Title

Github/sj-bs

Contents

§0 Acknowledgements

Thanks to Evan for creating the original stylesheet which I have since modified extensively to suit my needs and preferences. You can find the original preamble file here, and my rewritten version here.

Finally, I used an extract of Dexta Chua's Cambridge lecture notes as the basis to showcase the preamble. Note that since I did just copy and paste their TeX with minimal modification, the product may be suboptimal in parts.

§1 Introduction

In IA Differential Equations, we studied calculus in a non-rigorous setting. While we did define differentiation (properly) as a limit, we did not define what a limit was. We didn't even have a proper definition of integral, and just mumbled something about it being an infinite sum.

In Analysis, one of our main goals is to put calculus on a rigorous foundation. We will provide unambiguous definitions of what it means to take a limit, a derivative and an integral. Based on these definitions, we will prove the results we've had such as the product rule, quotient rule and chain rule. We will also rigorously prove the fundamental theorem of calculus, which states that the integral is the inverse operation to the derivative.

However, this is not all Analysis is about. We will study all sorts of limiting ("infinite") processes. We can see integration as an infinite sum, and differentiation as dividing two infinitesimal quantities. In Analysis, we will also study infinite series such as $1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots$, as well as limits of infinite sequences.

Another important concept in Analysis is *continuous functions*. In some sense, continuous functions are functions that preserve limit processes. While their role in this course is just being "nice" functions to work with, they will be given great importance when we study metric and topological spaces.

This course is continued in IB Analysis II, where we study uniform convergence (a stronger condition for convergence), calculus of multiple variables and metric spaces.

§2 The real number system

We are all familiar with real numbers — they are "decimals" consisting of infinitely many digits. When we really want to study real numbers, while this definition is technically valid, it is not a convenient definition to work with. Instead, we take an *axiomatic* approach. We define the real numbers to be a set that satisfies certain properties, and, if we really want to, show that the decimals satisfy these properties. In particular, we define real numbers to be an ordered field with the least upper bound property.

We'll now define what it means to be an "ordered field with the least upper bound property".

Definition 2.1 (Field)

A *field* is a set $\mathbb{F}$ with two binary operations $+$ and $\times$ that satisfies all the familiar properties satisfied by addition and multiplication in $\mathbb{Q}$, namely

1. $(\forall a, b, c)\, a + (b + c) = (a + b) + c$ and $a \times (b \times c) = (a \times b) \times c$ (associativity)

2. $(\forall a, b)\, a + b = b + a$ and $a \times b = b \times a$ (commutativity)

3. $\exists 0, 1$ such that $(\forall a) a + 0 = a$ and $a \times 1 = a$ (identity)

4. $\forall a\, \exists (-a)$ such that $a + (-a) = 0$. If $a \neq 0$, then $\exists a^{-1}$ such that $a \times a^{-1} = 1$.   (inverses)

5. $(\forall a, b, c)\, a \times (b + c) = (a \times b) + (a \times c)$ (distributivity)

Example 2.1

$\mathbb{Q}, \mathbb{R}, \mathbb{C}$, integers mod p, $\{a + b\sqrt{2} : a, b \in \mathbb{Z}\}$ are all fields.

Definition 2.2 (Totally ordered set)

An *(totally) ordered set* is a set X with a relation $<$ that satisfies

1. If $x, y, z \in X$, $x < y$ and $y < z$, then $x < z$ (transitivity)

2. If $x, y \in X$, exactly one of $x < y, x = y, y < x$ holds   (trichotomy)

We call it a *totally* ordered set, as opposed to simply an ordered set, when we want to emphasize the order is total, i.e. every pair of elements are related to each other, as opposed to a *partial order*, where "exactly one" in trichotomy is replaced with "at most one".

Now to define an ordered field, we don't simply ask for a field with an order. The order has to interact nicely with the field operations.

Definition 2.3 (Ordered field)

An *ordered field* is a field $\mathbb{F}$ with a relation $<$ that makes $\mathbb{F}$ into an ordered set such that

1. if $x, y, z \in \mathbb{F}$ and $x < y$, then $x + z < y + z$

2. if $x, y, z \in \mathbb{F}$, $x < y$ and $z > 0$, then $xz < yz$

Lemma 2.2

Let $\mathbb{F}$ be an ordered field and $x \in \mathbb{F}$. Then $x^2 \geq 0$.

Proof. By trichotomy, either $x < 0$, $x = 0$ or $x > 0$. If $x = 0$, then $x^2 = 0$. So $x^2 \geq 0$. If $x > 0$, then $x^2 > 0 \times x = 0$. If $x < 0$, then $x - x < 0 - x$. So $0 < -x$. But then $x^2 = (-x)^2 > 0$. $\qquad\square$

Definition 2.4 (Least upper bound)

Let X be an ordered set and let $A \subseteq X$. An *upper bound* for A is an element $x \in X$ such that $(\forall a \in A)\, a \leq x$. If A has an upper bound, then we say that A is *bounded above*.

    An upper bound x for A is a *least upper bound* or *supremum* if nothing smaller that x is an upper bound. That is, we need

1. $(\forall a \in A)\, a \leq x$

2. $(\forall y < x)(\exists a \in A)\, a \geq y$

    We usually write $\sup A$ for the supremum of A when it exists. If $\sup A \in A$, then we call it $\max A$, the maximum of A.

Example 2.3

Let $X = \mathbb{Q}$. Then the supremum of $(0, 1)$ is 1. The set $\{x : x^2 < 2\}$ is bounded above by 2, but has no supremum (even though $\sqrt{2}$ seems like a supremum, we are in $\mathbb{Q}$ and $\sqrt{2}$ is non-existent!).

    $\max[0, 1] = 1$ but $(0, 1)$ has no maximum because the supremum is not in $(0, 1)$.

We can think of the supremum as a point we can get arbitrarily close to in the set but cannot pass through.

Definition 2.5 (Least upper bound property)

An ordered set X has the *least upper bound property* if every non-empty subset of X that is bounded above has a supremum.

Obvious modifications give rise to definitions of lower bound, greatest lower bound (or infimum) etc. It is simple to check that an ordered field with the least upper bound property has the greatest lower bound property.

Definition 2.6 (Real numbers)

The *real numbers* is an ordered field with the least upper bound property.

Of course, it is *very* important to show that such a thing exists, or else we will be studying nothing. It is also nice to show that such a field is unique (up to isomorphism). However, we will not prove these in the course.

    In a field, we can define the "natural numbers" to be $2 = 1 + 1$, $3 = 1 + 2$ etc. Then an important property of the real numbers is

---

Lemma 2.4 (Archimedean property v1))

Let $\mathbb{F}$ be an ordered field with the least upper bound property. Then the set $\{1, 2, 3, \cdots\}$ is not bounded above.

Proof. If it is bounded above, then it has a supremum x. But then $x - 1$ is not an upper bound. So we can find $n \in \{1, 2, 3, \cdots\}$ such that $n > x - 1$. But then $n + 1 > x$, but x is supposed to be an upper bound. $\qquad\square$

Is the least upper bound property required to prove the Archimedean property? It seems like *any* ordered field should satisfy this even if they do not have the least upper bound property. However, it turns out there are ordered fields in which the integers are bounded above.

Consider the field of rational functions, i.e. functions in the form $\frac{P(x)}{Q(x)}$ with $P(x), Q(x)$ being polynomials, under the usual addition and multiplication. We order two functions $\frac{P(x)}{Q(x)}, \frac{R(x)}{S(x)}$ as follows: these two functions intersect only finitely many times because $P(x)S(x) = R(x)Q(x)$ has only finitely many roots. After the last intersection, the function whose value is greater counts as the greater function. It can be checked that these form an ordered field.

In this field, the integers are the constant functions $1, 2, 3, \cdots$, but it is bounded above since the function x is greater than all of them.

§3   **Convergence of sequences**

Having defined real numbers, the first thing we will study is sequences. We will want to study what it means for a sequence to *converge*. Intuitively, we would like to say that $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots$ converges to 0, while $1, 2, 3, 4, \cdots$ diverges. However, the actual formal definition of convergence is rather hard to get right, and historically there have been failed attempts that produced spurious results.

§3.1 **Definitions**

Definition 3.1 (Sequence)

A *sequence* is, formally, a function $a : \mathbb{N} \to \mathbb{R}$ (or $\mathbb{C}$). Usually (i.e. always), we write a_n instead of $a(n)$. Instead of a, we usually write it as (a_n), $(a_n)_1^\infty$ or $(a_n)_{n=1}^\infty$ to indicate it is a sequence.

Definition 3.2 (Convergence of sequence)

Let (a_n) be a sequence and $\ell \in \mathbb{R}$. Then a_n *converges to* ℓ, *tends to* ℓ, or $a_n \to \ell$, if for all $\varepsilon > 0$, there is some $N \in \mathbb{N}$ such that whenever $n > N$, we have $|a_n - \ell| < \varepsilon$. In symbols, this says

$$(\forall \varepsilon > 0)(\exists N)(\forall n \geq N)\, |a_n - \ell| < \varepsilon.$$

We say ℓ is the *limit* of (a_n).

One can think of $(\exists N)(\forall n \geq N)$ as saying "eventually always", or as "from some point on". So the definition means, if $a_n \to \ell$, then given any ε, there eventually, everything in the sequence is within ε of ℓ.

We'll now provide an alternative form of the Archimedean property. This is the form that is actually useful.

Lemma 3.1 (Archimedean property v2)

$1/n \to 0$.

