x_0)+f'(x_0)h) = o(h). \tag{8.7}$$

This is clearly visible in Figure 8.6.

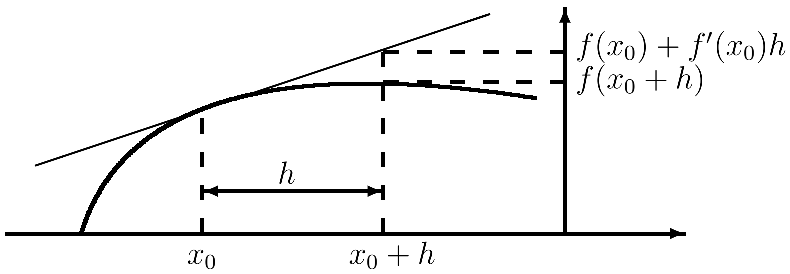


Figure 8.6: Linear approximation at a point: the single-variable case.

The most important fact that Equation (8.7) expresses and the figure shows is that differentiability implies that the line

$$L(h) := f(x_0)+f'(x_0)h,$$

[3]Multi-variable Taylor series will be a topic of Calculus & Analysis III.

gives a convenient approximation for f near x_0. We can equivalently express this in terms of the variable x:

$$L(x) := f(x_0) + f'(x_0)(x - x_0). \tag{8.8}$$

The process of approximating f by L is an instance of the procedure of LINEARIZATION that is arguably the most ubiquitous tool of mathematics!

Application: 1D Newton's method. The method of Newton is a very powerful tool for solving equations. Suppose we want to solve the (nonlinear!) equation $f(x) = 0$. So the problem is:

$$\text{find} \quad a \in \mathbb{R} \quad : \quad f(a) = 0. \tag{8.9}$$

If f is differentiable near the root a, than we can find an approximation to the root a as follows:

1) Approximate f by the line tangent to the graph of f at some point $(x_0, f(x_0))$, that is the line $y = L(x)$ with L given by (8.8).

2) Solve the **linear** equation $L(x) = 0$ to find the approximate root:

$$f(x_0) + f'(x_0)(x - x_0) = 0 \qquad \Rightarrow x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

The above steps can now be repeated for x_1 to find a hopefully better approximation x_2, and so on, until the result has the desired accuracy. This process is illustrated in Figure 8.7. Here we can see that each iterate is closer to the root than the previous one, and it is reasonable to expect that *in the limit*, the sequence $\{x_n\}_{n=0}^{\infty}$ converges to a. Be aware that this is **not** always the case. The convergence of the Newton method is guaranteed only if the function f is regular enough and the *initial guess* x_0 is close enough to the root...

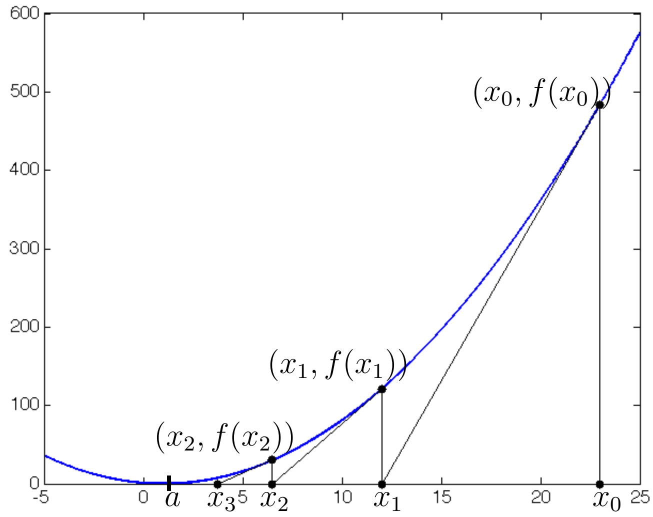


Figure 8.7: Few interations of the Newton algorithm applied to find the unique root $a = 1$ of the function $f(x) = x^2 - 2x + 1$.

Question: How can we generalise the concept of differentiability and the idea of linear approximation, to functions of several variables?

Definition 8.19. We say that the function f is differentiable at $\mathbf{x}_0$ if there exists a vector $\nabla f(\mathbf{x}_0)$ such that

$$\lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) - \nabla f(x_0)\cdot \mathbf{h}}{\|\mathbf{h}\|} = 0 \tag{8.10}$$

or, in little-o notation,

$$f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0)\cdot \mathbf{h} = o(\|\mathbf{h}\|). \tag{8.11}$$

As in the single-variable case, differentiability implies continuity.

Theorem 8.20. *If f is differentiable at $\mathbf{x}_0$, then f is continuous at $\mathbf{x}_0$.*

Proof. From (8.11) we have

$$\begin{aligned} &f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0)\cdot \mathbf{h} + o(\|\mathbf{h}\|) \\ \Rightarrow 0 \le\ &|f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0)| = |\nabla f(\mathbf{x}_0)\cdot \mathbf{h} + o(\|\mathbf{h}\|)| \le |\nabla f(\mathbf{x}_0)\cdot \mathbf{h}| + |o(\|\mathbf{h}\|)|, \end{aligned}$$

having used the Schwarz inequality, and thus $\lim_{\mathbf{h}\to\mathbf{0}}(f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0)) = 0$ by the Pinching Theorem as the two terms on the right-hand side clearly tend to zero as $\mathbf{h} \to \mathbf{0}$. □

Theorem 8.21. *If f is differentiable at $\mathbf{x}_0$, the* **vector** *$\nabla f(\mathbf{x}_0)$ satisfying* (8.10) *is* **unique**. *Such vector is called the gradient of f at $\mathbf{x}_0$.*
The gradient vector $\nabla f(\mathbf{x}_0)$ has component equal to the corresponding partial derivative. In particular:

- *if f is function of two variables then*

$$\nabla f(\mathbf{x}_0) = (f_x(\mathbf{x}_0) \quad f_y(\mathbf{x}_0)) = f_x(\mathbf{x}_0)\mathbf{i} + f_y(\mathbf{x}_0)\mathbf{j}, \tag{8.12}$$

- *if f is function of three variables then*

$$\nabla f(\mathbf{x}_0) = (f_x(\mathbf{x}_0) \quad f_y(\mathbf{x}_0) \quad f_z(\mathbf{x}_0)) = f_x(\mathbf{x}_0)\mathbf{i} + f_y(\mathbf{x}_0)\mathbf{j} + f_z(\mathbf{x}_0)\mathbf{k}, \tag{8.13}$$

where $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ are the coordinate vectors.

Proof. The proof is behind the scope of these notes. □

Checking for differentiability and calculating the gradient using the definition of differentiability is a rather laborious process as it involves finding a limit in n-variables. Theorem 8.21 shows that the gradient can be easily calculated from the partial derivatives. The next theorem provides sufficient (but not necessary!) condition for differentiability. It can be used to check for differentiability in most cases.

Theorem 8.22. *If f has continuous first partial derivatives in a* **neighborhood** *of $\mathbf{x}_0$, then f is differentiable at $\mathbf{x}_0$.*

Proof. The proof is behind the scope of these notes. □

Remark 8.23. *Bear in mind that our definition of the gradient is that given in Theorem 8.21; by definition, in these notes the gradient exists only if the function is differentiable*[4].

