MATH1026 Sets, Sequences and Series, 2021/2022

Notes, A. Strohmaier 2022

Contact Details

Lecturer:	Professor Alexander Strohmaier
Office:	Mathematics 11.10
E-mail:	a.strohmaier@leeds.ac.uk
Phone:	343 8884
Web:	http://www1.maths.leeds.ac.uk/∼pmtast

Note: During the pandemic contact is by e-mail only or after the Lecture on Zoom. There will be time reserved for discussions after the lecture.

All handouts will be made available on the Minerva.

Assessment arrangements

There is normally be a $2\frac{1}{2}$ hour exam during the May exam period, counting 80% towards your final module grade. During the pandemic this exam is likely to be online. **You must pass the final exam** (score at least 40% in the exam irrespective of the homework) **to pass the module. Single honours maths students must pass the module to progress to year 2.**

There will also be 10 homework assignments, collectively counting 15% towards your final grade. These will be due at **by 2pm** on the following dates:

- Monday 1 February 2022
- Monday 8 February 2022
- Monday 15 February 2022
- Monday 22 February 2022
- Monday 01 March 2022
- Monday 08 March 2022
- Monday 15 March 2022
- Monday 22 March 2022
- Monday 26 April 2022
- Monday 03 May 2022

Homework will be submitted to gradescope. You will be required to give a short presentation, which will count 5% towards your module grade. Further details will be announced nearer to the time.

Each homework assignment will be supported by a tutorial session. **It is very important that you attend these.**

Late submission

This course does not accept late submissions. If you submit late without a valid reason you will not be awarded any marks.

Recap of important notions from MATH1025

Set theory

A **set** is a collection of objects, real or of the imagination. Given a set A, we write $x \in A$ to denote that object x is in set A. This is usually translated into words (when reading the expression) as "x is an element of A". The expression $x \notin A$ means x is *not* an element of A.

Another set B is a **subset** of A if every element of B is an element of A. This is denoted $B \subseteq A$. Note that every set is a subset of itself, $A \subseteq A$. Two sets A, B are equal if they have precisely the same elements, in which case $A \subseteq B$ and $B \subseteq A$.

There are a few sets which occur frequently in mathematics, and so have been given special, commonly agreed symbols.

- The **empty set** is denoted $\emptyset$. This is the set containing no objects. Do not confuse it with $\{\emptyset\}$, which certainly isn't the empty set since it contains an object (namely, the empty set)!
- The set of **natural numbers** $\mathbb{N} = \{1, 2, 3, \ldots\}$.
- The set of **natural numbers and zero** $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}$.
- The set of **integers** $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
- The set of **rational numbers** $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z},\, b \neq 0\}$.
- The set of **real numbers** $\mathbb{R}$.
- The set of **complex numbers** $\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}$.

This list illustrates the two most common ways of specifying a set, namely, listing its elements,

$$A = \{1, 2, 3, 4, 5, 6, 7, 8\},$$

and giving a condition which elements must satisfy,

$$A = \{x \mid x \in \mathbb{N},\, x \leq 8\}.$$

In general $A = \{x \mid P(x)\}$ means the set of objects x for which condition $P(x)$ is true. If part of condition $P(x)$ is that x should be in a particular set, we often move that part to the left of the colon. For example, the set above could be written

$$A = \{x \in \mathbb{N} \mid x \leq 8\}.$$

Given real numbers a, b, with $a < b$, we define several different **intervals**, as follows:

$$\begin{aligned}
(a,b) &= \{x \in \mathbb{R} \mid a < x < b\} \\
[a,b] &= \{x \in \mathbb{R} \mid a \leq x \leq b\} \\
[a,b) &= \{x \in \mathbb{R} \mid a \leq x < b\} \\
(a,b] &= \{x \in \mathbb{R} \mid a < x \leq b\} \\
(a,\infty) &= \{x \in \mathbb{R} \mid x > a\} \\
[a,\infty) &= \{x \in \mathbb{R} \mid x \geq a\} \\
(-\infty,b) &= \{x \in \mathbb{R} \mid x < b\} \\
(-\infty,b] &= \{x \in \mathbb{R} \mid x \leq b\}.
\end{aligned}$$

So a round bracket denotes that the end point is omitted from the interval, while a square bracket denotes that it is included. Note that ∞ is *not* a number, so expressions like $(a, \infty]$ do not make sense. Intervals which *include* all their endpoints, that is, $[a, b]$, $[a, \infty)$, $(-\infty, b]$ are said to be **closed**. Those which exclude all their endpoints, that is, (a, b), (a, ∞), $(-\infty, b)$ are said to be **open**.

Given two sets A, B, there are several useful ways of constructing new sets from them.

- The **intersection** of A and B is

  $$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

- The **union** of A and B is

  $$A \cup B = \{x \mid x \in A \text{ or } x \in B\}.$$

  Note that, in mathematics, we use "or" in the non-exclusive sense, that is, "P or Q" means P is true, or Q is true (or both).

- The **complement** of B in A is

  $$A \backslash B = \{x \mid x \in A \text{ and } x \notin B\}.$$

- The **cartesian product** of A and B is

  $$A \times B = \{(x, y) \mid x \in A \text{ and } y \in B\},$$

  that is, the set of **ordered pairs** of objects (x, y), where x comes from A and y comes from B. Do not confuse the symbol $\times$, used in the context of sets, with multiplication.

Example -1.1.1

$$\begin{aligned}
\{1,2,3,4,5,6,7,8\} &= [1,8]\cap\mathbb{Z} \\
&= (-\infty,9)\cap\mathbb{N} \\
&= \{1,3,5,7\}\cup\{2,4,6,8\} \\
\mathbb{R}\backslash\{-1,1\} &= (-\infty,-1)\cup(-1,1)\cup(1,\infty) \\
\{\mathit{me},\mathit{you}\}\times\{1,2,3\} &= \{(\mathit{me},1),(\mathit{me},2),(\mathit{me},3),(\mathit{you},1),(\mathit{you},2),(\mathit{you},3)\} \\
(-2,\infty)\cap(-\infty,0] &= (-2,0]
\end{aligned}$$

□

Functions

Given two sets A, B, a **function** (or **mapping**) f from A to B is a rule which assigns to each element of A an element of B. The shorthand for this is $f : A \to B$, often read as "f maps A to B". The set A is called the **domain** of f and the set B is called the **codomain** of f. For each $x \in A$, we denote by $f(x)$ the element in B which the function f assigns to the element x, and call this element the **image** of x under f. There is no reason why two different elements in A cannot have the same image under a function.

It is often helpful to think of f as a machine which converts things in A to things in B. Given an input $x \in A$, the machine spits out an output $f(x) \in B$. The set of all outputs is called the **range** of f, denoted $f(A)$. In set notation, the range is

$$f(A) = \{f(x) \mid x \in A\}.$$

By definition, this is a subset of the codomain B.

Definition -1.1.2 *A function $f : A \to B$ is* **injective** *(or* **one-to-one***) if $f(x_1) = f(x_2)$ implies that $x_1 = x_2$. It is* **surjective** *(or* **onto***) if $f(A) = B$. It is* **bijective** *if it is both injective and surjective.*

We may restate these conditions as follows:

- f is injective if whenever $x_1 \neq x_2$, $f(x_1) \neq f(x_2)$, that is, distinct inputs always produce distinct outputs.
- f is surjective if its range is the whole codomain, and hence given any $y \in B$ there is some $x \in A$ such that $f(x) = y$.

Example -1.1.3 *The function $f : \mathbb{R} \to \mathbb{R}$ which maps x to $f(x) = x^3$ is injective:*

$$\begin{aligned} f(x_1) = f(x_2) &\Rightarrow x_1^3 = x_2^3 \\ &\Rightarrow x_1^3 - x_2^3 = 0 \\ &\Rightarrow (x_1 - x_2)(x_1^2 + x_1x_2 + x_2^2) = 0 \end{aligned}$$

which implies that $x_1 = x_2$, or $x_1^2 + x_1x_2 + x_2^2 = 0$. But

$$x_1^2 + x_1x_2 + x_2^2 = (x_1 + \frac{1}{2}x_2)^2 + \frac{3}{4}x_2^2$$

which is the sum of two non-negative numbers and, hence, is zero if and only if $x_2 = 0$ and $x_1 + \frac{1}{2}x_2 = 0$. But then $x_1 = x_2 = 0$. So in every case, $f(x_1) = f(x_2)$ implies that $x_1 = x_2$. Hence f is injective.
In fact, f is also surjective, though we are not (yet) in a position to prove this. What this means is that, given any real number y, there is some real number x such that $x^3 = y$. You may say "this is obvious: $x = y^{\frac{1}{3}}$", but what is $y^{\frac{1}{3}}$? It is, by definition, the real number whose cube is y. How do you know that such a real number exists? You don't (yet). $\square$

Definition -1.1.4 *Given functions $f : A \to B$ and $g : B \to C$, their* **composition** *is the function $g \circ f : A \to C$ which assigns to each $x \in A$ the image $(g \circ f)(x) = g(f(x))$.*

So $g \circ f$ is the function obtained by first "doing f" then taking the output and feeding it into g.[1] Note that we need the codomain of f to match the domain of g in order for $g(f(x))$ to make sense, otherwise $f(x)$ might not be in the domain of g, in which case $g(f(x))$ would be undefined.

Example -1.1.5 *Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x^2$ and $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = x + 1$. Then $g \circ f : \mathbb{R} \to \mathbb{R}$ such that*

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1.$$

In this case, the composition in the opposite order (first g then f) is also well defined, $f \circ g : \mathbb{R} \to \mathbb{R}$,

$$(f \circ g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2.$$

[1] The fact that function composition is ordered from right to left is a legacy of algebra's roots in the Arabic-speaking world.

Clearly $g \circ f \neq f \circ g$. □

Function composition preserves both injectivity and surjectivity. now prove.

Proposition -1.1.6 *Let $f : A \to B$ and $g : B \to C$.*

(i) If f and g are both injective, so is $g \circ f$.

(ii) If f and g are both surjective, so is $g \circ f$.

Proof:

(i) Assume f and g are injective, and that $(g \circ f)(x_1) = (g \circ f)(x_2)$ for some $x_1, x_2 \in A$. Then $g(f(x_1)) = g(f(x_2))$. But g is injective, so it follows that $f(x_1) = f(x_2)$. But f is injective, so it follows that $x_1 = x_2$. Hence $g \circ f$ is injective.

(ii) Assume f and g are surjective, and let $z \in C$. Since g is surjective, z is in the range of g, so there exists $y \in B$ such that $g(y) = z$. But f is also surjective, so y is in the range of f, and hence there exists $x \in A$ such that $f(x) = y$. Now $(g \circ f)(x) = g(f(x)) = g(y) = z$. Hence z is in the range of $g \circ f$. This argument works for any $z \in C$, so $g \circ f$ is surjective.

□

The size of sets

Definition -1.1.7 *Two sets A and B are said to have the same cardinality if there exists a bijection $\Phi : A \to B$.*

If either set is finite this means that the other is finite as well. In that case the statement just means that they have the same number of elements.

If a set has the cardinality of $\mathbb{N}$ it is said to be *countable*. If it is infinite but not countable, then we say it is *uncountable*.

Example -1.1.8 *The sets $\mathbb{Z}$, $\mathbb{Q}$ are countable. $\mathbb{R}$ as introduced in MATH1025 is uncountable.*

Chapter 0

What is real Analysis?

Mathematics is a science of precise argument. It assumes only some very basic axioms in logics and develops these into a powerful machine; a machine so potent that it is quite remarkably the engine behind the exact natural sciences. I would like to point out two things that distinguish mathematics from other sciences.

- mathematical statements are correct beyond any doubt. Mathematical truth is independent of the person checking and remains true for all times.
- apart from the very few basic axioms mathematics builds upon, there are no unexplained notions.

The first point is clearly appealing and the second can be a bit tedious. But it is the second point that makes the first possible.

A lot of first year mathematics revolves around turning the intuitive mathematics you have met at school into precise statements. Analysis is the precise mathematical theory behind calculus. The main purpose of Analysis is to provide you with the mathematical tools (armour of you wish) to do calculus without the doubt 'how can I absolutely sure that what I am doing is correct?" and without circular definitions.

Here are two examples that illustrate this well.

Statement A: If a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $f'(a) > 0$ for some point $a \in \mathbb{R}$. Then near that point the function is strictly increasing.

Even though Statement A sounds correct it is actually incorrect. We will see functions that behave in ways that are counterintuitive. Analysis will provide you with enough statements and counterexamples to be confident in either proving or disproving statements like this.

Definition B: The derivative of a function f at a point a is the slope of the graph of f at a.

Definition B sounds good at first, until you think further and ask yourself what the slope is. The answer is, of course, that the slope is defined as the derivative of the function at a. Unless you come up with a different definition of slope this definition is circular and therefore meaningless.

Definition C: The derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at $a \in \mathbb{R}$ is defined as

$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$ if the limit exists. Sounds much better. But is vacuous as long as we do not explain what the symbol $\lim$ means.

This course will fill this gap, give precise definitions, and provide you with the basic tools of Analysis: sets, sequences, series, and ultimately properties of functions. These tools are absolutely essential for doing calculus on a higher level. The methods you learn in this course are also more universally applicable. They will appear on other forms in later courses.

Chapter 1

The set of real numbers

1.1 Axiomatic characterisation of real numbers

Recall that in MATH1025 we have met the set of rational numbers $\mathbb{Q}$. This set was constructed from the set of natural numbers $\mathbb{N}$. The course MATH1025 was a bit vague about what the set of real numbers was and how it is constructed. At this stage we are going to assume that the set of real numbers exists and we will specify the properties that we are going to use. This "axiomatic approach" feels a bit like cheating and it leaves us with a certain degree of dissatisfaction. However, the construction of $\mathbb{R}$ is tedious. It would take at least a week of lecture time during which we would not gain too much insight into the structure of the real numbers. We will however, at the end of this course, give a sketch of the construction of the real numbers to fill the gap. With the notions that are available to us later in the semester this will be much easier than it would be now.

In order to explain the properties that we assume of the reals it is useful to define the notion of an ordered field.

Definition 1.1.1 (Ordered Field) *A set $\mathbb{K}$ with two binary operations $\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$, $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ and a relation $\leq$ is called an ordered field if for all $x, y, z \in \mathbb{K}$ we have*

(A1) $(x + y) + z = x + (y + z)$ *and* $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.

(A2) $x + y = y + x$ *and* $x \cdot y = y \cdot x$.

(A3) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

(A4) there exists a unique element $0 \in \mathbb{K}$ *such that* $0 + x = x$ *for all* $x \in \mathbb{K}$.

(A5) there exists a unique element $1 \in \mathbb{K}$ *such that* $1 \neq 0$ *and* $1 \cdot x = x$ *for all* $x \in \mathbb{K}$.

(A6) for each $x \in \mathbb{K}$ *there exists a unique* $y \in \mathbb{K}$ *such that* $x + y = 0$. *We will write* $-x$ *for* y.

(A7) for each $x \in \mathbb{K}$ *with* $x \neq 0$ *there exists a unique* $y \in \mathbb{K}$ *such that* $x \cdot y = 1$. *We will write* x^{-1} *or* $\frac{1}{x}$ *for* y.

(A8) $x \leq y$ *implies* $x + z \leq y + z$.

(A9) $x \leq y$ *and* $y \leq z$ *implies* $x \leq z$.

(A10) $x \leq y$ *and* $y \leq x$ *implies* $x = y$.

(A11) $x \leq y$ *and* $0 \leq z$ *implies* $x \cdot z \leq y \cdot z$.

(A12) given x *and* y *in* $\mathbb{K}$ *we have either* $x \leq y$ *or* $y \leq x$.

Notations: We will write xy for $x \cdot y$, x/y and $\frac{x}{y}$ for $x \cdot y^{-1}$. We also write $x < y$ if $x \leq y$ and $x \neq y$. Moreover, we will use the notations $x > y$ if $y < x$ and $x \geq y$ if $y \leq x$. Finally, if $n \in \mathbb{N}$ we will write a^n for $\underbrace{a \cdots a}_{n \text{ times}}$.

A set with binary operations $+, \cdot$ satisfying (A1) - (A7) is called a *field*. The fact that an ordering relation exists and is compatible with multiplication and addition makes the field into an ordered field. Hence the name.

Examples 1.1.2 *The following example illustrate the main points of this definition.*

- $\mathbb{Q}$ *is an ordered field.*
- $\mathbb{Z}$ *is not a field because (A7) fails.*
- $\mathbb{N}$ *is not a field because (A6) fails to hold.*
- $\mathbb{Z}_4$ *is not a field because there is no multiplicative inverse of* 2. *Hence, (A7) does not hold.*
- *One can show that* $\mathbb{Z}_5$ *is a field. It is not an ordered field because (A8) fails:* $3 \leq 4$, *but* $3+1$ *is not smaller than* $0 = 4+1$.
- *the set of real numbers as introduced in MATH1025 is an ordered field.*
- *the set of complex numbers as introduced in MATH1025 is not an ordered field because there is no obvious ordering relation on complex numbers. One can show there is none that is compatible with the field structure.*

Given an ordered field $\mathbb{K}$ one can then show that the usual computational rules follow from these axioms.

Theorem 1.1.3 *For any* $a, b, c \in \mathbb{K}$ *we have as consequences of the field properties*

(i) $a + c = b + c$ *implies* $a = b$.

(ii) $a \cdot 0 = 0$.

(iii) $(-a)b = -ab$.

(iv) $(-a)(-b) = ab$.

(v) $ac = bc$ *and* $c \neq 0$ *implies* $a = b$.

(vi) $ab = 0$ *implies that either* $a = 0$ *or* $b = 0$.

Proof:

(i) is proved by adding $-c$ to the right of both sides of the equation and using the axioms (A1), (A4) and (A6) to simplify.

(ii) note that $0 + a \cdot 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Now use (i) to obtain

$0 = a \cdot 0$.

(iii) follows from $(-a)b + ab = (-a + a)b = 0 \cdot b = 0$.

the other properties are proved along the same lines and the straightforward proofs are left as an exercise. $\square$

Theorem 1.1.4 *For any $a, b, c \in \mathbb{K}$ we have as consequences of the properties of an ordered field*

(i) $a \leq b$ implies $-b \leq -a$.

(ii) $a \leq b$ and $c \leq 0$ implies $bc \leq ac$.

(iii) $a \leq 0$ and $b \leq 0$ implies $ab \geq 0$.

(iv) $a^2 = a \cdot a \geq 0$.

(v) $0 < 1$.

(vi) $a > 0$ implies $a^{-1} > 0$.

(vii) $0 < a < b$ implies $0 < b^{-1} < a^{-1}$.

Proof:

(i) add $-a + (-b)$ on both sides and use (A8).

(ii) use the fact that $-c \geq 0$ and combine (i) and (A11).

(iii) Use (A11) with $y = 0$, $x = a$, $z = -b$ to obtain $-ba \leq 0$. Now use (i).

(iv) distinguish the cases $a \geq 0$ and $a \leq 0$. In the former the statement follows from (A11). In the latter case if follows from (iii).

(v) suppose by contradiction that $1 \leq 0$. Then, using (iii), one obtains $1 \cdot 1 \geq 0$ and therefore $1 \geq 0$. This violates (A12).

(vi) suppose, by contradiction, that $a > 0$, but $a^{-1} \leq 0$. Then by (iii) we have $1 = aa^{-1} \leq 0$, thus arriving at a contradiction.

the other properties are straightforward and left as an exercise. $\square$

Remark 1.1.5 *There are some other properties that are straightforward consequences of these axioms. For example it also follows from $x < y$ that $x + z < y + z$ for all z. In case $z > 0$ and $x < y$ one has $xz < yz$. This can be seen directly by using the axioms and excluding equality using Theorem 1.1.3.*

As you already anticipate $\mathbb{R}$ will also be an ordered field. We need however one more property that distinguishes it from $\mathbb{Q}$. In order to explain this property we need a bit more notation.

Definition 1.1.6 (Boundedness and Bounds) *Let $\mathbb{K}$ be an ordered field. Then a subset $S \subseteq \mathbb{K}$ is called bounded above if there is an element $M \in \mathbb{K}$ such that $x \leq M$ for all $x \in S$. In this case M is called an upper bound for S. A subset $S \subseteq \mathbb{K}$ is called bounded below if there is an element $M \in \mathbb{K}$ such that $x \geq M$ for all $x \in S$. Then M is called a lower bound for S.*

Remark 1.1.7 *A subset $S \subseteq K$ is bounded below if and only if $-S$ is bounded above.*

Examples 1.1.8 *An upper bound does not necessarily have to be contained in the set S, but it has to be in the ordered field. Here are some examples that illustrate this.*

- *consider the set $S_1 = \{x \in \mathbb{Q} \mid x < 3\}$ as a subset in the ordered field $\mathbb{Q}$. This set is clearly not bounded from below. It is bounded from above and any rational number larger or equal to 3 is an upper bound. Note that something like π or other irrational numbers are not upper bounds as these are not in the ordered field $\mathbb{Q}$.*
- *Consider the set $S_2 = \{x \in \mathbb{Q} \mid x^2 \leq 2\}$ in the ordered field $\mathbb{Q}$. This set has an upper bound. For example 3. Note that there is no upper bound that is contained in S_2 since there is no rational number that squares to 2.*
- *The set $S_3 = \{x \in \mathbb{Q} \mid x^2 \leq 4\}$ is bounded above. Here 2 is an upper bound that is in S. There is no smaller upper bound than 2 here.*

Definition 1.1.9 (Least Upper Bound) *Let S be a subset in a ordered field $\mathbb{K}$. Then M is said to be a least upper bound if it is an upper bound and there is no upper bound that is smaller than M.*

Another way of saying this is that every other upper bound is larger or equal to M. Therefore, such a least upper bound must be unique.

Remark 1.1.10 *If $S \subseteq \mathbb{K}$ is finite a least upper bound always exists, namely the largest element in S.*

Definition 1.1.11 (Completeness) *An ordered field $\mathbb{K}$ is called complete if every non-empty subset that is bounded from above has a least upper bound.*

As we have seen in the example above $\mathbb{Q}$ is not complete as $\{x \in \mathbb{Q} \mid x^2 \leq 2\}$ does not have a least upper bound.

We are now going to assume the existence of a complete ordered field $\mathbb{R}$ that contains $\mathbb{Q}$ as a subset such that the operations and the relation are compatible with that of $\mathbb{Q}$. This field $\mathbb{R}$ will be called the field of real numbers.

The property laid out in definition 1.1.11 is also called the completeness axiom for the set of real numbers. One can show that any complete ordered field is isomorphic (equivalent) to the set of real numbers. So there are basically no other complete ordered fields.

Remark 1.1.12 *The completeness axiom may also be stated equivalently by using lower bounds instead of upper bounds. A greatest lower bound of a set S is a lower bound that is greater or equal than any other lower bound. If M is an upper bound for S, then $-M$ is a lower bound for the set $-S$. Therefore, given a least upper bound of $-S$ we immediately get greatest lower bound, and vice versa.*

1.2 The structure of the set of real numbers

Given $a, b \in \mathbb{R}$ we are going to use the notations from MATH1025 for the intervals $(a, b), [a, b], (a, b], [a, b), (-\infty, a), (-\infty, a], (a, \infty), [a, \infty), (-\infty, \infty)$. For example

$$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}.$$

Definition 1.2.1 (Supremum and Infimum) *Let $S \subseteq \mathbb{R}$ be a non-empty subset. Then the supremum $\sup S$ of S is defined to be the least upper bound of S in case the set is bounded from above. If the set is not bounded from above we define $\sup S = \infty$.*
The infimum $\inf S$ of S is defined as $-\sup(-S)$.

Note that ∞ is not a number. It is not contained in $\mathbb{R}$. So the supremum is not necessarily a real number. It is however if the set is bounded above.

Examples 1.2.2 *We consider the following sets as subsets of $\mathbb{R}$.*

- $S_1 = (-2, \infty)$ *is bounded below and unbounded above. One gets* $\inf S_1 = -2$ *and* $\sup S_1 = \infty$.
- $S_2 = (-1, 1) \cup [2, 3]$. *In this case* $\inf S_2 = -1$ *and* $\sup S_2 = 3$.
- $S_3 = \mathbb{Z}$. *This set is unbounded from above and below. Therefore,* $\inf S_3 = -\infty$ *and* $\sup S_3 = \infty$.

Theorem 1.2.3 (Archimedean property of $\mathbb{R}$) *For every real number $x \in \mathbb{R}$ there exists a natural number $N \in \mathbb{N}$ such that $x < N$.*

Proof: Suppose, by contradiction, that $\mathbb{N}$ is bounded above by x. Then $\mathbb{N}$ admits a least upper bound $y \in \mathbb{R}$. Then $y - 1$ is not an upper bound and therefore there exists a natural number M such that $M \geq y - 1$. This means that $y \leq M + 1 \in \mathbb{N}$, so y is not an upper bound. This is a contradiction. □

Corollary 1.2.4 *For every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that $0 < \frac{1}{N} < \epsilon$.*

Proof: By the Archimedean property of $\mathbb{R}$ we can choose $N \in \mathbb{N}$ such that $N > \frac{1}{\epsilon}$. The rest follows from Theorem 1.1.4. □

Corollary 1.2.5 *Given $\delta > 0$ and $x > 0$ there exists an $N \in \mathbb{N}$ such that $N\delta \geq x$.*

Proof: Just find N such that $N > x\delta^{-1}$, using the Archimedean property of $\mathbb{R}$.

$\square$

Theorem 1.2.6 (Density of $\mathbb{Q}$ in $\mathbb{R}$) *Let $x \in \mathbb{R}$. Then, for any $\epsilon > 0$ there exists a rational number $q \in \mathbb{Q}$ such that $q \in (x - \epsilon, x + \epsilon)$.*

Proof: Choose $N \in \mathbb{N}$ such that $0 < \frac{1}{N} < \epsilon$. Now consider the set $X = \{y \in \mathbb{N} \mid \frac{y}{N} > x\}$. This set is non-empy by the corollaries above. The set of natural numbers is well-ordered (every non-empty subset has a smallest element) there exists a smallest element M. This means we have $\frac{M-1}{N} \leq x < \frac{M}{N}$. Adding $-\frac{M}{N}$ to both sides gives $-\frac{1}{N} \leq x - \frac{M}{N} < 0$. Therefore,

$$0 < \frac{M}{N} - x \leq \frac{1}{N} < \epsilon.$$

Therefore, the rational number $\frac{M}{N}$ is in the interval $(x, x+\epsilon) \subseteq (x - \epsilon, x + \epsilon)$. $\square$

1.3 Important Inequalities

We define the absolute value $|x| \in \mathbb{R}$ of a real number $x \in \mathbb{R}$ as

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

You can imagine $|x|$ as the distance the number x has from 0. Similarly $|x - y|$ is the *distance* between x and y.

Remark 1.3.1 *One could also define $|x| = \sup\{-x, x\}$. From this it is immediately clear that $x \leq |x|$ and $-x \leq |x|$.*

Theorem 1.3.2 *The modulus has the following properties for all for all $x, y \in \mathbb{R}$.*

- $|x| \geq 0$ *with equality if and only if $x = 0$ (Positivity).*
- $|-x| = |x|$.
- $|xy| = |x||y|$ *(Multiplicativity).*
- $|x + y| \leq |x| + |y|$ *(Triangle Inequality).*

Proof: Positivity and the fact $|-x| = |x|$ follow immediately from the definition. Multiplicativity is easily proved by considering the four different cases of possible signs for x and y. This is left to you, the reader, to check. Since $x \leq |x|$ and $y \leq |y|$ it follows that $x + y \leq |x| + |y|$. Similarly, $-x \leq |x|$ and $-y \leq |y|$ and therefore $-(x + y) \leq |x| + |y|$. Combining these two statements gives the triangle inequality. $\square$

Remark 1.3.3 *Note that (by replacing y by $-y$) the triangle inequality is equivalent to the inequality*

$$|x - y| \leq |x| + |y|$$

for all $x, y \in \mathbb{R}$. It implies

$$|x - y| \leq |x - a| + |y - a|$$

for $x, y, a \in \mathbb{R}$.

1.4 Supplementary material

This section provides a couple of more examples to deepen your understanding of inequalities.

1.4.1 Bernoulli inequality

If $x \geq 0$ and $n \in \mathbb{N}$ then we have

$$(1 + x)^n \geq 1 + nx.$$

The proof is by induction in n, see problem sheets.