Proof. Let $\varepsilon > 0$. We want to find an N such that $|1/N - 0| = 1/N < \varepsilon$. So pick N such that $N > 1/\varepsilon$. There exists such an N by the Archimedean property v1. Then for all $n \geq N$, we have $0 < 1/n \leq 1/N < \varepsilon$. So $|1/n - 0| < \varepsilon$. $\qquad\square$

Note that the red parts correspond to the *definition* of convergence of a sequence. This is generally how we prove convergence from first principles.

Definition 3.3 (Bounded sequence)

A sequence (a_n) is *bounded* if
$$(\exists C)(\forall n)\, |a_n| \leq C.$$

A sequence is *eventually bounded* if
$$(\exists C)(\exists N)(\forall n \geq N)\, |a_n| \leq C.$$

The definition of an *eventually bounded* sequence seems a bit daft. Clearly every eventually bounded sequence is bounded! Indeed it is:

> **Lemma 3.2**
>
> Every eventually bounded sequence is bounded.

Proof. Let C and N be such that $(\forall n \geq N) \, |a_n| \leq C$. Then $\forall n \in \mathbb{N}$, $|a_n| \leq \max\{|a_1|, \cdots, |a_{N-1}|, C\}$. $\quad\square$

The proof is rather trivial. However, most of the time it is simpler to prove that a sequence is eventually bounded, and this lemma saves us from writing that long line every time.

§3.2 Sums, products and quotients

Here we prove the things that we think are obviously true, e.g. sums and products of convergent sequences are convergent.

> **Lemma 3.3** (Sums of sequences)
>
> If $a_n \to a$ and $b_n \to b$, then $a_n + b_n \to a + b$.

Proof. Let $\varepsilon > 0$. We want to find a clever N such that for all $n \geq N$, $|a_n + b_n - (a+b)| < \varepsilon$. Intuitively, we know that a_n is very close to a and b_n is very close to b. So their sum must be very close to $a + b$.

Formally, since $a_n \to a$ and $b_n \to b$, we can find N_1, N_2 such that $\forall n \geq N_1$, $|a_n - a| < \varepsilon/2$ and $\forall n \geq N_2$, $|b_n - b| < \varepsilon/2$.

Now let $N = \max\{N_1, N_2\}$. Then by the triangle inequality, when $n \geq N$,

$$|(a_n + b_n) - (a+b)| \leq |a_n - a| + |b_n - b| < \varepsilon. \qquad\square$$

We want to prove that the product of convergent sequences is convergent. However, we will not do it in one go. Instead, we separate it into many smaller parts.

> **Lemma 3.4** (Scalar multiplication of sequences)
>
> Let $a_n \to a$ and $\lambda \in \mathbb{R}$. Then $\lambda a_n \to \lambda a$.

Proof. If $\lambda = 0$, then the result is trivial.

Otherwise, let $\varepsilon > 0$. Then $\exists N$ such that $\forall n \geq N$, $|a_n - a| < \varepsilon/|\lambda|$. So $|\lambda a_n - \lambda a| < \varepsilon$. $\quad\square$

> **Lemma 3.5**
>
> Let (a_n) be bounded and $b_n \to 0$. Then $a_n b_n \to 0$.

Proof. Let $C \neq 0$ be such that $(\forall n) \, |a_n| \leq C$. Let $\varepsilon > 0$. Then $\exists N$ such that $(\forall n \geq N) \, |b_n| < \varepsilon/C$. Then $|a_n b_n| < \varepsilon$. $\quad\square$

> **Lemma 3.6**
>
> Every convergent sequence is bounded.

Proof. Let $a_n \to l$. Then there is an N such that $\forall n \geq N$, $|a_n - l| \leq 1$. So $|a_n| \leq |l| + 1$. So a_n is eventually bounded, and therefore bounded. $\quad\square$

> **Lemma 3.7** (Product of sequences)
>
> Let $a_n \to a$ and $b_n \to b$. Then $a_n b_n \to ab$.

Proof. Let $a_n = a + \varepsilon_n$. Then $a_n b_n = (a + \varepsilon_n)b_n = ab_n + \varepsilon_n b_n$.

Since $b_n \to b$, $ab_n \to ab$. Since $\varepsilon_n \to 0$ and b_n is bounded, $\varepsilon_n b_n \to 0$. So $a_n b_n \to ab$. $\qquad\square$

Proof. (alternative) Observe that $a_n b_n - ab = (a_n - a)b_n + (b_n - b)a$. We know that $a_n - a \to 0$ and $b_n - b \to 0$. Since (b_n) is bounded, so $(a_n - a)b_n + (b_n - b)a \to 0$. So $a_n b_n \to ab$. $\qquad\square$

Note that in this proof, we no longer write "Let $\varepsilon > 0$". In the beginning, we have no lemmas proven. So we must prove everything from first principles and use the definition. However, after we have proven the lemmas, we can simply use them instead of using first principles. This is similar to in calculus, where we use first principles to prove the product rule and chain rule, then no longer use first principles afterwards.

Lemma 3.8 (Quotient of sequences)

Let (a_n) be a sequence such that $(\forall n)\, a_n \neq 0$. Suppose that $a_n \to a$ and $a \neq 0$. Then $1/a_n \to 1/a$.

Proof. We have
$$\frac{1}{a_n} - \frac{1}{a} = \frac{a - a_n}{aa_n}.$$
We want to show that this $\to 0$. Since $a - a_n \to 0$, we have to show that $1/(aa_n)$ is bounded.

Since $a_n \to a$, $\exists N$ such that $\forall n \geq N$, $|a_n - a| \leq a/2$. Then $\forall n \geq N$, $|a_n| \geq |a|/2$. Then $|1/(a_n a)| \leq 2/|a|^2$. So $1/(a_n a)$ is bounded. So $(a - a_n)/(aa_n) \to 0$ and the result follows. $\qquad\square$

Corollary 3.9

If $a_n \to a, b_n \to b$, $b_n, b \neq 0$, then $a_n/b_n = a/b$.

Proof. We know that $1/b_n \to 1/b$. So the result follows by the product rule. $\qquad\square$

Lemma 3.10 (Sandwich rule)

Let (a_n) and (b_n) be sequences that both converge to a limit x. Suppose that $a_n \leq c_n \leq b_n$ for every n. Then $c_n \to x$.

Proof. Let $\varepsilon > 0$. We can find N such that $\forall n \geq N$, $|a_n - x| < \varepsilon$ and $|b_n - x| < \varepsilon$.

Then $\forall n \geq N$, we have $x - \varepsilon < a_n \leq c_n \leq b_n < x + \varepsilon$. So $|c_n - x| < \varepsilon$. $\qquad\square$

Example 3.11

$1/2^n \to 0$. For every n, $n < 2^n$. So $0 < 1/2^n < 1/n$. The result follows from the sandwich rule.

Example 3.12

We want to show that
$$\frac{n^2 + 3}{(n+5)(2n-1)} \to \frac{1}{2}.$$

We can obtain this by
$$\frac{n^2 + 3}{(n+5)(2n-1)} = \frac{1 + 3/n^2}{(1 + 5/n)(2 - 1/n)} \to \frac{1}{2},$$

by sum rule, sandwich rule, Archimedean property, product rule and quotient rule.

> **Example 3.13**
>
> Let $k \in \mathbb{N}$ and let $\delta > 0$. Then
> $$\frac{n^k}{(1 + \delta)^n} \to 0.$$
>
> This can be summarized as "exponential growth beats polynomial growth eventually".
>   By the binomial theorem,
> $$(1 + \delta)^n \geq \binom{n}{k+1} \delta^{k+1}.$$
>
> Also, for $n \geq 2k$,
> $$\binom{n}{k+1} = \frac{n(n-1)\cdots(n-k)}{(k+1)!} \geq \frac{(n/2)^{k+1}}{(k+1)!}.$$
>
> So for sufficiently large n,
> $$\frac{n^k}{(1 + \delta)^n} \leq \frac{n^k 2^{k+1}(k+1)!}{n^{k+1}\delta^{k+1}} = \frac{2^{k+1}(k+1)!}{\delta^{k+1}} \cdot \frac{1}{n} \to 0.$$

§3.3 Monotone-sequences property

Recall that we characterized the least upper bound property. It turns out that there is an alternative characterization of real number using sequences, known as the *monotone-sequences property*. In this section, we will show that the two characterizations are equivalent, and use the monotone-sequences property to deduce some useful results.

> **Definition 3.4** (Monotone sequence)
>
> A sequence (a_n) is *increasing* if $a_n \leq a_{n+1}$ for all n.
>   It is *strictly increasing* if $a_n < a_{n+1}$ for all n. *(Strictly) decreasing* sequences are defined analogously.
>   A sequence is *(strictly) monotone* if it is (strictly) increasing or (strictly) decreasing.

> **Definition 3.5** (Monotone-sequences property)
>
> An ordered field has the *monotone sequences property* if every increasing sequence that is bounded above converges.

We want to show that the monotone sequences property is equivalent to the least upper bound property.

> **Lemma 3.14**
>
> Least upper bound property $\Rightarrow$ monotone-sequences property.

Proof. Let (a_n) be an increasing sequence and let C an upper bound for (a_n). Then C is an upper bound for the set $\{a_n : n \in \mathbb{N}\}$. By the least upper bound property, it has a supremum s. We want to show that this is the limit of (a_n).
   Let $\varepsilon > 0$. Since $s = \sup\{a_n : n \in \mathbb{N}\}$, there exists an N such that $a_N > s - \varepsilon$. Then since (a_n) is increasing, $\forall n \geq N$, we have $s - \varepsilon < a_N \leq a_n \leq s$. So $|a_n - s| < \varepsilon$. $\qquad\square$

We first prove a handy lemma.