Let us now go back to the question of differentiability and linear approximation of a function in the neighbourhood of a point. Exactly as done for single-variable functions, we consider Equation (8.11) (the analog of (8.7)). This tells us that, if f is differentiable at x_0, then

$$L(\mathbf{x}_0) := f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0), \tag{8.14}$$

gives a convenient approximation for f near $\mathbf{x}_0$. This is the equation of the tangent hyperplane to f at $\mathbf{x}_0$. In particular for functions of two variables we have the following.

Equation of the tangent plane Given the surface $z = f(x, y)$, if f is differentiable at (x_0, y_0) then

$$z = L(x, y) := f(x_0, y_0) + \nabla f(x_0, y_0) \cdot ((x - x_0) \quad (y - y_0)), \tag{8.15}$$

is the equation of the tangent plane to $z = f(x, y)$ at (x_0, y_0). Using (8.13) we can write this in terms of the first partial derivatives:

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0). \tag{8.16}$$

A depiction of a tangent plane to a surface is given in Figure 8.8. The tangent plane:

- passes through the point in space

$$P = (x_0, y_0, f(x_0, y_0)),$$

- is parallel to the two vectors

$$\mathbf{t}_1 = \mathbf{i} + f_x(x_0, y_0)\mathbf{k} \quad \text{and} \quad \mathbf{t}_2 = \mathbf{j} + f_y(x_0, y_0)\mathbf{k}.$$

- has normal vector

$$\mathbf{n} = \mathbf{t}_2 \times \mathbf{t}_1 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 1 & f_y(x_0, y_0) \\ 1 & 0 & f_x(x_0, y_0) \end{vmatrix} = f_x(x_0, y_0)\mathbf{i} + f_y(x_0, y_0)\mathbf{j} - \mathbf{k}.$$

Example 8.24. Find the tangent plane and normal vector to $z = f(x, y) = 1 - \frac{1}{4}x^2 - y^2$ at $(x_0, y_0) = (1, 1/2)$.

Solution First of all, we notice that the function f is everywhere differentiable. To find the tangent plane and normal, we first calculate the function at the given point, so $f(1, 1/2) = 1/2$, and then the first partial derivatives at the given point:

$$\begin{aligned} f_x(x, y) &= -\tfrac{x}{2} \qquad \Rightarrow \qquad f_x(1, 1/2) = -\tfrac{1}{2} \\ f_y(x, y) &= -2y \qquad \Rightarrow \qquad f_y(1, 1/2) = -1. \end{aligned}$$

[4]Some texts define the gradient as just the collection of first partials.

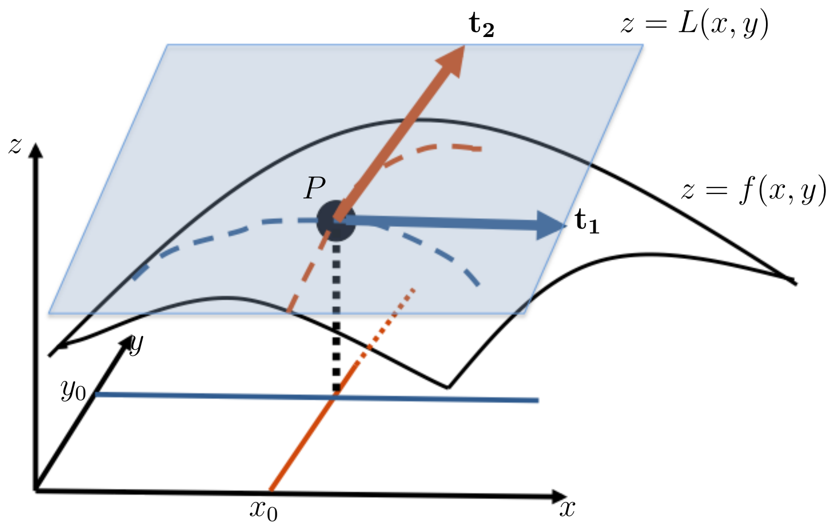


Figure 8.8: The tanget vector to the surface $z = f(x, y)$ at (x_0, y_0).

Thus, the equation of the tangent plane is

$$z = \frac{1}{2} - \frac{1}{2}(x-1) - (y - \frac{1}{2}) = \frac{3}{2} - \frac{1}{2}x - y.$$

The normal vector is

$$\mathbf{n} = -\frac{1}{2}\mathbf{i} - \mathbf{j} - \mathbf{k}.$$

$\square$

Exercises.

Example 8.25. The continuity of the first partial derivatives implies differentiability. Show that the opposite implication is not true by analysing the properties at the origin of the function given by

$$f(x, 0) = 0 \qquad \text{and} \qquad f(x, y) = y^2 \cos\frac{1}{y} \quad \text{if } y \neq 0.$$

Example 8.26. Find the maximum slope on the surface $z = x^2 y \ln(x + y)$ at the point $(2, 0)$ and the slope in the direction of the vector $\mathbf{v} = (3, 4)$. Find the tangent plane and the normal to the surface at the same point.

8.5 Chain rules

You are familiar with the chain rule for calculating the derivative of compositions of single-variable functions. Given to functions $f(x)$ and $g(t)$, if g is differentiable at some t and f is differentiable at $x = g(t)$, then the derivative of the composite function $f[g(t)]$ is given by the chain rule:

$$\frac{d}{dt}(f[g(t)]) = f'[g(t)]g'(t). \tag{8.17}$$

This can be re-written using the Leibniz notation, which is advantageous in our context. As f is function of x, which is function of t (through the law given by g), we can write (8.17) as

$$\frac{df}{dt} = \frac{df}{dx}\frac{dx}{dt} \quad \text{to mean} \quad \frac{df}{dt}[x(t)] = \frac{df}{dx}[x(t)]\frac{dx}{dt}(t). \tag{8.18}$$

Here you will learn generalisations of the chain rule for functions of several variables. To start with, let us motivate with an example, referring back to our bivariate function describing the elevation of a mountain.

Example 8.27. Let $z = f(x, y)$ be the function describing the elevation of a mountain above see level as in Example 8.1. Suppose to be walking along a trail climbing up the mountain and that the trail position in function of time on the xy-plain (the map) is given by

$$x = u(t) \quad \text{and} \quad y = v(t).$$

We call this *parametric equations of the trail on the xy-plane map*, with respect to the parameter t, here representing the time variable. At time t, the elevation reached is given by the composite function

$$z = f[u(t), v(t)] =: g(t).$$

The derivative of the function $g(t)$ tells us how fast we are climbing up the mountain, how fast is our elevation changing. To compute that we need a chain rule for the derivative of composite functions were one of the function (the outer one in this case) is bivariate.

By reading the example, you may have realised that we may think of a number of composition of functions. Here we will just consider two cases:

- composition of single-variable functions with a function of several variable:

$$z = f[u(t), v(t)],$$

- composition of functions of several variable with a function of several variable:

$$z = f[u(s, t), v(s, t)].$$

Here we used the example of bivariate functions, but both our chain rules will be generalised to n-variate functions.