1.4.2 Reversed Bernoulli inequality in case $n = \frac{1}{2}$

In this example we assume for a moment that there exists a positive square root for every $y > 0$. This will be proved later in the module. If $x \geq 0$ we have

$$\sqrt{1 + x} \leq 1 + \frac{1}{2}x.$$

Proof: Assume by contradiction that $\sqrt{1 + x} > 1 + \frac{1}{2}x$. Both sides are positive, so squaring it gives

$$1 + x > 1 + x + \frac{1}{4}x^2$$

and we arrive at $x^2 < 0$, which is a contradiction. □

1.4.3 If $x > 1$ then $x^2 > x$

Since $x > 1$ implies also $x > 0$ we have $x^2 > x$. If we assume for a moment that there exists a positive square root for every $y > 0$ this gives

$$\forall y > 1, \quad y > \sqrt{y}.$$

1.4.4 Triangle inequality

For all $x, y \in \mathbb{R}$ we have

$$|x+y| \leq |x| + |y|$$

and as a consequence

$$|x-y| \leq |x-z| + |y-z|$$

for all $x, y, z \in \mathbb{R}$.

1.4.5 The AMGM inequality

This is about arithmetic and geometric means. Turns out the arithmetic mean is always larger equal the geometric mean. For all $x, y \geq 0$ we have

$$\frac{x+y}{2} \geq \sqrt{xy}$$

or equivalently

$$\frac{(x+y)^2}{4} \geq xy.$$

Proof: $(x-y)^2 > 0$ and therefore $x^2 - 2xy + y^2 \geq 0$, which implies $x^2 + 2xy + y^2 \geq 4xy$ and finally

$$(x+y)^2 \geq 4xy.$$

Since $\frac{1}{4} > 0$ (as a simple consequence of $4 > 0$) we can multiply left and right hand side by $\frac{1}{4}$ top obtain the AMGM inequality. $\square$

Chapter 2

Sequences and Convergence

2.1 Sequences

We already met in MATH1025 the term *sequence*. A sequence of real numbers a is a family of numbers that is indexed by the natural numbers, i.e. we give ourselves a real number $a_n \in \mathbb{R}$ for every $n \in \mathbb{N}$. We will write $(a_n)_{n\in\mathbb{N}} = (a_1, a_2, a_3, \ldots)$ to specify a sequence.

Remark 2.1.1 *A sequence of real numbers can also be thought of as function $a : \mathbb{N} \to \mathbb{R}$. The implied notation $a(n)$ instead of a_n is however not costumary in mathematics.*

Examples 2.1.2 *Here are some examples and non-examples.*

(i) $(a_n)_{n\in\mathbb{N}}$ *defined by* $a_n = \frac{1}{n}$ *is a sequence. It equals* $(1, \frac{1}{2}, \frac{1}{3}, \cdots)$.

(ii) $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots)$ *is a sequence. (We assume here that the implied rule for forming the sequence is obvious).*

(iii) $(1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots)$ *is a sequence. (We assume here that the implied rule for forming the sequence is obvious).*

(iv) $\{1, 2, 3, 4, 5, 6, \ldots\}$ *is not a sequence.*

(v) $(1, 2, 3, 4)$ *is not a sequence.*

(vi) $(0, 1, 0, 1, 0, 1, 0, \ldots)$ *is a sequence. It can also be written explicitly as*

$$(a_n)_{n\in\mathbb{N}}, a_n = \frac{1 + (-1)^n}{2}.$$

(vii) the most boring sequence is the constant sequence $(0, 0, 0, 0, \ldots)$*, shortly followed by the constant sequence* $(1, 1, 1, 1, \ldots)$.

(viii) sequences may be recursively defined, for example

$$a_1 = 1, a_2 = 1,$$
$$a_{n+2} = a_{n+1} + a_n$$

defines the Fibonacci sequence $(1, 1, 2, 3, 5, 8, 13, \ldots)$.

You can think of a sequence as a machine that produces real numbers and never stops. Of course in the same way one can also consider sequences with values in any non-empty set (for example complex numbers). Sometimes sequences are also indexed by $\mathbb{N}_0$, in which case it would start with a_0 would be the first element.

Definition 2.1.3 (Monotonicity and Boundedness of Sequences) *A sequence is called*

- *constant if $a_{n+1} = a_n$ for all $n \in \mathbb{N}$.*
- *(monotonically) increasing if $a_{n+1} \geq a_n$ for all $n \in \mathbb{N}$.*
- *strictly (monotonically) increasing if $a_{n+1} > a_n$ for all $n \in \mathbb{N}$.*
- *(monotonically) decreasing if $a_{n+1} \leq a_n$ for all $n \in \mathbb{N}$.*
- *strictly (monotonically) decreasing if $a_{n+1} < a_n$ for all $n \in \mathbb{N}$.*
- *monotonic if it is either increasing or decreasing.*
- *bounded above if there exists $M \in \mathbb{R}$ such that $a_n \leq M$.*
- *bounded below if there exists $M \in \mathbb{R}$ such that $a_n \geq M$.*
- *bounded if it is bounded above and below.*

Remark 2.1.4 *In the last two cases above we say that M is an upper/lower bound for the sequence. A sequence (a_n) is bounded if and only if there exists $C \in \mathbb{R}$ such that $|a_n| \leq C$. The property of being bounded above and below does not depend on the order in the sequence. The corresponding property also makes sense for the set $\{a_n\}$. So, for example, a sequence (a_n) is bounded above if and only if the set $\{a_n\}$ is.*

Examples 2.1.5 *Here are some examples.*

- $(1, 2, 3, 4, \ldots)$ *is strictly increasing and bounded below.*
- $(1, 1, 2, 2, 3, 3, 4, 4, \ldots)$ *is increasing and bounded below.*
- $(1, 0, 1, 0, 1, 0, \ldots)$ *is neither increasing nor decreasing but bounded.*
- $(\frac{1}{n})_{n \in N}$ *is strictly decreasing and bounded.*
- $(-\frac{1}{n})_{n \in N}$ *is strictly increasing and bounded.*

Obviously a sequence (a_n) is decreasing if and only if $(-a_n)$ is increasing.

2.2 Definition of Limits

Definition 2.2.1 (Convergence) *We say the sequence (a_n) converges to $a \in \mathbb{R}$ if the following statement holds. For every $\epsilon > 0$ there exists a number $N \in \mathbb{N}$ such that for every $n > N$ we have*

$$|a_n - a| < \epsilon.$$

In this case we write $a_n \to a$, $\lim_{n\to\infty} a_n = a$, or simply $\lim a_n = a$. If a sequence converges to a we say that it converges.

This means starting from $N + 1$ all elements a_n are in the interval $(a - \epsilon, a + \epsilon)$, so their distance from a is less than ϵ.

Remark 2.2.2 *A sequence (a_n) converges to $a \in \mathbb{R}$ if and only if for every $\epsilon > 0$ there are only finitely many elements a_n outside the interval $(a - \epsilon, a + \epsilon)$.*

At the moment we just said that some limit a must exist for a sequence to converge. The following theorem makes sure this limit is always unique.

Theorem 2.2.3 (Uniqueness of Limits) *A convergent sequence has exactly one limit. In other words, if $\lim a_n = a$ and $\lim a_n = b$, then $a = b$.*

Proof: For any $\epsilon > 0$ we can find N_1 and N_2 such that for $n > N_1$ we have $|a_n - a| < \epsilon/2$ and for $n > N_2$ we have $|a_n - b| < \epsilon/2$. Taking the larger of these we get N such that $n > N$ implies that $|a_n - a| < \epsilon/2$ and $|a_n - b| < \epsilon/2$. Now just use the triangle inequality. Then for $n > N$ we get

$$|a - b| = |(a - a_n) - (b - a_n)| \leq |(a - a_n)| + |(b - a_n)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Since this is true for any $\epsilon > 0$ this implies $|a - b| = 0$ and hence $a = b$. $\square$

It is obvious from the definition that $a_n \to a$ if and only if $a_n - a \to 0$. In other words, the sequence (a_n) converges to a iff the shifted sequence $(a_n - a)$ converges to 0.

Examples 2.2.4 *Here are some examples.*

- *The constant sequence $(a, a, a, a, \ldots)$ converges to a.*
- *The sequence $(\frac{1}{n})_{n\in\mathbb{N}}$ converges to 0.*
- *The sequence $(1 + \frac{1}{n})$ converges to 1.*

Remark 2.2.5 *A sequence (a_n) converges to $a \in \mathbb{R}$ if and only if the sequence $(a_n - a)$ converges to 0. A sequence (a_n) converges to zero if and only if the sequence $(|a_n|)$ converges to 0.*

I will now explain how to prove that simple sequences converge. In order to illustrate this let us look at the sequence $(\frac{2n^2+1}{n^2+3})_{n\in\mathbb{N}}$. We expect this sequence to converge to 2. If we want to prove this we need to show that, given $\epsilon > 0$, for large enough n we have

$$|\frac{2n^2+1}{n^2+3} - 2| < \epsilon.$$

The best way is to simply write down a chain of inequalities starting with the term on the left and then simplify to something easier to handle.

$$|\frac{2n^2+1}{n^2+3} - 2| = |\frac{2n^2+1-2n^2-6}{n^2+3}| = |\frac{-5}{n^2+3}| = \frac{5}{n^2+3} \leq \frac{5}{n^2} \leq \frac{5}{n}.$$

This is always true. Now, given $\epsilon > 0$, if we achieve $\frac{5}{n} < \epsilon$ we will also have

$$|\frac{2n^2+1}{n^2+3} - 2| < \epsilon.$$

Suppose we are given $\epsilon > 0$. Simply find an integer N such that $N > \frac{5}{\epsilon}$. Then $n > N$ implies $|\frac{2n^2+1}{n^2+3} - 2| < \epsilon$. This then proves that $\lim_{n\to\infty} \frac{2n^2+1}{n^2+3} = 2$ by using only the definition of the limit. If the sequences are more complicated this becomes increasingly more painful. Therefore, we will develop theorems that will make our life easier in dealing with limits. Before we discuss some actual theorems let us make some simple observations.

Proposition 2.2.6 (Domination) *Suppose that $|a_n| \leq |b_n|$ and that $b_n \to 0$. Then $a_n \to 0$.*

Proof: We need to show that, given $\epsilon > 0$, we can find N such that for $n > N$ we have $|a_n| < \epsilon$. We know we can find N such that for $n > N$ we have $|b_n| < \epsilon$. But the hypothesis implies that $|a_n| \leq |b_n| < \epsilon$. $\square$

Another two simple observation is that the existence of a limit and its value do not depend on what the sequence looks like in finitely many places and if we shift the sequence by finitely many places.

Proposition 2.2.7 (Shift invariance of the limit) *Suppose that (a_k) converges to $a \in \mathbb{R}$ and let $M \in \mathbb{N}$. Then the sequence (a_{k+M}) converges to a as well.*

Proof: Suppose $\epsilon > 0$. Then we can find N such that for $k > N$ we have $|a_k - a| < \epsilon$. Therefore $|a_{k+M} - a| < \epsilon$ for $k > N$ because then in particular $n + M > N$. $\square$

Proposition 2.2.8 (Finitely many terms don't matter) *Suppose (a_k) and (b_k) are sequences and suppose that $a_k \to a$. Assume that $a_k = b_k$ for all but finitely many k. Then $b_k \to a$.*

Proof: Since the set $\{k \in \mathbb{N} \mid a_k \neq b_k\}$ is finite it has a maximal element M. Thus, for $n > M$ we have $a_n = b_n$. Now, given $\epsilon > 0$, we can find N_1 such that $n > N_1$ implies $|a_n - a| < \epsilon$. Therefore, for $n > N := \max\{N_1, M\}$ we have $|b_n - a| = |a_n - a| < \epsilon$. $\square$

2.3 Theorems about Limits

This section collects important theorems about limits that make their computation easier. We start with some simple observations.

Theorem 2.3.1 (Boundedness of convergent sequences) *Any convergent sequence is bounded.*

Proof: Suppose that $a_n \to a$. Now choose $\epsilon = 1$. We know, by the definition of convergence, that for some $N > 0$ we have $|a_n - a| < 1$ if $n > N$. In this case the triangle inequality gives us $|a_n| \leq |a| + 1$ for $n > N$. Therefore, choosing $C = \sup\{|a_n||n \leq N\} \cup \{|a| + 1\}$ we obtain $|a_n| \leq C$. $\square$

Given sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ we can form new sequences $(a_n + b_n)$ and $(a_n \cdot b_n)$ or even $(\lambda a_n + \mu b_n)$ if $\mu, \lambda \in \mathbb{R}$.

Theorem 2.3.2 *Suppose that (a_n) and (b_n) are sequences such that $a_n \to 0$ and such that (b_n) is bounded. Then $a_n b_n \to 0$.*

Proof: Let $C > 0$ such that $|b_n| \leq C$. Such a constant exists because (b_n) was assumed to be bounded. Given $\epsilon > 0$ we need to show that there exists $N > 0$ such that $|a_n b_n| < \epsilon$ if $n > N$. Since $a_n \to 0$ we can find N such that $n > N$ implies that $|a_n| < \epsilon \cdot C^{-1}$. Then, whenever $n > N$ we have $|a_n b_n| < \epsilon C^{-1} C = \epsilon$ which is what we wanted. $\square$

Theorem 2.3.3 (Algebra of Limits) *Suppose that $a_n \to a \in \mathbb{R}$ and $b_n \to b \in \mathbb{R}$. Then*

(i) $a_n + b_n \to a + b$.

(ii) $a_n \cdot b_n \to ab$.

(iii) $\lambda a_n + \mu b_n \to \lambda a + \mu b$ *for all* $\mu, \lambda \in \mathbb{R}$.

(iv) if $b_n \neq 0$ and $b \neq 0$ then $\frac{a_n}{b_n} \to \frac{a}{b}$

Proof:

(i) Choose $\epsilon > 0$ and $N > 0$ such that if $n > N$ implies $|a_n - a| < \frac{\epsilon}{2}$ and $|b_n - b| < \frac{\epsilon}{2}$. Then by the triangle inequality

$$|a_n + b_n - a - b| \leq |a_n - a| + |b_n - b| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Hence, for $n > N$ we have established the inequality above, which implies $|a_n + b_n - a - b| < \epsilon$. This shows that $a_n + b_n \to a + b$.

(ii) By assumption $a_n - a \to 0$. Since (b_n) converges it must be bounded. Therefore, $a_n b_n + (-a)b_n = (a_n - a)b_n \to 0$. Similarly, since the constant sequence (a) is bounded this implies $a(b_n - b) \to 0$. The statement then follows from (i) above as

$$a_n b_n - ab = (a_n - a)b_n + a(b_n - b) \to 0.$$

(iii) is a simple consequence of the two statements above combined.

(iv) It is enough to show that $\frac{1}{b_n} \to \frac{1}{b}$ as the more general statement follows in combination with (ii). First note that $|b_n|$ is bounded from below by a positive constant $\delta > 0$. In order to see that note that as $|b| > 0$ we can choose $N > 0$ such that for any $n > N$ we have $|b_n - b| < \frac{|b|}{2}$ which implies that $|b_n| > \frac{|b|}{2}$ Indeed,

$$|b| \leq |b - b_n + b_n| \leq |b_n - b| + |b_n| < \frac{|b|}{2} + |b_n|,$$

which implies the inequality. Now choose $\delta = \inf\{|b_n| \mid n \leq N\} \cup \{\frac{|b|}{2}\} > 0$. Therefore, the sequence $\frac{1}{b_n b}$ is bounded by δ^{-2}. Now simply observe that $|\frac{1}{b_n} - \frac{1}{b}| = |\frac{1}{b_n b}(b_n - b)|$ and use Theorem 2.3.2.

□

Theorem 2.3.4 (Squeeze Theorem) *Suppose that $(a_n), (b_n), (c_n)$ are sequences such that $a_n \to L$, $c_n \to L$ and $a_n \leq b_n \leq c_n$. Then $b_n \to L$.*

Proof: Given $\epsilon > 0$ we need to show that there is $N > 0$ such that for all $n > N$ we have

$$|b_n - L| < \epsilon.$$

We know that $a_n \to a$ and $c_n \to c$. Therefore, we can choose $N > 0$ such that for $n > N$ we have $|a_n - L| < \epsilon$ and $|c_n - L| < \epsilon$. In particular we have $a_n > L - \epsilon$ and $c_n < L + \epsilon$. Therefore for $n > N$ we get

$$b_n \geq a_n > L - \epsilon,$$
$$b_n \leq c_n < L + \epsilon.$$

We conclude that $|b_n - L| < \epsilon$ for $n > N$. $\square$

Sometimes it is not easy to guess the limit of a sequence. The sequence

$$(1, 1 + \frac{1}{2^2}, 1 + \frac{1}{2^2} + \frac{1}{3^2}, 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2}, \ldots)$$

is an example. In order to check from the definition that the sequence converges we would need to know what the limit is. But we don't at the moment. It will turn out much later in your studies that the sequence above converges to $\frac{\pi^2}{6}$. The methods in this course are not sufficient to prove this. We will however be able to show that the sequence converges and we will come back to this particular sequence at a later stage.

Theorem 2.3.5 (Monotone Convergence Theorem) *Let $(a_n)_{n\in\mathbb{N}}$ be an increasing sequence that is bounded from above. Then (a_n) converges to* $\sup\{a_n \mid n \in \mathbb{N}\}$.

Proof: Suppose we are given $\epsilon > 0$. Let $a = \sup\{a_n \mid n \in \mathbb{N}\}$. We need to show that for some $N \in \mathbb{N}$ we have $|a_n - a| < \epsilon$ for $n > N$. Note that $a_n \leq a$ since a is an upper bound, so $|a_n - a| = a - a_n$. So we need to show that $a - a_n < \epsilon$ for $n > N$. Since the sequence is increasing this would already follow from $a - a_N < \epsilon$. Thus, we only need to find $N \in \mathbb{N}$ such that $a - a_N < \epsilon$. We argue by contradiction that such an N always exists. If it did not exist, we would have $a_n \leq a - \epsilon$ for all $n \in \mathbb{N}$. But this means that $a - \epsilon$ is an upper bound smaller than a, contradicting the fact that a is the least upper bound. $\square$

Corollary 2.3.6 *Let $(a_n)_{n\in\mathbb{N}}$ be a decreasing sequence that is bounded from below. Then (a_n) converges to* $\inf\{a_n \mid n \in \mathbb{N}\}$.

Proof: Follows by applying theorem 2.3.5 to the sequence $(-a_n)$ and noting that $a_n \to a$ iff $-a_n \to -a$. $\square$

If $S \subseteq \mathbb{R}$ is a subset then we say a sequence (a_n) is *contained in S* iff $a_n \in S$ for all $n \in \mathbb{N}$. Recall that the supremum is not always contained in the set S. The following theorem clarifies that it can always be approximated by elements in S.

Theorem 2.3.7 *Suppose that $S \subseteq \mathbb{R}$ is non-empty and bounded above. Then there exists a sequence (a_n) in S that converges to* $\sup S$.

Proof: Let $a = \sup S$. For $n \in \mathbb{N}$ is, the number $a - \frac{1}{n}$ can not be an upper bound for S, because a is the least upper bound. Since a is an upper bound that implies that there must be an element of S in the interval $(a - \frac{1}{n}, a]$. Call this element a_n.

This way we obtain a sequence (a_n). By construction $0 \leq a - a_n < \frac{1}{n}$. By Theorem 2.2.6 $(a - a_n)$ therefore converges to 0. This implies that $a_n \to a$, as claimed. $\square$

We will say that a sequence (a_n) is *contained in* a subset $S \subseteq \mathbb{R}$ if $a_n \in S$ for all $n \in \mathbb{N}$.

Remark 2.3.8 *If a convergent sequence (a_n) is contained in a set $S \subseteq \mathbb{R}$ this does not mean that its limit is contained in S too. For example the sequence $(\frac{1}{n})$ is contained in $(0, 1]$, but its limit 0 is not.*

The above shows that limits do not respect the $<$ relation. It may be that $a_n > 0$ but in the limit we get $a_n \to a = 0 \not> 0$. This may not happen with the $\leq$ relation as the following discussion shows.

Theorem 2.3.9 (Stability of non-negativity under limits) *Suppose that (a_n) is converges to $a \in \mathbb{R}$ and $a_n \geq 0$ for all $n \in \mathbb{N}$. Then $a \geq 0$.*

Proof: Let us denote the limit of (a_n) by a. We need to show $a \geq 0$. Assume by contradiction that $a < 0$. Define $\epsilon := -a$, so that $|x - a| < \epsilon$ implies that $x < 0$. Since $a_n \to a$ we can choose N such that $n > N$ implies $|a_n - a| < \epsilon$. This implies that for $n > N$ the sequence elements $a_n < 0$ which is in contradiction with our assumption. $\square$

This implies that unlike $<$ the relation $\leq$ is *closed*. It survives taking limits.

Corollary 2.3.10 (Stability of closed inequalities under limits) *Suppose that $a_n \to a$, $b_n \to b$ and $a_n \leq b_n$. Then $a \leq b$.*

Proof: Simply apply the above theorem to the sequence $b_n - a_n$. $\square$

The same applies to the $\geq$ relation. This follows from the above by switching the role of a_n and b_n.

Corollary 2.3.11 (Closed intervals are closed) *Suppose $x_n \to x$ is contained in a closed interval $[a, b]$. Then $x \in [a, b]$.*

Definition 2.3.12 *A subset $S \in \mathbb{R}$ is called closed if for any convergent sequence (a_n) in S its limit a is also contained in S.*

Example 2.3.13 *By the above theorem the closed intervals are closed. $\mathbb{R}$ is closed. $\mathbb{Q} \subseteq \mathbb{R}$ is not closed. $(0, 1)$ is not closed. $\mathbb{R} \backslash \mathbb{Q}$ is not closed.*

Before we state the next theorem, let us call a sequence of intervals I_n *nested* if $I_{n+1} \subseteq I_n$. For example the sequence $I_n = (0, \frac{1}{n})$ is a nested sequence of open intervals. Note that for this particular nested sequence the intersection $\cap_n I_n$ is empty, i.e. there is no element that is contained in all of these intervals. The next theorem guarantees that this can not happen if the intervals are closed. For these purposes we allow closed intervals to have the form $[a, a]$ and mean by that the set $\{a\}$.

Theorem 2.3.14 (Nested Sequence Lemma) *Suppose that I_n be a nested sequence of non-empty closed intervals $I_n = [a_n, b_n]$. Then $\cap_n I_n \neq \emptyset$.*

Proof: By assumption a_n is an increasing sequence, and b_n is a decreasing sequence. Moreover,

$$a_n \leq b_m$$

for all $n, m \in \mathbb{N}$. In particular both sequences are monotonic and bounded and therefore converge, $a_n \to a$ and $b_n \to b$. Then we have that

$$a_n \leq a \leq b \leq b_n.$$

This means that the non-empty set $[a, b]$ is contained in I_n for every n. $\square$

As a nice application of this lemma we obtain another proof of what you have already discussed in MATH1025.

Theorem 2.3.15 *The set $(0, 1)$ is not countable.*

Proof: Suppose by contradiction that $(0, 1)$ is countable. This means we can find a sequence (x_n) such that for every $x \in (0, 1)$ there exists a unique $n \in \mathbb{N}$ with $x = x_n$. Now we inductively construct a nested sequence of closed intervals as follows. Start with I_1 a closed interval in $(0, 1)\backslash\{x_1\}$. Next, inside I_1 choose a closed interval I_2 in $I_1\backslash\{x_2\}$ (if x_2 is not contained in I_1 we set $I_1\backslash\{x_2\} = I_1$). Proceed inductively by choosing I_{k+1} in $I_k\backslash\{x_{k+1}\}$. We obtain a nested sequence of closed non-empty intervals inside $(0, 1)$. Hence, there must be a point y that is in all these intervals. By the above $y = x_m$ for some m. But by construction x_m is not contained in I_m. This is the contradiction that we needed to conclude the proof. $\square$

Corollary 2.3.16 *The set $\mathbb{R}$ is not countable.*

2.4 Supplementary material

This section provides extra examples to deepen your understanding.

It is quite common to have difficulties grasping the concept of the ϵ, N-definition of convergence at first. I provide below a couple of examples to show you how to

proof convergence of sequences using the definition only. I do hope you will find it helpful for your understanding.
There is however an important disclaimer: The ϵ, N-definition of convergence is rarely applied to examples but it is a theoretical concept that is used for proofs. Do not therefore approach these problems in an algorithmic manner. It is not a learning outcome of this module to manipulate expressions to give ϵ, N-proofs of the type below. The aim is to **understand** the concept and use it in proofs. This means you need to be able to apply the idea to new situations that you have not encountered before. Of course this means that you will not have seen a detailed worked solution, and the only way to reach this understanding is through independent hard work.
Therefore, please do not abuse these notes by simply just learning an algorithmic way to give ϵ, N-proofs. Merely use them to deepen your understanding of the theoretical concept. Also, please let me know if you find mistakes or typos.

2.4.1 The sequence $a_n = \frac{1}{n}$

The sequence is strictly decreasing and bounded above by 1 as well as bounded below by 0. The infimum $\inf\{a_n \mid n \in \mathbb{N}\}$ is 0, the supremum is 1. This sequence converges to zero and is therefore a Cauchy sequence.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \epsilon^{-1}$. As usual such an $N \in \mathbb{N}$ exists by the Archimedean property. Then

$$|a_n| = |\frac{1}{n}| < |\frac{1}{N}| < (\epsilon^{-1})^{-1} = \epsilon$$

for all $n > N$.
Direct proof of Cauchy-property Given $\epsilon > 0$ take $N > 2\epsilon^{-1}$, then

$$|a_n - a_m| = |\frac{1}{n} - \frac{1}{m}| \leq |\frac{1}{n}| + |\frac{1}{m}| < |\frac{1}{N}| + |\frac{1}{N}| = \frac{2}{N} < (\epsilon^{-1})^{-1} = \epsilon$$

for all $m, n > N$.

2.4.2 The sequence $a_n = \frac{1}{\sqrt{n}}$

The sequence is strictly decreasing and bounded above by 1 as well as bounded below by 0. We assume here for this example that there exists a unique positive root $\sqrt{x}$ of x for any $x \geq 0$ with the property $\sqrt{x}\sqrt{x} = x$. This is all we need. First note that $n > N > 0$ implies that $\sqrt{n} > \sqrt{N} > 0$. This can be proved very easily by contradiction. Uniqueness implies $\sqrt{x^2} = x$ for $x > 0$ and more generally $\sqrt{x^2} = |x|$.
This sequence converges to zero and is therefore a Cauchy sequence.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \epsilon^{-2}$. As usual such an $N \in \mathbb{N}$ exists by the Archimedean property. Then

$$a_n = |\frac{1}{\sqrt{n}}| < |\frac{1}{\sqrt{N}}| < (\sqrt{\epsilon^{-2}})^{-1} = \epsilon$$

for all $n > N$.
Direct proof of Cauchy-property Given $\epsilon > 0$ take $N > 2\epsilon^{-2}$, then

$$|a_n - a_m| = |\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{m}}| \leq |\frac{1}{\sqrt{n}}| + |\frac{1}{\sqrt{m}}| < |\frac{1}{\sqrt{N}}| + |\frac{1}{\sqrt{N}}| = \frac{2}{\sqrt{N}} < 2(2\epsilon^{-1})^{-1} = \epsilon$$

for all $m, n > N$.

2.4.3 The sequence $a_n = \sqrt{n+1} - \sqrt{n}$

The sequence is strictly decreasing (prove this). It is bounded above by 2 and bounded below by 0. This sequence converges to zero and is therefore a Cauchy sequence.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \epsilon^{-2}$. As usual such an $N \in \mathbb{N}$ exists by the Archimedian property. Then

$$|a_n| = \sqrt{n+1} - \sqrt{n}| = \frac{n+1-n}{\sqrt{n+1}+\sqrt{n}} \leq \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{N}} < \epsilon.$$

for all $n > N$.
Direct proof of Cauchy-property Given $\epsilon > 0$ take $N > 2\epsilon^{-2}$, then

$$\begin{aligned} |a_n - a_m| &= |\sqrt{n+1} - \sqrt{n} - (\sqrt{m+1} - \sqrt{m})| \\ = |\frac{n+1-n}{\sqrt{n+1}+\sqrt{n}} - \frac{m+1-m}{\sqrt{m+1}+\sqrt{m}}| &\leq \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{m}} < \frac{1}{\sqrt{N}} + \frac{1}{\sqrt{N}} < \epsilon. \end{aligned}$$

for all $m, n > N$.