> **Lemma 3.15**
>
> Let (a_n) be a sequence and suppose that $a_n \to a$. If $(\forall n)\, a_n \leq x$, then $a \leq x$.

Proof. If $a > x$, then set $\varepsilon = a - x$. Then we can find N such that $a_N > x$. Contradiction. $\qquad\square$

Before showing the other way implication, we will need the following:

Lemma 3.16

Monotone-sequences property $\Rightarrow$ Archimedean property.

Proof. We prove version 2, i.e. that $1/n \to 0$.

Since $1/n > 0$ and is decreasing, by MSP, it converges. Let δ be the limit. By the previous lemma, we must have $\delta \geq 0$.

If $\delta > 0$, then we can find N such that $1/N < 2\delta$. But then for all $n \geq 4N$, we have $1/n \leq 1/(4N) < \delta/2$. Contradiction. Therefore $\delta = 0$. $\qquad\square$

Lemma 3.17

Monotone-sequences property $\Rightarrow$ least upper bound property.

Proof. Let A be a non-empty set that's bounded above. Pick u_0, v_0 such that u_0 is not an upper bound for A and v_0 is an upper bound. Now do a repeated bisection: having chosen u_n and v_n such that u_n is not an upper bound and v_n is, if $(u_n + v_n)/2$ is an upper bound, then let $u_{n+1} = u_n$, $v_{n+1} = (u_n + v_n)/2$. Otherwise, let $u_{n+1} = (u_n + v_n)/2$, $v_{n+1} = v_n$.

Then $u_0 \leq u_1 \leq u_2 \leq \cdots$ and $v_0 \geq v_1 \geq v_2 \geq \cdots$. We also have

$$v_n - u_n = \frac{v_0 - u_0}{2^n} \to 0.$$

By the monotone sequences property, $u_n \to s$ (since (u_n) is bounded above by v_0). Since $v_n - u_n \to 0$, $v_n \to s$. We now show that $s = \sup A$.

If s is not an upper bound, then there exists $a \in A$ such that $a > s$. Since $v_n \to s$, then there exists m such that $v_m < a$, contradicting the fact that v_m is an upper bound.

To show it is the *least* upper bound, let $t < s$. Then since $u_n \to s$, we can find m such that $u_m > t$. So t is not an upper bound. Therefore s is the least upper bound. $\qquad\square$

Why do we need to prove the Archimedean property first? In the proof above, we secretly used the it. When showing that $v_n - u_n \to 0$, we required the fact that $\frac{1}{2^n} \to 0$. To prove this, we sandwiched it with $\frac{1}{n}$. But to show $\frac{1}{n} \to 0$, we need the Archimedean property.

Lemma 3.18

A sequence can have at most 1 limit.

Proof. Let (a_n) be a sequence, and suppose $a_n \to x$ and $a_n \to y$. Let $\varepsilon > 0$ and pick N such that $\forall n \geq N$, $|a_n - x| < \varepsilon/2$ and $|a_n - y| < \varepsilon/2$. Then $|x - y| \leq |x - a_N| + |a_N - y| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Since ε was arbitrary, x must equal y. $\qquad\square$

Lemma 3.19 (Nested intervals property)

Let $\mathbb{F}$ be an ordered field with the monotone sequences property. Let $I_1 \supseteq I_2 \supseteq \cdots$ be closed bounded non-empty intervals. Then $\bigcap_{n=1}^{\infty} I_n \neq \varnothing$.

Proof. Let $T_n = [a_n, b_n]$ for each n. Then $a_1 \leq a_2 \leq \cdots$ and $b_1 \geq b_2 \geq \cdots$. For each n, $a_n \leq b_n \leq b_1$. So the sequence a_n is bounded above. So by the monotone sequences property, it has a limit a. For each n, we must have $a_n \leq a$. Otherwise, say $a_n > a$. Then for all $m \geq n$, we have $a_m \geq a_n > a$. This implies that $a > a$, which is nonsense.

Also, for each fixed n, we have that $\forall m \geq n$, $a_m \leq b_m \leq b_n$. So $a \leq b_n$. Thus, for all n, $a_n \leq a \leq b_n$. So $a \in I_n$. So $a \in \bigcap_{n=1}^{\infty} I_n$. $\qquad\square$

We can use this to prove that the reals are uncountable:

> **Claim 3.20**
>
> $\mathbb{R}$ is uncountable.

Proof. Suppose the contrary. Let $x_1, x_2, \cdots$ be a list of all real numbers. Find an interval that does not contain x_1. Within that interval, find an interval that does not contain x_2. Continue *ad infinitum*. Then the intersection of all these intervals is non-empty, but the elements in the intersection are not in the list. Contradiction. $\qquad\square$

A powerful consequence of this is the *Bolzano-Weierstrass theorem*. This is formulated in terms of subsequences:

> **Definition 3.6** (Subsequence)
>
> Let (a_n) be a sequence. A *subsequence* of (a_n) is a sequence of the form $a_{n_1}, a_{n_2}, \cdots$, where $n_1 < n_2 < \cdots$.

> **Example 3.21**
>
> $1, 1/4, 1/9, 1/16, \cdots$ is a subsequence of $1, 1/2, 1/3, \cdots$.

> **Theorem 3.22** (Bolzano-Weierstrass theorem)
>
> Let $\mathbb{F}$ be an ordered field with the monotone sequences property (i.e. $\mathbb{F} = \mathbb{R}$).
>   Then every bounded sequence has a convergent subsequence.

Proof. Let u_0 and v_0 be a lower and upper bound, respectively, for a sequence (a_n). By repeated bisection, we can find a sequence of intervals $[u_0, v_0] \supseteq [u_1, v_1] \supseteq [u_2, v_2] \supseteq \cdots$ such that $v_n - u_n = (v_0 - u_0)/2^n$, and such that each $[u_n, v_n]$ contains infinitely many terms of (a_n).

By the nested intervals property, $\bigcap_{n=1}^{\infty}[u_n, v_n] \neq \varnothing$. Let x belong to the intersection. Now pick a subsequence $a_{n_1}, a_{n_2}, \cdots$ such that $a_{n_k} \in [u_k, v_k]$. We can do this because $[u_k, v_k]$ contains infinitely many a_n, and we have only picked finitely many of them. We will show that $a_{n_k} \to x$.

Let $\varepsilon > 0$. By the Archimedean property, we can find K such that $v_K - u_K = (v_0 - u_0)/2^K \leq \varepsilon$. This implies that $[u_K, v_K] \subseteq (x - \varepsilon, x + \varepsilon)$, since $x \in [u_K, v_K]$.

Then $\forall k \geq K$, $a_{n_k} \in [u_k, v_k] \subseteq [u_K, v_K] \subseteq (x - \varepsilon, x + \varepsilon)$. So $|a_{n_k} - x| < \varepsilon$. $\qquad\square$

§3.4   Cauchy sequences

The third characterization of real numbers is in terms of Cauchy sequences. Cauchy convergence is an alternative way of defining convergent sequences without needing to mention the actual limit of the sequence. This allows us to say $\{3, 3.1, 3.14, 3.141, 3.1415, \cdots\}$ is *Cauchy convergent* in $\mathbb{Q}$ even though the limit π is not in $\mathbb{Q}$.

> **Definition 3.7** (Cauchy sequence)
>
> A sequence (a_n) is *Cauchy* if for all ε, there is some $N \in \mathbb{N}$ such that whenever $p, q \geq N$, we have $|a_p - a_q| < \varepsilon$. In symbols, we have
> $$(\forall \varepsilon > 0)(\exists N)(\forall p, q \geq N)\, |a_p - a_q| < \varepsilon.$$

Roughly, a sequence is Cauchy if all terms are eventually close to each other (as opposed to close to a limit).

> **Lemma 3.23**
>
> Every convergent sequence is Cauchy.

Proof. Let $a_n \to a$. Let $\varepsilon > 0$. Then $\exists N$ such that $\forall n \geq N$, $|a_n - a| < \varepsilon/2$. Then $\forall p, q \geq N$, $|a_p - a_q| \leq |a_p - a| + |a - a_q| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. $\qquad\square$

> **Lemma 3.24**
>
> Let (a_n) be a Cauchy sequence with a subsequence (a_{n_k}) that converges to a. Then $a_n \to a$.

Proof. Let $\varepsilon > 0$. Pick N such that $\forall p, q \geq N$, $|a_p - a_q| < \varepsilon/2$. Then pick K such that $n_K \geq N$ and $|a_{n_K} - a| < \varepsilon/2$.

Then $\forall n \geq N$, we have

$$|a_n - a| \leq |a_n - a_{n_K}| + |a_{n_K} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \qquad \square$$

An important result we have is that in $\mathbb{R}$, Cauchy convergence and regular convergence are equivalent.

> **Theorem 3.25** (The general principle of convergence)
>
> Let $\mathbb{F}$ be an ordered field with the monotone-sequence property. Then every Cauchy sequence of $\mathbb{F}$ converges.

Proof. Let (a_n) be a Cauchy sequence. Then it is eventually bounded, since $\exists N$, $\forall n \geq N$, $|a_n - a_N| \leq 1$ by the Cauchy condition. So it is bounded. Hence by Bolzano-Weierstrass, it has a convergent subsequence. Then (a_n) converges to the same limit. $\qquad \square$

> **Definition 3.8** (Complete ordered field)
>
> An ordered field in which every Cauchy sequence converges is called *complete*.