Theorem 8.28 (The chain rule I). *If $z = f[x, y]$ has continuous first partial derivatives on an open set $U \subset \mathbb{R}^2$ and $x = u(t), y = v(t)$ are differentiable functions of t whose range is contained in U (so, whenever $(x, y) \in U$), then the composition function is differentiable in t and*

$$\frac{dz}{dt} = \nabla f[u(t), v(t)] \cdot (u'(t) \quad v'(t)), \tag{8.19}$$

or, in Leibniz notation,

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \tag{8.20}$$

Derivatives of composite functions can be calculated by chain rule or by direct substitution.

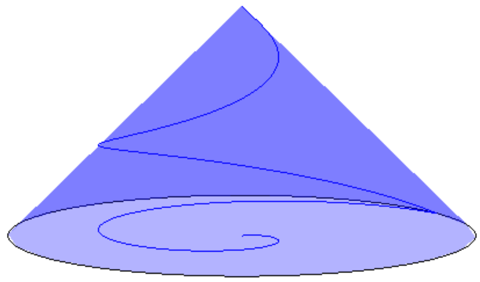


Figure 8.9: The trail of Example 8.29

Example 8.29. Let us go back to the mountain climbing example. Assume that the mountain elevation is given by $z = f(x, y) = 1 - \sqrt{x^2 + y^2}$ for $(x, y) \in B_1(0, 0)$. This is a cone, with the vertex on $(0, 0, 1)$ and the base being the unit disk, see Figure 8.9. Further, assume that the trail followed is given by $x = u(t) = (1 - t)\cos(2\pi t)$ and $y = v(t) = (1 - t)\sin(2\pi t)$ for $t \in [0, 1)$; notice that these are the parametric equations of a curve. Calculate the vertical speed.

Solution We use the chain rule (8.20):

$$\frac{\partial z}{\partial x} = -\frac{x}{\sqrt{x^2 + y^2}} = -\cos(2\pi t), \qquad \frac{\partial z}{\partial y} = -\frac{y}{\sqrt{x^2 + y^2}} = -\sin(2\pi t),$$

and

$$\frac{dx}{dt} = -\cos(2\pi t) - 2\pi(1 - t)\sin(2\pi t), \qquad \frac{dy}{dt} = -\sin(2\pi t) + 2\pi(1 - t)\sin(2\pi t).$$

Thus,

$$\frac{dz}{dt} = \cos^2(2\pi t) + 2\pi(1 - t)\cos(2\pi t)\sin(2\pi t) + \sin^2(2\pi t) - 2\pi(1 - t)\cos(2\pi t)\sin(2\pi t) = 1.$$

So, the vertical speed is constant!

Let us now repete the calculation by direct substitution. Using the expressions defining x and y in the definition of z we have:

$$z = f(u(t), v(t)) = f((1 - t)\cos(2\pi t), (1 - t)\sin(2\pi t)) = 1 - \sqrt{(1 - t)^2} = t.$$

Hence, rather simply in this case, we get $\dfrac{dz}{dt} = 1$ as before. □

As in the example, we see that the typical situation in which the chain rule (8.19) is used is to evaluate the rate of change of a function along a curve, given in parametric form. Indeed, we can interpret the two functions $x = u(t)$ and $y = v(t)$ as giving the x and y coordinate of a curve in the x, y-plane as a function of t (the *parameter*). The parametric curve is thus given by

$$\mathbf{r} : I \to \mathbb{R}^2, \qquad \mathbf{r}(t) = (x(t), y(t)), \tag{8.21}$$

where $I \subset \mathbb{R}$ is some interval, the domain of the parameter. If $x'(t) \neq 0 y'(t)$, the vector

$$\mathbf{r}'(t) = (x'(t), y'(t)),$$

is a vector tangent to the curve at the point $(x(t), y(t))$. With this in mind, the chain rule above can be rephrased as follows. Given a function f with continuous first partials on an open set U and a differentiable curve $\mathbf{r} = \mathbf{r}(t)$ lying in U, the composition $f \circ \mathbf{r}$ is differentiable and

$$\frac{d}{dt}(f[\mathbf{r}(t)]) = \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t). \tag{8.22}$$

Theorem 8.30 (The chain rule II)**.** *If $z = f[x, y]$ has continuous first partial derivatives on an open set $U \subset \mathbb{R}^2$ and $x = u(s,t), y = v(s,t)$ are differentiable functions of t whose range is contained in U (so, whenever $(x, y) \in U$), then the composition function admits first partial derivatives in s, and t and*

$$\begin{aligned} \frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}, \\ \frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}. \end{aligned} \tag{8.23}$$

Proof. This is an easy consequence of Chain rule I applied to the partial derivatives of $z = f[x, y]$ with respect to s and t. $\square$

Notice that (8.23) can be re-written in matrix form as follows:

$$\begin{pmatrix} \frac{\partial z}{\partial s} & \frac{\partial z}{\partial t} \end{pmatrix} = \begin{pmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial s} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial s} & \frac{\partial y}{\partial t} \end{pmatrix}. \tag{8.24}$$

Each term in (8.24) is named Jacobian matrix of the respective function. Notice, in particular, that the matrix in (8.24) collects the partial derivatives of a function of two variables and two components[5], that is $(s, t) \to (x(s, t), y(s, t))$.

Example 8.31. Calculate the Jacobian matrix of the transformation from polar to cartesian variables.

Solution The change of variables polar to cartesian is given by

$$\begin{cases} x(r, \theta) = r\cos(\theta), \\ y(r, \theta) = r\sin(\theta). \end{cases} \tag{8.25}$$

The Jacobian of the transformation $(r, \theta) \to (x(r, \theta), y(r, \theta))$ is given by

$$\begin{pmatrix} \frac{\partial x}{\partial r}(r, \theta) & \frac{\partial x}{\partial \theta}(r, \theta) \\ \frac{\partial y}{\partial r}(r, \theta) & \frac{\partial y}{\partial \theta}(r, \theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

$\square$

Another application of the chain rule: the mean-value theorem for functions of several variables.

[5]You will study vector functions and their Jacobians in Calculus & Analysis III.

Theorem 8.32 (Mean-value theorem for functions of several variables)**.** *Let $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}$. If f is defined and differentiable at each point of the line segment $\overline{\mathbf{ab}}$ joining point $\mathbf{a}$ and point $\mathbf{b}$, then*

$$f(\mathbf{b}) - f(\mathbf{a}) = \nabla f(\mathbf{c}) \cdot (\mathbf{b} - \mathbf{a}), \tag{8.26}$$

for some $\mathbf{c} \in \overline{\mathbf{ab}}$.

Proof. Here we just give a not completely rigorous (why?) sketch of the proof.