2.4.4 The sequence $a_n = \frac{3n+2}{4n-1}$

This sequence converges to $\frac{3}{4}$ and is therefore a Cauchy sequence.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \frac{1}{\epsilon}$. As usual such an $N \in \mathbb{N}$ exists by the Archimedian property. Then

$$|a_n - \frac{3}{4}| = |\frac{3n+2}{4n-1} - \frac{3}{4}| = |\frac{3n+2-3n+3/4}{4n-1}| = \frac{1}{4}\frac{11}{4n-1} \leq \frac{3}{4n-1} \leq \frac{1}{n} < \epsilon$$

for all $n > N$. Here we have used $n \geq 1$ and therefore $4n - 1 \leq 3n$.
Direct proof of Cauchy-property Given $\epsilon > 0$ take $N > 4\epsilon^{-1}$, then

$$\begin{aligned} |a_n - a_m| &= |\frac{3n+2}{4n-1} - \frac{3m+2}{4m-1}| = |\frac{(3n+2)(4m-1) - (3m+2)(4n-1)}{(4n-1)(4m-1)}| \\ &= |\frac{8m - 3n + 3m - 8n}{(4n-1)(4m-1)}| \leq \frac{11n + 11m}{9nm} \leq 2(\frac{1}{n} + \frac{1}{m}) \leq \frac{4}{N} < \epsilon \end{aligned}$$

for all $m, n > N$. Here we have used the triangle inequality for $|11n - 11m| \leq 11n + 11m$ and also that $\frac{11}{9} < 2$ as well as $4n - 1 \geq 3n$.

2.4.5 The sequence $a_n = \sqrt{1 + \frac{1}{n}}$

This sequence converges to 1 and is therefore a Cauchy sequence.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \frac{1}{\epsilon}$. As usual such an $N \in \mathbb{N}$ exists by the Archimedean property. Then

$$|a_n - 1| = |\sqrt{1 + \frac{1}{n}} - 1| = \frac{1 + \frac{1}{n} - 1}{\sqrt{1 + \frac{1}{n}} + 1} \leq \frac{1}{n} < \epsilon$$

for all $n > N$. We have used that $\sqrt{x} \geq 0$ for all $x \geq 0$.
Direct proof of Cauchy-property Given $\epsilon > 0$ take $N > 2\epsilon^{-1}$, then

$$|a_n - a_m| = |\sqrt{1 + \frac{1}{n}} - \sqrt{1 + \frac{1}{m}}| = \frac{\frac{1}{n} + \frac{1}{m}}{\sqrt{1 + \frac{1}{n}} + \sqrt{1 + \frac{1}{m}}} \leq \frac{1}{n} + \frac{1}{m} < \frac{2}{N} < \epsilon$$

for all $m, n > N$. Here we have used $\sqrt{1 + x} \geq 1$ for $x \geq 0$ (prove this by contradiction).

2.4.6 The alternating sequence $a_n = (-1)^n$

This sequence does not converge. The subsequence a_{2n} is the constant sequence 1, that converges to 1. The subsequence a_{2n+1} converges to -1. Thus, the sequence has two accumulation points -1 and $+1$. If follows that $\limsup a_n = +1$ and $\liminf a_n = -1$.

2.4.7 The sequence $a_n = q^n$ for $0 < |q| < 1$

This sequence converges to zero. To show this use $|a_n| = |q|^n \to 0$. We have $\frac{1}{|q|} > 1$ and therefore $\frac{1}{|q|} = 1+r$ for some $r > 0$. Next use Bernoulli's inequality $|q|^{-n} \leq 1+nr$.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \frac{1}{r\epsilon}$. Then

$$|a_n| = \frac{1}{|q|^n} \geq \frac{1}{1 + nr} \leq \frac{1}{nr} < \frac{1}{Nr} < \epsilon.$$

if $n > N$. The sequence is Cauchy since it converges. We will from now on omit the direct proof since it follows indirectly.

2.4.8 The sequence $a_n = 1 - \frac{1}{n}$ for n odd, $a_n = 1 - \frac{1}{n^2}$ for n odd

More precisely

$$a_n = \begin{cases} 1 - \frac{1}{n} & n \text{ odd} \\ 1 - \frac{1}{n^2} & n \text{ even} \end{cases}.$$

The sequence is bounded above and below by 1 and 0 respectively. This sequence is not monotonic but it converges to its least upper bound, namely 1.

Direct proof of convergence: Given $\epsilon > 0$ take $N > \frac{1}{\epsilon}$. Then, if n is odd and $n > N$ we have

$$|a_n - 1| = |\frac{1}{n}| < |\frac{1}{N}| < \epsilon.$$

If n is even and $n > N$ we have

$$|a_n - 1| = |\frac{1}{n^2}| < |\frac{1}{N^2}| < |\frac{1}{N}| < \epsilon.$$

Thus, for all $n > N$ we have $|a_n - 1| < \epsilon$. The sequence is Cauchy since it converges. We will from now on omit the direct proof since it follows indirectly.

2.4.9 The sequence $a_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$

This sequence is bounded above by 1 and bounded below by -1. This sequence can also be defined recursively as $a_1 = 1$, $a_n = a_{n-1} + \frac{1}{n}$ for $n \geq 2$. This sequence is not bounded above (see lecture notes) and therefore it does not converge. Note however that $|a_{n+1} - a_n| = \frac{1}{n+1} \to 0$. In fact for any $k \in \mathbb{N}$ we have

$$|a_{n+k} - a_n| \to 0.$$

This is still **not** a Cauchy sequence because the right hand side can not be made smaller than $\epsilon > 0$ for all $k \in \mathbb{N}$ simultaniously.

2.4.10 The sequence $a_n = \sqrt{n^2 + n} - n$

This sequence converges to $\frac{1}{2}$.
Direct proof of convergence: Given $\epsilon > 0$ take $N > \frac{1}{\epsilon}$. As usual such an $N \in \mathbb{N}$ exists by the Archimedian property. Then

$$|a_n - \frac{1}{2}| = |\sqrt{n^2+n} - n - \frac{1}{2}| = |\frac{n^2 + n - n^2 - \frac{1}{2}(\sqrt{n^2+n} + n)}{\sqrt{n^2+n} + n}|$$
$$= |\frac{n - \frac{1}{2}(\sqrt{n^2+n} + n)}{\sqrt{n^2+n} + n}| \leq |\frac{\frac{1}{2}(\sqrt{n^2+n} - n)}{n}| = \frac{1}{2}(\sqrt{1 + \frac{1}{n}} - 1) \leq \frac{1}{2n} < \epsilon.$$

for all $n > N$.

2.4.11 The sequence $a_n = (1 + (-1)^n)n + \frac{1}{n}$

This sequence is also given by

$$a_n = \begin{cases} \frac{1}{n} & n \text{ odd} \\ 2n + \frac{1}{n} & n \text{ even} \end{cases}.$$

Consider the sub-sequence $a_{2n} = 4n + \frac{1}{2n} \geq 4n > n$. By the Archimedian property for any $C > 0$ there exists $n \in \mathbb{N}$ with $a_{2n} > C$ and therefore there is no upper bound. Hence, a_{2n} is not bounded above and therefore a_n is also not bounded above. The sequence has a limit point 0. Indeed, the subsequence $a_{2n+1} = \frac{1}{2n+1}$ converges to 0. The sequence is not monotonic. We have $\limsup a_n = \infty$ (sequence is unbounded above) and $\liminf a_n = 0$.

2.4.12 The sequence $a_1 = \frac{1}{2}$, $a_{n+1} = \frac{1}{3}\left(a_n + \frac{1}{a_n}\right)$

We again assume the existence of $\sqrt{2}$. Then this sequence converges to $\frac{1}{\sqrt{2}}$. Here a direct proof might not be so easy any more. We first establish a bound for this sequence. We will show this bound by remarking that the function $f(x) = \frac{1}{3}(x + \frac{1}{x})$ maps the interval $[\frac{1}{2}, 1]$ to itself. Indeed, if $x \geq \frac{1}{2}$ and $x \leq 1$ then

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

Hence, all the a_n are contained in $[\frac{1}{2}, 1]$, which prove the sequence is bounded. Now note that for $x \in [\frac{1}{2}, 1]$ we have

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

using $\sqrt{2} \leq \frac{3}{2}$ which follows from $9/4 > 2$. This shows by induction that $|a_n - \frac{1}{\sqrt{2}}| \leq (\frac{2}{3})^n$. Since $(\frac{2}{3})^n \to 0$ this implies the claim. Note that this proof relies on the existence of $\sqrt{2}$. The sequence is not monotonic. Otherwise this proof would have been much quicker using the monotone convergence theorem.

Chapter 3

Subsequences

3.1 Definition and convergence properties

Definition 3.1.1 (Subsequence) *A sequence (b_k) is a subsequence of a sequence (a_n) if there exists a strictly increasing sequence of positive integers (n_k) such that $b_k = a_{n_k}$ for all $k \in \mathbb{N}$.*

In other words, all the terms of (b_k) must occur in (a_n), and they must occur in the same order, but we are allowed to omit some (possibly infinitely many) terms from (a_n).

Examples 3.1.2 *(i) The constant sequence $(b_k) = (1, 1, 1, 1, \ldots)$ is a subsequence of $a_n = (-1)^n$, with $n_k = 2k$. It's the subsequence consisting of only the even-numbered terms of (a_n). From this example we see that a divergent (that is non-convergent) sequence can have a convergent subsequence.*

(ii) $b_k = \dfrac{1}{2^k}$ is a subsequence of $a_n = 1/n$, with $n_k = 2^k$. In this case, both (a_n) and (b_k) converge, and their limits are equal.

(iii) $b_k = \dfrac{1}{|3k-11|}$ is not *a subsequence of $a_n = 1/n$ even though every term of (b_k) is the reciprocal of a positive integer, so appears somewhere in the sequence (a_n). The problem is that the terms don't occur in the right order:*

$$\begin{aligned}(b_k) &= \left(\frac{1}{8}, \frac{1}{5}, \frac{1}{2}, 1, \frac{1}{4}, \frac{1}{7}, \ldots\right) \\ &= (a_8, a_5, a_2, a_1, a_4, a_7, \ldots)\end{aligned}$$

(iv) Every sequence (a_n) is a subsequence of itself! We take n_k to be the increasing sequence $n_k = k$.

(v) *Given any sequence (a_n), and any positive integer p, $b_k = a_{k+p}$ is a subsequence of (a_n), with $n_k = k + p$. This is just the subsequence obtained by omitting the first p terms of (a_n).*

$\square$

As we have seen, it is possible for a divergent sequence to have a convergent subsequence. On the other hand, if a sequence is convergent, with limit L, one expects that all its subsequences should also converge to L. This turns out to be true.

Theorem 3.1.3 (Convergence of Subsequences) *If $a_n \to L$ and (b_k) is a subsequence of (a_n), then $b_k \to L$.*

Proof: Let $\varepsilon > 0$ be given. Since $a_n \to L$, there exists $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for all $n \geq N$. Since (b_k) is a subsequence of (a_n), $b_k = a_{n_k}$ where n_k is a strictly increasing sequence of positive integers. Now $n_1 \geq 1$, and if $n_k \geq k$ then $n_{k+1} \geq n_k + 1 \geq k + 1$. Hence, by induction, $n_k \geq k$ for all k. Hence, for all $k \geq N$, $n_k \geq n_N \geq N$, so $|b_k - L| = |a_{n_k} - L| < \varepsilon$. Hence $(b_k) \to L$. $\square$

We can use this to give a sneaky proof that $a_n = (-1)^n$ doesn't converge, without having to resort to an ε–N argument:

Example 3.1.4 *Claim: $a_n = (-1)^n$ doesn't converge to any limit.*

Proof: Assume, to the contrary, that $a_n \to L$. Then, by Theorem 3.1.3, the subsequences a_{2k} and a_{2k+1} also converge to L. But $a_{2k} = 1 \to 1$ and $a_{2k+1} = -1 \to -1$, so $1 = L = -1$, a contradiction. $\square$

3.2 The Bolzano-Weierstrass Theorem

We have seen that all subsequences of a convergent sequence are convergent, and that non-convergent sequences can sometimes have convergent subsequences. The goal of this section is to prove a famous and very powerful result called the Bolzano-Weierstrass Theorem, which says that every *bounded* sequence has at least one convergent subsequence. It won't be immediately clear why this fact is such a big deal. Later we will use it to prove the Extreme Value Theorem, which is a crucial result when one comes to define integration carefully, for example.

The strategy of proof is quite straightforward: we first prove that *every* sequence has a monotonic subsequence. Having done this, the Bolzano-Weierstrass Theorem follows immediately from the Monotone Convergence Theorem.

Lemma 3.2.1 *Every sequence has a monotonic subsequence.*

Proof: Let the sequence in question be (a_n). We make the following definition: a term a_m is *dominant* if $a_n \leq a_m$ for all $n > m$ (that is, a_m is at least as big as all subsequent terms). Consider the set of all dominant terms of the sequence (a_n). There are exactly two possibilities: either (i) the set of dominant terms is infinite, or (ii) the set of dominant terms is finite (which includes the case where there are no dominant terms).
Case (i): if the set of dominant terms is infinite, then the subsequence consisting of just these terms is, by definition, decreasing. Hence (a_n) has a decreasing subsequence.
Case (ii): if the set of dominant terms is finite (or empty), there is a term in the sequence, a_m say, beyond which there are no dominant terms. Let $n_1 = m + 1$. Since a_{n_1} is *not* dominant, there exists $n > n_1$ such that $a_n > a_{n_1}$. Choose such an n and call it n_2. Then a_{n_2} is also not dominant (since $n_2 > n_1 > m$), so there exists a positive integer greater than n_2, let's call it n_3, such that $a_{n_3} > a_{n_2}$. In this way we construct a strictly increasing sequence of positive integers $n_1, n_2, n_3, \ldots$ such that $a_{n_{k+1}} > a_{n_k}$ for all k. But then (a_{n_k}) is an increasing subsequence of (a_n). Hence (a_n) has an increasing subsequence.
In either case, we see that (a_n) has a monotonic subsequence, as was to be proved.
□

Theorem 3.2.2 (The Bolzano-Weierstrass Theorem) *Let (a_n) be a bounded real sequence. Then (a_n) has a convergent subsequence.*

Proof: By Lemma 3.2.1, (a_n) has a monotonic subsequence (a_{n_k}). But the range of (a_{n_k}) is a subset of the range of (a_n), so (a_{n_k}) is also bounded (by the same upper and lower bounds as (a_n)), and hence converges by the Monotone Convergence Theorem (Theorem 2.3.5). □

Of course, the converse of this Theorem is false: just because a sequence has a convergent subsequence, it doesn't follow that it's bounded. For example

$$a_n = \begin{cases} n & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$$

has a convergent subseqence, $a_{2k} = 0 \to 0$, but is clearly unbounded above.

Example 3.2.3 (revisited) *Recall we defined the sequence $a_n = \sin n$ and noted that its terms "bounce around" in the interval $(-1, 1)$ seemingly at random. By Theorem 3.2.2 we know that a_n has at least one convergent subsequence, although we have absolutely no idea how to write it down, or what its limit might be. (In fact, given any $L \in [-1, 1]$ it turns out that this sequence has a subsequence converging to L, but proving this goes somewhat beyond the scope of this course).*

3.3 The Cauchy property

One slightly inconvenient thing about the definition of convergence is that, to prove a sequence converges from first principles, we must first know what its limit is. The Monotone Convergence Theorem allows us to prove convergence of some sequences without knowing their limits, but this only works in the special case where the sequence happens to be monotonic (increasing or decreasing). In this section we will study a criterion which turns out to be *equivalent* to convergence, for *all* sequences, but which makes no mention of limits.

Definition 3.3.1 (Cauchy sequence) *A real sequence (a_n) is Cauchy (or has the Cauchy property) if, for each $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that, for all $n, m \geq N$, $|a_n - a_m| < \varepsilon$.*

Remark This looks very similar to the definition of convergence. It says that, given any $\varepsilon > 0$ (no matter how small), there is a point in the sequence beyond which *all terms lie closer than distance ε from one another*. Note that it makes no mention of any limit.

We will show that real sequences converge if and only if they have the Cauchy property.

Lemma 3.3.2 *If (a_n) converges then (a_n) is Cauchy.*

Proof: Assume $a_n \to L$. Let $\varepsilon > 0$ be given. Then $\varepsilon/2 > 0$, so there exists $N \in \mathbb{N}$ such that, for all $n \geq N$,

$$|a_n - L| < \frac{\varepsilon}{2}.$$

But then, for all $n, m \geq N$,

$$|a_n - a_m| = |(a_n - L) - (a_m - L)| \leq |a_n - L| + |a_m - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

by the Triangle Inequality. Hence (a_n) is Cauchy. □

Lemma 3.3.3 *If (a_n) is Cauchy then (a_n) is bounded.*

Proof: Exercise. Choose $\epsilon = 1, m = N + 1$ and then use the triangle inequality. $\square$

Lemma 3.3.4 *Let (a_n) be Cauchy, and assume some subsequence of (a_n) converges to L. Then (a_n) converges to L.*

Proof: Let (a_{n_k}) be the subsequence which converges to L. Let $\varepsilon > 0$ be given. Since $a_{n_k} \to L$, there exists $N_1 \in \mathbb{N}$ such that, for all $k \geq N_1$, $|a_{n_k} - L| < \varepsilon/2$. Further, since (a_n) is Cauchy, there exists $N_2 \in \mathbb{N}$ such that, for all $n, m \geq N_2$, $|a_n - a_m| < \varepsilon/2$. Let $N = \max\{N_1, N_2\}$. Then, for all $n \geq N$,

$$\begin{aligned}
|a_n - L| &= |(a_n - a_{n_N}) + (a_{n_N} - L)| \\
&\leq |a_n - a_{n_N}| + |a_{n_N} - L| && \text{(Triangle Inequality)} \\
&< \frac{\varepsilon}{2} + |a_{n_N} - L| && \text{(since } n \geq N_2\text{, and } n_N \geq N \geq N_2\text{)} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} && \text{(since } N \geq N_1\text{).}
\end{aligned}$$

Hence, $a_n \to L$. $\square$

Theorem 3.3.5 (Convergence of Cauchy sequencies) *A real sequence converges if and only if it is Cauchy.*

Proof: If (a_n) converges, it is Cauchy, by Lemma 3.3.2. Conversely, if (a_n) is Cauchy, it is bounded (Lemma 3.3.3), so has a convergent subsequence (by the Bolzano-Weierstrass Theorem) and so converges (Lemma 3.3.4). $\square$

3.4 Accumulation points, $\limsup$, $\liminf$

Definition 3.4.1 *Let (a_n) be a sequence. Then $a \in \mathbb{R}$ is called an accumulation point of (a_n) if the set*

$$\{k \in \mathbb{N} \mid |a_k - a| < \epsilon\}$$

is infinite for any $\epsilon > 0$.

Note the difference to convergence, where it is required that the set of points

$$\{k \in \mathbb{N} \mid |a_k - a| \geq \epsilon\}$$

is finite for any $\epsilon > 0$. Sometimes accumulation points also called *limit point* or *subsequential limit* of the sequence.

Theorem 3.4.2 *A real number $a \in \mathbb{R}$ is an accumulation point of (a_n) if and only of $\lim a_{n_k} = a$ for some subsequence of (a_n).*

Proof: Will be done in the lecture. □

Assume that (a_n) is a bounded sequence. Then we can define the tail of the sequence as the sequence (c_k) by

$$c_k = \sup\{a_k \mid k > n\}.$$

Since the set over which the supremum is taken becomes smaller as k increases the supremum decreases. Hence, the sequence (c_k) is descreasing. Since it is also bounded it converges to its infimum, by the monotone convergence theorem.

Definition 3.4.3 *Suppose that a_n is bounded. Then $\limsup a_n$ is defined as the limit of c_k, as defined above.*

If a sequence is not bounded above we write $\limsup a_n = \infty$. If the sequence is only bounded above (but not necessarily below) the tail of the sequence is still defined an decreasing, it may however be unbounded below. In that case we write $\limsup a_n = -\infty$. Please bear in mind that ∞ is not a number. Most of the applications concern bounded sequences, so this is not a very important point.

There is the same concept as $\limsup$ when considering the infimum of the tail and showing that this is an increasing sequence that then converges to its supremum. Easier is simply

Definition 3.4.4 *The limes inferior is defined as*

$$\liminf a_n = -(\limsup(-a_n)).$$

Theorem 3.4.5 *Let (a_n) be a bounded sequence, then $\limsup a_n$ is the largest accumulation point of (a_n), and $\liminf a_n$ is the smallest accumulation point.*

The very existence of $\limsup$ together with the above theorems gives a simple proof of the Bolzano-Weierstrass theorem, that any bounded sequence has a convergent subsequence.

Chapter 4

Series

4.1 Definition and convergence

Informally, a series is an infinite sum

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots ,$$

for example

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots .$$

Before proceeding any further, we should ask ourselves what, exactly, the expression on the right hand side of these equations really means. In fact, it turns out that a series is really nothing more than a sequence – but *not* the sequence (a_n), rather the sequence of *partial sums*.

Definition 4.1.1 (Series and their convergence) *Given a series $\sum_{n=1}^{\infty} a_n$ we define a_n to be the n^{th} term of the series and*

$$s_k = \sum_{n=1}^{k} a_n$$

to be its k^{th} partial sum. The series converges if the sequence (s_k) converges. In this case we use the same symbol, $\sum_{n=1}^{\infty} a_n$, to denote its limit. If (s_k) does not converge we say the series diverges.

Remark Do not confuse the k^{th} *term* of a series, a_k, with its k^{th} *partial sum*, s_k. In particular, convergence of the series means convergence of (s_k), *not* convergence of (a_k).

Example 4.1.2 (Harmonic series) *Claim: the series* $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n}$ *diverges.*

Proof: *We will show that the sequence of partial sums (s_k) has an unbounded, and hence divergent, subsequence. It then follows from Theorem 3.1.3 that (s_k) itself diverges. Consider s_{2^p}:*

$$\begin{aligned}
s_{2^p} &= 1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)+\cdots+\left(\frac{1}{2^{p-1}+1}+\cdots+\frac{1}{2^p}\right)\\
&> 1+\frac{1}{2}+\left(\frac{1}{4}+\frac{1}{4}\right)+\left(\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}\right)+\cdots+\left(\frac{1}{2^p}+\cdots+\frac{1}{2^p}\right)\\
&= 1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\cdots\frac{1}{2}\\
&= 1+\frac{p}{2}.
\end{aligned}$$

This is unbounded, so (s_k) is divergent. □

Remark This is the most famous of all divergent series, called the *harmonic series*. Note that the *terms* of the series $a_n = 1/n$ converge to 0. Nonetheless, the series itself does not converge.

Example 4.1.3 *Claim: the series* $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n(n+1)}$ *converges to* 1.

Proof: *Note that*

$$a_n = \frac{1}{n(n+1)} = \frac{1}{n}-\frac{1}{n+1}$$

so

$$\begin{aligned}
s_k &= \sum_{n=1}^{k}\left(\frac{1}{n}-\frac{1}{n+1}\right)\\
&= \sum_{n=1}^{k}\frac{1}{n}-\sum_{n=1}^{k}\frac{1}{n+1}\\
&= \sum_{n=1}^{k}\frac{1}{n}-\sum_{m=2}^{k+1}\frac{1}{m} \qquad \text{(where } m=n+1\text{)}\\
&= 1-\frac{1}{k+1}.
\end{aligned}$$

Hence $s_k \to 1$. □

Example 4.1.4 *Claim: the series* $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^2}$ *converges.*

Proof: *The partial sums satisfy the inquality*

$$s_k = \sum_{n=1}^{k} \frac{1}{n^2} \le 1 + \sum_{n=2}^{k} \frac{1}{n(n-1)} = 1 + \sum_{n=1}^{k-1} \frac{1}{n(n+1)} \le 2.$$

Moreover, s_k is strictly increasing as

$$s_{k+1} - s_k = \frac{1}{(k+1)^2} > 0.$$

Therefore the sequence (s_k) is increasing and bounded by one. It therefore converges. □

Another important example where we can compute the limit exactly is the *geometric series*:

Example 4.1.5 (Geometric series) *Let $q \in (-1, 1)$ and consider the series*

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + \cdots .$$

This is the geometric series with common ratio q. We claim that it converges to $\dfrac{1}{1-q}$.

Proof:

$$\begin{aligned} s_k &= 1 + q + q^2 + \cdots + q^k \\ \Rightarrow \qquad qs_k &= q + q^2 + q^3 + \cdots + q^{k+1} \\ \Rightarrow \qquad (1-q)s_k &= 1 - q^{k+1} \\ \Rightarrow \qquad s_k &= \frac{1-q^{k+1}}{1-q}. \end{aligned}$$

Now $q^{k+1} \to 0$, so $s_k \to 1/(1-q)$. □

Usually it's not possible to find a simple formula for s_k, so we need to develop ways of determining whether a series converges by dealing directly with its *terms*, a_n, rather than its partial sums, s_k. In the next section we will construct several convergence tests for this purpose. It's important to realize that these tests will only tell us *whether* a series converges, not *what limit* it converges to (if it does converge).

4.2 Convergence tests for series

Recall that a series $\sum_{n=1}^{\infty} a_n$ converges if and only if its sequence of partial sums, $s_k = \sum_{n=1}^{k} a_n$ converges. Our first convergence test gives a very simple **necessary** but **not sufficient** condition for convergence.

Theorem 4.2.1 (Divergence Test) *If $\sum_{n=1}^{\infty} a_n$ converges, then (a_n) converges to 0.*

Proof: By assumption, $s_k = \sum_{n=1}^{k} a_n$ converges to some limit L. Hence, $s_{k+1} \to L$ also by Theorem 3.1.3 (it's a subsequence of (s_k)). Hence $a_{k+1} = s_{k+1} - s_k \to L - L = 0$, so the sequence $a_n \to 0$. $\square$

Warning: this theorem says that **if** $\sum a_n$ converges, then $a_n \to 0$. It does **NOT** say that if $a_n \to 0$ then $\sum a_n$ converges. We have already seen a counterexample, the harmonic series: $a_n = 1/n \to 0$ but $\sum_{n=1}^{\infty} \frac{1}{n}$ does **not** converge!

Example 4.2.2 $\sum_{n=1}^{\infty} \frac{n^2-1}{n^2+1}$ *does not converge, by the Divergence Test (since its terms a_n converge to 1, not 0).*

If the terms a_n of a series are non-negative, i.e. $a_n \geq 0$, then the sequence of partial sums is increasing, since $s_{k+1} = s_k + a_{k+1} > s_k$. Hence, such a series is convergent if (and only if) its sequence of partial sums is **bounded**, by the Monotone Convergence Theorem. We will use the notation

$$\sum_{n=1}^{\infty} a_n < \infty$$

in that case to indicate that the sequence of partial sums is bounded above. If it is bounded above by a constant C we will write

$$\sum_{n=1}^{\infty} a_n \leq C.$$

Note that it is important that this notation can be used only of $a_n \geq 0$.

In case we have $0 \leq a_n \leq b_n$ we then have the conclusion that

$$\sum_{n=1}^{\infty} b_n \leq C$$

implies

$$\sum_{n=1}^{\infty} a_n \leq C.$$

This observation allows us to formulate several useful tests which apply in the special case $a_n > 0$ for all n. Summarising, if $a_n \geq 0$ and we have

$$\sum_{n=1}^{\infty} a_n < \infty$$

this means that $\sum_{n=1}^{\infty} a_n$ converges.

Theorem 4.2.3 (The Comparison Test) *Let $a_n > 0$ and $b_n > 0$ for all n.*

(i) If the sequence $\dfrac{a_n}{b_n}$ is bounded above and $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.

(ii) If the sequence $\dfrac{b_n}{a_n}$ is bounded above and $\sum_{n=1}^{\infty} b_n$ diverges, then $\sum_{n=1}^{\infty} a_n$ diverges.

Proof:
(i) By assumption $0 \leq \dfrac{a_n}{b_n} \leq C$ for some $C > 0$. This means $0 \leq a_n \leq Cb_n$. If $\sum_{n=1}^{\infty} b_n$ converges we have in particular $\sum_{n=1}^{\infty} b_n < \infty$ and therefore

$$\sum_{n=1}^{\infty} a_n \leq C \sum_{n=1}^{\infty} b_n < \infty.$$

(ii) By (i) convergence of $\sum_{n=1}^{\infty} a_n$ implies convergence of $\sum_{n=1}^{\infty} b_n$. The contrapositive is the claimed statement.