Hence we say that $\mathbb{R}$ is a complete ordered field.

However, not every complete ordered field is (isomorphic to) $\mathbb{R}$. For example, we can take the rational functions as before, then take the Cauchy completion of it (i.e. add all the limits we need). Then it is already too large to be the reals (it still doesn't have the Archimedean property) but is a complete ordered field.

To show that completeness implies the monotone-sequences property, we need an additional condition: the Archimedean property.

> **Lemma 3.26**
>
> Let $\mathbb{F}$ be an ordered field with the Archimedean property such that every Cauchy sequence converges. The $\mathbb{F}$ satisfies the monotone-sequences property.

Proof. Instead of showing that every bounded monotone sequence converges, and is hence Cauchy, We will show the equivalent statement that every increasing non-Cauchy sequence is not bounded above.

Let (a_n) be an increasing sequence. If (a_n) is not Cauchy, then

$$(\exists \varepsilon > 0)(\forall N)(\exists p, q > N) |a_p - a_q| \geq \varepsilon.$$

wlog let $p > q$. Then

$$a_p \geq a_q + \varepsilon \geq a_N + \varepsilon.$$

So for any N, we can find a $p > N$ such that

$$a_p \geq a_N + \varepsilon.$$

Then we can construct a subsequence $a_{n_1}, a_{n_2}, \cdots$ such that

$$a_{n_{k+1}} \geq a_{n_k} + \varepsilon.$$

Therefore

$$a_{n_k} \geq a_{n_1} + (k-1)\varepsilon.$$

So by the Archimedean property, (a_{n_k}), and hence (a_n), is unbounded. $\qquad \square$

Note that the definition of a convergent sequence is

$$(\exists l)(\forall \varepsilon > 0)(\exists N)(\forall n \geq N)\, |a_n - l| < \varepsilon,$$

while that of Cauchy convergence is

$$(\forall \varepsilon > 0)(\exists N)(\forall p, q \geq N)\, |a_p - a_q| < \varepsilon.$$

In the first definition, l quantifies over all real numbers, which is uncountable. However, in the second definition, we only have to quantify over natural numbers, which is countable (by the Archimedean property, we only have to consider the cases $\varepsilon = 1/n$).

Since they are equivalent in $\mathbb{R}$, the second definition is sometimes preferred when we care about logical simplicity.

§3.5 Limit supremum and infimum

Here we will define the limit supremum and infimum. While these are technically not part of the course, eventually some lecturers will magically assume students know this definition. So we might as well learn it here.

Definition 3.9 (Limit supremum/infimum)

Let (a_n) be a bounded sequence. We define the *limit supremum* as

$$\limsup_{n \to \infty} a_n = \lim_{n \to \infty} \left(\sup_{m \geq n} a_m \right).$$

To see that this exists, set $b_n = \sup_{m \geq n} a_m$. Then (b_n) is decreasing since we are taking the supremum of fewer and fewer things, and is bounded below by any lower bound for (a_n) since $b_n \geq a_n$. So it converges.

Similarly, we define the *limit infimum* as

$$\liminf_{n \to \infty} a_n = \lim_{n \to \infty} \left(\inf_{m \geq n} a_m \right).$$

Example 3.27

Take the sequence

$$2, -1, \frac{3}{2}, -\frac{1}{2}, \frac{4}{3}, -\frac{1}{3}, \cdots$$

Then the limit supremum is 1 and the limit infimum is 0.

Lemma 3.28

Let (a_n) be a sequence. The following two statements are equivalent:

- $a_n \to a$

- $\limsup a_n = \liminf a_n = a$.

Proof. If $a_n \to a$, then let $\varepsilon > 0$. Then we can find an n such that

$$a - \varepsilon \leq a_m \leq a + \varepsilon \text{ for all } m \geq n$$

It follows that

$$a - \varepsilon \leq \inf_{m \geq n} a_m \leq \sup_{m \geq n} a_m \leq a + \varepsilon.$$

Since ε was arbitrary, it follows that

$$\liminf a_n = \limsup a_n = a.$$

Conversely, if $\liminf a_n = \limsup a_n = a$, then let $\varepsilon > 0$. Then we can find n such that

$$\inf_{m \geq n} a_m > a - \varepsilon \text{ and } \sup_{m \geq n} a_m < a + \varepsilon.$$

It follows that $\forall m \geq n$, we have $|a_m - a| < \varepsilon$. $\qquad\square$

§4 Convergence of infinite sums

In this chapter, we investigate which infinite *sums*, as opposed to sequences, converge. We would like to say $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2$, while $1 + 1 + 1 + 1 + \cdots$ does not converge. The majority of the chapter is coming up with different tests to figure out if infinite sums converge.

§4.1 Infinite sums

> **Definition 4.1** (Convergence of infinite sums and partial sums)
>
> Let (a_n) be a real sequence. For each N, define
> $$S_N = \sum_{n=1}^{N} a_n.$$
>
> If the sequence (S_N) converges to some limit s, then we say that
> $$\sum_{n=1}^{\infty} a_n = s,$$
>
> and we say that the series $\sum_{n=1}^{\infty} a_n$ *converges*.
>   We call S_N the Nth *partial sum*.

There is an immediate necessary condition for a series to converge.

> **Lemma 4.1**
>
> If $\sum_{n=1}^{\infty} a_n$ converges. Then $a_n \to 0$.

Proof. Let $\sum_{n=1}^{\infty} a_n = s$. Then $S_n \to s$ and $S_{n-1} \to s$. Then $a_n = S_n - S_{n-1} \to 0$. □

However, the converse is false!

> **Example 4.2** (Harmonic series)
>
> If $a_n = 1/n$, then $a_n \to 0$ but $\sum a_n = \infty$.
>   We can prove this as follows:
> $$S_{2^n} - S_{2^{n-1}} = \frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^n} \geq \frac{2^{n-1}}{2^n} = \frac{1}{2}.$$
>
> Therefore $S_{2^n} \geq S_1 + n/2$. So the partial sums are unbounded.

> **Example 4.3** (Geometric series)
>
> Let $|\rho| < 1$. Then
> $$\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho}.$$
>
> We can prove this by considering the partial sums:
> $$\sum_{n=0}^{N} \rho^n = \frac{1-\rho^{N+1}}{1-\rho}.$$
>
> Since $\rho^{N+1} \to 0$, this tends to $1/(1-\rho)$.

> **Example 4.4**
>
> $\sum_{n=2}^{\infty} \dfrac{1}{n(n-1)}$ converges. This is since
> $$\frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}.$$
>
> So
> $$\sum_{n=2}^{N} \frac{1}{n(n-1)} = 1 - \frac{1}{N} \to 1.$$

> **Lemma 4.5**
>
> Suppose that $a_n \geq 0$ for every n and the partial sums S_n are bounded above. Then $\sum_{n=1}^{\infty} a_n$ converges.

Proof. The sequence (S_n) is increasing and bounded above. So the result follows form the monotone sequences property. $\square$

The simplest convergence test we have is the *comparison test*. Roughly speaking, it says that if $0 \leq a_n \leq b_n$ for all n and $\sum b_n$ converges, then $\sum a_n$ converges. However, we will prove a much more general form here for convenience.

> **Lemma 4.6** (Comparison test)
>
> Let (a_n) and (b_n) be non-negative sequences, and suppose that $\exists C, N$ such that $\forall n \geq N$, $a_n \leq Cb_n$. Then if $\sum b_n$ converges, then so does $\sum a_n$.

Proof. Let $M > N$. Also for each R, let $S_R = \sum_{n=1}^{R} a_n$ and $T_R = \sum_{n=1}^{R} b_n$. We want S_R to be bounded above.

$$S_M - S_N = \sum_{n=N+1}^{M} a_n \leq C \sum_{n=N+1}^{M} b_n \leq C \sum_{n=N+1}^{\infty} b_n.$$

So $\forall M \geq N$, $S_M \leq S_n + C \sum_{n=N+1}^{\infty} b_n$. Since the S_M are increasing and bounded, it must converge. $\square$