The line segment $\overline{\mathbf{ab}}$ can be described by the parametric curve

$$\mathbf{r}(t) := \mathbf{a} + t(\mathbf{b} - \mathbf{a}) \qquad t \in [0, 1].$$

Indeed the graph of $\mathbf{r}(t)$ is a straight line and $\mathbf{a}$ corresponds to $t = 0$ and $\mathbf{b}$ to $t = 1$. Further, we can view the function f restricted on the line segment $\overline{\mathbf{ab}}$ as a function of one variable:

$$g(t) := f[\mathbf{a} + t(\mathbf{b} - \mathbf{a})] = f[\mathbf{r}(t)] \qquad t \in [0, 1].$$

The mean-value theorem for single-variable functions applied to g tells us that there must exist $t^* \in [0, 1]$ such that

$$f(\mathbf{b}) - f(\mathbf{a}) = g(1) - g(0) = g'(t^*)(1 - 0) = g'(t^*). \tag{8.27}$$

On the other hand, the chain rule (8.22) yields:

$$g'(t^*) = \frac{dg}{dt}(t^*) = \frac{d}{dt}(f[\mathbf{r}(t^*)]) = \nabla f[\mathbf{r}(t^*)] \cdot \mathbf{r}'(t^*) = \nabla f(\mathbf{a} + t^*(\mathbf{b} - \mathbf{a})) \cdot (\mathbf{b} - \mathbf{a}), \tag{8.28}$$

as $\mathbf{r}'(t) = \mathbf{b} - \mathbf{a}$, for all t. Combining (8.27) with (8.28) readily yields (8.26) with $\mathbf{c} = \mathbf{a} + t^*(\mathbf{b} - \mathbf{a})$.

□

Notice that in the mean-value theorem above, the gradient takes the place of the derivative, and the scalar product between vectors takes the place of the product between scalars. The mean value theorem is of great importance, both in theory and in practice. For instance, Theorem 8.22 showing that a function is differentiable if it has continuous first partial derivatives is proven by using the mean-value theorem.

Sometimes, when, say, y is function of x and their relationship is given implicitly, it is possible to calculate the rate of change of y with respect to x, that is to get $\frac{dy}{dx}$ without finding first the explicit dependence of y with respect to x. This technique is called implicit differentiation and can be easily derived using the chain rule.

Theorem 8.33 (Implicit differentiation)**.** *If $z = u(x, y)$ is continuously differentiable and y is a continuously differentiable function of x that satisfies the equation $u(x, y(x)) = 0$, then at all points z where $\frac{\partial z}{\partial y} \neq 0$,*

$$\frac{dy}{dx} = -\frac{\partial z}{\partial x} \Big/ \frac{\partial z}{\partial y}. \tag{8.29}$$

Proof. The proof is based on using the chain rule (8.20). In order to use the chain rule, we introduce a new variable t and set $x = t$, in such a way that

$$z = u(x(t), y(t)) \quad \text{with} \quad x = t \quad \text{and} \quad y = y(t).$$

Now, since $z = u(x(t), y(t)) = 0$ for all t by hypothesis, we have that $dz/dt = 0$. Moreover, $dx/dt = 1$ and $dy/dt = dy/dx$. Using these expression in Equation (8.20) we get:

$$0 = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y}\frac{dy}{dx}$$

Thus, for all those points (x, y) for which $\frac{\partial z}{\partial y} \neq 0$, we have (8.29). □

Example 8.34. Suppose that $x^2 + y^2 = 1$. Find dy/dx using implicit differentiation and by direct calculation.

Solution The function $z = u(x, y) = x^2 + y^2 - 1$ defines the equation relating x to y, that is $u(x, y) = 0$. Thus, the implicit differentiation method gives:

$$\frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y}.$$

To get the same result by direct calculation, we first need to find y explicitly in function of f. Clearly, $y = \sqrt{1 - x^2}$, at least for $x \in [-1, 1]$. Then, $\frac{dy}{dx} = -\frac{x}{\sqrt{1-x^2}}$, which coincide with the previous result if you consider that $y = \sqrt{1 - x^2}$. □

Exercises.

Example 8.35. Let $z = u(x, y)$ be a twice continuously differentiable[6] function and $y = y(x)$ be a twice continuously differentiable function that satisfies the equation $u(x, y(x)) = 0$. Obtain an implicit differentiation formula for the second derivative of the function $y(x)$ by differentiation of the implicit differentiation formula seen in the lectures. State the range of validity of your formula.
[Hint: it is best to start from the implicit differentiation formula in the form $y'(x) = -z_x(x, y(x))/z_y(x, y(x))$*].*

Example 8.36. Prove the Implicit differentiation formula for functions of more than one variable. That is, given that $u = u(x, y, z)$ is continuously differentiable and $z = z(x, y)$ is a differentiable function that satisfies the equation $u(x, y, z) = 0$, show that

$$\frac{\partial z}{\partial x} = -\frac{\partial u}{\partial x}\Big/\frac{\partial u}{\partial z} \qquad \text{and} \qquad \frac{\partial z}{\partial y} = -\frac{\partial u}{\partial y}\Big/\frac{\partial u}{\partial z}$$

at all points $(x, y, z(x, y))$ such that $\frac{\partial u}{\partial z} \neq 0$.

[6] A function is twice continuously differentiable if all its first and second partial derivatives exists and are continuous.

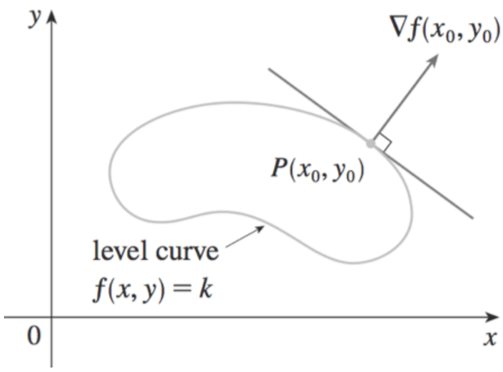


Figure 8.10: The gradient is orthogonal to the level curve.

8.6 Gradient and directional derivatives

Let us start by recalling the definition of the gradient. We have seen that the gradient is defined if f is differentiable and it is a vector whose components are the first partials. The following are the properties that make the gradient so important.

Our first result is yet another application of the chain rule! It gives a geometric interpretation of the gradient as a vector orthogonal to the level curve, as depicted in Figure 8.10.

Theorem 8.37. *If $f(x, y)$ is differentiable at the point (x_0, y_0) and $\nabla f(x_0, y_0) \neq 0$, then $\nabla f(x_0, y_0)$ is a vector normal to the level curve of f passing through $\mathbf{x}_0 = (x_0, y_0)$.*

Proof. Let $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ be a parametrisation of the level curve through (x_0, y_0) such that $(x_0, y_0) = \mathbf{r}(0)$. Thus, by definition, we have

$$f(x(t), y(t)) = f(x_0, y_0) \qquad \forall t.$$

As the right-hand side of the above equation is a constant, differentiating the above equation with respect to t gives

$$\begin{aligned} &= \tfrac{d}{dt} f(x(t), y(t)) = 0 \\ &| \\ &= \nabla f(x(t), y(t)) \cdot \left(\frac{dx}{dt} \quad \frac{dy}{dt}\right) = \nabla f(x(t), y(t)) \cdot \frac{d\mathbf{r}}{dt}(t), \end{aligned}$$

having used, once again, the chain rule to go from the first to the second line. Using the above with $t = 0$ gives

$$\nabla f(x_0, y_0) \cdot \left.\frac{d\mathbf{r}}{dt}\right|_{t=0} = 0,$$

expressing the fact that the vector $\nabla f(x_0, y_0)$ is orthogonal to the vector $\left.\frac{d\mathbf{r}}{dt}\right|_{t=0}$. The latter vector is tangent to the curve (it expresses the direction (slope) of the curve in correspondence to $t = 0$, that is at (x_0, y_0)). Hence the gradient is orthogonal to the level curve as required. □

Example 8.38. Calculate the gradient of the function $f(x, y) = x^2 + y^2$ at the point $(x_0, y_0) = (\sqrt{2}/2, \sqrt{2}/2)$ and verify that it is indeed orthogonal to the level curve through that point.