□

We can use this theorem to show that a series of interest converges/diverges by comparing it to a series which we already know converges/diverges.

Example 4.2.4 (Revisited) *Claim: $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.*

Proof: *Let $a_n = \dfrac{1}{n^2}$ and $b_n = \dfrac{1}{n(n+1)}$. Then, as shown in Example 4.1.3, $\sum_{n=1}^{\infty} b_n$ converges. Now*

$$\frac{a_n}{b_n} = \frac{1}{n^2} \times n(n+1) = 1 + \frac{1}{n} \to 1.$$

Since a_n/b_n converges, it is bounded, so by part (i) of the Comparison Test, $\sum_{n=1}^{\infty} a_n$ also converges. □

Remark Note that we didn't show that a_n/b_n was bounded directly. Instead, we used the Algebra of Limits to show that it converges. Every convergent sequence is bounded, so it follows that a_n/b_n is bounded, as the Comparison Test requires.

Example 4.2.5 *Let $p \in \mathbb{N}$ and $p > 2$. Then $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges.*

Proof: Let $a_n = 1/n^p$ and $b_n = 1/n^2$. Then $\sum_{n=1}^{\infty} b_n$ converges, by Example 4.2.4. Now

$$\frac{a_n}{b_n} = \frac{1}{n^p} \times n^2 = \frac{1}{n^{p-2}} \to 0$$

since $p > 2$. Since a_n/b_n converges, it is bounded, so by part (i) of the Comparison Test, $\sum_{n=1}^{\infty} a_n$ also converges. □

Example 4.2.6 $\displaystyle\sum_{n=1}^{\infty} \frac{n}{2n^2 + \sin n}$ *diverges.*

Proof: Let $a_n = n/(2n^2 + \sin n)$ and $b_n = 1/n$. Then $\sum_{n=1}^{\infty} b_n$ diverges, by Example 4.1.2. Now

$$\frac{b_n}{a_n} = \frac{2n^2 + \sin n}{n^2} = 2 + \frac{\sin n}{n^2}.$$

But $\quad -\dfrac{1}{n^2} \le \dfrac{\sin n}{n^2} \le \dfrac{1}{n^2} \quad \Rightarrow \quad \dfrac{\sin n}{n^2} \to 0 \quad$ (by the Squeeze Rule)

so $b_n/a_n \to 2$. Since b_n/a_n converges, it is bounded, so by part (ii) of the Comparison Test, $\sum_{n=1}^{\infty} a_n$ also diverges. □

To use the Comparison Test, you have to have some intuition about whether the series converges. If you suspect it does, compare it with a series you know converges. If you suspect it doesn't, compare it with a series you know doesn't. In Example 4.2.6 I recognized that, for large n, the $\sin n$ part is essentially irrelevant, so the terms look roughly like $n/(2n^2) = 1/(2n)$. So I compared the series to the divergent series $\sum(1/n)$.

Our next convergence test (everyone's favourite) involves only the terms of the series under consideration, and is very easy to apply. Its only disadvantage is that it very often gives an inconclusive answer.

Theorem 4.2.7 (The Ratio Test) *Let $a_n > 0$ for all n and assume that the sequence of ratios of consecutive terms converges, that is, $\dfrac{a_{n+1}}{a_n} \to L$ for some $L \in \mathbb{R}$.*

(i) If $L < 1$ then $\sum_{n=1}^{\infty} a_n$ converges.

(ii) If $L > 1$ then $\sum_{n=1}^{\infty} a_n$ diverges.

Proof: Before we proof this let us remark that if (a_n) is a sequence with $a_n \to L > 1$. Then there exists $N \in \mathbb{N}$ such that for all $n > N$ we have

$$a_n > 1 + \delta,$$

where $\delta = \frac{L-1}{2} > 0$. To see this simply choose $\epsilon = \delta$ in the definition of convergence and note that $|a_n - L| < \delta$ implies that $a_n > 1 + \delta$. Similarly, if $a_n \to L < 1$ then there exists $N \in \mathbb{N}$ such that for all $n > N$ we have

$$a_n < 1 - \delta,$$

where $\delta = \frac{1-L}{2} > 0$. Now let's prove the two statements.
(i) Again, we only need to show that $\sum_{n=1}^{\infty} a_n < \infty$. By the above observation there exists $N \in \mathbb{N}, \delta > 0$ such that for $n > N$ we have $0 \leq \frac{a_{n+1}}{a_n} < 1 - \delta$. Therefore, $0 \leq a_{n+1} < (1-\delta)a_n$. By induction we have

$$a_{N+1+n} \leq (1-\delta)^n a_{N+1}$$

for all $n \in \mathbb{N}$. Therefore,

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{N+1} a_n + \sum_{n=N+2}^{\infty} a_n \leq \sum_{n=1}^{N+1} a_n + a_{N+1} \sum_{n \in \mathbb{N}} (1-\delta)^n < \infty.$$

We have used convergence of the geometric series, namely

$$\sum_{n \in \mathbb{N}} (1-\delta)^n = \frac{(1-\delta)}{\delta} < \infty.$$

(ii) By the above there exists $N \in \mathbb{N}, \delta > 0$ such that $\frac{a_{n+1}}{a_n} > 1 + \delta$. By induction we obtain $a_{N+1+n} \geq (1+\delta)^n a_{N+1}$. Since $(1+\delta)^n \geq 1 + n\delta$ is unbounded above this implies that the sequence a_{N+1+n} is unbounded. Since this is an subsequence of (a_n) this means that (a_n) is not bounded. Hence, a_n does not converge (to zero). By the Divergence Test Theorem 4.2.1 the series must diverge. □

The ratio test is very handy, but it gives no useful information in the case $L = 1$. For example, if you try to use it for $\sum_{n=1}^{\infty} n^{-p}$ it will be inconclusive.

Example 4.2.8 *Claim: the series* $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!}$ *converges.*

Proof: *Let* $a_n = 1/n!$. *Then* $a_{n+1}/a_n = n!/(n+1)! = 1/(n+1) \to 0 < 1$, *so the series converges by the Ratio Test.* □

The limit of this series is a very famous number, called Euler's number e. *Note that the series starts with* $n = 0$, *and recall that* $0! = 1$. *Another way to define* e *is as the limit of the* sequence $(1 + 1/n)^n$. *However, proving that this sequence* *converges is much trickier than proving that the series $\sum(1/n!)$ converges.* $\square$

Example 4.2.9 (Exponentials beat polynomials) *Let $p \in \mathbb{N}$ and $\alpha \in (0,1)$ be constant. Then $a_n = n^p \alpha^n \to 0$.*

Proof: Consider the series $\sum_{n=1}^{\infty} a_n$. This has

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)^p \alpha^{n+1}}{n^p \alpha^n} = \left(1 + \frac{1}{n}\right)^p \alpha \to \alpha < 1.$$

Hence $\sum_{n=1}^{\infty} a_n$ converges, by the Ratio Test, so $a_n \to 0$ by the Divergence Test. $\square$

4.3 Alternating series

Apart from the Divergence Test, all the convergence tests we studied in section 4.2 assumed that the terms of the series are strictly positive. There is another special case which arises quite commonly and which can be handled using a convenient convergence test.

Definition 4.3.1 (Alternating Series) *A real series $\sum_{n=1}^{\infty} b_n$ is said to be alternating if $b_n \neq 0$ and $b_{n+1}/b_n < 0$ for all n.*

The point is that consecutive terms of the series alternate in sign.

Example 4.3.2 *$\sum_{n=1}^{\infty}(-1)^n$ is an alternating series. Its partial sums are*

$$s_k = \begin{cases} -1 & k \text{ odd} \\ 0 & k \text{ even} \end{cases}$$

Hence, this series diverges. □

Clearly every alternating series can be written in the form

$$\sum_{n=1}^{\infty}(-1)^{n+1}a_n \qquad \text{or} \qquad -\sum_{n=1}^{\infty}(-1)^{n+1}a_n$$

where $a_n > 0$ for all n. We will state and prove our convergence test for series of the left hand type, but it applies equally well to those of the right hand type (by the Algebra of Limits).

Theorem 4.3.3 (The Alternating Series Test) *Let (a_n) be a decreasing sequence of positive numbers which converges to 0. Then*

$$\sum_{n=1}^{\infty}(-1)^{n+1}a_n$$

converges.

To prove this, we will need a simple Lemma concerning subsequences. Recall that if a sequence (a_n) converges to L, then every subsequence of (a_n) converges to L. Recall also, that if one, or two, or twenty-thousand subsequences of (a_n) converge to L it does not, in general, follow that (a_n) converges to L. The next Lemma shows that if the "odd" and "even" subsequences of (a_n) both converge to L, then (a_n) converges to L.

Lemma 4.3.4 *Let (a_n) be a sequence such that $a_{2k} \to L$ and $a_{2k+1} \to L$. Then $a_n \to L$.*

Proof: Let $\varepsilon > 0$ be given. Since $a_{2k} \to L$, there exists $K_1 \in \mathbb{N}$ such that for all $k \geq K_1$, $|a_{2k} - L| < \varepsilon$. Since $a_{2k+1} \to L$, there exists $K_2 \in \mathbb{N}$ such that for all $k \geq K_2$, $|a_{2k+1} - L| < \varepsilon$. Let $N = \max\{2K_1, 2K_2 + 1\}$. Then for all $n \geq N$, n is either even, in which case $|a_n - L| < \varepsilon$ because $n \geq 2K_1$, or n is odd, in which case $|a_n - L| < \varepsilon$ because $n \geq 2K_2 + 1$. $\square$

Proof of Theorem 4.3.3: As usual let s_k denote the k^{th} partial sum of the series. Consider the subsequence of even numbered partial sums (s_{2m}). Now

$$
\begin{aligned}
s_{2m} &= (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2m-1} - a_{2m}) \\
\Rightarrow \qquad s_{2m+2} - s_{2m} &= a_{2m+1} - a_{2m+2} \geq 0
\end{aligned}
$$

so the sequence (s_{2m}) is increasing. Further,

$$
s_{2m} = a_1 - (a_2 - a_3) - \cdots - (a_{2m-2} - a_{2m-1}) - a_{2m} < a_1
$$

so (s_{2m}) is bounded above, by a_1. Hence, by the Monotone Convergence Theorem, (s_{2m}) converges to some limit L.

Consider now the subsequence of odd numbered partial sums (s_{2m+1}). Now

$$
s_{2m+1} = s_{2m} + a_{2m+1}.
$$

We have shown that $s_{2m} \to L$, and a_{2m+1} is a subsequence of (a_n) which, by assumption, converges to 0, so $a_{2m+1} \to 0$ also (Theorem 3.1.3). Hence $s_{2m+1} \to L + 0 = L$ by the Algebra of Limits.

Since $s_{2m} \to L$ and $s_{2m+1} \to L$, it follows from Lemma 4.3.4 that $s_k \to L$, as was to be proved.

$\square$

Example 4.3.5 (The alternating harmonic series)

$$
\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}
$$

converges, by the Alternating Series Test, since $a_n = 1/n$ is positive, decreasing, and converges to 0. $\square$

The condition that (a_n) should be *decreasing* is a crucial part of the Alternating Series Test. Without it, the conclusion can be false.

Example 4.3.6 *Consider the sequence*

$$
a_n = \begin{cases} 1/n^2 & n \text{ odd} \\ 1/n & n \text{ even} \end{cases}
$$

This is positive and converges to 0*, but is* not *decreasing. I claim that the alternating series* $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ *diverges.*

Proof: *Consider the even subsequence of partial sums*

$$\begin{aligned} s_{2m} &= 1 - \frac{1}{2} + \frac{1}{3^2} - \frac{1}{4} + \cdots + \frac{1}{(2m-1)^2} - \frac{1}{2m} \\ &= \sum_{n=1}^{m} \frac{1}{(2n-1)^2} - \frac{1}{2}\sum_{n=1}^{m}\frac{1}{n} \\ &=: t_m - \frac{1}{2}u_m. \end{aligned}$$

The sequence (t_m) *converges, by the Comparison Test (compare with the convergent series* $\sum(1/n^2)$*), so if* (s_{2m}) *converges, then* (u_m) *converges by the Algebra of Limits. But* u_m *is the* m^{th} *partial sum of the harmonic series, which is known to be divergent. Hence* (s_{2m}) *diverges, so* (s_k) *diverges.* □

Equally, one can write down alternating series for which (a_n) is not decreasing, but which do, nevertheless converge (e.g. modify Example 4.3.6 by taking $a_n = 1/n^3$ for n even). The point is that the Alternating Series Test is a **sufficient** but **not necessary** test for convergence.

4.4 Absolute convergence

One rather crude way to show that some alternating series converge is to throw away the extra information about the sign of the terms altogether.

Definition 4.4.1 (Absolute Convergence) *A real series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely if* $\sum_{n=1}^{\infty}|a_n|$ *converges in the usual sense (that is, the sequence of partial sums* $s_k = \sum_{n=1}^{k}|a_n|$ *converges).*

Remark This definition applies to all series, whether alternating or not.

Theorem 4.4.2 (Absolute convergence is stronger than convergence) *If a real series* $\sum_{n=1}^{\infty} a_n$ *converges absolutely, then it converges.*

Proof: Let $s_k = \sum_{n=1}^{k} a_n$ be the sequence of partial sums for the series $\sum_{n=1}^{\infty} a_n$. Then (s_k) converges if and only if it is a Cauchy sequence. This means we need to show that for any $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for $m > n > N$ we have

$$|s_m - s_n| = |\sum_{k=n+1}^{m} a_k| < \epsilon.$$

We know that the series $\sum_{n=1}^{k} |a_n|$ converges, so its sequence (t_k) of partial sums is a Cauchy sequence. Hence, given $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that for $m > n > N$ we have

$$|t_m - t_n| = \sum_{k=n+1}^{m} |a_k| < \epsilon.$$

Now using the triangle inequality shows that for $m > n > N$ we have

$$|s_m - s_n| = |\sum_{k=n+1}^{m} a_k| \leq \sum_{k=n+1}^{m} |a_k| < \epsilon.$$

□

Remark The converse of this Theorem is false. For example, the alternating harmonic series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

is convergent, but is not absolutely convergent, since in this case

$$\sum_{n=1}^{\infty} \left| \frac{(-1)^{n+1}}{n} \right| = \sum_{n=1}^{\infty} \frac{1}{n},$$

which we have proved is divergent (see Examples 4.1.2 and 4.3.5).

Example 4.4.3 *Claim:* $\displaystyle\sum_{n=1}^{\infty} \frac{\sin n}{n^3}$ *converges.*

Note that this series is not alternating, so we can't use the Alternating Series Test. Nor are the terms all positive, so we can't use the Comparison or Ratio Tests. What we can do is show that it converges absolutely and use Theorem 4.4.2

Proof: *Let* $a_n = n^{-3} \sin n$ *and* $b_n = 1/n^2$*. Then* $|a_n| > 0$ *and* $b_n > 0$ *and*

$$\frac{|a_n|}{b_n} = \frac{|\sin n|}{n} \leq \frac{1}{n} \leq 1$$

Since we know that $\sum_{n=1}^{\infty} b_n$ *converges (Example 4.2.4) it follows that* $\sum_{n=1}^{\infty} |a_n|$ *converges also, by the Comparison Test, part (i). Hence* $\sum_{n=1}^{\infty} a_n$ *converges absolutely, and so converges by Theorem 4.4.2.* □

Example 4.4.4 *Consider the sequence*

$$a_n = \begin{cases} \frac{1}{n^3} & n \text{ odd} \\ -\frac{1}{n^2} & n \text{ even} \end{cases}$$

I claim that the series $\sum_{n=1}^{\infty} a_n$ *converges.*

Proof: *Let* $b_n = 1/n^2$. *Then*

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

$$\leq 1.$$

Since $|a_n|/b_n$ *is bounded above, and* $\sum b_n$ *is known to converge (Example 4.2.4, we conclude that* $\sum |a_n|$ *converges also, by part (i) of the Comparison Test, that is,* $\sum a_n$ *converges* absolutely. *Hence* $\sum a_n$ *converges by Theorem 4.4.2.* $\square$

Remark This is an alternating series, but we can't use the Alternating Series Test, because $|a_n|$ is not a decreasing sequence.

4.5 The importance of absolute convergence

Why are we so concerned about these different notions of convergence? The following example illustrates that quite a bit of care is necessary when manipulating series.

Example 4.5.1 *As you will see the following manipulation of a series gives a contradiction. So there is a mistake somewhere. Let* S *be the value of the convergent series* $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$. *Let us write down the first few terms*

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \frac{1}{11} - \frac{1}{12} + \frac{1}{13} - \ldots.$$

Now multiply this by two to obtain

$$2S = 2 - 1 + \frac{2}{3} - \frac{1}{2} + \frac{2}{5} - \frac{1}{3} + \frac{2}{7} - \frac{1}{4} + \frac{2}{9} - \frac{1}{5} + \frac{2}{11} - \frac{1}{6} + \frac{2}{13} - \ldots.$$

Now re-arrange that and collect the odd common denominators to obtain

$$2S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \ldots.$$

So we end up with $2S = S$. *Since* $S \neq 0$ *(collecting positive and subsequent negative terms one gets the* $S > \frac{1}{2}$*), one concludes that* $2 = 1$. *This is obviously nonsense. What has gone wrong?*

What has gone wrong here is that the re-arrangement of a series is not permitted in general if we only have convergence. Let us be a bit more precise. If $f : \mathbb{N} \to \mathbb{N}$ is an bijective map, then the series

$$\sum_{n=1}^{\infty} a_{f(n)}$$

can be understood as a re-arrangement of the terms in a different way. The above example shows that such a re-arrangement does not necessarily result in the same value. If a series is however absolutely convergent we can be sure that the series can be re-arranged.

In fact there are some more properties that one expects to be true formally, but are true only for absolutely convergent series.

Theorem 4.5.2 *Suppose that $I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots$ is an increasing sequence of* **finite** *subsets $I_k \subseteq \mathbb{N}$ such that*

$$\cup_{k\in\mathbb{N}} I_k = \mathbb{N}.$$

Suppose that (a_n) is a sequence such that $\sum_{k=1}^{\infty} |a_k| < \infty$, i.e. the series $\sum_{k=1}^{\infty} a_k$ converges absolutely. Then

$$\lim_{n\to\infty} \sum_{k\in I_n} a_k = \sum_{k=1}^{\infty} a_k.$$

Remark 4.5.3 *In case $I_n = \{1, 2, \ldots, n\}$ the left hand side and the right hand side of this equation are the same by definition.*

Proof: Denote $L = \sum_{k=1}^{\infty} a_k$. Given $\epsilon > 0$ we can choose $M \in \mathbb{N}$ such that

$$(n \geq M) \implies (\sum_{k=1}^{n} a_k - L) < \epsilon/2,$$

$$(m > n > M) \implies \sum_{k=n}^{m} |a_k| < \epsilon/2.$$

Here we used the fact that $\sum_{k=1}^{n} |a_k|$ is a Cauchy sequence. Now choose $N \in \mathbb{N}$ such that $\{1, \ldots, M\} \subseteq I_N$. Then, if $n > N$ we have

$$|\sum_{k\in I_n} a_k - L| = |\sum_{k=1}^{M} a_k - L + \sum_{k\in I_n, k>M} a_k|$$

$$\leq |\sum_{k=1}^{M} a_k - L| + \sum_{k\in I_n, k>M} |a_k| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

$\square$

Theorem 4.5.4 (Absolute convergenet series may be re-arranged) *Suppose that* $L = \sum_{n=1}^{\infty} a_n$ *is an absolutely convergent series and let* $f : \mathbb{N} \to \mathbb{N}$ *be a bijective map. Then the series*

$$\sum_{n=1}^{\infty} a_{f(n)}$$

is absolutely convergent and has the same limit L*.*

Proof: This follows from the above theorem by simply taking $I_n = f(\{1, \ldots, n\})$. Absolute convergence follows by replacing a_n by $|a_n|$ in the statement. □

Now suppose that I is an infinite countable index set and $(a_\alpha)_{\alpha \in I}$ is a family of real numbers $a_\alpha \in \mathbb{R}$ indexed by I. Since I is countable there exists a bijective map $f : \mathbb{N} \to I$. Suppose that

$$\sum_{k=1}^{\infty} |a_{f(k)}| < \infty,$$

and assume $g : \mathbb{N} \to I$ is another bijective map. Then, since $f^{-1} \circ g : \mathbb{N} \to \mathbb{N}$ is bijective, we have

$$\sum_{k=1}^{\infty} |a_{f(k)}| = \sum_{k=1}^{\infty} |a_{f(f^{-1}(g(k)))}| = \sum_{k=1}^{\infty} |a_{g(k)}|,$$

and

$$\sum_{k=1}^{\infty} a_{f(k)} = \sum_{k=1}^{\infty} a_{f(f^{-1}(g(k)))} = \sum_{k=1}^{\infty} a_{g(k)}.$$

This means that the following definition is independent of the chosen f.

Definition 4.5.5 *We say* $\sum_{\alpha \in I} a_\alpha$ *is absolutely convergent, and write*

$$\sum_{\alpha \in I}^{\infty} |a_\alpha| < \infty,$$

if

$$\sum_{k=1}^{\infty} |a_{f(k)}| < \infty$$

for some (and hence any) bijective map $f : \mathbb{N} \to I$*. In this case we define*

$$\sum_{\alpha \in I}^{\infty} a_\alpha = \sum_{k=1}^{\infty} a_{f(k)}.$$

The same proof as in Theorem 4.5.2 now shows that if there is a family of (not necessarily finite) subsets $(I_n)_{n\in\mathbb{N}}$ such that $I_1 \subseteq I_2 \subseteq I_3 \subseteq \ldots$ with $\cup_{n\in\mathbb{N}} = I$ then we have

$$\sum_{\alpha\in I} a_\alpha = \lim_{n\to\infty} \sum_{\alpha\in I_n} a_\alpha.$$

The fact that the finiteness assumption is not necessary in the proof is left as an exercise to the reader. On may just go literally through the proof of Theorem 4.5.2 once absolute convergence is established.

An important example is when the index set I equals the Cartesian product $\mathbb{N}\times\mathbb{N}$, which is a countable set. Hence, our family of numbers (a_α) is indexed by pairs $\alpha = (n, m)$ of natural numbers. We write $a_{(n,m)} = a_{nm}$. One can imagine this as an infinite matrix of real numbers. The sequence of sets $I_n = \{(k, m) \mid m, k \in \mathbb{N}, k \leq n$ satisfies the above assumptions (is increasing and exhausts $\mathbb{N} \times \mathbb{N}$). Similarly, the sequence of sets $K_n = \{(m, k) \mid m, k \in \mathbb{N}, k \leq n$ is increasing an exhausts $\mathbb{N} \times \mathbb{N}$. Hence, we have the following theorem.

Theorem 4.5.6 (Fubini's theorem for sums) *Suppose that $(a_{nm})_{n,m\in\mathbb{N}}$ is a family of real numbers $a_{nm} \in \mathbb{R}$ as assume that*

$$\sum_{n,m\in\mathbb{N}} |a_{nm}| < \infty.$$

Then for each $n \in N$ the sums $\sum_{m=1}^{\infty} a_{mn}$ and $\sum_{m=1}^{\infty} a_{nm}$ converge absolutely, and

$$\sum_{n,m\in\mathbb{N}} a_{nm} = \sum_{n=1}^{\infty}\left(\sum_{m=1}^{\infty} a_{mn}\right) = \sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty} a_{mn}\right).$$

Chapter 5

Functions and Continuity

5.1 Sequential continuity

Given a function $f : \mathbb{R} \to \mathbb{R}$, what does it mean to say that the function is *continuous*? It's common, when first introduced to the notion of continuity, to be told that a function is continuous "if you can draw its graph without taking your pencil off the paper". Clearly this is *not* an acceptable mathematical definition (what if you have no pencil, or paper, or hands? Does the continuity or otherwise of f depend on your skill as a draughtsman?). So our first job will be to define continuity precisely. In fact it's quite easy to give a precise (and rather elegant) definition of continuity using nothing more than the notion of convergence of sequences. The definition works equally well when the domain of f is a subset of $\mathbb{R}$, not necessarily the whole of $\mathbb{R}$, so that is how we will formulate it.

Definition 5.1.1 (Continuity of a function) *Let $D \subseteq \mathbb{R}$. A function $f : D \to \mathbb{R}$ is continuous at a point $a \in D$ if, for all sequences (x_n) in D such that $x_n \to a$, $f(x_n) \to f(a)$. If f is not continuous at $a \in D$, we say that f is discontinuous at a. The function is continuous if it is continuous at a for all $a \in D$.*

Example 5.1.2 *(i) $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = c$, constant, is continuous.*

Proof: Let $a \in \mathbb{R}$ and $x_n \to a$. Then $f(x_n) = c \to c = f(a)$. $\square$

(ii) $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x$ is continuous.

Proof: Let $a \in \mathbb{R}$ and $x_n \to a$. Then $f(x_n) = x_n \to a = f(a)$. $\square$

(iii) $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ such that $f(x) = 1/x$ is continuous.

If you've been taught in the past that a continuous function is "one whose graph you can draw without taking your pen off the paper" then this example will be surprising at first sight. Nonetheless, this function really is continuous *according to our (precise) definition:*

> *Proof:* *Let $a \in \mathbb{R}\backslash\{0\}$ (i.e., let a be a nonzero real number) and x_n be a sequence in $\mathbb{R}\backslash\{0\}$ such that $x_n \to a$. Then $x_n \neq 0$, so*
>
> $$f(x_n) = \frac{1}{x_n} \to \frac{1}{a} = f(a)$$
>
> *by Limit Algebra.* □

(iv) *$f : \mathbb{Z} \to \mathbb{R}$ such that $f(x) = (-1)^x$ is continuous.*

This one is also counterintuitive. In fact, it gets worse:

(v) Every *function $f : \mathbb{Z} \to \mathbb{R}$ is continuous!*

> *Proof:* *Let $a \in \mathbb{Z}$ and (x_n) be a sequence in $\mathbb{Z}$ converging to a. Then, by the definition of convergence, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - a| < 1/2$. But x_n and a are both integers, so if they lie distance less than $1/2$ away from one another, they must be equal. Hence, for all $n \geq N$, $x_n = a$. Now let $\varepsilon > 0$ be given. Then for all $n \geq N$, $|f(x_n) - f(a)| = |f(a) - f(a)| = 0 < \varepsilon$. Hence, $f(x_n) \to f(a)$.* □

Don't worry, however: there are *discontinuous functions:*

(vi) *Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = 0$ if $x < 0$ and $f(x) = 1$ for $x \geq 0$. Then f is discontinuous at 0.*

> *Proof:* *The sequence $x_n = -1/n$ converges to 0 but*
>
> $$f(x_n) = f(-1/n) = 0 \to 0 \neq f(0).$$
>
> □

This is an example of a function which is discontinuous *at a single isolated point. It's also possible to construct functions which are* continuous *at a single isolated point (and discontinuous everywhere else):*

(vii) *Let $f : \mathbb{R} \to \mathbb{R}$ such that*

$$f(x) = \begin{cases} 0 & x \in \mathbb{Q} \\ x & x \notin \mathbb{Q}. \end{cases}$$

Then f is continuous at 0 and discontinuous everywhere else.