> **Example 4.7**
>
> 1. $\sum \frac{1}{n2^n}$ converges, since $\sum \frac{1}{2^n}$ converges.
>
> 2. $\sum \frac{n}{2^n}$ converges.
>
>   If $n \geq 4$, then $n \leq 2^{n/2}$. That's because $4 = 2^{4/2}$ and for $n \geq 4$, $(n+1)/n < \sqrt{2}$, so when we increase n, we multiply the right side by a greater number by the left. Hence by the comparison test, it is sufficient to show that $\sum 2^{n/2}/2^n = \sum 2^{-n/2}$ converges, which it does (geometric series).
>
> 3. $\sum \frac{1}{\sqrt{n}}$ diverges, since $\frac{1}{\sqrt{n}} \geq \frac{1}{n}$. So if it converged, then so would $\sum \frac{1}{n}$, but $\sum \frac{1}{n}$ diverges.
>
> 4. $\sum \frac{1}{n^2}$ converges, since for $n \geq 2$, $\frac{1}{n^2} \leq \frac{1}{n(n-1)}$, and we have proven that the latter converges.
>
> 5. Consider $\displaystyle\sum_{n=1}^{\infty} \frac{n+5}{n^3 - 7n^2/2}$. We show this converges by noting
>
> $$n^3 - \frac{7n^2}{2} = n^2\left(n - \frac{7}{2}\right).$$
>
>   So if $n \geq 8$, then
>
> $$n^3 - \frac{7n^2}{2} \geq \frac{n^3}{2}.$$
>
>   Also, $n + 5 \leq 2n$. So
>
> $$\frac{n+5}{n^3 - 7n^2/2} \leq 4/n^2.$$
>
>   So it converges by the comparison test.
>
> 6. If $\alpha > 1$, then $\sum 1/n^\alpha$ converges.
>
>   Let $S_N = \sum_{n=1}^{N} 1/n^\alpha$. Then
>
> $$\begin{aligned}
S_{2^n} - S_{2^{n-1}} &= \frac{1}{(2^{n-1}+1)^\alpha} + \cdots + \frac{1}{(2^n)^\alpha} \\
&\leq \frac{2^{n-1}}{(2^{n-1})^\alpha} \\
&= (2^{n-1})^{1-\alpha} \\
&= (2^{1-\alpha})^{n-1}.
\end{aligned}$$
>
>   But $2^{1-\alpha} < 1$. So
>
> $$S_{2^n} = (S_{2^n} - S_{2^{n-1}}) + (S_{2^{n-1}} - S_{2^{n-2}}) + \cdots (S_2 - S_1) + S_1$$
>
>   and is bounded above by comparison with the geometric series $1 + 2^{1-\alpha} + (2^{1-\alpha})^2 + \cdots$

§4.2 Absolute convergence

Here we'll consider two stronger conditions for convergence — absolute convergence and unconditional convergence. We'll prove that these two conditions are in fact equivalent.

> **Definition 4.2** (Absolute convergence)
>
> A series $\sum a_n$ *converges absolutely* if the series $\sum |a_n|$ converges.

Example 4.8

The series $\sum \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ converges, but not absolutely.

To see the convergence, note that

$$a_{2n-1} + a_{2n} = \frac{1}{2n-1} - \frac{1}{2n} = \frac{1}{2n(2n-1)}.$$

It is easy to compare with $1/n^2$ to get that the partial sums S_{2n} converges. But $S_{2n+1} - S_{2n} = 1/(2n+1) \to 0$, so the S_{2n+1} converges to the same limit.

It does not converge absolutely, because the sum of the absolute values is the harmonic series.

Lemma 4.9

Let $\sum a_n$ converge absolutely. Then $\sum a_n$ converges.

Proof. We know that $\sum |a_n|$ converges. Let $S_N = \sum_{n=1}^{N} a_n$ and $T_N = \sum_{n=1}^{N} |a_n|$.

We know two ways to show random sequences converge, without knowing what they converge to, namely monotone-sequences and Cauchy sequences. Since S_N is not monotone, we shall try Cauchy sequences.

If $p > q$, then

$$|S_p - S_q| = \left| \sum_{n=q+1}^{p} a_n \right| \leq \sum_{n=q+1}^{p} |a_n| = T_p - T_q.$$

But the sequence T_p converges. So $\forall \varepsilon > 0$, we can find N such that for all $p > q \geq N$, we have $T_p - T_q < \varepsilon$, which implies $|S_p - S_q| < \varepsilon$. $\qquad \square$

Definition 4.3 (Unconditional convergence)

A series $\sum a_n$ *converges unconditionally* if the series $\sum_{n=1}^{\infty} a_{\pi(n)}$ converges for every bijection $\pi : \mathbb{N} \to \mathbb{N}$, i.e. no matter how we re-order the elements of a_n, the sum still converges.

Theorem 4.10

If $\sum a_n$ converges absolutely, then it converges unconditionally.

Proof. Let $S_n = \sum_{n=1}^{N} a_{\pi(n)}$. Then if $p > q$,

$$|S_p - S_q| = \left| \sum_{n=q+1}^{p} a_{\pi(n)} \right| \leq \sum_{n=q+1}^{\infty} |a_{\pi(n)}|.$$

Let $\varepsilon > 0$. Since $\sum |a_n|$ converges, pick M such that $\sum_{n=M+1}^{\infty} |a_n| < \varepsilon$.

Pick N large enough that $\{1, \cdots, M\} \subseteq \{\pi(1), \cdots, \pi(N)\}$.

Then if $n > N$, we have $\pi(n) > M$. Therefore if $p > q \geq N$, then

$$|S_p - S_q| \leq \sum_{n=q+1}^{p} |a_{\pi(n)}| \leq \sum_{n=M+1}^{\infty} |a_n| < \varepsilon.$$

Therefore the sequence of partial sums is Cauchy. $\qquad \square$

The converse is also true.

Theorem 4.11

If $\sum a_n$ converges unconditionally, then it converges absolutely.

Proof. We will prove the contrapositive: if it doesn't converge absolutely, it doesn't converge unconditionally.

Suppose that $\sum |a_n| = \infty$. Let (b_n) be the subsequence of non-negative terms of a_n, and (c_n) be the subsequence of negative terms. Then $\sum b_n$ and $\sum c_n$ cannot both converge, or else $\sum |a_n|$ converges.

wlog, $\sum b_n = \infty$. Now construct a sequence $0 = n_0 < n_1 < n_2 < \cdots$ such that $\forall k$,

$$b_{n_{k-1}+1} + b_{n_{k-1}+2} + \cdots + b_{n_k} + c_k \geq 1,$$

This is possible because the b_n are unbounded and we can get it as large as we want.

Let π be the rearrangement

$$b_1, b_2, \cdots b_{n_1}, c_1, b_{n_1+1}, \cdots b_{n_2}, c_2, b_{n_2+1}, \cdots b_{n_3}, c_3, \cdots$$

So the sum up to c_k is at least k. So the partial sums tend to infinity. □

We can prove an even stronger result:

Lemma 4.12

Let $\sum a_n$ be a series that converges absolutely. Then for any bijection $\pi : \mathbb{N} \to \mathbb{N}$,

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\pi(n)}.$$

Proof. Let $\varepsilon > 0$. We know that both $\sum |a_n|$ and $\sum |a_{\pi(n)}|$ converge. So let M be such that $\sum_{n>M} |a_n| < \frac{\varepsilon}{2}$ and $\sum_{n>M} |a_{\pi(n)}| < \frac{\varepsilon}{2}$.

Now N be large enough such that

$$\{1, \cdots, M\} \subseteq \{\pi(1), \cdots, \pi(N)\},$$

and

$$\{\pi(1), \cdots, \pi(M)\} \subseteq \{1, \cdots, N\}.$$

Then for every $K \geq N$,

$$\left| \sum_{n=1}^{K} a_n - \sum_{n=1}^{K} a_{\pi(n)} \right| \leq \sum_{n=M+1}^{K} |a_n| + \sum_{n=M+1}^{K} |a_{\pi(n)}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

We have the first inequality since given our choice of M and N, the first M terms of the $\sum a_n$ and $\sum a_{\pi(n)}$ sums are cancelled by some term in the huge sum.

So $\forall K \geq N$, the partial sums up to K differ by at most ε. So $|\sum a_n - \sum a_{\pi(n)}| \leq \varepsilon$.

Since this is true for all ε, we must have $\sum a_n = \sum a_{\pi(n)}$. □

§4.3 Convergence tests

We'll now come up with a *lot* of convergence tests.

Lemma 4.13 (Alternating sequence test)

Let (a_n) be a decreasing sequence of non-negative reals, and suppose that $a_n \to 0$. Then $\sum_{n=1}^{\infty} (-1)^{n+1} a_n$ converges, i.e. $a_1 - a_2 + a_3 - a_4 + \cdots$ converges.

Proof. Let $S_N = \sum_{n=1}^{N} (-1)^{n+1} a_n$. Then

$$S_{2n} = (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2n-1} - a_{2n}) \geq 0,$$

and (S_{2n}) is an increasing sequence.

Also,

$$S_{2n+1} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n} - a_{2n+1}),$$

and (S_{2n+1}) is a decreasing sequence. Also $S_{2n+1} - S_{2n} = a_{2n+1} \geq 0$.

Hence we obtain the bounds $0 \leq S_{2n} \leq S_{2n+1} \leq a_1$. It follows from the monotone sequences property that (S_{2n}) and (S_{2n+1}) converge.

Since $S_{2n+1} - S_{2n} = a_{2n+1} \to 0$, they converge to the same limit. $\qquad\square$

Example 4.14

$$1 - \frac{1}{\sqrt[3]{2}} + \frac{1}{\sqrt[3]{3}} - \frac{1}{\sqrt[3]{4}} + \cdots \text{converges.}$$

Lemma 4.15 (Ratio test)

We have three versions:

1. If $\exists c < 1$ such that

$$\frac{|a_{n+1}|}{|a_n|} \leq c,$$

    for all n, then $\sum a_n$ converges.

2. If $\exists c < 1$ and $\exists N$ such that

$$(\forall n \geq N) \frac{|a_{n+1}|}{|a_n|} \leq c,$$

    then $\sum a_n$ converges. Note that just because the ratio is always less than 1, it doesn't necessarily converge. It has to be always less than a fixed number c. Otherwise the test will say that $\sum 1/n$ converges.

3. If $\exists \rho \in (-1, 1)$ such that

$$\frac{a_{n+1}}{a_n} \to \rho,$$

    then $\sum a_n$ converges. Note that we have the *open* interval $(-1, 1)$. If $\frac{|a_{n+1}|}{|a_n|} \to 1$, then the test is inconclusive!

Proof.

1. $|a_n| \leq c^{n-1}|a_1|$. Since $\sum c^n$ converges, so does $\sum |a_n|$ by comparison test. So $\sum a_n$ converges absolutely, so it converges.