Solution We have $\nabla f(x,y) = 2x\mathbf{i} + 2y\mathbf{j}$, and thus $\nabla f(\sqrt{2}/2, \sqrt{2}/2) = \sqrt{2}\mathbf{i} + \sqrt{2}\mathbf{j}$. The level curve through (x_0, y_0) is given by

$$x^2 + y^2 = f(x,y) = f(x_0, y_0) = 1,$$

so it is the unit circle centred in the origin. A parametric representation of this curve is $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j}$, $t \in [0, 2\pi)$, with x_0, y_0 corresponding to $t = \pi/4$. The tangent to the curve at this point is $\mathbf{r}'(\pi/4) = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j}$ and is indeed orthogonal to the gradient vector. $\square$

First partials give the rate of change of a function in the direction of the (positive) axis. And what about the rate of change in *any* direction?

Definition 8.39 (Directional derivative)**.** The directional derivative of a function of several variable f at $\mathbf{x}_0$ in the direction of the unit[7] vector $\mathbf{u}$ is given by

$$f_{\mathbf{u}}(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{u}) - f(\mathbf{x}_0)}{h}. \tag{8.30}$$

Directional derivatives are also denoted by:

$$f_{\mathbf{u}}(\mathbf{x}_0) = \frac{\partial f}{\partial \mathbf{u}}(\mathbf{x}_0) = D_{\mathbf{u}}(\mathbf{x}_0).$$

They give the rate of change in the direction given by the vector $\mathbf{u}$, in the same way partial derivatives give the rate of change in the direction of the coordinate axis.

In particular, for bivariate functions, the definition can be written

$$f_{\mathbf{u}}(\mathbf{x}_0) = \lim_{h \to 0} \frac{f(x_0 + hu, y_0 + hv) - f(x_0, y_0)}{h}$$

where $\mathbf{u} = u\mathbf{i} + v\mathbf{j}$ and $\mathbf{x}_0 = (x_0, y_0)$. From which it is apparent that the partial derivatives are directional derivatives (*cf.* Definition 8.13). Indeed $f_x(\mathbf{x}_0) = f_{\mathbf{i}}(\mathbf{x}_0)$ and $f_y(\mathbf{x}_0) = f_{\mathbf{j}}(\mathbf{x}_0)$.

As usual, we want to avoid the evaluation of a derivative based on its definition. The gradient can be used to calculate directional derivatives, as the following theorem shows.

Theorem 8.40. *If f is* **differentiable** *in $\mathbf{x}$ then the directional derivative of f at $\mathbf{x}$ in the direction of the unit vector $\mathbf{u}$ is given by*

$$f_{\mathbf{u}}(\mathbf{x}) = \mathbf{u} \cdot \nabla f(\mathbf{x}). \tag{8.31}$$

Proof. If f is differentiable, than the composition of function $f(\mathbf{x} + h\mathbf{u})$ is a differentiable function of h and, clearly,

$$f_{\mathbf{u}}(\mathbf{x}) = \left.\frac{df(\mathbf{x} + h\mathbf{u})}{dh}\right|_{h=0} = \mathbf{u} \cdot \nabla f(\mathbf{x}),$$

with the last equality due to the chain rule I. $\square$

[7] That is a vector of length 1, so $||\mathbf{u}||^2 = u_1^2 + \cdots + u_n^2 = 1$.

In conclusion, although the gradient collects the rate of change in the coordinate directions, it can be used to obtain the rate of change in any direction.

Directional derivative when the vector does not have unit length. If you are given *any* nonzero vector $\mathbf{u}$, we can still calculate the rate of change of a function f in the direction given by $\mathbf{u}$ as the directional derivative with respect to the unit vector having the same direction as $\mathbf{u}$:

$$f_{\mathbf{u}/|\mathbf{u}|}(\mathbf{x}) = \frac{\mathbf{u}}{|\mathbf{u}|} \cdot \nabla f(\mathbf{x}).$$

So, once again, it is the unit vectors that are relevant. Note that all possible directions (all unit vectors) are given by the vectors pointing at the unit circle, which are given by

$$\mathbf{u}_\phi = \cos\phi \mathbf{i} + \sin\phi \mathbf{j} \qquad \phi \in [0, 2\pi).$$

Hence the following example gives a formula to caclulate *all* the directional derivatives.

Example 8.41. Deduce a formula for the directional derivative of a bivariate function in the direction of the vector $\mathbf{u}_\phi = \cos\phi \mathbf{i} + \sin\phi \mathbf{j}$ for $\phi \in [0, 2\pi)$.

Solution Applying (8.31) we get:

$$f_{\mathbf{u}_\phi}(\mathbf{x}) = \mathbf{u}_\phi \cdot \nabla f(\mathbf{x}) = f_x(\mathbf{x}) \cos\phi + f_y(\mathbf{x}) \sin\phi.$$

□

Let θ be the angle between $\nabla f(\mathbf{x})$ and $\mathbf{u}_\phi$ and assume that $\nabla f(\mathbf{x}) \neq 0$. Then,

$$f_{\mathbf{u}_\phi}(\mathbf{x}) = ||\mathbf{u}_\phi|| ||\nabla f(\mathbf{x})|| \cos\theta = ||\nabla f(\mathbf{x})|| \cos\theta.$$

(as $||\mathbf{u}_\phi|| = 1$). As $\cos\phi$ takes values between -1 and 1, this implies that

$$-||\nabla f(\mathbf{x})|| \leq f_{\mathbf{u}_\phi}(\mathbf{x}) \leq ||\nabla f(\mathbf{x})||$$

and at $\mathbf{x}$. In particular, if $\mathbf{u}_\phi$ is tangent to the level curve, than $\cos\theta = 0$ and thus $f_{\mathbf{u}_\phi}(\mathbf{x}) = 0$. We can summarize all our findings as follows.

Proposition 1 (Geometric properties of the gradient)**.** Let f be differentiable at $\mathbf{x}$. Then,

- The gradient of f at $\mathbf{x}$ is a vector normal to the level curve (set) passing through $\mathbf{x}$;
- f increases most rapidly in the direction of $\nabla f(\mathbf{x})$;
- f decreases most rapidly in the direction of $-\nabla f(\mathbf{x})$;
- the maximum slope (rate of change) of f at $\mathbf{x}$ is given by $||\nabla f(\mathbf{x})||$;
- the slope (rate of change) of f is zero in the directions tangent to the level curve of f passing through $\mathbf{x}$.