Proof: Let $a \in \mathbb{R}$ and assume that f is continuous at a. Since $\mathbb{Q}$ is dense, and the irrational numbers are dense, for each $n \in \mathbb{N}$, there exist $r_n \in \mathbb{Q}$ and $i_n \in \mathbb{R}\backslash\mathbb{Q}$ such that $a < r_n < a + 1/n$ and $a < i_n < a + 1/n$. Clearly $r_n \to a$ and $i_n \to a$ by the Squeeze Rule. Hence, by the definition of continuity, $f(r_n) \to f(a)$ and $f(i_n) \to f(a)$. But r_n is rational, so $f(r_n) = 0 \to 0$, and i_n is irrational, so $f(i_n) = i_n \to a$. Hence $a = 0$.
Aside: so what have we proved? Have we proved that f is continuous at 0? No we flipping well have **NOT!** *We have proved that* <u>*if*</u> *f is continuous at a* <u>*then*</u> *$a = 0$. We have* **not** *proved that* <u>*if*</u> *$a = 0$* <u>*then*</u> *f is continuous at a! Another way to say what we have proved is that* <u>*if*</u> *$a \neq 0$* <u>*then*</u> *f is* <u>*not*</u> *continuous at a. That is, up to this point, we have proved that f is* **discontinuous** *at every $a \neq 0$.*
Let (x_n) be any sequence in $\mathbb{R}$ such that $x_n \to 0$. Let $\varepsilon > 0$ be given. Then, since $x_n \to 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - 0| < \varepsilon$. But then, for all $n \geq N$,

$$|f(x_n) - 0| = |f(x_n)| \leq |x_n| < \varepsilon$$

since $f(x_n) = x_n$ if x_n is irrational, and is 0 otherwise. Hence $f(x_n) \to 0 = f(0)$. Hence, f is continuous at 0. $\square$

The definition of continuity that we have used here is sometimes called *sequential continuity*. We have used it for pedagogical reasons as it is easier to understand than the more fundamental definition that uses more logical quantifies. The following characterisation is usually used as a definition of continuity and you will see it in later years.

Theorem 5.1.3 (Continuity: Alternative definition) *A function $f : D \to \mathbb{R}$ is continuous at $a \in D$ if and only if for any $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(a)| < \epsilon$ for all $x \in D$ with $|x - a| < \delta$.*

Proof: Suppose that f has the above $\epsilon{-}\delta$-property. We show that f is continuous. Let x_n be a sequence in D with $x_n \to a$. We need to show that $f(x_n) \to f(a)$. Thus, given $\epsilon > 0$ need to find $N \in \mathbb{N}$ such that $n > N$ implies $|f(x_n) - f(a)| < \epsilon$. By the $\epsilon{-}\delta$-property we can find a $\delta > 0$ such that $|f(x_n) - f(a)| < \epsilon$ if $|x_n - a| < \delta$. Since $x_n \to a$ we can find $N \in \mathbb{N}$ such that $|x_n - a| < \delta$ for all $n > N$, and hence also $|f(x_n) - f(a)| < \epsilon$ for all $n > N$.

Conversely, suppose that does not have the $\epsilon - \delta$-property. Then there exists $\epsilon > 0$ such that for any $\delta > 0$ there is an $x \in D$ with $|x - a| < \delta$ but $|f(x)| - f(a)| \geq \epsilon$. Choose $\delta = 1/n$ and let $x_n \in D$ be the corresponding element x. Since $|x_n - a| < \frac{1}{n}$ the sequence (x_n) converges to a, but since $|f(x_n) - f(a)| > \epsilon$ the sequence $f(x_n)$ does not converge to $f(a)$. Hence, the function is not continuous. $\square$

5.2 Basic properties of continuous functions

If we are given a complicated function such as

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = \frac{x^3 - 7x}{x^2 + 1}$$

we can show directly from Definition 5.1.1 that f is continuous, by considering an arbitrary sequence (x_n) converging to a. To do this we will have to appeal at some point to the algebra of limits, to argue that

$$f(x_n) = \frac{x_n^3 - 7x_n}{x_n^2 + 1} \qquad \text{converges to} \qquad \frac{a^3 - 7a}{a^2 + 1} = f(a).$$

The alternative to this is to prove, once and for all, that continuity is "preserved" under sums, products and (under suitable restrictions) quotients.

Theorem 5.2.1 (Algebra Property of Continuous Functions) *Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be continuous at $a \in D$. Then*

(i) $f + g$ is continuous at a.

(ii) fg is continuous at a.

If, in addition, $f(x) \neq 0$ for all $x \in D$, then

(iii) $1/f$ is continuous at a.

Proof: Given any sequence (x_n) in D such that $x_n \to a$, we know that $f(x_n) \to f(a)$ and $g(x_n) \to g(a)$ (since f, g are continuous at a). Hence

$$\begin{aligned} (f+g)(x_n) &= f(x_n) + g(x_n) \to f(a) + g(a) \\ (fg)(x_n) &= f(x_n)g(x_n) \to f(a)g(a) \\ \text{and} \qquad (1/f)(x_n) &= \frac{1}{f(x_n)} \to \frac{1}{f(a)} \end{aligned}$$

where, in the final line, we have assumed $f(x_n) \neq 0$ and $f(a) \neq 0$. Hence $f + g$, fg and $1/f$ are continuous at a. $\square$

Recall that a *polynomial* function is a function of the form

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

where $a_0, \ldots, a_m$ are (real) constants. If $a_m \neq 0$, we say that p has *degree* m. Polynomials are a very nice class of functions. In particular they are always continuous:

Proposition 5.2.2 (Continuity of polynomials) *Every polynomial function $p : \mathbb{R} \to \mathbb{R}$ is continuous.*

Proof: We prove this by induction on m, the degree of the polynomial p. So, if $m = 0$, then $p(x)$ is a degree 0 polynomial, that is, $p(x) = a_0$, constant. Hence $p(x)$ is continuous (see Example 5.1.2). Now assume that every polynomial of degree k is continuous. Let $p(x)$ be a degree $k+1$ polynomial. Then $p(x) = xq(x) + a_0$ where $q(x)$ is a polynomial of degree k and so is, by assumption, continuous. Now $f(x) = x$ is continuous (see Example 5.1.2), so $p(x) = xq(x) + a_0$ is continuous by Theorem 5.2.1(i),(ii). Hence, by induction every polynomial is continuous □

Example 5.2.3 (Rational functions) *Let $p(x)$, $q(x)$ be polynomials, and define $D = \{x \in \mathbb{R} : q(x) \neq 0\}$. Then the function*

$$f : D \to \mathbb{R}, \qquad f(x) = \frac{p(x)}{q(x)}$$

is said to be rational (it is a quotient of polynomials). By Proposition 5.2.2 and Theorem 5.2.1(iii), every rational function is continuous. We've already seen an example of this:

$$f : \mathbb{R}\backslash\{0\} \to \mathbb{R}, \qquad f(x) = \frac{1}{x}$$

is continuous. □

A nice thing about the set of polynomials is that it is *closed* under function composition. That is, if p and q are polynomials, so is $q \circ p$ (recall that $(q \circ p)(x) = q(p(x))$). For example:

$$\begin{aligned}
p(x) &= x^3 + 1, \qquad q(x) = 3x^2 - 2 \\
\Rightarrow (q \circ p)(x) &= q(x^3 + 1) \\
&= q(x^3 + 1) \\
&= 3(x^6 + 2x^3 + 1) - 2 \\
&= 3x^6 + 6x^3 + 1.
\end{aligned}$$

So given a pair of polynomials (certainly continuous), their composition is a polynomial (and hence is continuous). Actually, rational functions are also closed under composition (though we have to be careful about the domain of the composite functions to make this precise). So, given a pair of rational functions (certainly continuous), their composition is rational (and hence is continuous). Is this a special feature of polynomials and rational functions? No: it follows almost immediately from our definition of continuity that the composition of two continuous functions is continuous.

Theorem 5.2.4 (Continuity of Compositions) *Let $D, E \subseteq \mathbb{R}$, $f : D \to E$ and $g : E \to \mathbb{R}$. If f is continuous at $a \in D$ and g is continuous at $f(a) \in E$, then $g \circ f$ is continuous at a.*

Proof: Let (x_n) be any sequence in D which converges to a. Then $f(x_n) \to f(a)$ (since f is continuous at a), so $(g \circ f)(x_n) = g(f(x_n)) \to g(f(a)) = (g \circ f)(a)$ (since g is continuous at $f(a)$). Hence $g \circ f$ is continuous at a. $\square$

At the moment, we don't have many functions which we know (rigorously) are continuous (polynomials and rational functions are pretty much it). We will see shortly that the functions

$$f : [0, \infty) \to \mathbb{R}, \qquad f(x) = x^{1/p}$$

are well-defined and continuous (for any positive integer p). Once we know this, it will follow immediately from Theorem 5.2.4 that a complicated beast like

$$f : [0, \infty) \to \mathbb{R}, \qquad f(x) = \frac{\sqrt{x + \sqrt{x} + 1}}{1 + x^{1/3} + x^{5/6}}$$

is continuous. Let's ask ourselves, for a given $x \geq 0$, what does $x^{1/p}$ actually mean? It is, by definition, that (non-negative) real number whose p^{th} power is x, that is, $x^{1/p} = y$ if $y \geq 0$ and $y^p = x$. To prove that such a number y exists we will appeal to a big theorem: the Intermediate Value Theorem.

5.3 The Intermediate Value Theorem

The following theorem was stated without proof in MATH1005. To say that it is "important" would be something of an understatement. It is the basis of a useful method for approximately solving equations (the bisection method) and, as we will see, implies the existence of square roots and other similar functions.

Theorem 5.3.1 (Intermediate Value Theorem) *Let $f : [a, b] \to \mathbb{R}$ be continuous, and y be any real number between $f(a)$ and $f(b)$. Then there exists $c \in [a, b]$ such that $f(c) = y$.*

Proof: The case $f(a) = f(b)$ is trivial since then the only number between $f(a)$ and $f(b)$ is $f(a)$ for which $c = a$ solves the problem. The case $f(a) > f(b)$ is reduced to the case $f(a) < f(b)$ by replacing f with $-f$. Therefore, we can assume without loss of generality that $f(a) < f(b)$. Define $X := \{x \in [a, b] \mid f(x) \leq y\}$. Then X is non-empty (it contains a) and bounded above (by b). Hence, there is a supremum c. Our aim is to show that $f(c) = y$. Since the interval is closed the supremum will be contained in $[a, b]$. Now take a sequence $x_n \in X$ that converges to c. Such a sequence always exits because of Theorem 2.3.7. Then, $f(x_n) \leq y$ since the sequence is in X. The limit of $f(x_n)$ exists and is equal to $f(c)$ since we assumed f to be continuous. Taking the limit, using that $\geq$ is a closed relation, we obtain

$$f(c) = \lim f(x_n) \le y < f(b).$$

Since c is the supremum of X the numbers $c + \frac{1}{n}$ are not in X. Since $c < b$ these numbers are in $[a, b]$ for n sufficiently large. Hence, for n sufficiently large we have $f(c + \frac{1}{n}) > y$. Since $c + \frac{1}{n} \to c$ this gives $f(c) \ge y$. This gives $f(c) = y$. $\square$

Note that the assumption that the domain of f is an interval is crucial:

Example 5.3.2 *$f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, $f(x) = 1/x$ is continuous (by Example 5.1.2, or Example 5.2.3). Let $a = -1$ and $b = 1$. Then $f(a) = -1$ and $f(b) = 1$ and so 0 is a number between $f(a)$ and $f(b)$. But there is no c such that $f(c) = 0$ ($1/c = 0$ has no solution). This is not a counterexample to Theorem 5.3.1, since f is* not *continuous on $[-1, 1]$.* $\square$

Example 5.3.3 *Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x^5 - 2$. Then f is continuous. In particular, it is continuous on $[0, 2]$. Now $f(0) = -2$ and $f(2) = 30$, so 0 is a number between $f(0)$ and $f(2)$. Hence, by the Intermediate Value Theorem, there exists $c \in [0, 2]$ such that $f(c) = 0$. That is, there exists $c \in [0, 2]$ such that $c^5 = 2$. Note that the Intermediate Value Theorem only tells us that c exists, not what it is, or whether it is unique (maybe there are hundreds of different c's!). However, by repeatedly applying the theorem, we can find an approximate solution to $f(c) = 0$ to arbitrarily high accuracy. We know that $f(0) < 0$ and $f(2) > 0$, so 0 is a number between $f(0)$ and $f(2)$, and hence, there is some $c \in [0, 2]$ such that $f(c) = 0$. Now test the midpoint of $[0, 2]$:*

$$f(1) = 1 - 2 < 0$$

So 0 is a number between $f(1)$ and $f(2)$, and f is continuous on $[1, 2]$, so by the IVT, there exists $c \in [1, 2]$ such that $f(c) = 0$. (Had $f(1)$ been positive, we'd have deduced that $c \in [0, 1]$. Had it been 0, we'd have deduced that $c = 1$ exactly!) Now test the midpoint of $[1, 2]$, and so on:

$$f(1.5) = 5.59375 > 0 \quad \Rightarrow \quad c \in [1, 1.5]$$

$$f(1.25) = 1.051757812 > 0 \quad \Rightarrow \quad c \in [1, 1.25]$$

$$f(1.125) = -0.197967529 < 0 \quad \Rightarrow \quad c \in [1.125, 1.25]$$

This is called the "bisection method". Each time we test the midpoint and apply the IVT, we halve the width of the interval in which we've "trapped" our solution c. We initially know c is in $[0, 2]$ an interval of width 2, so after n iterations of the method, we know that c is in some interval of width $2 \times (1/2)^n = 1/2^{n-1}$ (check: after $n = 4$ iterations, we should have c trapped in an interval of width $1/8$ —

correct!). Let's say we need to know c to within an error of 5×10^{-5}. Then we need to trap it in an interval of width less than 0.00005. How many iterations of the method do we need? The answer is any n for which

$$\frac{1}{2^{n-1}} < 0.00005.$$

will do. We know that such an n exists because the sequence $1/2^{n-1}$ converges to 0. If we allow ourselves to use logarithms, we can solve the above inequality. Otherwise, trial and error (substituting values of n) shows that $n = 16$ is the smallest positive integer satisfying the inequality. So, not only will the bisection method give us the (approximate) solution to the accuracy we require, we can figure out before we start how many iterations we'll need to make to get this accuracy. □

By the way, we've just proved that 2 has a fifth root! Our next endeavour will be to generalize this argument so as to prove the existence of square, and other, roots of any non-negative number. For this, we will need the following simple definition and fact.

Definition 5.3.4 (Stricly Increasing) *Let $D \subseteq \mathbb{R}$. A function $f : D \to \mathbb{R}$ is strictly increasing if, for all $x, y \in D$, if $x < y$ then $f(x) < f(y)$.*

Lemma 5.3.5 *Let $p \in \mathbb{N}$ and $f : [0, \infty) \to \mathbb{R}$ such that $f(x) = x^p$. Then f is strictly increasing.*

Proof: Let $x > 0$ and $y > x$. Then

$$\begin{aligned} f(y) - f(x) &= y^p - x^p \\ &= (y - x)(y^{p-1} + y^{p-2}x + \cdots + yx^{p-2} + x^{p-1}). \end{aligned}$$

Now $y - x > 0$ and

$$y^{p-1} + y^{p-2}x + \cdots + yx^{p-2} + x^{p-1} \geq y^{p-1} > 0$$

so $f(y) - f(x) > 0$. □

Proposition 5.3.6 (Existence of Roots) *Let $p \in \mathbb{N}$. Given any $y \geq 0$, there exists a unique $x \geq 0$ such that $x^p = y$. We denote this real number by $y^{1/p}$ (or, if $p = 2$, $\sqrt{y}$) and call it the p^{th} root of y.*

Proof: Consider the function $f : [0, 1 + y] \to \mathbb{R}$, $f(x) = x^p$. This, being a polynomial, is continuous. Now $f(0) = 0 \le y$. Further, I claim that $f(1 + y) > y$ since, if $y \ge 1$ then $f(1 + y) > f(y) = y^p \ge y$ (by Lemma 5.3.5), while if $y < 1$ then $f(1 + y) \ge f(1) = 1 > y$ (again by Lemma 5.3.5). So y is a real number between $f(0)$ and $f(1 + y)$. Hence, by the Intermediate Value Theorem, there exists $x \in [0, 1 + y]$ such that $f(x) = y$.
Let $z \in [0, \infty)$ also satisfy $f(z) = y$. Then $f(z) = f(x)$. Hence, z is not less than x (since this would contradict Lemma 5.3.5), nor is x less than z (as this, too would contradict Lemma 5.3.5). Hence $z = x$, that is, the real number in $[0, \infty)$ whose p^{th} power is y is unique. $\square$

We've now proved that there exists a function

$$g : [0, \infty) \to \mathbb{R}, \qquad g(x) = x^{1/p}$$

for any positive integer p. You won't be surprised to learn that this function is, like x^p, strictly increasing on $[0, \infty)$.

Lemma 5.3.7 *Let $g : [0, \infty) \to \mathbb{R}$ such that $g(x) = x^{1/p}$. Then g is strictly increasing.*

Proof: Assume, to the contrary, that $0 \le x < y$ but $g(x) \ge g(y)$. Let $f(x) = x^p$. Then, by Lemma 5.3.5,

$$f(g(y)) \le f(g(x)).$$

But $f(g(y)) = y$ and $f(g(x)) = x$, so $y \le x$, a contradiction. $\square$

You will also not be surprised to learn that g is continuous. Proving this is slightly tricky. We'll do the case $p = 2$ first, since this illustrates the required trick, then return to the general case.

Proposition 5.3.8 (Continuity of the square root) *The function $g : [0, \infty) \to \mathbb{R}$, $g(x) = \sqrt{x}$ is continuous.*

Proof: Let $a \in (0, \infty)$ (we will handle the case $a = 0$ separately). We must show that, given any sequence $x_n \in [0, \infty)$ converging to a, $g(x_n)$ converges to $g(a)$. So, let (x_n) be such a sequence, and let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that, for all $n \ge N$, $|x_n - a| < \varepsilon\sqrt{a}$. Now for all $n \ge N$,

$$\begin{aligned} |g(x_n) - g(a)| &= |\sqrt{x_n} - \sqrt{a}| \\ &= \frac{|x_n - a|}{\sqrt{x_n} + \sqrt{a}} \\ &\le \frac{1}{\sqrt{a}}|x_n - a| < \varepsilon. \end{aligned}$$

Hence $g(x_n) \to g(a)$.

Finally, let (x_n) be a sequence in $[0,\infty)$ such that $x_n \to 0$. We must show that $g(x_n) \to g(0) = 0$. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $0 \leq x_n < \varepsilon^2$. Now for all $n \geq N$,

$$\begin{aligned} |g(x_n) - g(0)| &= |\sqrt{x_n}| \\ &= \sqrt{x_n} \\ &< \sqrt{\varepsilon^2} \qquad \text{by Lemma 5.3.7} \\ &= \varepsilon. \end{aligned}$$

Hence $g(x_n) \to g(0)$. □

The trick here was to rewrite $\sqrt{x_n} - \sqrt{a}$ using the factorization formula for a difference of two squares. In the general p case, the analogue of the difference of two squares is the formula

$$x^p - y^p = (x-y)(x^{p-1} + x^{p-2}y + \cdots + xy^{p-2} + y^{p-1})$$

which we already used to prove Lemma 5.3.5.

Proposition 5.3.9 (Continuity of roots) *The function $g : [0,\infty) \to \mathbb{R}$, $g(x) = x^{1/p}$ is continuous.*

Proof: Let $a \in (0,\infty)$ (we will handle the case $a = 0$ separately). We must show that, given any sequence $x_n \in [0,\infty)$ converging to a, $g(x_n)$ converges to $g(a)$. So, let (x_n) be such a sequence, and let $\varepsilon > 0$ be given. Let $y_n = g(x_n) = x_n^{\frac{1}{p}}$ and $b = g(a) = a^{\frac{1}{p}}$. Then

$$\begin{aligned} y_n^p - b^p &= (y_n - b)(y_n^{p-1} + y_n^{p-2}b + \cdots + y_n b^{p-2} + b^{p-1}) \\ \Rightarrow \quad y_n - b &= \frac{y_n^p - b^p}{y_n^{p-1} + y_n^{p-2}b + \cdots + y_n b^{p-2} + b^{p-1}} \\ \Rightarrow \quad |y_n - b| &\leq \frac{|y_n^p - b^p|}{b^{p-1}} \\ \Rightarrow \quad |g(x_n) - g(a)| &\leq \frac{|x_n - a|}{b^{p-1}}. \end{aligned}$$

Since $x_n \to a$, there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $|x_n - a| < \varepsilon b^{p-1}$. Now for all $n \geq N$,

$$|g(x_n) - g(a)| \leq \frac{|x_n - a|}{b^{p-1}} < \varepsilon.$$

Hence $g(x_n) \to g(a)$.
Finally, let (x_n) be a sequence in $[0,\infty)$ such that $x_n \to 0$. We must show that $g(x_n) \to g(0) = 0$. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that, for all

$n \geq N$, $|x_n - 0| < \varepsilon^p$. Now for all $n \geq N$,

$$|g(x_n) - g(0)| = g(x_n) < g(\varepsilon^p) = \varepsilon$$

by Lemma 5.3.7. Hence $g(x_n) \to g(0)$. $\square$

Of course, having defined $x^{1/p}$, we can now define x^r for any rational number $r = p/q$:

$$x^{p/q} = (x^{1/q})^p.$$

Strictly speaking, we should check that this definition doesn't depend on which integers p, q we choose to "represent" $r = p/q$. For example, we should check that $x^{2/6}$ is the same as $x^{1/3}$. While we're at it, we should also probably check that $(x^{1/q})^p$ is the same as $(x^p)^{1/q}$, and that $x^{r+r'} = x^r x^{r'}$ and $x^{rr'} = (x^r)^{r'}$ etc. etc. Having established that $x^{1/p}$ *exists*, and is *unique*, verifying all these facts is an exercise in algebra which I invite the interested reader to attempt. We have bigger fish to fry.

5.4 The Extreme Value Theorem

In this section we will prove another very important property of continuous functions on closed bounded intervals: they are bounded (above and below) and, furthermore, they attain both a maximum and a minimum value.

Definition 5.4.1 *Let $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$. We say that f is bounded above if its range $f(D) \subseteq \mathbb{R}$ is bounded above, that is, if there exists $K \in \mathbb{R}$ such that, for all $x \in D$, $f(x) \leq K$. In this case, we define the supremum of f, $\sup f$, to be the supremum of its range $f(D)$ (that is, the least upper bound on the set $f(D)$). Similarly, we say that f is bounded below if its range $f(D) \subseteq \mathbb{R}$ is bounded below, that is, if there exists $L \in \mathbb{R}$ such that, for all $x \in D$, $f(x) \geq L$. In this case, we define the infimum of f, $\inf f$, to be the infimum of its range $f(D)$ (that is, the greatest lower bound on the set $f(D)$). The function is bounded if it is bounded above and below.*
If there exists $a \in D$ such that $f(x) \leq f(a) = M$ for all $x \in D$, we say that f attains a maximum value of M at a. Similarly, if there exists $b \in D$ such that $f(x) \geq f(b) = L$, we say that f attains a minimum value of L at b. $\square$

Remark If a function attains a maximum value, M, say, then it is bounded above, since $f(x) \leq M$ for all $x \in D$. Hence, by the Axiom of Completeness, f has a supremum. In fact, $\sup f = M$. (Why? Since f attains a maximum of M, there exists $a \in D$ such that $f(a) = M$. Hence, every $K < M$ is less than $f(a)$, so is not an upper bound on f. Hence, M is the *least* upper bound on f.) Note, however, that *not every function which is bounded above attains a maximum value!*

Example 5.4.2 *Let $f : (0, \infty) \to \mathbb{R}$ such that $f(x) = -1/x$. Then f is bounded above, by 0 for example (since $x > 0$, $-1/x < 0$). But it does not attain a maximum* *value because, given any $a \in (0, \infty)$, $f(a+1) = -1/(a+1) > -1/a = f(a)$. Of course, this function (or rather, its range) has a* supremum*: it must do, by the Axiom of Completeness. We've seen that 0 is an upper bound, and given any $K < 0$, $-2/K \in (0, \infty)$ and*

$$f(-2/K) = \frac{-1}{(-2/K)} = \frac{K}{2} > K \qquad \textit{(careful - remember } K < 0!\textit{)}$$

so K is not an upper bound on f. Hence 0 is the supremum of f. But I emphasize again: 0 is **not** *the maximum of f, and, in fact, f has no maximum value!*
Clearly f has no minimum value either, since it isn't even bounded below. Let's prove it: Assume to the contrary that K is a lower bound on f. Since $f(1) = -1$, $K \leq -1$. But then $x = -1/(2K) \in (0, \infty)$ and $f(x) = 2K < K$, a contradiction.
□

The function above is continuous, but bad things happen at the "edges" of its domain (as x gets big, $f(x)$ gets close to 0, but never reaches it, and as x gets close to 0, $f(x)$ grows unbounded below). We now prove that such bad things cannot happen for a continuous function on a closed, bounded interval.

Theorem 5.4.3 *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is bounded above. Furthermore, there exists $c \in [a, b]$ such that $f(c) = \sup f$.*

Proof: Assume, to the contrary, that f is unbounded above. Then for each $n \in \mathbb{N}$, n is *not* an upper bound on f, so there exists $x_n \in [a, b]$ such that $f(x_n) > n$. The sequence (x_n) is bounded, and hence, by the Bolzano-Weierstrass Theorem, has a convergent subsequence, (x_{n_k}), converging to x, say. Since $a \leq x_{n_k} \leq b$ for all k, we know that $x \in [a, b]$. Now f is continuous, so $f(x_{n_k}) \to f(x)$, and hence $f(x_{n_k})$ is bounded (convergent sequences are bounded). But, by definition $f(x_{n_k}) > n_k \geq k$, so $f(x_{n_k})$ is unbounded, a contradiction. Hence, f is bounded above.
Since $A = \{f(t) : t \in [a, b]\}$ is nonempty and bounded above, it has a supremum L say. It remains to show that there exists $c \in [a, b]$ such that $f(c) = L$. Given any $n \in \mathbb{N}$, $L - 1/n < L$ so is *not* an upper bound on A. Hence, there exists $y_n \in [a, b]$ such that $L - 1/n < f(y_n) \leq L$. Since y_n is bounded, it has a convergent subsequence y_{n_k} (by the Bolzano-Weierstrass Theorem). Call its limit c. Then, $c \in [a, b]$. Further $L - 1/n_k < f(y_{n_k}) \leq L$, so $f(y_{n_k}) \to L$ by the Squeeze Rule. But f is continuous, so $f(y_{n_k}) \to f(c)$. Hence $f(c) = L$. □

Corollary 5.4.4 (Extreme Value Theorem (EVT)) *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is bounded and attains both a maximum and a minimum value.*

Proof: By Theorem 5.4.3, f is bounded above, and there exists $c \in [a, b]$ such that $f(c) = M$ where M is the supremum of f. By definition (of supremum),

$f(x) \leq f(c)$ for all $x \in [a, b]$, so f attains a maximum value (of M at c).
The function $g : [a, b] \to \mathbb{R}$, $g(x) = -f(x)$ is continuous, so, by Theorem 5.4.3 is also bounded above and there exists $d \in [a, b]$ such that $g(d) = L$, where L is the supremum of g. By definition (of supremum), $g(x) \leq g(d)$ for all $x \in [a, b]$. Hence, $f(x) \geq f(d)$ for all $x \in [a, b]$, so f attains a minimum value (of $-L$ at d). $\square$

Example 5.4.5 *The rational function*

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

is continuous (see Example 5.2.3) and hence, by the Extreme Value Theorem, is bounded and attains both a maximum and a minimum value. In this case, it's not hard to show this directly: for all $x \in [-1, 1]$, $x^2 + 1 \geq 1$, so $f(x) \leq f(0) = 1$, and hence f attains a maximum value of 1 at 0. Similarly, for all $x \in [-1, 1]$, $1 + x^2 \leq 2$, so $f(x) \geq f(1) = 1/2$, and hence f attains a minimum value of $1/2$ at 1. $\square$

Example 5.4.6 *The rational function*

$$f : [-13, -1] \to \mathbb{R}, \qquad f(x) = \frac{x^{27} - 99x^{20} + 45x^{11} - 14x^7 - 5}{x^{12}(x-1)^{15}(x^{80} + 1)^{99}}$$

is continuous (see Example 5.2.3) and hence, by the Extreme Value Theorem, is bounded and attains both a maximum and a minimum value. Showing this directly would be quite fiddly. $\square$

5.5 Supplementary material

First of all a general remark. Continuity is a **local property**. This means to check if a function $f : D \to \mathbb{R}$ is continuous at $a \in D$ it is sufficient to know the function near a in the following sense. Suppose that $\delta > 0$ is any number and assume that the function g is obtained by restricting f to $(a - \delta, a + \delta) \cap D$. Then, if g is continuous at a, then so is f. This follows easily by noting that any sequence in D converging to a will eventually have to be contained in $(a - \delta, a + \delta) \cap D$, i.e. there exists an $N \in \mathbb{N}$ such that $a_n \in (a - \delta, a + \delta) \cap D$ for $n > N$. The first N terms do however not change convergence of the sequence. I will show below how to use this observation in arguments.