2. For all $k \geq 0$, we have $|a_{N+k}| \leq c^k |a_N|$. So the series $\sum |a_{N+k}|$ converges, and therefore so does $\sum |a_k|$.

3. If $\frac{a_{n+1}}{a_n} \to \rho$, then $\frac{|a_{n+1}|}{|a_n|} \to |\rho|$. So (setting $\varepsilon = (1 - |\rho|)/2$) there exists N such that $\forall n \geq N$, $\frac{|a_{n+1}|}{|a_n|} \leq \frac{1+|\rho|}{2} < 1$. So the result follows from (ii). $\qquad\square$

> **Example 4.16**
>
> If $|b| < 1$, then $\sum nb^n$ converges, since
>
> $$\frac{a_{n+1}}{a_n} = \frac{(n+1)b^{n+1}}{nb^n} = \left(1 + \frac{1}{n}\right) b \to b < 1.$$
>
> So it converges.
>
> We can also evaluate this directly by considering $\sum\limits_{i=1}^{\infty} \sum\limits_{n=i}^{\infty} b^n$.

The following two tests were taught at the end of the course, but are included here for the sake of completeness.

> **Theorem 4.17** (Condensation test)
>
> Let (a_n) be a decreasing non-negative sequence. Then $\sum_{n=1}^{\infty} a_n < \infty$ if and only if
>
> $$\sum_{k=1}^{\infty} 2^k a_{2^k} < \infty.$$

Proof. This is basically the proof that $\sum \frac{1}{n}$ diverges and $\sum \frac{1}{n^\alpha}$ converges for $\alpha < 1$ but written in a more general way.

We have

$$a_1 + a_2 + (a_3 + a_4) + (a_5 + \cdots + a_8) + (a_9 + \cdots + a_{16}) + \cdots$$
$$\geq a_1 + a_2 + 2a_4 + 4a_8 + 8a_{16} + \cdots$$

So if $\sum 2^k a_{2^k}$ diverges, $\sum a_n$ diverges.

To prove the other way round, simply group as

$$a_1 + (a_2 + a_3) + (a_4 + \cdots + a_7) + \cdots$$
$$\leq a_1 + 2a_2 + 4a_4 + \cdots . \qquad \square$$

> **Example 4.18**
>
> If $a_n = \frac{1}{n}$, then $2^k a_{2^k} = 1$. So $\sum_{k=1}^{\infty} 2^k a_{2^k} = \infty$. So $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

After we formally define integrals, we will prove the integral test:

> **Theorem 4.19** (Integral test)
>
> Let $f : [1, \infty] \to \mathbb{R}$ be a decreasing non-negative function. Then $\sum_{n=1}^{\infty} f(n)$ converges iff $\int_1^{\infty} f(x) \, \mathrm{d}x < \infty$.

§4.4 Complex versions

Most definitions in the course so far carry over unchanged to the complex numbers. e.g. $z_n \to z$ iff $(\forall \varepsilon > 0)(\exists N)(\forall n \geq N) |z_n - z| < \varepsilon$.

Two exceptions are least upper bound and monotone sequences, because the complex numbers do not have an ordering! (It cannot be made into an ordered field because the square of every number in an ordered field has to be positive) Fortunately, Cauchy sequences still work.

We can prove the complex versions of most theorems so far by looking at the real and imaginary parts.

> **Example 4.20**
>
> Let (z_n) be a Cauchy sequence in $\mathbb{C}$. Let $z_n = x_n + iy_n$. Then (x_n) and (y_n) are Cauchy. So they converge, from which it follows that $z_n = x_n + iy_n$ converges.

Also, the Bolzano-Weierstrass theorem still holds: If (z_n) is bounded, let $z_n = x_n + y_n$, then (x_n) and (y_n) are bounded. Then find a subsequence (x_{n_k}) that converges. Then find a subsequence of (y_{n_k}) that converges.

Then nested-intervals property has a "nested-box" property as a complex analogue.

Finally, the proof that absolutely convergent sequences converge still works. It follows that the ratio test still works.

> **Example 4.21**
>
> If $|z| < 1$, then $\sum nz^n$ converges. Proof is the same as above.

However, we do have an extra test for complex sums.

> **Lemma 4.22** (Abel's test)
>
> Let $a_1 \geq a_2 \geq \cdots \geq 0$, and suppose that $a_n \to 0$. Let $z \in \mathbb{C}$ such that $|z| = 1$ and $z \neq 1$. Then $\sum a_n z^n$ converges.

Proof. We prove that it is Cauchy. We have

$$
\sum_{n=M}^{N} a_n z^n = \sum_{n=M}^{N} a_n \frac{z^{n+1} - z^n}{z - 1}
$$

$$
= \frac{1}{z-1} \sum_{n=M}^{N} a_n (z^{n+1} - z^n)
$$

$$
= \frac{1}{z-1} \left(\sum_{n=M}^{N} a_n z^{n+1} - \sum_{n=M}^{N} a_n z^n \right)
$$

$$
= \frac{1}{z-1} \left(\sum_{n=M}^{N} a_n z^{n+1} - \sum_{n=M-1}^{N-1} a_{n+1} z^{n+1} \right)
$$

$$
= \frac{1}{z-1} \left(a_N z^{N+1} - a_M z^M + \sum_{n=M}^{N-1} (a_n - a_{n+1}) z^{n+1} \right)
$$

We now take the absolute value of everything to obtain

$$
\left| \sum_{n=M}^{N} a_n z^n \right| \leq \frac{1}{|z-1|} \left(a_N + a_M + \sum_{n=M}^{N-1} (a_n - a_{n+1}) \right)
$$

$$
= \frac{1}{|z-1|} \left(a_N + a_M + (a_M - a_{M+1}) + \cdots + (a_{N-1} - a_N) \right)
$$

$$
= \frac{2a_M}{|z-1|} \to 0.
$$

So it is Cauchy. So it converges $\qquad\square$

Note that here we transformed the sum $\sum a_n(z^{n+1} - z^n)$ into $a_N z^{N+1} - a_M z^M + \sum (a_n - a_{n+1}) z^{n+1}$. What we have effectively done is a discrete analogue of integrating by parts.

> **Example 4.23**
>
> The series $\sum z^n/n$ converges if $|z| < 1$ or if $|z| = 1$ and $z \neq 1$, and it diverges if $z = 1$ or $|z| > 1$.
>
> The cases $|z| < 1$ and $|z| > 1$ are trivial from the ratio test, and Abel's test is required for the $|z| = 1$ cases.

§5 Continuous functions

§5.1 Continuous functions

Definition 5.1 (Continuous function)

Let $A \subseteq \mathbb{R}$, $a \in A$, and $f : A \to \mathbb{R}$. Then f is *continuous at* a if for any $\varepsilon > 0$, there is some $\delta > 0$ such that if $y \in A$ is such that $|y - a| < \delta$, then $|f(y) - f(a)| < \varepsilon$. In symbols, we have

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall y \in A) \, |y - a| < \delta \Rightarrow |f(y) - f(a)| < \varepsilon.$$

f is *continuous* if it is continuous at every $a \in A$. In symbols, we have

$$(\forall a \in A)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall y \in A) \, |y - a| < \delta \Rightarrow |f(y) - f(a)| < \varepsilon.$$

Intuitively, f is continuous at a if we can obtain $f(a)$ as accurately as we wish by using more accurate values of a (the definition says that if we want to approximate $f(a)$ by $f(y)$ to within accuracy ε, we just have to get our y to within δ of a for some δ).

For example, suppose we have the function

$$f(x) = \begin{cases} 0 & x \leq \pi \\ 1 & x > \pi \end{cases}.$$

Suppose that we don't know what the function actually is, but we have a computer program that computes this function. We want to know what $f(\pi)$ is. Since we cannot input π (it has infinitely many digits), we can try 3, and it gives 0. Then we try 3.14, and it gives 0 again. If we try 3.1416, it gives 1 (since $\pi = 3.1415926\cdots < 3.1416$). We keep giving more and more digits of π, but the result keeps oscillating between 0 and 1. We have no hope of what $f(\pi)$ might be, even approximately. So this f is discontinuous at π.

However, if we have the function $g(x) = x^2$, then we *can* find the (approximate) value of $g(\pi)$. We can first try $g(3)$ and obtain 9. Then we can try $g(3.14) = 9.8596$, $g(3.1416) = 9.86965056$ etc. We can keep trying and obtain more and more accurate values of $g(\pi)$. So g is continuous at π.

Example 5.1

- Constant functions are continuous.

- The function $f(x) = x$ is continuous (take $\delta = \varepsilon$).

The definition of continuity of a function looks rather like the definition of convergence. In fact, they are related by the following lemma:

Lemma 5.2

The following two statements are equivalent for a function $f : A \to \mathbb{R}$.

- f is continuous

- If (a_n) is a sequence in A with $a_n \to a$, then $f(a_n) \to f(a)$.

Proof. (i)$\Rightarrow$(ii) Let $\varepsilon > 0$. Since f is continuous at a,

$$(\exists \delta > 0)(\forall y \in A) \, |y - a| < \delta \Rightarrow |f(y) - f(a)| < \varepsilon.$$

We want N such that $\forall n \geq N$, $|f(a_n) - f(a)| < \varepsilon$. By continuity, it is enough to find N such that $\forall n \geq N$, $|a_n - a| < \delta$. Since $a_n \to a$, such an N exists.