8.7 Extreme values

Whatever you have in mind to do, you would like to do it *optimally*, that is minimising your effort, or maximising your income, or taking a minimum time, etc. Thus, whenever what you are trying to do can be described as a mathematical function of some variables (time, costs, distance, energy, etc), the goal is to find a maximum (*the* maximum, if possible) or a minimum (*the* maximum or minimum, if possible) of such function. Determining maximum and minimum values for functions of several variables is a hugely important task in applied mathematics. What follows is the basic *theory* on how to study the problem of extreme (maximum or minimum) values[8]. It is an application of the theory of partial derivatives developed so far.

Note: as before, most results will be written for functions of *any* number of variables, while the application will be in the simpler case of bivariate functions. As usual, let us start with the definition.

Definition 8.42 (Extreme values)**.** Let $f : D \to \mathbb{R}$ be a function of several variables and $\mathbf{x}_0 \in D$. We say that

- The function f has a local maximum at $\mathbf{x}_0$ if
$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad \text{for all } x \in D \text{ in some neighborhood of } x_0;$$
- The function f has a local minimum at $\mathbf{x}_0$ if
$$f(\mathbf{x}_0) \leq f(\mathbf{x}) \quad \text{for all } x \in D \text{ in some neighborhood of } x_0;$$
- The function f has a global maximum at $\mathbf{x}_0$ if
$$f(\mathbf{x}_0) \geq f(\mathbf{x}) \quad \text{for all } x \in D;$$
- The function f has a global minimum at $\mathbf{x}_0$ if
$$f(\mathbf{x}_0) \leq f(\mathbf{x}) \quad \text{for all } x \in D\ .$$

Maxima and minima values of f are collectively named extreme values of f.

Notice that the definition is actually the same seen for functions of one variables. Examples of functions with global maximum and minimum are shown in Figure 8.11. You can easily verify that, in both cases, the gradient is zero at the extreme point (in this case, the origin). This is not by chance, as the following theorem shows. The characterisation of the point at which an extreme value can be attained is also the same with the gradient taking the role of the derivative.

Theorem 8.43. *If f has an extreme value at a point $\mathbf{x}_0$ which is an interior point of the domain of f, then*

(i) *either* $\nabla f(\mathbf{x}_0) = 0$*,*

[8]You have see the *practice* of it in the computer practicals.

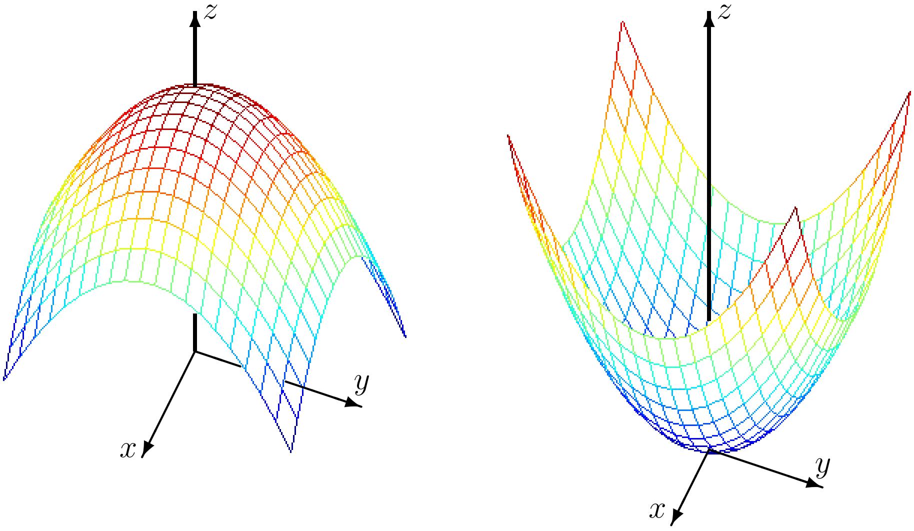


Figure 8.11: Graph of $f(x) = 1 - x^2 - y^2$ (left) and $f(x) = x^2 + y^2$ (right).

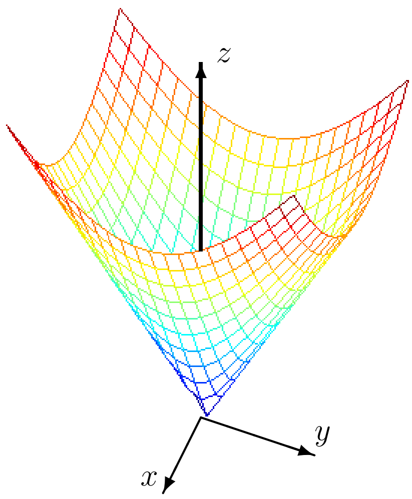


Figure 8.12: Graph of $f(x) = \sqrt{x^2 + y^2}$. The gradient is does not exist at the origin which is a point were the absolute minimum is attained.

(ii) *or* $\nabla f(\mathbf{x}_0)$ *does not exist.*

Proof. We shaw that if (i) and (ii) are both negated, then $\mathbf{x}_0$ cannot be an extreme point. So, assume that there exists $\nabla f(\mathbf{x}_0) \neq 0$. Hence f has (strictly) positive directional derivative in the direction of $\nabla f(\mathbf{x}_0)$. It follows that at $\mathbf{x}_0$ long the line with direction given by $\nabla f(\mathbf{x}_0)$ the function f is increasing and thus it cannot have at $\mathbf{x}_0$ neither a maximum or a minimum value[9]. □

An example of a function with an extreme value at a point were the gradient does not exist is given in Figure 8.12

Definition 8.44. An interior point $\mathbf{x}_0$ of the domain of f is named a

- stationary point if $\nabla f(\mathbf{x}_0) = 0$;

[9] This proof is essentially an application of the first derivative test for functions of one variable.

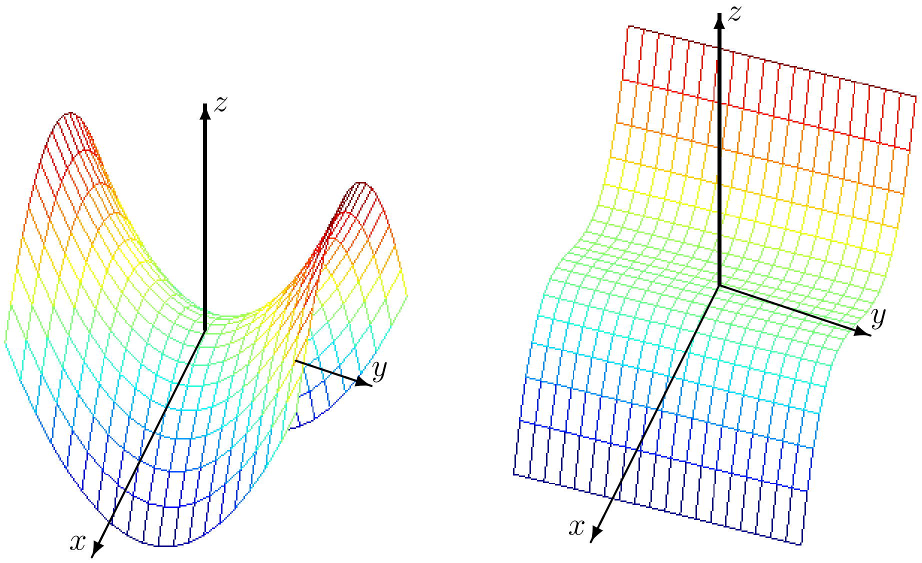


Figure 8.13: Examples of saddle points: graph of $f(x) = -x^2 + y^2$ (left) and $f(x) = -x^3$ (right).