5.5.1 The modulus is continuous

Consider the function $|\cdot| : \mathbb{R} \to \mathbb{R}, x \mapsto |x|$. This function agrees with x and $[0, \infty)$ and with $-x$ on $(-\infty, 0)$. By the above remark the function $|\cdot|$ is therefore continuous at $a \in \mathbb{R}$ for all $a \neq 0$. To see that $|\cdot|$ is continuous on $\mathbb{R}$ it is therefore sufficient to check continuity at zero. To see how to do this assume that a_n is any sequence converging to zero. This means also that $|a_n| \to 0$ (since the definition of the limit to zero only involves $|a_n|$ anyway). This shows that $|\cdot|$ is continuous at zero.

5.5.2 Continuity of piecewise other functions

Assume that $f : \mathbb{R} \to \mathbb{R}$ is given by

$$f(x) = \begin{cases} g(x) & x < 0, \\ h(x) & x \geq 0, \end{cases}$$

where $g, h : \mathbb{R} \to \mathbb{R}$ are continuous. This means the function agrees with g on the set $(0, \infty)$ and is therefore continuous on this set because g is. The function agrees with h on the set $(-\infty, 0)$ and is therefore continuous on this set because h is. The only remaining point is the point $x = 0$. There are two possibilities. Either

$$g(0) = h(0) \text{ or } g(0) \neq h(0).$$

We will now show that f is continuous at 0 in case $g(0) = h(0)$ and it is not continuous at zero in case $g(0) \neq h(0)$.

Suppose that $g(0) \neq h(0)$. We will show that then f is not continuous. Indeed, simply choose the sequence $x_n = (-1)^n \frac{1}{n}$. This converges to zero and therefore $g(x_n)$ converges to $g(0)$ and $h(x_n)$ converges to $h(0)$. We have used continuity of g and h here. If n is odd then $x_n < 0$ and if n is even we have $x_n > 0$. Therefore,

$$f(x_n) = \begin{cases} g(x_n) & n \text{ odd} \\ h(x_n) & n \text{ even.} \end{cases}$$

Thus, $f(x_n)$ has two convergent subsequences $f(x_{2n+1})$, which converges to $g(0)$, and $f(x_{2n})$, which converges to $h(0)$. Since these are distinct the sequence $f(x_n)$ does not converge, and therefore f is not continuous at zero.

Now assume that $g(0) = h(0)$. To show f is continuous it is not sufficient to simply use the same sequence above, but we need to show that $f(x_n) \to f(0) = g(0) = h(0)$ for all possible sequences (x_n) with $x_n \to 0$. Suppose we have an arbitrary sequence x_n that converges to 0. We will show that $f(x_n)$ has only one accumulation point and is bounded, and therefore converges. First note that x_n is bounded since it converges. Therefore x_n is contained in some interval $[-M, M]$. Next use the EVT to conclude that both g and h are bounded functions on $[-M, M]$. Therefore, so is f (even though we do not know it is continuous yet), because $\sup f \leq \max\{\sup g, \sup h\}$ and $\inf f \geq \min\{\inf g, \inf h\}$. Hence, $f(x_n)$ is bounded. Pick a convergent subsequence $f(x_{n_k})$. Then either there are infinitely many terms $x_{n_k} < 0$ or infinitely many terms $x_{n_k} \geq 0$. In the former case we have a subsequence of x_{n_k} which converges to zero from the left, in the latter case we have a subsequence that converges to zero from the right. In the former case $f(x_{n_k})$ must therefore converge to $g(0)$. In the latter case it must converge to $h(0)$. Since these two are equal there is only one accumulation point. This implies $\liminf f(x_n) = \limsup f(x_n) = f(0)$ and therefore the sequence $f(x_n)$ converges to $f(0)$. Hence the function is continuous at zero.

The same argument works for points other than zero.

Example 5.5.1 *The function* $f_1 : \mathbb{R} \to \mathbb{R}$ *defined by*

$$f_1(x) = \begin{cases} x^2 & x < 0, \\ -x & x \geq 0, \end{cases}$$

is continuous. The function $f_2 : \mathbb{R} \to \mathbb{R}$ *defined by*

$$f_2(x) = \begin{cases} 1 + x^2 & x < 0, \\ -x & x \geq 0, \end{cases}$$

is not continuous.

Chapter 6

Some symbolic logic

6.1 Statements and their symbolic manipulation

In this module we have dealt with many mathematical statements, some of which were quite logically subtle. For example, consider the statement

> "The real sequence (a_n) converges to L."

We learned that this statement really means

> "For each positive real number ε there exists a positive integer N such that, for all $n \geq N$, $|a_n - L| < \varepsilon$."

What does the statement

> "The real sequence (a_n) converges"

mean? It means that (a_n) converges to L for some real number L. That is,

> "There exists a real number L such that, for each positive number ε, there exists a positive integer N such that, for all $n \geq N$, $|a_n - L| < \varepsilon$."

Recall that a sequence *diverges* if it doesn't converge. So what does the statement

> "The real sequence (a_n) diverges"

really mean? We will see that it means precisely

> "For each real number L there exists a positive real number ε such that for all positive integers N there exists $n \geq N$ such that $|a_n - L| \geq \varepsilon$."

This is, to put it mildly, *not* obvious! In this section we will introduce some basic ideas from symbolic logic which will allow us to manipulate and analyze complicated mathematical statements like "(a_n) converges," and obtain from them related statements like "(a_n) diverges."

The key idea is to represent mathematical statements by symbols which can be combined and algebraically manipulated. For our purposes, a **statement** is any declaration which is unambiguously either true or false. For example, the declarations

P:	$(-2)^2 = 4$
Q:	Leeds is the third largest city in the UK by population
R:	$-2 > 2$
S:	Every even integer greater than 2 can be expressed as the sum of two primes

are all statements: P and Q are true, R is false, and no-one knows (at the time of writing) whether S is true or false, but it is clearly one or the other[1]. On the other hand, declarations like

P':	Tough on crime, tough on the causes of crime!
Q':	Lucy in the sky with diamonds

are not statements because neither can be meaningfully said to be true, or false.

Given one or more statements $P, Q, \ldots$, we can produce new statements from them by three basic constructions:

- **Negation**: not P, denoted by $\neg P$, is the statement which is true if and only if P is false. For the examples above, we have

$\neg P$:	$(-2)^2 \neq 4$	False
$\neg Q$:	Leeds is not the third largest city in the UK by population	False
$\neg R$	$-2 \leq 2$	True

- **Conjunction**: P and Q, denoted by $P \wedge Q$, is the statement which is true if both P and Q are true, and false otherwise. So, for the examples above, $P \wedge Q$ is true, but $P \wedge R$ is false (because R is false).

- **Disjunction**: P or Q, denoted by $P \vee Q$, is the statement which is true if P is true, or Q is true, or both, but is false if both P and Q are false. So, for the examples above, $P \vee Q$ is true, as is $P \vee R$ (because P is true). However, $(\neg P) \vee R$ is false (because both $\neg P$ and R are false).

You should bear in mind that in mathematics we use the word "or" in its non-exclusive sense. That is "P or Q" means "P or Q (or both)," which is different from its use in everyday English. For example, a child who is told

"Finish your broccoli or you can't have any ice cream"

could reasonably expect to get some ice cream if he/she finishes his/her broccoli. If the nagging adult is a mathematician, the child would be wise to seek clarification.

It is useful to summarize these definitions using **truth tables**. These represent the truth/falsity of a constructed statement in terms of the truth/falsity of its constituent pieces. We represent "false" by the value 0 and "true" by the value 1:

P	$\neg P$
0	1
1	0

P	Q	$P \wedge Q$
0	0	0
0	1	0
1	0	0
1	1	1

P	Q	$P \vee Q$
0	0	0
0	1	1
1	0	1
1	1	1

[1]This is the infamous Goldbach Conjecture.

Two statements P, Q are **logically equivalent** if one is true if and only if the other is true. We represent this symbolically by $P \iff Q$. For example

P	$\neg P$	$\neg(\neg P)$
0	1	0
1	0	1

so for any statement,

$$\neg(\neg P) \iff P \tag{6.1}$$

which is sometimes called the Principle of Double Negation. We can use truth tables to verify many other logical equivalences. For example:

P	Q	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P$	$\neg Q$	$(\neg P) \vee (\neg Q)$
0	0	0	1	1	1	1
0	1	0	1	1	0	1
1	0	0	1	0	1	1
1	1	1	0	0	0	0

from which we deduce that for all statements P, Q,

$$\neg(P \wedge Q) \iff (\neg P) \vee (\neg Q). \tag{6.2}$$

As a consequence of (6.1) and (6.2),

$$\begin{aligned} \neg(P \vee Q) &\iff \neg[(\neg\neg P) \vee (\neg\neg Q)] && \text{by (6.1)} \\ &\iff \neg[\neg((\neg P) \wedge (\neg Q))] && \text{by (6.2)} \\ &\iff (\neg P) \wedge (\neg Q) && \text{by (6.1)} \end{aligned}$$

that is, for all statements P, Q,

$$\neg(P \vee Q) \iff (\neg P) \wedge (\neg Q). \tag{6.3}$$

Exercise 6.1.1 *Verify this explicitly using a truth table.*

The rules (6.2) and (6.3) are known as **de Morgan's laws**.

Example 6.1.2 *Find the negation of the statement* $P \vee (Q \wedge R)$.
Solution:

$$\begin{aligned} \neg[P \vee (Q \wedge R)] &\iff (\neg P) \wedge \neg(Q \wedge R) && \textit{by (6.3)} \\ &\iff (\neg P) \wedge [(\neg Q) \vee (\neg R)] && \textit{by (6.2)} \end{aligned}$$

$\square$

6.2 Implications

It is convenient to define another connective between two statements P, Q in addition to conjunction ($P \wedge Q$) and disjunction ($P \vee Q$), namely **implication**, denoted by $P \Rightarrow Q$, pronounced "if P then Q", or "P implies Q". (Actually, the standard notation is $P \to Q$, and $P \Rightarrow Q$ is reserved for something subtly different. However, we are already making heavy use of the symbol $\to$ to denote convergence of sequences, so we will stick to $\Rightarrow$ for implication.) The truth table for this statement is

P	Q	$P \Rightarrow Q$
0	0	1
0	1	1
1	0	0
1	1	1

As an example, consider the statement

> "If everyone in this class passes the final exam, I will eat my hat."

This is clearly false if everyone passes the exam, and I refuse to eat my hat. On the other hand, it's clearly true if everyone passes and I *do* eat my hat. What if not everyone passes? Then the question of whether I eat my hat becomes moot: the statement is true if I don't, but it's also true if I do (I never said I'd eat my hat *only* if everyone passes - maybe I just like eating hats!). So $P \Rightarrow Q$ is false only when P is true and Q is false. In every other case, it is true.

Unlike $P \wedge Q$ and $P \vee Q$, the order of the statements P, Q in $P \Rightarrow Q$ matters, as can be seen immediately from its truth table. For example, the statement

> "If you are vegetarian then you don't eat kangaroo"

is true, whereas

> "If you don't eat kangaroo then you are vegetarian"

is false. Given an implication $P \Rightarrow Q$, the implication $Q \Rightarrow P$ is called its **converse**. As we have just seen, an implication and its converse are **not** logically equivalent. We have seen many examples of this fact in this module. For example

> "If a real sequence converges then it is bounded"

is true, but its converse

> "If a real sequence is bounded then it converges"

is false (e.g. $a_n = (-1)^n$). Similarly

> "If a real sequence is bounded then it has a convergent subsequence"

is true (this is the Bolzano-Weierstrass Theorem), but its converse

> "If a real sequence has a convergent subsequence then it is bounded"

is false (e.g. $a_n = (1 + (-1)^n)n$). Another important, and often misused, example is

"If $\sum_{n=1}^{\infty} a_n$ converges then $a_n \to 0$"

which is true (this is the Divergence Test), but its converse

"If $a_n \to 0$ then $\sum_{n=0}^{\infty} a_n$ converges"

is false (e.g. $\sum_{n=1}^{\infty} 1/n$).

It is possible to rewrite $P \Rightarrow Q$ using only $\neg$ and $\vee$, and this is often useful when analyzing the precise meaning of an implication. In fact

$$[P \Rightarrow Q] \quad \iff \quad [(\neg P) \vee Q]. \tag{6.4}$$

We can check this by computing the truth table of the right hand side:

P	Q	$\neg P$	$(\neg P) \vee Q$
0	0	1	1
0	1	1	1
1	0	0	0
1	1	0	1

From this, we deduce that

$$\begin{aligned} [P \Rightarrow Q] &\iff [(\neg P) \vee Q] \\ &\iff [Q \vee (\neg P)] \\ &\iff [(\neg(\neg Q)) \vee (\neg P)] \qquad \text{by (6.1)} \\ &\iff [(\neg Q) \Rightarrow (\neg P)] \qquad \text{by (6.4)} \end{aligned}$$

So, for any pair of statements P, Q,

$$[P \Rightarrow Q] \quad \iff \quad [(\neg Q) \Rightarrow (\neg P)].$$

Exercise 6.2.1 *Check this by computing the truth table of* $(\neg Q) \Rightarrow (\neg P)$.

Given an implication $P \Rightarrow Q$, the implication $(\neg Q) \Rightarrow (\neg P)$ is called its **contrapositive**. We have just shown that an implication and its contrapositive are logically equivalent. Looking back at the above examples, we see that the statement

"A convergent sequence is bounded"

could equally well have been stated

"If a real sequence is unbounded then it does not converge,"

and the Bolzano-Weierstrass Theorem is equivalent to

"If a real sequence has no convergent subsequence then it is unbounded,"

and the Divergence Test is equivalent to

"If a_n does not converge to 0 then $\sum_{n=1}^{\infty} a_n$ diverges,"

which is, in fact, the form in which one actually uses it.

Example 6.2.2 *Write down the converse and contrapositive of each of the following implications.*

(i) *If $x^2 = 2$ then x is irrational*

Converse: If x is irrational then $x^2 = 2$

Contrapositive: If x is rational then $x^2 \neq 2$

(ii) *If $a_n \to A$ and $b_n \to B$ then $a_n + b_n \to A + B$*

Converse: If $a_n + b_n \to A + B$ then $a_n \to A$ and $b_n \to B$

Contrapositive: If $a_n + b_n \nrightarrow A + B$ then $a_n \nrightarrow A$ or $b_n \nrightarrow B$

(iii) *If $x > 2$ then $x^2 > 4$*

Converse: If $x^2 > 4$ then $x > 2$

Contrapostive: If $x^2 \leq 4$ then $x \leq 2$ □

If you are asked to prove an implication, you should always consider whether proving its contrapositive will be easier. Since the contrapositive is logically equivalent to the original implication, this will suffice.

Example 6.2.3 *Claim: Let n be an integer. If n^2 is divisible by 3 then n is divisible by 3.*

Proof: *We will prove the contrapositive: if n is* not *divisible by 3 then n^2 is* not *divisible by 3. So, assume that n is not divisible by 3. Then its remainder on division by 3 is either 1 or 2, that is, either $n = 3k + 1$ or $n = 3k + 2$, where k is an integer. In the first case*

$$n^2 = 9k^2 + 6k + 1 = 3k' + 1$$

where $k' = 3k^2 + 2k \in \mathbb{Z}$, and in the second case

$$n^2 = 9k^2 + 12k + 4 = 3k' + 1$$

where $k' = 3k^2 + 4k + 1 \in \mathbb{Z}$. In either case, we see that n^2 is not divisible by 3 (in fact, its remainder is 1 in both cases), as was to be proved. □

Another important fact follows immediately from (6.4): **the negation of an implication is NOT an implication!**

$$\neg[P \Rightarrow Q] \quad \Longleftrightarrow \quad \neg[Q \vee \neg P] \quad \Longleftrightarrow \quad (\neg Q) \wedge (\neg\neg P) \quad \Longleftrightarrow \quad P \wedge \neg Q.$$

Note that this is entirely different from $(\neg P) \Rightarrow (\neg Q)$. For example, the negation of the statement "if n is prime then n is odd" is the statement "n is prime and n is not odd", **not** the statement "if n is not prime then n is not odd."

6.3 Quantifiers

We have been a little bit sloppy up to now in dealing with statements which contain variables. Consider, for example, the statement

$$x > 2.$$

Is this really a statement? Until one specifies a particular value for x, one cannot say whether it is true. So it's true for $x = 3$, and false for $x = -2$, for example. What about if $x = 1 + 3i$? The "statement" $1 + 3i > 2$ is pretty meaningless, and it isn't very helpful to declare it either true or false (e.g. if you decide it's false, then you'd expect $1 + 3i \leq 2$ to be true, and this "statement" is equally meaningless). If we're being careful, we should make explicit the condition that x is a real number, and if we denote the statement $x > 2$ with a symbol, we should choose one like $P(x)$, rather than P, to remind ourselves that whether $P(x)$ is true depends on the particular x under consideration. So $P(3)$ is true, in this case, while $P(2)$ is false.

There are two very useful ways to turn a statement with a variable into an unambiguous true-or-false statement, using **quantifiers**. Let $P(x)$ be any statement which makes sense for all values of x in some specified set A (e.g. $A = \mathbb{R}$ and $P(x) : x > 2$, as above). Then

$\forall x\, P(x)$ means "for all $x \in A$, $P(x)$ is true"

and

$\exists x\, P(x)$ means "there exists some $x \in A$ such that $P(x)$ is true".

The symbol $\forall$ is called the **universal quantifier** and the symbol $\exists$ is called the **existential quantifier**. For the example above, $\forall x\, P(x)$ is false, because, for example, $P(-1)$ is false (since -1 is not greater than 2), but $\exists x\, P(x)$ is true because, for example, $P(3)$ is true. Note that, once we've coupled it with a quantifier, the statement no longer contains a variable, and is a simple true or false assertion. Note also that it's entirely irrelevant what we *call* the variable. So $\forall y\, P(y)$ is precisely the same statement as $\forall x\, P(x)$.

In practice, rather than specifying the set A in advance, we often specify it explicitly with the quantifier. In the above example, we would write

$$\forall x \in \mathbb{R},\ x > 2$$

which is false, and

$$\exists x \in \mathbb{R},\ x > 2 \tag{6.5}$$

which is true. You should feel free to add punctuation and/or ordinary English to expressions like this, if you think it clarifies the meaning. For example, I think

$$\exists x \in \mathbb{R} \text{ such that } x > 2$$

is clearer than (6.5). Indeed, you should feel free to avoid using the symbols $\forall$ and $\exists$ entirely if you think they obscure the mathematical meaning of the sentences you are writing. As we will see, they can be extremely useful, but you shouldn't chuck them into your mathematical prose willy-nilly. They won't make your proofs any more rigorous, elegant or impressive.

Example 6.3.1 *Translate the following sentences into symbols:*

(i) *The square of any even integer is even.*

In symbols: $\forall n \in \mathbb{Z}, [(n/2) \in \mathbb{Z} \Rightarrow (n^2/2) \in \mathbb{Z}]$

(ii) *Given any real number K there is a positive integer N such that $N > K$.*

In symbols: $\forall K \in \mathbb{R}, \exists N \in \mathbb{N}\, N > K$

(iii) *There exists a rational number whose square is 2.*

In symbols: $\exists x \in \mathbb{Q},\, x^2 = 2$

(iv) *There exists a real number which is bigger than all rational numbers.*

In symbols: $\exists x \in \mathbb{R}, \forall y \in \mathbb{Q}\, x > y$

(v) *The sequence (a_n) converges to L.*

In symbols: $\forall \varepsilon \in (0, \infty), \exists N \in \mathbb{N}, \forall n \in \mathbb{N} \cap [N, \infty), |a_n - L| < \varepsilon$

(vi) *The sequence (a_n) converges.*

In symbols: $\exists L \in \mathbb{R}, \forall \varepsilon \in (0, \infty), \exists N \in \mathbb{N}, \forall n \in \mathbb{N} \cap [N, \infty), |a_n - L| < \varepsilon$ □

Example 6.3.2 *Translate the following "sentences" into English:*

(i) $\forall x \in \mathbb{Z},\ x^2 > x$

Translation: The square of every integer x exceeds x.

Better translation: Every integer is less than its square.

(ii) $\exists x \in \mathbb{R},\ \forall y \in \mathbb{R}\ x^2 > y$

Translation: There exists a real number whose square is greater than every real number.

(iii) $\forall y \in \mathbb{R},\ \exists x \in \mathbb{R}\ x^2 > y$

Translation: For each real number y, there exists a real number x whose square exceeds y.

Better translation: The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, is unbounded above.
$\square$

In the above example, we had a statement depending on two variables, $P(x, y)$, to which we applied a pair of quantifiers, $\exists x$ and $\forall y$. An important point to notice is that *it matters which order we apply the quantifiers in!* That is

$$\exists x\ \forall y\ P(x, y) \qquad \text{is not the same as} \qquad \forall y\ \exists x\ P(x, y).$$

In Example 6.3.2, we had $P(x, y) : x^2 > y$, from which we constructed

$$S_1 : \qquad \exists x \in \mathbb{R}\, \forall y \in \mathbb{R}\, P(x, y) \qquad\qquad \text{FALSE!}$$

and

$$S_2 : \qquad \forall y \in \mathbb{R}\, \exists x \in \mathbb{R}\, P(x, y) \qquad\qquad \text{TRUE!}$$

Since S_2 is true and S_1 is false, they're clearly not equivalent.

Example 6.3.3 *Here's a non-mathematical example illustrating the same point. Let X be the set of all people (in this room, say), C be the set of all colours and $P(x, c)$ be the statement "x's favourite colour is c". Then*

$$\forall x \in X\ \exists c \in C\ \ P(x, c) \quad \Longleftrightarrow \quad \text{Everyone has a favourite colour,}$$

whereas

$$\exists c \in C\ \forall x \in X\ \ P(x, c) \quad \Longleftrightarrow \quad \text{There is a colour which is everyone's favourite,}$$

which is clearly a very different statement. $\square$

The situation in which quantifiers really come into their own is when you want to understand the *negation* of a statement involving several quantifiers. Let's ask ourselves, what does it mean to say that $\forall x\, P(x)$ is *false*? It means that it is *not* true that for all x in A, $P(x)$ is true. That is, there is at least one element x in A for which $P(x)$ is *not* true. In other words, there exists $x \in A$ such that $\neg P(x)$ is true. That is

$$\neg[\forall x\, P(x)] \qquad \Longleftrightarrow \qquad \exists x\, \neg P(x). \tag{6.6}$$

Example 6.3.4 *Consider the assertion that "all politicians are corrupt." I think this is false. In saying that, I am certainly* not *asserting that "all politicians are not corrupt" (which is equivalent to "no politician is corrupt"). Rather, I am asserting that "there exists at least one politician who is not corrupt." In symbols, let X*

denote the set of politicians and $P(x)$ be the statement "x is corrupt." Then the negation of $\forall x \in X\, P(x)$ is **not** *$\forall x \in X\, \neg P(x)$, but rather $\exists x \in X\, \neg P(x)$.* $\square$

Note that to prove that $[\forall x \in A\ P(x)]$ is **false**, we need to show that $[\exists x \in A\, \neg P(x)]$ is true, that is, we need to show that there exists (at least) **one** element x in A such that $P(x)$ is **false**. Such an element x is called a **counterexample** to $[\forall x \in A\ P(x)]$.

Example 6.3.5 *Consider the statement $\forall x \in \mathbb{R}\ x^2 < 10$. This is clearly false. To prove that it is false, all we need to do is exhibit a* **single** *real number whose square is not less than 10. For example, $x = 10$ will do (since $10^2 \geq 10$). To prove the statement is false we do* **not** *have to find* **all** *real numbers x such that $x^2 \geq 10$. Giving a single counterexample is quicker, clearer and more elegant.* $\square$

The rule for negating $\exists x\, P(x)$ actually follows from (6.6) and rule (6.1), but let's derive it by thinking. What does it mean to say that $\exists x\, P(x)$ is *false*? It means that there does *not* exist x in A such that $P(x)$ is true. Hence, for each and every $x \in A$, $P(x)$ must be *false*. That is

$$\neg[\exists x\, P(x)] \qquad \Longleftrightarrow \qquad \forall x\, \neg P(x).$$

Example 6.3.6 *Consider the assertion "someone in this room speaks fluent Mandarin." Let's assume this is false. What does this mean? It certainly does* not *merely mean "someone in this room does not speak fluent Mandarin." Rather it means "no-one in this room speaks fluent Mandarin," or, equivalently, "everyone in this room does not speak fluent Mandarin." In symbols, let X denote the set of people in this room and $P(x)$ be the statement "x speaks fluent Mandarin." Then the negation of $\exists x \in X\, P(x)$ is* **not** *$\exists x \in X\, \neg P(x)$, but rather $\forall x \in X\, \neg P(x)$.* $\square$

Example 6.3.7 *Negate the following statements. Determine which is true, the statement or its negation.*

(i)

$$S \quad : \qquad \forall x \in \mathbb{Z}\ x^2 > x$$

$$\neg S \quad : \qquad \exists x \in \mathbb{Z}\ x^2 \leq x$$

$\neg S$ is true (for example $0^2 \leq 0$).

(ii)

$$S \quad : \qquad \exists x \in \mathbb{R}\ x^2 + 2x + 2 = 0$$

$$\neg S \quad : \qquad \forall x \in \mathbb{R}\ x^2 + 2x + 2 \neq 0$$

$\neg S$ is true, since $x^2 + 2x + 2 = (x+1)^2 + 1 \geq 1$.

(iii)

$$\begin{aligned}
S &: \quad \forall x \in (0,\infty)\, \exists y \in \mathbb{Q}\; 0 < y < 1/x \\
\neg S &: \quad \exists x \in (0,\infty)\, \neg[\exists y \in \mathbb{Q}\; 0 < y < 1/x] \\
&\iff \exists x \in (0,\infty)\, \forall y \in \mathbb{Q}\; \neg[0 < y < 1/x] \\
&\iff \exists x \in (0,\infty)\, \forall y \in \mathbb{Q}\; \neg[0 < y \,\wedge\, y < 1/x] \\
&\iff \exists x \in (0,\infty)\, \forall y \in \mathbb{Q}\; [y \le 0 \,\vee\, y \ge 1/x]
\end{aligned}$$

S is true by density of rationals in the reals given any $x > 0$, $1/x > 0$, so there exists a rational number y between 0 and $1/x$. □

You've already seen these negation rules in many different contexts. For example, consider the Archimedean Property of $\mathbb{R}$. Our first statement of this property was simply that "$\mathbb{N}$ is unbounded above". We then reformulated this as

> "Given any real number K, there is some positive integer n such that $n > K$."