(ii)$\Rightarrow$(i) We prove the contrapositive: Suppose f is not continuous at a. Then

$$(\exists \varepsilon > 0)(\forall \delta > 0)(\exists y \in A) \, |y - a| < \delta \text{ and } |f(y) - f(a)| \geq \varepsilon.$$

For each n, we can therefore pick $a_n \in A$ such that $|a_n - a| < \frac{1}{n}$ and $|f(a_n) - f(a)| \geq \varepsilon$. But then $a_n \to a$ (by Archimedean property), but $f(a_n) \not\to f(a)$. □

Example 5.3

1. Let $f(x) = \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$. Then f is not continuous because $-\frac{1}{n} \to 0$ but $f(-\frac{1}{n}) \to -1 \neq f(0)$.

2. Let $f : \mathbb{Q} \to \mathbb{R}$ with
$$f(x) = \begin{cases} 1 & x^2 > 2 \\ 0 & x^2 < 2 \end{cases}$$
Then f is continuous. For every $a \in \mathbb{Q}$, we can find an interval about a on which f is constant. So f is continuous at a.

3. Let
$$f(x) = \begin{cases} \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$
Then $f(a)$ is discontinuous. For example, let $a_n = 1/[(2n + 0.5)\pi]$. Then $a_n \to 0$ and $f(a_n) \to 1 \neq f(0)$.

We can use this sequence definition as the definition for continuous functions. This has the advantage of being cleaner to write and easier to work with. In particular, we can reuse a lot of our sequence theorems to prove the analogous results for continuous functions.

Lemma 5.4

Let $A \subseteq \mathbb{R}$ and $f, g : A \to \mathbb{R}$ be continuous functions. Then

1. $f + g$ is continuous

2. fg is continuous

3. if g never vanishes, then f/g is continuous.

Proof.

1. Let $a \in A$ and let (a_n) be a sequence in A with $a_n \to a$. Then
$$(f + g)(a_n) = f(a_n) + g(a_n).$$
But $f(a_n) \to f(a)$ and $g(a_n) \to g(a)$. So
$$f(a_n) + g(a_n) \to f(a) + g(a) = (f + g)(a).$$

(ii) and (iii) are proved in exactly the same way. □

With this lemma, from the fact that constant functions and $f(x) = x$ are continuous, we know that all polynomials are continuous. Similarly, rational functions $P(x)/Q(x)$ are continuous except when $Q(x) = 0$.

Lemma 5.5

Let $A, B \subseteq \mathbb{R}$ and $f : A \to B$, $g : B \to \mathbb{R}$. Then if f and g are continuous, $g \circ f : A \to \mathbb{R}$ is continuous.

Proof. We offer two proofs:

1. Let (a_n) be a sequence in A with $a_n \to a \in A$. Then $f(a_n) \to f(a)$ since f is continuous. Then $g(f(a_n)) \to g(f(a))$ since g is continuous. So $g \circ f$ is continuous.

2. Let $a \in A$ and $\varepsilon > 0$. Since g is continuous at $f(a)$, there exists $\eta > 0$ such that $\forall z \in B, |z - f(a)| < \eta \Rightarrow |g(z) - g(f(a))| < \varepsilon$.

   Since f is continuous at a, $\exists \delta > 0$ such that $\forall y \in A, |y - a| < \delta \Rightarrow |f(y) - f(a)| < \eta$. Therefore $|y - a| < \delta \Rightarrow |g(f(y)) - g(f(a))| < \varepsilon$. $\qquad\square$

There are two important theorems regarding continuous functions — the maximum value theorem and the intermediate value theorem.

> **Theorem 5.6** (Maximum value theorem)
>
> Let $[a, b]$ be a closed interval in $\mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is bounded and attains its bounds, i.e. $f(x) = \sup f$ for some x, and $f(y) = \inf f$ for some y.

Proof. If f is not bounded above, then for each n, we can find $x_n \in [a, b]$ such that $f(x_n) \geq n$ for all n.

By Bolzano-Weierstrass, since $x_n \in [a, b]$ and is bounded, the sequence (x_n) has a convergent subsequence (x_{n_k}). Let x be its limit. Then since f is continuous, $f(x_{n_k}) \to f(x)$. But $f(x_{n_k}) \geq n_k \to \infty$. So this is a contradiction.

Now let $C = \sup\{f(x) : x \in [a, b]\}$. Then for every n, we can find x_n such that $f(x_n) \geq C - \frac{1}{n}$. So by Bolzano-Weierstrass, (x_n) has a convergent subsequence (x_{n_k}). Since $C - \frac{1}{n_k} \leq f(x_{n_k}) \leq C$, $f(x_{n_k}) \to C$. Therefore if $x = \lim x_{n_k}$, then $f(x) = C$.

A similar argument applies if f is unbounded below. $\qquad\square$

> **Theorem 5.7** (Intermediate value theorem)
>
> Let $a < b \in \mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Suppose that $f(a) < 0 < f(b)$. Then there exists an $x \in (a, b)$ such that $f(x) = 0$.

Proof. We have several proofs:

1. Let $A = \{x : f(x) < 0\}$ and let $s = \sup A$. We shall show that $f(s) = 0$ (this is similar to the proof that $\sqrt{2}$ exists in Numbers and Sets). If $f(s) < 0$, then setting $\varepsilon = |f(s)|$ in the definition of continuity, we can find $\delta > 0$ such that $\forall y, |y - s| < \delta \Rightarrow f(y) < 0$. Then $s + \delta/2 \in A$, so s is not an upper bound. Contradiction.

   If $f(s) > 0$, by the same argument, we can find $\delta > 0$ such that $\forall y, |y - s| < \delta \Rightarrow f(y) > 0$. So $s - \delta/2$ is a smaller upper bound.

2. Let $a_0 = a$, $b_0 = b$. By repeated bisection, construct nested intervals $[a_n, b_n]$ such that $b_n - a_n = \frac{b_0 - a_0}{2^n}$ and $f(a_n) < 0 \leq f(b_n)$. Then by the nested intervals property, we can find $x \in \cap_{n=0}^{\infty}[a_n, b_n]$. Since $b_n - a_n \to 0$, $a_n, b_n \to x$.

   Since $f(a_n) < 0$ for every n, $f(x) \leq 0$. Similarly, since $f(b_n) \geq 0$ for every n, $f(x) \geq 0$. So $f(x) = 0$. $\qquad\square$

It is easy to generalize this to get that, if $f(a) < c < f(b)$, then $\exists x \in (a, b)$ such that $f(x) = c$, by applying the result to $f(x) - c$. Also, we can assume instead that $f(b) < c < f(a)$ and obtain the same result by looking at $-f(x)$.

> **Corollary 5.8**
>
> Let $f : [a, b] \to [c, d]$ be a continuous strictly increasing function with $f(a) = c$, $f(b) = d$. Then f is invertible and its inverse is continuous.

Proof. Since f is strictly increasing, it is an injection (suppose $x \neq y$. wlog, $x < y$. Then $f(x) < f(y)$ and so $f(x) \neq f(y)$). Now let $y \in (c, d)$. By the intermediate value theorem, there exists $x \in (a, b)$ such that $f(x) = y$. So f is a surjection. So it is a bijection and hence invertible.

Let g be the inverse. Let $y \in [c,d]$ and let $\varepsilon > 0$. Let $x = g(y)$. So $f(x) = y$. Let $u = f(x - \varepsilon)$ and $v = f(x + \varepsilon)$ (if $y = c$ or d, make the obvious adjustments). Then $u < y < v$. So we can find $\delta > 0$ such that $(y - \delta, y + \delta) \subseteq (u, v)$. Then $|z - y| < \delta \Rightarrow g(z) \in (x - \varepsilon, x + \varepsilon) \Rightarrow |g(z) - g(y)| < \varepsilon$. $\qquad \square$

With this corollary, we can create more continuous functions, e.g. $\sqrt{x}$.

§5.2 Continuous induction*

Continuous induction is a generalization of induction on natural numbers. It provides an alternative mechanism to prove certain results we have shown.

> **Claim 5.9** (Continuous induction v1)
>
> Let $a < b$ and let $A \subseteq [a, b]$ have the following properties:
>
> 1. $a \in A$
>
> 2. If $x \in A$ and $x \neq b$, then $\exists y \in A$ with $y > x$.
>
> 3. If $\forall \varepsilon > 0$, $\exists y \in A : y \in (x - \varepsilon, x]$, then $x \in A$.
>
> Then $b \in A$.

Proof. Since $a \in A$, $A \neq \varnothing$. A is also bounded above by b. So let $s = \sup A$. Then $\forall \varepsilon > 0$, $\exists y \in A$ such that $y > s - \varepsilon$. Therefore, by (iii), $s \in A$.

If $s \neq b$, then by (ii), we can find $y \in A$ such that $y > s$. $\qquad \square$

It can also be formulated as follows:

> **Claim 5.10** (Continuous induction v2)
>
> Let $A \subseteq [a, b]$ and suppose that
>
> 1. $a \in A$
>
> 2. If $[a, x] \subseteq A$ and $x \neq b$, then there exists $y > x$ such that $[a, y] \subseteq A$.
>
> 3. If $[a, x) \subseteq A$, then $[a, x] \subseteq A$.
>
> Then $A = [a, b]$

Proof. We prove that version 1 $\Rightarrow$ version 2. Suppose A satisfies the conditions of v2. Let $A' = \{x \in [a, b] : [a, x] \subseteq A\}$.