- singular point if $\nabla f(\mathbf{x}_0)$ does not exist;
- critical point if it is a stationary or a singular point;
- saddle point if it is critical but it is not of extreme.

The terminology *saddle point* comes from the typical shape of the graph of a function near a critical point that is not of extreme, see Figure 8.11 (left). There are, tough, saddle points that do not look like a saddle at all. For instance, all the points in the y-axis are saddle points for the function $f(x, y) = -x^3$ depicted in Figure 8.13 (right).

In conclusion, internal extreme points must be must be critical points. An extreme point, though, could also be located on the boundary of the domain. We can summarise our findings as follows.

Proposition 2 (Location of extreme points)**.** Points were extreme values are attained are either critical (stationary or singular) points or boundary points.

So, now we know were to look for extreme values. Before exploring techniques for actually find them, it is useful to have an *a priori* knowledge of their existence.

Theorem 8.45 (The extreme-value theorem)**.** *Let $f : D \to \mathbb{R}$ with $D \subset \mathbb{R}^n$. If*

- *f is continuous, and*
- *D is closed and bounded,*

then f takes on an absolute maximum and an absolute minimum on D.

The practical relevance of the extreme-value theorem is demonstrated by the following examples.

Example 8.46. Find the global maximum and minimum value of the function $f(x, y) = x^2 + y^2$ on the unit disk $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

Solution A representation of the function f is given in Figure 8.11 (right). As the hypothesis of Theorem 8.45 are satisfied, there *must* be points were the function attains global maximum and minimum values. Further, Proposition 2 tells us that we can restrict our search to the collection of the critical and boundary points.

- *Singular points.* The function t does not have singular points.
- *Stationary points.* These are found by equating the gradient $\nabla f(x, y) = 2x\mathbf{i} + 2y\mathbf{j}$ to zero, yielding the origin as the only stationary point. We have $f(0, 0) = 0$.
- *Boundary points.* At the boundary $x^2 + y^2 = 1$ and thus $f(x, y) = 1$ (in this case, the boundary corresponds to a level curve).

As we knew already, the origin must be the point were the function has its global minimum equal to 0. The global maximum is attained at all points of the boundary of the domain D and is equal to 1. □

Example 8.47. Find the hottest and coolest point of a triangular hot plate with vertices $V_1 = (0, -4)$, $V_2 = (6, -4)$, and $V_3 = (0, 8)$ if its temperature is given by $t(x, y) = x^2 + xy + 2y^2 - 3x + 2y$.

Solution It is understood that the domain occupied by the hot plate is comprehensive of its boundary. Hence it is closed (and bounded). Further, the function t is clearly continuous. As the hypothesis of Theorem 8.45 are satisfied, there *must* be points were the temperature is the lowest and the highest. We look for them among the critical and boundary points.

- *Singular points.* The function t does not have singular points.
- *Stationary points.* These are found by equating the gradient $\nabla t(x, y) = (2x + y - 3)\mathbf{i} + (x + 4y + 2)\mathbf{j}$ to zero yielding the system of equations

  $$\begin{cases} 2x + y - 3 = 0 \\ x + 4y + 2 = 0 \end{cases} \qquad \Longleftrightarrow \qquad (x, y) = (2, -1).$$

  The point $P = (2, -1)$ is the only stationary point of t. It belongs to the region occupied by the hot plate and the temperature in it is $t(2, -1) = 4 - 2 + 2 - 6 - 2 = -4$.
- *Boundary points.* The boundary is made of the three sides of the triangle. We study them one by one.
  - The side $\overline{V_1V_2} = \{(x, -4) : x \in [0, 6]\}$. The restriction of t on $\overline{V_1V_2}$ is the function of one variable $g(x) = x^2 - 7x + 24$. This is a parabola *concave up*. Its minimum is achieved at the point such that $g'(x) = 2x - 7 = 0$, that is $x = 7/2$. The maximum is achieved at both or either end point. We have $f(0, -4) = g(0) = 24$, $f(7/2, -4) = g(7/2) = 47/4 = 11.75$, and $f(6, -4) = g(6) = 18$.

- On the side $\overline{V_2V_3} = \{(x, -2x+8) : x \in [0,6]\}$ we get the function of one variable $g(x) = 7x^2 - 63x + 144$. This is again a parabola *concave up*. Its minimum is achieved at the point such that $g'(x) = 14x - 63 = 0$, that is $x = 63/14$. The maximum is achieved at both or either end point. We have $f(0,8) = g(0) = 144$, $f(63/14, -1) = g(63/14) = 2.25$, and $f(6,-4) = g(6) = 18$.
- On the side $\overline{V_1V_3} = \{(0, y) : y \in [-4, 8]\}$ we get the function of one variable $g(y) = 2y^2 + 2y$. This is again a parabola *concave up*. Its minimum is achieved at the point such that $g'(y) = 4y + 2 = 0$, that is $y = -1/2$. The maximum is achieved at both or either end point. We have $f(0,-4) = g(0) = 24$, $f(0,-1/2) = g(-1/2) = -1/2$, and $f(0,8) = g(8) = 144$.

By comparing all the values found, we conclude that the hottest point is the vertex $V_3 = (0,8)$ with $t(0,8) = 144$ and the coolest point is the internal point $(2,-1)$ with $t(2,-1) = -4$. $\square$

8.8 Classification of stationary points of bivariate functions

Here we give a formal procedure of deciding if a stationary point is a maximum, a minimum or a saddle point. Recall the second derivative test for functions of one variable. Suppose that $f = f(x)$ has a stationary point in x_0, namely $f'(x_0) = 0$, and it has the second derivative there. Then, according to the second-derivative test,

- if $f''(x_0) > 0$, then f has a local minimum at x_0;
- if $f''(x_0) < 0$, then f has a local maximum at x_0;
- if $f''(x_0) = 0$ the test gives no information (f can have a maximum, a minimum, or an inflection point).

For functions of several variables, the *Hessian* takes the role of the second derivative.

Definition 8.48. Let f be a function of n variables and $\mathbf{x}$ a point in the domain of f. The Hessian matrix is defined as the matrix whose entries are all second partials of f at $\mathbf{x}$:

$$H_f(\mathbf{x}) = D^2 f(\mathbf{x}) := \begin{pmatrix} f_{11}(\mathbf{x}) & f_{12}(\mathbf{x}) & \cdots & f_{1n}(\mathbf{x}) \\ f_{21}(\mathbf{x}) & f_{22}(\mathbf{x}) & \cdots & f_{2n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}(\mathbf{x}) & f_{n2}(\mathbf{x}) & \cdots & f_{nn}(\mathbf{x}) \end{pmatrix}. \tag{8.32}$$

Note that, whenever the hypothesis of Theorem 8.17 (Equality of mixed partials) are satisfied, the Hessian matrix is symmetric.