We can now see symbolically that these two statements are logically equivalent:

$$\begin{aligned}
\mathbb{N} \text{ is unbounded above} &\iff \neg\,[\mathbb{N} \text{ is bounded above}] \\
&\iff \neg[\exists K \in \mathbb{R}\; \forall n \in \mathbb{N}\; n \le K] \\
&\iff \forall K \in \mathbb{R} \neg[\forall n \in \mathbb{N}\; n \le K] \\
&\iff \forall K \in \mathbb{R}\; \exists n \in \mathbb{N} \neg[n \le K] \\
&\iff \forall K \in \mathbb{R}\; \exists n \in \mathbb{N}\, n > K.
\end{aligned}$$

In this case, using symbolic logic was really overkill: we could figure out the negation just by thinking carefully. A good example where symbolic logic really helps is the statement "(a_n) does not converge to L". To illustrate, consider the following question, which is taken from an old tutorial problem set:

Example 6.3.8 *Let (a_n) be a real sequence with the property that every subsequence of (a_n) has a subsequence which converges to 0. Show that (a_n) converges to 0.*

Proof: *We will prove this by contradiction. So, let (a_n) be a sequence with the specified property, but assume that (a_n) does* **not** *converge to 0. Our first task is to understand precisely what this means:*

$$\begin{aligned}
a_n \not\to 0 \quad &\iff \quad \neg[a_n \to 0] \\
&\iff \quad \neg[\forall \varepsilon \in (0,\infty)\, \exists N \in \mathbb{N}\, \forall n \in \mathbb{N} \cap [N,\infty)\ |a_n - 0| < \varepsilon] \\
&\iff \quad \exists \varepsilon \in (0,\infty) \neg[\exists N \in \mathbb{N}\, \forall n \in \mathbb{N} \cap [N,\infty)\ |a_n| < \varepsilon] \\
&\iff \quad \exists \varepsilon \in (0,\infty)\, \forall N \in \mathbb{N} \neg[\forall n \in \mathbb{N} \cap [N,\infty)\ |a_n| < \varepsilon] \\
&\iff \quad \exists \varepsilon \in (0,\infty)\, \forall N \in \mathbb{N}\, \exists n \in \mathbb{N} \cap [N,\infty)\ \neg[|a_n| < \varepsilon] \\
&\iff \quad \exists \varepsilon \in (0,\infty)\, \forall N \in \mathbb{N}\, \exists n \in \mathbb{N} \cap [N,\infty)\ |a_n| \geq \varepsilon.
\end{aligned}$$

Translating back into English, $a_n \not\to 0$ means

> *there exists $\varepsilon > 0$ such that, for all $N \in \mathbb{N}$ there is some integer $n \geq N$ such that $|a_n| \geq \varepsilon$.*

Let ε_* be this magic positive real number. Then we know that, for each and every $N \in \mathbb{N}$ there is some $n \geq N$ such that $|a_n| \geq \varepsilon_*$. So this is true for $N = 1$, for example: there exists an integer n_1 say such that $n_1 \geq 1$ and $|a_{n_1}| \geq \varepsilon_*$. It's also true for $N = n_1 + 1$: there exists $n_2 \geq n_1 + 1$ such that $|a_{n_2}| \geq \varepsilon_*$. It's also true for $N = n_2 + 1$, and so on. In this way we generate a subsequence *$b_k = a_{n_k}$ of (a_n) such that $|b_k| = |a_{n_k}| \geq \varepsilon_*$ for all $k \in \mathbb{N}$. By assumption the subsequence (b_k) itself has a subsequence which* does *converge to 0, $c_m = b_{k_m}$ say. Since $c_m \to 0$, given any postive number, for example our magic number ε_*, there is some positive integer M such that, for all $m \geq M$, $|c_m - 0| < \varepsilon_*$. In particular, $|c_M| < \varepsilon_*$. But $c_M = b_{k_M}$, and the sequence (b_k) was chosen specifically so that $|b_k| \geq \varepsilon_*$ for all $k \in \mathbb{N}$, including k_M. Hence, we have reached a contradiction, so our intial assumption that $a_n \not\to 0$ must be false.* $\square$

Remark This result is quite subtle. It does *not* say that

> If every susbequence of (a_n) converges to 0 then (a_n) converges to 0

although this is trivially true, since (a_n) is a subsequence of itself! On the other hand, by Theorem 3.1.3, Example 6.3.8 has the following corollary:

> If every subsequence of (a_n) has a subsequence which converges to 0, then every subsequence of (a_n) converges to 0.

This is true despite the fact that, in general, having a subsequence which converges to 0 does *not* imply that a sequence converges to 0.

Here are a couple of questions for you to think about.

(i) Consider the statement

> If every subsequence of (a_n) has a subsequence which has a subsequence which converges to 0 then (a_n) converges to 0.

Is this true?

(ii) In Example 6.3.8, we assumed that all subsequences have a subsequence converging to 0. Presumably there's nothing special about 0. We could equally well have assumed that all subsequences have a subsequence converging to L for any other fixed $L \in \mathbb{R}$. (If this isn't clear, try adapting the above proof to deal with the case of general L.) Now consider the following statement:

> If every subsequence of (a_n) has a convergent subsequence then (a_n) is convergent.

Is this true? $\square$

Chapter 7

Limits of functions

7.1 The main definition

We now have a very precise (and powerful) definition of *limit* for real sequences. In this section we will show how to use this to give a precise definition of the limit of a real function of a real variable. We will also see that basic properties of limits of functions, which you assumed without proof in MATH1005, follow very easily from the corresponding limit theorems for sequences.

Informally, $\lim_{x\to a} f(x) = L$ means that $f(x)$ is "arbitrarily close" to L for all x "sufficiently close to, but different from" a. So limits concern the behaviour of x arbitrarily close to a, but *not at* a. There is no reason why a has to be in the domain of f, that is, $f(a)$ may, or may not, be defined. If it is, its value should be irrelevant to $\lim_{x\to a} f(x)$. For example,

$$f(x) = \frac{\sin x}{x}$$

is clearly undefined at 0 (its domain is $(-\infty, 0) \cup (0, \infty)$), but $\lim_{x\to 0} f(x)$ is well defined (it's 1).

In order to probe the behaviour of $f(x)$ for values of x close to, but different from, a, we can consider the sequence $f(x_n)$ where (x_n) is any sequence *converging* to a with $x_n \neq a$. First we need to make sure that there exists such a sequence.

Definition 7.1.1 *Let $D \subseteq \mathbb{R}$. Then $a \in \mathbb{R}$ is a* **limit point** *of D if there exists a sequence (x_n) in $D\backslash\{a\}$ such that $x_n \to a$.*

Given a function $f : D \to \mathbb{R}$ it only makes sense to try to define its limit at a, as in Definition 7.1.2, if a is a limit point of D. This leads us to:

Definition 7.1.2 (Limits of functions) *Let $D \subseteq \mathbb{R}$, and a be a limit point of D. Suppose $f : D \to \mathbb{R}$, Then f has* **limit** *L* **at** *a if, for all sequences (x_n) in $D\backslash\{a\}$ such that $x_n \to a$, $f(x_n) \to L$. In this case, we write $\lim_{x\to a} f(x) = L$.*

Remarks

(i) It's important to note that the definition says that for **all** sequences (x_n) in $D\backslash\{a\}$ converging to a, $f(x_n) \to L$. It's not enough to show that $f(x_n) \to L$ for just one specific sequence.

(ii) A sequence (x_n) in $D\backslash\{a\}$ which converges to a gets arbitrarily close to a (since $x_n \to a$), but never *equals* a (since it takes values in $D\backslash\{a\}$). So the value of f at a, if this exists, is irrelevant.

An important fact about limits is that, if they exist, they're unique:

Proposition 7.1.3 *If* $\lim_{x\to a} f(x)$ *exists, then it is unique.*

Proof: Let $f : D \to \mathbb{R}$ and assume, to the contrary, that K and L are non-equal real numbers satisfying the definition of $\lim_{x\to a} f(x)$. Let (x_n) be a sequence in $D\backslash\{a\}$ such that $x_n \to a$. Then $f(x_n) \to K$ and $f(x_n) \to L$ and $f(x_n) \to L \neq K$. But this contradicts the uniqueness of limits. $\square$

So it makes sense to speak of *the* limit of a function at a.

Example 7.1.4 *(i) Let* $f : \mathbb{R}\backslash\{1\} \to \mathbb{R}$ *such that*

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

Let x_n *be any sequence in* $\mathbb{R}\backslash\{1\}$ *converging to* 1*. Then*

$$f(x_n) = \frac{(x_n - 1)(x_n + 1)}{x_n - 1} = x_n + 1 \to 2.$$

So $\lim_{x\to 1} f(x) = 2$*, as you would expect.*

(ii) Let $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ *such that*

$$f(x) = \frac{x}{|x|}.$$

Then $\lim_{x\to 0} f(x)$ *does not exist. To prove this, note that* $x_n = (-1)^n/n$ *is a sequence converging to* 0*, but*

$$f(x_n) = \frac{(-1)^n/n}{1/n} = (-1)^n$$

which does not converge at all.

(iii) Let $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ *such that*

$$f(x) = \sin\frac{1}{x}.$$

We can prove that $\lim_{x\to 0} f(x)$ *does not exist using a similar argument. Assume that* $\lim_{x\to 0} f(x)$ *exists, and call it* L*. Let*

$$x_n = \frac{1}{\pi n}.$$

Then (x_n) *is a sequence in* $D\backslash\{0\} = D$ *converging to* 0*, so* $f(x_n) \to L$*. But*

$$f(x_n) = \sin \pi n = 0 \to 0,$$

so $L = 0$*. Let*

$$y_n = \frac{1}{\pi(2n + \frac{1}{2})}.$$

Again, this is a sequence in $D\backslash\{0\}$ *converging to* 0*, so* $f(y_n) \to L = 0$*. But*

$$f(y_n) = \sin\left(\frac{\pi}{2} + 2n\pi\right) = 1 \to 1,$$

a contradiction. □

Having defined limits purely in terms of convergence of sequences, it is very easy to establish many of their basic properties, since these follow directly from theorems about sequences which we've already proved. We've already seen an example of this: limits of functions are unique (Proposition 7.1.3) because limits of convergent sequences are unique. Another very useful example is the Algebra of Limits for functions:

Theorem 7.1.5 (The Algebra of Limits for functions) *Let* $f : D \to \mathbb{R}$*,* $g : D \to \mathbb{R}$*,* $\lim_{x\to a} f(x) = L$ *and* $\lim_{x\to a} g(x) = K$*. Then*

(i) $\displaystyle\lim_{x\to a}(f(x) + g(x)) = L + K$*,*

(ii) $\displaystyle\lim_{x\to a} f(x)g(x) = LK$*,*

and if, in addition, $f(x) \neq 0$ *for all* $x \in D\backslash\{a\}$ *and* $L \neq 0$*, then*

(iii) $\displaystyle\lim_{x\to a} \frac{1}{f(x)} = \frac{1}{L}$*.*

Proof: Let (x_n) be any sequence in $D\backslash\{a\}$ such that $x_n \to a$. Then $f(x_n) \to L$ and $g(x_n) \to K$ by assumption. Hence

(i) $f(x_n) + g(x_n) \to L + K$.

(ii) $f(x_n)g(x_n) \to LK$.

(iii) $f(x_n) \neq 0$ and $L \neq 0$ so $1/f(x_n) \to 1/L$, using limit algebra. $\square$

You stated and used this theorem in MATH1010, but didn't prove it.[1]

7.2 An alternative definition for limits of functions

The above definition of a limit of a function at a point uses sequences. Many textbooks use an equivalent definition that manages without sequences but merely mimics the definition for sequences. We state this here as a theorem.

Theorem 7.2.1 (Alternative definition for limits of functions) *Suppose that a is a limit point of D and $f : D \to \mathbb{R}$ is a function. Then $\lim_{x\to a} f(x) = L$ if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in D$ we have*

$$0 < |x - a| < \delta \implies |f(x) - L| < \epsilon.$$

Proof: As usual we prove the logical equivalence by showing the two implications $\Rightarrow$ and $\Leftarrow$ separately. In order to show that $\lim_{x\to a} f(x) = L$ implies the above mentioned $\epsilon - \delta-$property we use the contrapositive. That is suppose that $\epsilon - \delta-$property does not hold. Then there exists an $\epsilon > 0$ such that for every $\delta > 0$ we can find $x \in D$ with $0 < |x - a| < \delta$, but $|f(x) - L| \geq \epsilon$. For that $\epsilon > 0$ and every $n \in \mathbb{N}$ choose $\delta = \frac{1}{n}$ and the corresponding x_n such that $0 < |x_n - a| < \frac{1}{n}$ and $|f(x_n) - L| \geq \epsilon$. This means that (x_n) is a sequence in $D\backslash\{a\}$ that converges to a, but $f(x_n)$ does not converge to L. Hence it is not true that $\lim_{x\to a} f(x) = L$.

Now suppose that the $\epsilon - \delta-$property holds. Let (x_n) be any sequence in $D\backslash\{a\}$ that converges to a. We need to show that $f(x_n) \to L$. Thus, given $\epsilon > 0$ we need to show that there exists $N \in \mathbb{N}$ such that for all $n > N$ we have $|f(x_n) - L| < \epsilon$. Given $\epsilon > 0$ we know we can find $\delta > 0$ such that for all $x \in D$ we have

$$0 < |x - a| < \delta \implies |f(x) - L| < \epsilon.$$

Since $x_n \to a$ there exists $N \in \mathbb{N}$ such that $n > N$ implies $|x_n - a| < \delta$, and since x_n never equals a we also have $0 < |x_n - a| < \delta$. Therefore, for $n > N$ we have $|f(x_n) - L| < \epsilon$. $\square$

Combining this with Theorem 5.1.3 (the alternative definition of continuity) one obtains a characterisation of continuity in terms of limits.

Corollary 7.2.2 *Let $D \subseteq \mathbb{R}$ and $a \in D$. Then a function $f : D \to \mathbb{R}$ is continuous at a if and only if either a is not a limit point or the limit $\lim_{x\to a} f(x)$ exists and equals $f(a)$.*

[1] For good reason: it's impossible to prove a theorem about limits if you haven't defined what limits are!

7.3 Limits at infinity

Another important type of limit which you studied informally in MATH1010 was $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$. Again, we can make these precise using sequences. First, we have to define what it means for a sequence to "diverge to infinity" (or minus infinity).

Definition 7.3.1 *A real sequence (a_n)* **diverges to infinity** *if, for each $K \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $a_n > K$ for all $n \geq N$. In this case we write $a_n \to \infty$. Similarly, (a_n)* **diverges to minus infinity** *if, for each $K \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $a_n < K$ for all $n \geq N$. In this case we write $a_n \to -\infty$.*

Remarks

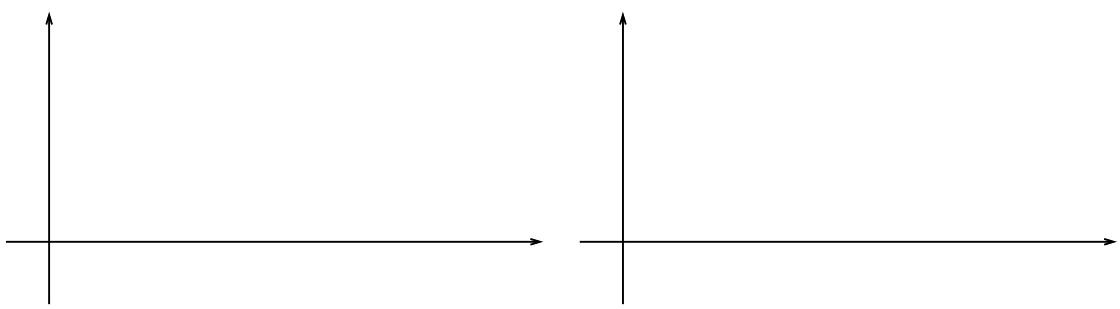
(i) Having learned about quantifiers, we can now give a symbolic version of these definitions:

$$a_n \to \infty \quad \Leftrightarrow \quad \forall K \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N} \cap [N, \infty), a_n > K$$

$$a_n \to -\infty \quad \Leftrightarrow \quad \forall K \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n \in \mathbb{N} \cap [N, \infty), a_n < K$$

(ii) We can visualize these definitions geometrically, in the same way as we did for the definition of convergence. Imagine graphing the sequence (a_n) with n along the x-axis and a_n along the y-axis. Assume $a_n \to \infty$. Then, given any $K \in \mathbb{R}$, draw the horizontal line $y = K$. The definition says that there is a point in the sequence (the Nth term) to the right of which all points on the graph lie above the line $y = K$. If we choose a different, larger K, the corresponding N may need to be larger too. But no matter how big we make K, there is a point to the right of which all points on the graph lie above the line $y = K$.

(iii) It is easy to see that $a_n \to -\infty$ if and only if $-a_n \to \infty$.



Example 7.3.2 *Let $a_n = n$. Then $a_n \to \infty$ by the Archimedean Principle: given any $K \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that $N > K$, and then for all $n \geq N$,*

$a_n = n \geq N > K$.

Warning! It's clear (hopefully) that if $a_n \to \infty$ then (a_n) is unbounded above. (Given any $K \in \mathbb{R}$, there exists $N \in \mathbb{N}$ such that $a_N > K$.) However the converse is false: not every sequence which is unbounded above diverges to infinity.

Example 7.3.3 *Let $a_n = (-1)^n n$. This is certainly unbounded above. But it does not diverge to infinity because, for example, there is no N such that $a_n > 0$ for all $n \geq N$ (because there are always odd numbers n bigger than N, and $a_n < 0$ whenever n is odd). By similar reasoning, this sequence does* not *diverge to minus infinity, although it is unbounded below.*

Example 7.3.4 *$a_n = (1 + (-1)^n)n$ is unbounded above, but the subsequence $a_{2k+1} = 0$ so $a_n \not\to \infty$.*

Proposition 7.3.5 (Sweep Rule) *If $a_n \to \infty$ and $b_n \geq a_n$ for all n then $b_n \to \infty$.*

Proof: Let $K \in \mathbb{R}$ be given. Since $a_n \to \infty$, there exists $N \in \mathbb{N}$ such that $a_n > K$ for all $n \geq N$. Hence, for all $n \geq N$, $b_n \geq a_n > K$. Hence $b_n \to \infty$. $\square$

We can also prove theorems about sequences which diverge to infinity which look rather similar to the Algebra of Limits. Here's a couple of examples:

Proposition 7.3.6 *If $a_n \to \infty$ and (b_n) is bounded, then $a_n + b_n \to \infty$.*

Proof: Let $K \in \mathbb{R}$ be given. We must show that there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $a_n + b_n > K$. Since (b_n) is bounded there exists $L \in \mathbb{R}$ with $|b_n| \leq L$, and thus $b_n \geq -L$. Since $a_n \to \infty$ there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $a_n > K + L$. Then, for $n \geq N$ we have

$$a_n + b_n > K + L + b_n > K + L - L = K.$$

$\square$

Proposition 7.3.7 *If $a_n \to \infty$ and $a_n \neq 0$, then $\frac{1}{a_n} \to 0$.*

Proof: Given $\epsilon > 0$ find $N \in \mathbb{N}$ such that $n > N$ implies that $a_n > 1/\epsilon$. Then, for $n > N$ we have $|1/a_n| = 1/a_n < \epsilon$. $\square$

Returning to the subject of limits of *functions*, we can use a similar characterisation for limits to infinity.

Definition 7.3.8 *Let $f : D \to \mathbb{R}$ where D is unbounded above. Then we say that $\lim_{x\to\infty} f(x) = L$ if, given any sequence $x_n \in D$ which diverges to ∞, $f(x_n) \to L$.*
Similarly, let $f : D \to \mathbb{R}$ where D is unbounded below. Then we say that $\lim_{x\to-\infty} f(x) = L$ if, given any sequence $x_n \in D$ which diverges to $-\infty$, $f(x_n) \to L$.

Now in the same way as before there is an alternative definiton that works without sequences and that sometimes is easier to handle.

Theorem 7.3.9 *Let $f : D \to \mathbb{R}$ where D is unbounded above. Then $\lim_{x\to\infty} f(x) = L$ if and only if for every $\epsilon > 0$ there exists $K \in \mathbb{R}$ such that for all $x \in D$ with $x > K$ we have $|f(x) - L| < \epsilon$.*

Proof: The proof is analoguous to the one of Theorems 7.2.1 and 5.1.3 and is left as an exercise. □

Example 7.3.10 *Claim:* $\displaystyle\lim_{x\to\infty} \frac{1}{x} = 0$.

Proof: *Let $x_n \to \infty$. We must prove that $f(x_n) = 1/x_n \to 0$. So, let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that $x_n > 1/\varepsilon$ for all $n \geq N$ (since $x_n \to \infty$). Hence, for all $n \geq N$,*

$$0 < \frac{1}{x_n} < \varepsilon,$$

so $|f(x_n) - 0| < \varepsilon$. Hence $f(x_n) \to 0$. □

Example 7.3.11 *Claim:* $\displaystyle\lim_{x\to-\infty} \frac{x+1}{x-1} = 1$

Proof: Let $x_n \to -\infty$. We must prove that $f(x_n) = (x_n+1)/(x_n-1) \to 1$. So, let $\varepsilon > 0$ be given. Then, since $x_n \to -\infty$, there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $x_n < -2/\varepsilon$. Hence, for all $n \geq N$,

$$\begin{aligned}
|f(x_n) - 1| &= \left|\frac{x_n+1}{x_n-1} - 1\right| = \frac{2}{|x_n-1|} \\
&< \frac{2}{-x_n+1} \qquad \text{(since } x_n < 1\text{, so } x_n - 1 < 0\text{)} \\
&< \frac{2}{-x_n} \\
&< \varepsilon
\end{aligned}$$

□

Continuous functions $f : \mathbb{R} \to \mathbb{R}$ which have limits at both plus and minus infinity are, in some ways, analogous to continuous functions on closed bounded intervals $f : [a,b] \to \mathbb{R}$. In particular, we can prove something rather similar to the Extreme Value Theorem (Theorem 5.4.4).

Theorem 7.3.12 (Boundedness of continuous functions with limits at $\pm\infty$) *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and have limits at plus and minus infinity. Then f is bounded (above and below).*

Proof: Let $L = \lim_{x\to\infty} f(x)$ and $K = \lim_{x\to-\infty} f(x)$. Then, because of Theorem 7.3.9 we can find $R_1 > 0$ such that for $x > R_1$ we have $|f(x) - L| < 1$. This implies that for $x > R_1$ we have $|f(x)| < |L|+1$. Similarly, we can find $R_2 > 0$ such that for $x < -R_2$ we have $|f(x)| < |K| + 1$. Here one can also use Theorem 7.3.9 directly bearing in mind that $\lim_{x\to-\infty} f(x) = \lim_{x\to\infty} f(-x)$. Since f is continuous on the finite interval $[-R_2, R_1]$ it is bounded there by some constant M_1. Summarising we have with $M := \sup\{|L|+1, |K|+1, M_1\}$ that $|f(x)| \leq M$ for all $x \in \mathbb{R}$. □

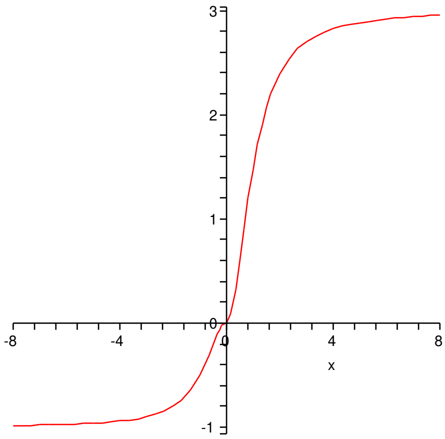
Example 7.3.13 *Consider the function*

$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = \frac{x^2 + 2|x|x}{x^2+1}.$$

This is continuous and has limits

$$\lim_{x\to\infty} f(x) = 3, \qquad \lim_{x\to-\infty} f(x) = -1.$$

Hence, by Theorem 7.3.12, f is bounded above and below. Here is a picture of its graph:



Note that this function attains neither a maximum nor a minimum value. □

Hey! You now know what limits are!

It's worthwhile to pause and reflect on what we've achieved so far. We now have a completely precise and rigorous definition of $\lim_{x\to a} f(x)$ for functions $f : D \to \mathbb{R}$, where D is a subset of $\mathbb{R}$. It follows, that we now know, with no hand-waving or weasel words, precisely what the limit

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

means, that is, we have put the definition of the *derivative* of a function on a proper mathematical footing. Note that this is not a limit of $f : D \to \mathbb{R}$, but of the associated function

$$h : D\backslash\{a\} \to \mathbb{R}, \qquad h(x) = \frac{f(x) - f(a)}{x - a}.$$

The interesting, and somewhat curious, thing about this is that we managed to do this by studying *sequences* and, in particular, making precise what it means for sequences to converge. So from the rather modest and unassuming definition of convergence of sequences something very large and powerful has grown.

Chapter 8

Complex sequences and series

8.1 Convergence of complex sequences

Recall that a complex number z is an expression of the form $z = x + iy$ where x, y are real and $i^2 = -1$. We call x the **real part** of z and y (not iy!) the **imaginary part** of z, and use $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ as shorthand for these. So, for example

$$\mathrm{Re}\,(2 - 3i) = 2, \qquad \mathrm{Im}\,(2 - 3i) = -3$$

We can add, subtract and multiply complex numbers in the obvious way, and even divide them, provided the denominator is not 0 (this, is not so obvious). So the set of complex numbers $\mathbb{C}$ is a *field* (though not an *ordered* field).

We define the **modulus** of a complex number $z = x + iy$ to be

$$|z| = \sqrt{x^2 + y^2},$$

so

$$|2 - 3i| = \sqrt{4 + 9} = \sqrt{13}$$

for example. Note that this non-negative real number exists, and is unique (see Proposition 5.3.6), so we know (rigorously) that this definition makes sense. If we visualize a complex number $z = x + iy$ as a point (x, y) in the Cartesian plane, then $|z|$ is the length of the straight line segment joining z to 0 (apply Pythagoras' Theorem).

Similarly, $|z - w|$ is the length of the line segment joining z and w, that is, the distance between z and w. Check: let $z = x + iy$, $w = u + iv$. Then

$$|z - w| = |x + iy - (u + iv)| = |(x - u) + i(y - v)| = \sqrt{(x - u)^2 + (y - v)^2}.$$

It's often useful to note that $|z|^2 = \overline{z}z$, where $\overline{z} = \overline{x + iy} = x - iy$ is the **complex conjugate** of z. Check:

$$\overline{z}z = (x - iy)(x + iy) = x^2 - (iy)^2 = x^2 + y^2 = |z|^2.$$

Note that $|\overline{z}| = |z|$. We will use several basic properties of the modulus which (I hope) you proved in MATH1025. For reference, let's collect them together:

Proposition 8.1.1 *For all complex numbers z and w,*

(i) $|zw| = |z||w|$

(ii) $|z| \geq \operatorname{Re} z$

(iii) $|z| \geq \operatorname{Im} z$

(iv) $|z + w| \leq |z| + |w|$.

The last of these is often called the Triangle Inequality.

Proof: (i), (ii), (iii) I leave as an exercise.
(iv) Assume, to the contrary, that $|z + w| > |z| + |w|$. Then

$$\begin{array}{rrcl} & |z+w|^2 & > & (|z|+|w|)^2 \\ \Rightarrow & (z+w)(\overline{z}+\overline{w}) & > & |z|^2 + 2|z||w| + |w|^2 \\ \Rightarrow & z\overline{w} + \overline{z}w & > & 2|z||w| = 2|z||\overline{w}| = 2|z\overline{w}| \\ \Rightarrow & 2\mathrm{Re}\,(z\overline{w}) & > & 2|z\overline{w}| \end{array}$$

which contradicts part (ii). $\square$

We use the same symbol, $|\cdot|$, to denote modulus (of a complex number) and absolute value (of a real number), because modulus is a strict generalization of absolute value. That is, any *real* number x is also a *complex* number $x + i0$ whose modulus is

$$|x + i0| = \sqrt{x^2 + 0^2} = \sqrt{x^2} = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

which is precisely equal to the absolute value of x.

Our aim in this section is to generalize the ideas of convergence and limit which we've developed for real sequences and series to complex sequences and series. The definition of sequence itself immediately generalizes: a complex sequence is a mapping $a : \mathbb{N} \to \mathbb{C}$, that is, a rule which assigns to each positive integer n, some complex number $a(n)$. As with real sequences, we denote the number associated to n by a_n (instead of $a(n)$) and call it the n^{th} term of the sequence. Having recognized that modulus is a generalization of absolute value, and $|z - w|$ measures the distance between z and w, it's pretty clear how to modify our definition of convergence to deal with complex sequences:

Definition 8.1.2 (Convergence of a complex sequence) *Let (z_n) be a complex sequence. Then (z_n)* **converges** *to a complex number L if, for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that, for all $n \geq N$, $|z_n - L| < \varepsilon$. In this case we write $z_n \to L$ for short.*

Remarks

- The definition is almost identical to the one for real sequences. Note, however, that in Definition 8.1.2, $|z_n - L|$ means the *modulus* of the *complex* number $z_n - L$.
- The set of points in $z \in \mathbb{C}$ satisfying $|z - L| < \varepsilon$ is a disk of radius ε centred on L. So the definition says that, given any radius $\varepsilon > 0$, no matter how small, there's a point in the sequence (call it the N^{th} term) beyond which all terms lie in the disk of radius ε centred on L. If we make ε smaller, the disk shrinks and we may have to increase N. But however small ε is, the sequence "eventually" gets stuck entirely in the disk of radius ε.

Example 8.1.3 *Claim* $z_n = \dfrac{n+i}{n-i} \to 1$

Proof: *Let* $\varepsilon > 0$ *be given. Then, by the Archimedean Property of* $\mathbb{R}$*, there exists* $N \in \mathbb{N}$ *such that* $N > 2/\varepsilon$*. Now for all* $n \geq N$*,*

$$\begin{aligned}
|z_n - 1| &= \left|\frac{n+i}{n-i} - 1\right| \\
&= \left|\frac{n+i-(n-i)}{n-i}\right| \\
&= \frac{|2i|}{|n-i|} \\
&= \frac{2}{\sqrt{n^2+1}} \\
&< \frac{2}{n} \\
&\leq \frac{2}{N} \\
&< \varepsilon.
\end{aligned}$$

□

We can reduce the question of whether a complex sequence converges to a question involving only *real* sequences in two different ways:

Proposition 8.1.4 *Let* $z_n = x_n + iy_n$ *and* $L = A + iB$*, where* x_n, y_n, A, B *are real. Then the following statements are equivalent:*

(i) $z_n \to L$;

(ii) $x_n \to A$ *and* $y_n \to B$;

(iii) $r_n = |z_n - L| \to 0$.

Proof: (i) $\Rightarrow$ (ii): Assume $z_n \to L$. We must show that $x_n \to A$ and $y_n \to B$. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|z_n - L| < \varepsilon$. But

$$\begin{aligned} |z_n - L| &= |(x_n - A) + i(y_n - B)| = \sqrt{(x_n - A)^2 + (y_n - B)^2} \\ &\geq \max\{\sqrt{(x_n - A)^2}, \sqrt{(y_n - B)^2}\} \qquad \text{(since the function } f(x) = \sqrt{x} \text{ is increasing)} \\ &= \max\{|x_n - A|, |y_n - B|\} \end{aligned}$$

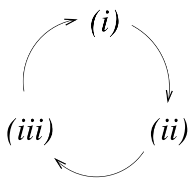
Hence, for all $n \geq N$,

$$|x_n - A| \leq |z_n - L| < \varepsilon, \qquad \text{and} \qquad |y_n - B| \leq |z_n - L| < \varepsilon$$

so $x_n \to A$ and $y_n \to B$.
(ii) $\Rightarrow$ (iii): Assume $x_n \to A$ and $y_n \to B$. We must show that $r_n = |z_n - L| \to 0$. So, let $\varepsilon > 0$ be given. Then, since $x_n \to A$, there exists $N_1 \in \mathbb{N}$ such that, for all $n \geq N_1$, $|x_n - A| < \varepsilon/2$. Similarly, since $y_n \to B$, there exists $N_2 \in \mathbb{N}$ such that for all $n \geq N_2$, $|y_n - B| < \varepsilon/2$. Hence, for all $n \geq \max\{N_1, N_2\}$,

$$\begin{aligned} |r_n - 0| &= r_n = |(x_n - A) + i(y_n - B)| \leq |x_n - A| + |y_n - B| \qquad \text{(Triangle inequality)} \\ &< \frac{\varepsilon}{2} + |y_n - B| \qquad \text{(since } n \geq N_1\text{)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \qquad \text{(since } n \geq N_2\text{)} \\ &= \varepsilon. \end{aligned}$$

Hence $r_n \to 0$.
(iii) $\Rightarrow$ (i): Assume $r_n \to 0$. We must show that $z_n \to L$. So, let $\varepsilon > 0$ be given. then, since $r_n \to 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|r_n - 0| < \varepsilon$. Hence, for all $n \geq N$, $|z_n - L| = r_n = |r_n| < \varepsilon$. Hence $z_n \to L$.
Having proved the circle of implications



all the other implications immediately follow (e.g. (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i), so (ii) $\Rightarrow$ (i)).
$\square$

Proposition 8.1.4 has many important immediate consequences:

- The limit of a convergent complex sequence is unique.
- Every convergent complex sequence (z_n) is bounded, meaning there exists $K > 0$ such that $|z_n| \leq K$ for all n.
- If $z_n \to Z \in \mathbb{C}$ and $w_n \to W \in \mathbb{C}$ then $z_n + w_n \to Z + W$, $z_n w_n \to ZW$, and, if $w_n \neq 0$ and $W \neq 0$, $z_n/w_n \to Z/W$.

Example 8.1.5 *Claim: $z_n = \alpha^n \to 0$ for any constant $\alpha \in \mathbb{C}$, with $|\alpha| < 1$.*

Proof:

$$|z_n - 0| = |\alpha^n| = |\alpha|^n \to 0,$$

so $z_n \to 0$ by Proposition 8.1.4. □

Here's an interesting question: what's the behaviour of $z_n = \alpha^n$ when $|\alpha| \geq 1$? Certainly it doesn't converge to 0 (since $|z_n - 0| = |\alpha|^n \geq |\alpha| \geq 1$). □

Example 8.1.6 (revisited) *Claim: $z_n = \dfrac{n+i}{n-i} \to 1$*
Proof 1:

$$0 \leq |z_n - 1| = \left|\frac{n + i - (n-1)}{n+i}\right| = \frac{2}{|n+i|} = \frac{2}{\sqrt{n^2+1}} < \frac{2}{n},$$

so $|z_n - 1| \to 0$ by the Squeeze Rule. Hence $z_n \to 1$ by Proposition 8.1.4. □
Proof 2:

$$z_n = \frac{n+i}{n-i} \times \frac{n+i}{n+i} = \frac{n^2 - 1 + 2in}{n^2+1} = x_n + iy_n$$

where $x_n = \frac{n^2-1}{n^2+1}$ and $y_n = \frac{2n}{n^2+1}$. Now

$$x_n = \frac{1 - 1/n^2}{1 + 1/n^2} \to 1$$

and

$$y_n = \frac{2/n}{1 + 1/n^2} \to 0,$$

so $z_n \to 1 + i0 = 1$ by Proposition 8.1.4. □

Just as for real sequences, we can define complex sequences inductively, by iterating a function $f : \mathbb{C} \to \mathbb{C}$. That is, we take $z_0 = \alpha \in \mathbb{C}$, some "starting value", and define z_n for all $n \geq 1$ by the rule

$$z_n = f(z_{n-1}).$$

Example 8.1.7 *Let $f : \mathbb{C} \to \mathbb{C}$ be the function $f(z) = z^2$ and $z_0 = i$, the inductive* *sequence defined by f is*

$$\begin{aligned} z_1 &= f(i) = -1, \\ z_2 &= f(-1) = 1 \\ z_3 &= f(1) = 1, \ldots \end{aligned}$$

so $z_n = 1$ for all $n \geq 2$. Clearly this converges to 1. What if $z = 2i$? Then

$$\begin{aligned} z_1 &= f(2i) = -4, \\ z_2 &= f(-4) = 16, \\ z_3 &= f(16) = 256, \ldots \end{aligned}$$

In fact z_n is unbounded, so z_n cannot converge to anything.

We can define continuity for functions $f : \mathbb{C} \to \mathbb{C}$ just as we did for real functions:

Definition 8.1.8 (Continuity of a complex function) *A function $f : \mathbb{C} \to \mathbb{C}$ is* **continuous at** $w \in \mathbb{C}$ *if, for all complex sequences (z_n) such that $z_n \to w$, $f(z_n) \to f(w)$. The function is* **continuous** *if it is continuous at w for all $w \in \mathbb{C}$.*

It follows from Proposition 8.1.4 and limit calculus that all complex polynomial functions, for example, are continuous.

Proposition 8.1.9 *Let $f : \mathbb{C} \to \mathbb{C}$ be continuous and (z_n) be the inductive sequence defined by iterating f on some starting value $z_0 \in \mathbb{C}$ (so $z_n = f(z_{n-1})$). If $z_n \to L$, then $f(L) = L$, that is, L is a fixed point of f.*

Proof: Assume $z_n \to L$. Then, by Proposition 8.1.4 and Theorem 3.1.3, $z_{n+1} \to L$ also. But $z_{n+1} = f(z_n)$ and f is continuous (at L) so $f(z_n) \to f(L)$. But limits are unique, so $f(L) = L$. $\square$

Example 8.1.10 (revisited) *$f(z) = z^2$, being polynomial, is continuous. Hence, any sequence obtained by iterating f, if it converges, must converge to a fixed point of f. L is a fixed point of f if and only if $f(L) = L$, i.e. $L^2 = L$, that is $L = 0$ or 1. Hence such a sequence cannot converge to any limit other than 0 or 1.*

8.2 Complex series

The definition of convergence for a complex *series* is the obvious generalization of the definition for real series.

Definition 8.2.1 (Complex Series and their convergence) *Given a complex series $\sum_{n=1}^{\infty} a_n$ (so $a_n \in \mathbb{C}$), we define its k^{th}* **partial sum** *to be*

$$s_k = \sum_{n=1}^{k} a_n.$$

We say the series **converges** *to $L \in \mathbb{C}$ if the complex sequence (s_k) converges to L (in the sense of Definition 8.1.2). We say the series* **converges absolutely** *if the* real *series $\sum_{n=1}^{\infty} |a_n|$ converges.*

We developed lots of useful tests for convergence of real series (the Ratio Test, Comparison Test, Divergence Test, Alternating Series Test). By Proposition 8.1.4, a complex series

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} (x_n + iy_n)$$

converges if and only if the pair of real series

$$\sum_{n=1}^{\infty} x_n, \quad \sum_{n=1}^{\infty} y_n$$

both converge.

Example 8.2.2 *Claim: the series $\displaystyle\sum_{n=1}^{\infty} \frac{i^n}{n}$ converges.*

Proof: *Consider the sequence i^n. It cycles through exactly 4 values:*

$$(i^n) = (i, -1, -i, 1, i, -1, -i, 1, \dots)$$

So the real part of $a_n = i^n/n$ is

$$x_n = \begin{cases} 0 & n \text{ odd} \\ \dfrac{(-1)^{n/2}}{n} & n \text{ even} \end{cases}$$

and the imaginary part of a_n is

$$y_n = \begin{cases} \dfrac{(-1)^{(n-1)/2}}{n} & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$$

Hence, the real part of the series is

$$\sum_{n=1}^{\infty} x_n = \sum_{k=1}^{\infty} x_{2k} = \sum_{k=1}^{\infty} \frac{(-1)^k}{2k}$$

which converges by the Alternating Series Test (since $1/(2k)$ is decreasing and converges to 0). Similarly, the imaginary part of the series is

$$\sum_{n=1}^{\infty} y_n = \sum_{k=1}^{\infty} y_{2k-1} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{2k-1}$$

which converges by the Alternating Series Test (since $1/(2k-1)$ is decreasing and converges to 0). Hence, the (complex) series converges by Proposition 8.1.4. □

Note that the series considered in Example 8.2.2 certainly does **not** converge abolutely, since

$$\sum_{n=1}^{\infty} \left| \frac{i^n}{n} \right| = \sum_{i=1}^{\infty} \frac{1}{n}$$

which diverges (see Example 4.1.2). So convergence of a complex series does **not** imply absolute convergence of the series. On the other hand, just as for real series, absolute convergence of a complex series **does** imply convergence.

Proposition 8.2.3 *If a complex series $\sum_{n=1}^{\infty} a_n$ converges absolutely, then it converges.*

Proof: Let $a_n = x_n + iy_n$, and assume that $\sum_{n=1}^{\infty} a_n$ converges absolutely, that is, the real sequence

$$s_k = \sum_{n=1}^{k} |a_n|$$

converges. Then (s_k) is bounded above. Now $|a_n| = \sqrt{x_n^2 + y_n^2} \geq |x_n|$, so

$$u_k := \sum_{n=1}^{k} |x_n| \leq s_k.$$

Hence u_k is also bounded above. But (u_k) is increasing, and hence convergent (by the Monotone Convergence Theorem). Hence the real series $\sum_{n=1}^{\infty} x_n$ converges absolutely, and so converges, by Theorem 4.4.2. Similarly, $|a_n| \geq |y_n|$, so

$$v_k := \sum_{n=1}^{k} |y_n| \leq s_k$$

is bounded above. Since (v_k) is increasing, it converges (by the Monotone Convergence Theorem), so the real series $\sum_{n=1}^{\infty} y_n$ converges absolutely, and hence converges (Theorem 4.4.2). Since $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ converge, $\sum_{n=1}^{\infty} a_n$ converges by Proposition 8.1.4. $\square$

Remark 8.2.4 (a useful notation) *The sequence of partial sums of $\sum_{n=1}^{\infty} |a_n|$ is increasing because*

$$s_{k+1} - s_k = |a_{k+1}| \geq 0.$$

For an increasing sequence there are only two possibilities. Either it is bounded above (and hence convergent), or it diverges to ∞. Therefore, absolute convergence is equivalent to the sequence

$$s_k = \sum_{n=1}^{k} |a_n|$$

being bounded above. It is customary to write this as

$$\sum_{n=1}^{\infty} |a_n| < \infty.$$

Hence, absolute convergence means $\sum_{n=1}^{\infty} |a_n| < \infty$. It is however important to always keep in mind what this means. This notation somewhat simplifies the proof above. For example, if $\sum_{n=1}^{\infty} |a_n| < \infty$ then the series $\sum_{n=1}^{\infty} a_n$ converges absolutely. Indeed, considering real and imaginary parts separately, and using that $|\mathrm{Re}\, z| \leq |z|, |\mathrm{Im}\, z| \leq$

$|z|$, *one has*

$$\sum_{n=1}^{\infty} |\mathrm{Re}\, a_n| \le \sum_{n=1}^{\infty} |a_n| < \infty,$$
$$\sum_{n=1}^{\infty} |\mathrm{Im}\, a_n| \le \sum_{n=1}^{\infty} |a_n| < \infty.$$

Note that this is exactly the same proof as above, but the notation makes it shorter.

When faced with a complex series $\sum_{n=1}^{\infty} a_n$, it is sometimes notationally shorter to show that it converges *absolutely*. The advantage here is that $|a_n| \ge 0$ so the tests for real convergence which rely on having only positive terms (Comparison and Ratio Tests) may work. If the series converges absolutely, then it converges (Proposition 8.2.3). But be careful! Showing that a series does *not* converge absolutely does *not* show that it diverges (see Example 8.2.2 for a counterexample).

Example 8.2.5 *Claim:* $\displaystyle\sum_{n=1}^{\infty} \frac{(3-2i)^n}{n!}$ *converges.*

Proof: *Let* $a_n = \dfrac{(3-2i)^n}{n!}$. *Then*

$$|a_n| = \frac{|3-2i|^n}{n!} = \frac{13^{n/2}}{n!}$$

Hence

$$\frac{|a_{n+1}|}{|a_n|} = \frac{13^{(n+1)/2}}{13^{n/2}} \frac{n!}{(n+1)!} = \frac{\sqrt{13}}{n+1} \to 0 < 1 \qquad ,$$

so the real series $\sum_{n=1}^{\infty} |a_n|$ *converges, by the Ratio Test. Hence* $\sum_{n=1}^{\infty} a_n$ *converges, by Theorem 8.2.3.* $\square$

8.3 Power series

A power series is a complex series of the form

$$\sum_{n=0}^{\infty} a_n z^n$$

where z is interpreted as a complex *variable*. The series depends on the choice of value for $z \in \mathbb{C}$. The terms of the series are $a_n z^n$, and the k^{th} partial sum is

$$s_k = \sum_{n=0}^{k} a_n z^n = a_0 + a_1 z + a_2 z^2 + \cdots + a_k z^k,$$

just as for series in general. Note that, if $z = 0$, then every partial sum is $s_k = a_0$, so the series certainly converges (to a_0) in that case. In general, a power series may converge for some values of z but diverge for others. If the series converges for a particular choice of z, its limit will, in general, depend on z. So a power series defines a complex-valued function on (perhaps only part of) $\mathbb{C}$.

Example 8.3.1 *Consider the series $\sum_{n=0}^{\infty} z^n$. Its k^{th} partial sum is*

$$\begin{aligned}
s_k &= 1 + z + z^2 + \cdots + z^k \\
\Rightarrow \qquad z s_k &= z + z^2 + z^3 + \cdots + z^{k+1} \\
\Rightarrow \qquad (1-z) s_k &= 1 - z^{k+1} \\
\Rightarrow \qquad s_k &= \frac{1 - z^{k+1}}{1 - z}
\end{aligned}$$

If $|z| < 1$ then $|z|^{k+1} \to 0$ (Example 8.1.5) so $s_k \to 1/(1-z)$. Hence, for $|z| < 1$, the series converges to $1/(1-z)$. Note that this argument is almost a verbatim repetition of Example 4.1.5. The only difference is that now the "common ratio" in the geometric series is allowed to be complex.

For $|z| \geq 1$, the terms of the series z^n have $|z^n| \geq 1$. Let $z^n = x_n + iy_n$. Since $x_n^2 + y_n^2 \geq 1$ either x_n or y_n does not converge to 0. But then either $\sum_{n=0}^{\infty} x_n$ of $\sum_{n=0}^{\infty} y_n$ diverges (by the Divergence Test) and so $\sum_{n=0}^{\infty} z^n$ diverges (by Proposition 8.1.4).

In summary, if we use the power series to define a function

$$f(z) = \sum_{n=0}^{\infty} z^n$$

then $f : D \to \mathbb{C}$ where $D = \{z \in \mathbb{C} \,:\, |z| < 1\}$ and for all $z \in D$,

$$f(z) = \frac{1}{1-z}.$$

$\square$

Representing a function by a power series can be an extremely smart move indeed (as you will find out in detail if you choose the level 3 module MATH3017 Calculus in the Complex Plane). But for the time being, let's just note that, for some functions, a definition in terms of power series is really the only definition we have.

Definition 8.3.2 *The* **exponential, sine** *and* **cosine functions** *are defined by*

$$\begin{aligned} \exp : \mathbb{C} \to \mathbb{C}, \qquad \exp(z) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} \\ \sin : \mathbb{C} \to \mathbb{C}, \qquad \sin(z) &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \\ \cos : \mathbb{C} \to \mathbb{C}, \qquad \cos(z) &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} \end{aligned}$$

Of course, we should check that these functions are well-defined, that is, that the power series defining them converge for all $z \in \mathbb{C}$ (unlike the power series for $f(z) = 1/(1-z)$). We'll check $\exp$, and leave $\sin$ and $\cos$ as an exercise:

Example 8.3.3 *Claim:* $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$ *converges (absolutely) for all* $z \in \mathbb{C}$.

Proof: *As we've already noted, all power series converge at* $z = 0$ *(to* $a_0 = 1/0! = 1$*, in this case). So assume* $z \neq 0$*. Let* $b_n = z^n/n!$*. Then* $|b_n| > 0$ *and*

$$\frac{|b_{n+1}|}{|b_n|} = \frac{|z|}{n+1} \to 0 < 1$$

so $\sum_{n=0}^{\infty} |b_n|$ *converges, by the Ratio Test. Hence* $\sum_{n=0}^{\infty} b_n$ *converges (Theorem 8.2.3).* $\square$

So, unlike the power series $\sum_{n=0}^{\infty} z^n$, which converges only for z in the disk of radius 1, the power series defining $\exp$, $\sin$ and $\cos$ converge for all $z \in \mathbb{C}$. On the other hand, it's not hard to come up with power series which converge only at $z = 0$.

Example 8.3.4 *Claim* $\sum_{n=1}^{\infty} n^n z^n$ *diverges for all* $z \neq 0$.

Proof: Let $b_n = n^n z^n$. By the Archimedean Property of $\mathbb{R}$, there exists $N \in \mathbb{N}$ such that $N \geq 2/|z|$. Then for all $n \geq N$

$$
\begin{aligned}
\frac{|b_{n+1}|}{|b_n|} &= \frac{(n+1)^{n+1}}{n^n}\frac{|z|^{n+1}}{|z|^n} \\
&= |z|(n+1)\left(1+\frac{1}{n}\right)^n \\
&> |z|(n+1) \\
&> |z|n \\
&\geq 2.
\end{aligned}
$$

Hence, for all $n \geq N$, $|b_n| \geq |b_N|2^{n-N}$ *(check by induction). So* $|b_n|$ *is unbounded above, and hence either* $\operatorname{Re} b_n$ *or* $\operatorname{Im} b_n$ *does not converge to* 0. *Hence, either* $\sum_{n=1}^{\infty} \operatorname{Re} b_n$ *or* $\sum_{n=1}^{\infty} \operatorname{Im} b_n$ *diverges (by the Divergence Test), so* $\sum_{n=1}^{\infty} b_n$ *diverges by Proposition 8.1.4.* $\square$

Given a power series

$$\sum_{n=0}^{\infty} a_n z^n$$

it's interesting to ask "what is the subset $D \subseteq \mathbb{C}$ of values z for which the series converges?" Since the coefficients a_n could be absolutely any complex numbers you like, it may seem that this set could be arbitrarily complicated. But actually, the examples we've already looked at illustrate (essentially) all possibilities: either

(i) the series converges for all z, or

(ii) it converges only for $z = 0$, or

(iii) there is some constant $R > 0$ such that the series converges for all $|z| < R$ and diverges for all $|z| > R$.

The only subtlety occurs in case (iii): the series may converge for some, none or all values z with $|z| = R$.

Definition 8.3.5 (Radius of Convergence) *The* **radius of convergence** *of a power series* $\sum_{n=0}^{\infty} a_n z^n$ *is*

$$R := \sup\{|z| : \sum_{n=0}^{\infty} |a_n z^n| \text{ converges}\}.$$

Note the set $\{|z| : \sum_{n=0}^{\infty} |a_n z^n| \text{ converges}\} \subseteq \mathbb{R}$ *always contains* 0 *(since* $\sum_{n=0}^{\infty} |a_n z^n|$ *converges to* $|a_0|$ *at* $z = 0$*), so is certainly nonempty. If the set is* *unbounded, we say $R = \infty$.*

Lemma 8.3.6 *Let $\sum_{n=0}^{\infty} a_n z^n$ converge at $z = z_1$. Then it converges* absolutely *for all z such that $|z| < |z_1|$.*

Proof: Since $\sum a_n z_1^n$ converges, so do its real and imaginary parts (Proposition 8.1.4), and hence their sequences of terms, $\mathrm{Re}(a_n z_1^n)$ and $\mathrm{Im}(a_n z_1^n)$ must converge to 0 (by the Divergence Test). Thus $a_n z_1^n \to 0$ (Proposition 8.1.4), and so is bounded: there exists $K > 0$ such that for all n $|a_n z_1^n| < K$. But then for all n,

$$|a_n z^n| = |a_n z_1^n| \frac{|z|^n}{|z_1|^n} < K \left(\frac{|z|}{|z_1|} \right)^n$$

so if $|z| < |z_1|$, $\sum_{n=0}^{\infty} |a_n z^n|$ converges by comparison with the convergent (geometric) series $\sum_{n=0}^{\infty} (|z|/|z_1|)^n$. □

Theorem 8.3.7 *Let $\sum_{n=0}^{\infty} a_n z^n$ have radius of convergence R. Then*

$$\sum_{n=0}^{\infty} a_n z^n \begin{cases} \text{converges } \textbf{absolutely} \text{ for } |z| < R, \\ \text{diverges for } |z| > R. \end{cases}$$

Proof: Let $A := \{|z| : \sum_{n=0}^{\infty} |a_n z^n| \text{ converges}\} \subseteq \mathbb{R}$, so that $R = \sup A$.
If $|z| < R$ then there exists z_1 with $|z| < |z_1| < R$ such that $\sum_{n=0}^{\infty} a_n z_1^n$ is absolutely convergent (else R isn't the *least* upper bound on A, $|z|$ being smaller), hence convergent, so $\sum_{n=0}^{\infty} a_n z^n$ is absolutely convergent by Lemma 8.3.6.
If $|z| > R$ and $\sum_{n=0}^{\infty} a_n z^n$ *converges* then $\sum_{n=0}^{\infty} a_n z_2^n$, where $z_2 = (|z| + R)/2$, converges absolutely by Lemma 8.3.6, since $|z_2| < |z|$. But $|z_2| > R$, a contradiction (R isn't an upper bound on A at all!), so we conclude that $\sum_{n=0}^{\infty} a_n z^n$ *diverges*. □

So, given a power series, if we can figure out its radius of convergence R, we know immediately that it converges absolutely on the disk

$$D = \{z \in \mathbb{C} \; : \; |z| < R\}$$

and diverges for all $|z| > R$. This is very useful information, and it's important to be able to compute radii of convergence. Luckily, this is usually possible by a simple application of the ratio test.

Example 8.3.8 *(i) What is the radius of convergence of $\displaystyle\sum_{n=0}^{\infty} \frac{n^2}{n+1} z^{2n}$?*

Solution: Let $b_n = \dfrac{n^2}{n+1} z^{2n}$. Then, for all $z \neq 0$ and $n > 0$, $|b_n| > 0$, so we *can (try to) use the Ratio Test to see whether $\sum |b_n|$ converges. Now*

$$\begin{aligned}\frac{|b_{n+1}|}{|b_n|} &= \frac{(n+1)^2}{n+2}|z|^{2n+2}\frac{n+1}{n^2|z|^{2n}} \\ &= \frac{(n+1)^3}{n^2(n+2)}|z|^2 \\ &= \frac{(1+(1/n))^3}{1+(2/n)}|z|^2 \\ &\to |z|^2\end{aligned}$$

So if $|z| < 1$, $|b_{n+1}|/|b_n|$ converges to a limit less than 1, so $\sum |b_n|$ converges (by the Ratio test), that is $\sum b_n$ converges absolutely. On the other hand, for $|z| > 1$, $\sum b_n$ does not *converge absolutely (again, by the Ratio Test). Comparing with Definition 8.3.5, we deduce that $R = 1$.*

(ii) What is the radius of convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{n^3}{8^n} z^{3n-1}$?

Solution: Let $b_n = \dfrac{n^3}{8^n} z^{3n-1}$. Then, for all $z \neq 0$, $|b_n| > 0$, so we can (try to) use the Ratio Test to see whether $\sum |b_n|$ converges. Now

$$\begin{aligned}\frac{|b_{n+1}|}{|b_n|} &= \frac{(n+1)^3}{n^3}\frac{8^n}{8^{n+1}}|z|^3 \\ &= \left(1+\frac{1}{n}\right)^3 \frac{1}{8}|z|^3 \\ &\to \frac{|z|^3}{8}\end{aligned}$$

So by the Ratio Test, $\sum b_n$ converges absolutely if $|z| < 2$, but not if $|z| > 2$. Comparing with Definition 8.3.5, we deduce that $R = 2$.

Warning! Example 8.3.8 describes a method of *finding* the radius of convergence, not the *definition* of the radius of convergence. Some people may try and tell you that the radius of convergence of a power series $\sum_{n=0}^{\infty} a_n z^n$ is $R = \lim_{n\to\infty} |a_n|/|a_{n+1}|$, a formula which results from (blindly) applying the method of Example 8.3.8 to a "general" power series, without thinking carefully about whether it works. Have no truck with such charlatans! Note that, for both the power series we just considered, the formula generates gibberish:

$$\sum_{n=0}^{\infty} \frac{n^2}{n+1} z^{2n} \qquad \text{has} \qquad a_n = \begin{cases} 0 & n \text{ odd} \\ \frac{(n/2)^2}{n/2+1} & n \text{ even} \end{cases}$$

so for all even n, $|a_n|/|a_{n+1}|$ is undefined, and it is meaningless to speak of the limit of this sequence. Similarly

$$\sum_{n=0}^{\infty} \frac{n^3}{8^n} z^{3n-1} \qquad \text{has} \qquad a_n = \begin{cases} 0 & n+1 \text{ not divisible by } 3 \\ \frac{((n+1)/3)^3}{8^{(}(n+1)/3)} & n+1 \text{ divisible by } 3 \end{cases}$$

so again the sequence $|a_n|/|a_{n+1}|$ is undefined. But of course, both these power series had perfectly well-defined radii of convergence, and the method of Example 8.3.8, applied carefully, allowed us to compute them. The moral of the story is this: don't confuse the *definition* of a mathematical object with a *method* used to find that mathematical object. The definition of radius of convergence is Definition 8.3.5 and it makes no mention of the sequence $|a_n|/|a_{n+1}|$, or the Ratio Test.

Remark If all the coefficients a_n in a power series $\sum_{n=0}^{\infty} a_n z^n$ are real, then the series itself is real whenever z is real. In this case, the power series defines a *real* function $f : (-R, R) \to \mathbb{R}$, where R is its radius of convergence. For example, the series defining $\exp$, $\sin$ and $\cos$ have only real a_n's, and each has radius of convergence $R = \infty$, so they define real functions

$$\exp : \mathbb{R} \to \mathbb{R}, \qquad \sin : \mathbb{R} \to \mathbb{R}, \qquad \cos : \mathbb{R} \to \mathbb{R}.$$

Of course, these coincide with the real functions of the same names with which you have long been familiar.