Theorem 8.49 (Second derivative test)**.** *Suppose that* $\mathbf{x}_0$ *is a stationary point for* f *and that it is internal to the domain of* f*. Further, assume that the Hessian is defined and continuous in a neighbourhood of* $\mathbf{x}_0$ *(hence the Hessian is symmetric at* $\mathbf{x}_0$*).*

- *If $H_f(\mathbf{x}_0)$ is a positive definite matrix*[10]*, then f has a local minimum at $\mathbf{x}_0$;*
- *If $H_f(\mathbf{x}_0)$ is a negative definite matrix, then f has a local maximum at $\mathbf{x}_0$;*
- *If $H_f(\mathbf{x}_0)$ is indefinite matrix, then f has a saddle point at $\mathbf{x}_0$;*
- *If $H_f(\mathbf{x}_0)$ is neither positive nor negative definite nor indefinite, then this test gives no information.*

Let us specialise the second derivative test in the case of bivariate functions. If $f = f(x,y)$, then the Hessian is given by

$$H_f(x,y) = D^2 f(x,y) = \begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{pmatrix} \qquad \text{THE HESSIAN in 2D.} \tag{8.33}$$

An easy way to remember the Hessian is by noticing that the first and second row of the Hessian matrix are the gradient of the functions f_x and f_y, respectively, that is

$$H_f(x,y) = \begin{pmatrix} \nabla f_x(x,y) \\ \nabla f_y(x,y) \end{pmatrix}.$$

Now assume that the Hessian of f is defined and continuous in a neighbourhood of $\mathbf{x}_0 = (x_0, y_0)$ (hence the Hessian is symmetric at (x_0,y_0)). Then,

$$H_f(x_0,y_0) = \begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix},$$

where $A = f_{xx}(x_0,y_0)$, $B = f_{xy}(x_0,y_0) = f_{yx}(x_0,y_0)$, and $C = f_{yy}(x_0,y_0)$. The determinant of the Hessian matrix, that is the number $f_{xx}f_{yy} - f_{xy}^2 = AC - B^2$ is known as the discriminant. Applying the criteria of definiteness to the Hessian matrix we get the following practical version of the second derivative test:

Second derivative test (bivariate functions)

- If $AC - B^2 > 0$ and $A > 0$, then f has a local minimum at $\mathbf{x}_0$;
- If $AC - B^2 > 0$ and $A < 0$, then f has a local maximum at $\mathbf{x}_0$;
- If $AC - B^2 < 0$, then f has a saddle point at $\mathbf{x}_0$;
- If $AC - B^2 = 0$, the test gives no information.

Example 8.50. Find and classify the stationary points of $f(x,y) = x^4 + x^2y^2 - 2x^2 + 2y^2 - 8$, $(x,y) \in \mathbb{R}^2$.

[10] You have see the definition of definiteness in terms of bilinear forms at Linear Algebra II. Definiteness can be characterised by looking at the eigenvalues or at the leading principal minors which are defined as the determinants of the $k \times k$ upper left corners of the matrix, for $K = 1, \dots, n$. Search for the Sylvester's Criterion. A real symmetric $n \times n$ matrix A is positive definite iff all its eigenvalues are positive and iff all the leading principal minors are positive, negative definite iff all its eigenvalues are negative and iff the sign of the leading principal minors alternate starting with negative, and indefinite iff it has both positive and negative eigenvalues and iff it is neither positive nor negative definite.

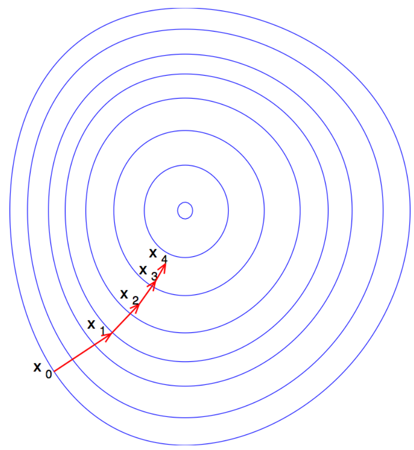


Figure 8.14: Few iterates of the gradient descent method. In blue the level curves of the target function.

Solution Clearly f is infinitely many times continuously differentiable. So we can apply the second derivative test. First of all, let us find the stationary points. These are found by solving the system

$$\begin{cases} f_x(x,y) = 4x^3 + 2xy^2 - 4x = 2x(2x^2 + y^2 - 2) = 0 \\ f_y(x,y) = 2x^2y + 4y = 2y(x^2+2) = 0 \end{cases} \Leftrightarrow \begin{cases} x(2x^2-2) = 0 \\ y = 0 \end{cases},$$

whose solutions are the points $(0,0)$, $(1,0)$, and $(-1,0)$. To classify them, we calculate the hessian matrix:

$$H_f(x,y) = \begin{pmatrix} 12x^2 + 2y^2 - 4 & 4xy \\ 4xy & 2x^2 + 4 \end{pmatrix},$$

and, in particular,

$$H_f(0,0) = \begin{pmatrix} -4 & 0 \\ 0 & 4 \end{pmatrix}, \qquad H_f(\pm 1, 0) = \begin{pmatrix} 8 & 0 \\ 0 & 6 \end{pmatrix}.$$

At $(0,0)$ the discriminant is equal to $-16 < 0$, hence $(0,0)$ is a saddle point. At $(\pm 1, 0)$ the discriminant is equal to $48 > 0$ and $A = 8 > 0$, hence $(\pm 1, 0)$ are both minimum points. $\square$

It is not always possible to solve the (nonlinear) equations yielding the stationary points. In these cases, a numerical procedure should be employed in order to find extreme values.

The steepest (or gradient) descent method is one such iterative procedure to approximate extreme values. It is based on the fact that the gradient gives the steepest descent. Suppose that we want to find the maximum of a function f. Starting from any initial guess point $\mathbf{x}_0$, look for $\mathbf{x}_1$ along the direction of steepest descent $\equiv$ the direction of the gradient at which $f(\mathbf{x}_1) > f(\mathbf{x}_2)$, then iterate, see Figure 8.14.

Exercises.

Example 8.51. Verify that the cube has the most volume among all cuboids whose dimensions sum equals $3a$.

Example 8.52. Find the distance between the point $(1,0,-2)$ and the plane $x+2y+z=4$.

Example 8.53. We want to build a rectangular cardboard box with no top and of given volume V using the least quantity of cardboard. Find the length of the box's edges by finding the minimum of the total surface function. Verify that the point you found is indeed a minimum applying the second derivatives test on such function.

Example 8.54. Consider the function $f(x,y) = \dfrac{xy}{1+x^2+y^2}$.

- Find all stationary points of f.
- Classify all stationary points of f as maximum, minimum, or saddle point. Justify your arguments.
- Find the absolute maximum and minimum values of the function f on the unit disk centred in the origin. List all points (x,y) at which the extreme values are attained.

Index