Then $a \in A'$. If $x \in A'$ with $x \neq b$, then $[a, x] \subseteq A$. So $\exists y > x$ such that $[a, y] \subseteq A$. So $\exists y > x$ such that $y \in A'$.

If $\forall \varepsilon > 0, \exists y \in (x - \varepsilon, x]$ such that $[a, y] \subseteq A$, then $[a, x) \subseteq A$. So by (iii), $[a, x] \subseteq A$, so $x \in A'$. So A' satisfies properties (i) to (iii) of version 1. Therefore $b \in A'$. So $[a, b] \subseteq A$. So $A = [a, b]$. $\qquad \square$

We reprove intermediate value theorem here:

> **Theorem 5.11** (Intermediate value theorem)
>
> Let $a < b \in \mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Suppose that $f(a) < 0 < f(b)$. Then there exists an $x \in (a, b)$ such that $f(x) = 0$.

Proof. Assume that f is continuous. Suppose $f(a) < 0 < f(b)$. Assume that $(\forall x)\, f(x) \neq 0$, and derive a contradiction.

Let $A = \{x : f(x) < 0\}$ Then $a \in A$. If $x \in A$, then $f(x) < 0$, and by continuity, we can find $\delta > 0$ such that $|y - x| < \delta \Rightarrow f(y) < 0$. So if $x \neq b$, then we can find $y \in A$ such that $y > x$.

We prove the contrapositive of the last condition, i.e.

$$x \notin A \Rightarrow (\exists \delta > 0)(\forall y \in A)\, y \notin (x - \delta, x].$$

If $x \notin A$, then $f(x) > 0$ (we assume that f is never zero. If not, we're done). Then by continuity, $\exists \delta > 0$ such that $|y - x| < \delta \Rightarrow f(y) > 0$. So $y \notin A$.

Hence by continuous induction, $b \in A$. Contradiction. $\qquad \square$

Now we prove that continuous functions in closed intervals are bounded.

Theorem 5.12

Let $[a, b]$ be a closed interval in $\mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is bounded.

Proof. Let $f : [a, b]$ be continuous. Let $A = \{x : f$ is bounded on $[a, x]\}$. Then $a \in A$. If $x \in A, x \neq b$, then $\exists \delta > 0$ such that $|y - x| < \delta \Rightarrow |f(y) - f(x)| < 1$. So $\exists y > x$ (e.g. take $\min\{x + \delta/2, b\}$) such that f is bounded on $[a, y]$, which implies that $y \in A$.

Now suppose that $\forall \varepsilon > 0, \exists y \in (x, -\varepsilon, x]$ such that $y \in A$. Again, we can find $\delta > 0$ such that f is bounded on $(x - \delta, x + \delta)$, and in particular on $(x - \delta, x]$. Pick y such that f is bounded on $[a, y]$ and $y > x - \delta$. Then f is bounded on $[a, x]$. So $x \in A$.

So we are done by continuous induction. $\qquad \square$

Finally, we can prove a theorem that we have not yet proven.

Definition 5.2 (Cover of a set)

Let $A \subseteq \mathbb{R}$. A *cover* of A by open intervals is a set $\{I_\gamma : \gamma \in \Gamma\}$ where each I_γ is an open interval and $A \subseteq \bigcup_{\gamma \in \Gamma} I_\gamma$.

A *finite subcover* is a finite subset $\{I_{\gamma_1}, \cdots, I_{\gamma_n}\}$ of the cover that is still a cover.

Not every cover has a finite subcover. For example, the cover $\{(\frac{1}{n}, 1) : n \in \mathbb{N}\}$ of $(0, 1)$ has no finite subcover.

Theorem 5.13 (Heine-Borel*)

Every cover of a closed, bounded interval $[a, b]$ by open intervals has a finite subcover. We say closed intervals are *compact* (cf. Metric and Topological Spaces).

Proof. Let $\{I_\gamma : \gamma \in \Gamma\}$ be a cover of $[a, b]$ by open intervals. Let $A = \{x : [a, x]$ can be covered by finitely many of the $I_\gamma\}$.

Then $a \in A$ since a must belong to some I_γ.

If $x \in A$, then pick γ such that $x \in I_\gamma$. Then if $x \neq b$, since I_γ is an open interval, it contains $[x, y]$ for some $y > x$. Then $[a, y]$ can be covered by finitely many I_γ, by taking a finite cover for $[a, x]$ and the I_γ that contains x.

Now suppose that $\forall \varepsilon > 0, \exists y \in A$ such that $y \in (x - \varepsilon, x]$.

Let I_γ be an open interval containing x. Then it contains $(x - \varepsilon, x]$ for some $\varepsilon > 0$. Pick $y \in A$ such that $y \in (x - \varepsilon, x]$. Now combine I_γ with a finite subcover of $[a, y]$ to get a finite subcover of $[a, x]$. So $x \in A$.

Then done by continuous induction. $\qquad \square$

We can use Heine-Borel to prove that continuous functions on $[a, b]$ are bounded.

Theorem 5.14

Let $[a, b]$ be a closed interval in $\mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is bounded and attains it bounds, i.e. $f(x) = \sup f$ for some x, and $f(y) = \inf f$ for some y.

Proof. Let $f : [a, b] \to \mathbb{R}$ be continuous. Then by continuity,

$$(\forall x \in [a, b])(\exists \delta_x > 0)(\forall y)\, |y - x| < \delta_x \Rightarrow |f(y) - f(x)| < 1.$$

Let $\gamma = [a, b]$ and for each $x \in \gamma$, let $I_x = (x - \delta_x, x + \delta_x)$. So by Heine-Borel, we can find $x_1, \cdots, x_n$ such that $[a, b] \subseteq \bigcup_1^n (x_i - \delta_{x_i}, x_i + \delta_{x_i})$.

But f is bounded in each interval $(x_i - \delta_{x_i}, x_i + \delta_{x_i})$ by $|f(x_i)| + 1$. So it is bounded on $[a, b]$ by $\max |f(x_i)| + 1$. $\qquad\square$

§6 Differentiability

In the remainder of the course, we will properly develop calculus, and put differentiation and integration on a rigorous foundation. Every notion will be given a proper definition which we will use to prove results like the product and quotient rule.

§6.1 Limits

First of all, we need the notion of limits. Recall that we've previously had limits for *sequences*. Now, we will define limits for functions.

> **Definition 6.1** (Limit of functions)
>
> Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$. We say
> $$\lim_{x \to a} f(x) = \ell,$$
> or "$f(x) \to \ell$ as $x \to a$", if
> $$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in A)\, 0 < |x - a| < \delta \Rightarrow |f(x) - \ell| < \varepsilon.$$
> We couldn't care less what happens when $x = a$, hence the strict inequality $0 < |x - a|$. In fact, f doesn't even have to be defined at $x = a$.

> **Example 6.1**
>
> Let
> $$f(x) = \begin{cases} x & x \neq 2 \\ 3 & x = 2 \end{cases}$$
> Then $\lim_{x \to 2} = 2$, even though $f(2) = 3$.

> **Example 6.2**
>
> Let $f(x) = \frac{\sin x}{x}$. Then $f(0)$ is not defined but $\lim_{x \to 0} f(x) = 1$.
>
> We will see a proof later after we define what $\sin$ means.

We notice that the definition of the limit is suspiciously similar to that of continuity. In fact, if we define
$$g(x) = \begin{cases} f(x) & x \neq a \\ \ell & x = a \end{cases}$$

Then $f(x) \to \ell$ as $x \to a$ iff g is continuous at a.

Alternatively, f is continuous at a if $f(x) \to f(a)$ as $x \to a$. It follows also that $f(x) \to \ell$ as $x \to a$ iff $f(x_n) \to \ell$ for every sequence (x_n) in A with $x_n \to a$.

The previous limit theorems of sequences apply here as well

> **Claim 6.3**
>
> If $f(x) \to \ell$ and $g(x) \to m$ as $x \to a$, then $f(x) + g(x) \to \ell + m$, $f(x)g(x) \to \ell m$, and $\frac{f(x)}{g(x)} \to \frac{\ell}{m}$ if g and m don't vanish.

§6.2 Differentiation

Similar to what we did in IA Differential Equations, we define the derivative as a limit.

Definition 6.2 (Differentiable function)

f is *differentiable* at a with derivative λ if

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

Equivalently, if

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

We write $\lambda = f'(a)$.

Here we see why, in the definition of the limit, we say that we don't care what happens when $x = a$. In our definition here, our function is $0/0$ when $x = a$, and we can't make any sense out of what happens when $x = a$.

Alternatively, we write the definition of differentiation as

$$\frac{f(x + h) - f(x)}{h} = f'(x) + \varepsilon(h),$$

where $\varepsilon(h) \to 0$ as $h \to 0$. Rearranging, we can deduce that

$$f(x + h) = f(x) + h f'(x) + h \varepsilon(h),$$

Note that by the definition of the limit, we don't have to care what value ε takes when $h = 0$. It can be 0, π or $10^{10^{10}}$. However, we usually take $\varepsilon(0) = 0$ so that ε is continuous.

Using the small-o notation, we usually write $o(h)$ for a function that satisfies $\frac{o(h)}{h} \to 0$ as $h \to 0$. Hence we have

---

Claim 6.4

$$f(x + h) = f(x) + h f'(x) + o(h).$$

We can interpret this as an approximation of $f(x + h)$:

$$f(x + h) = \underbrace{f(x) + h f'(x)}_{\text{linear approximation}} + \underbrace{o(h)}_{\text{error term}}.$$

And differentiability shows that this is a very good approximation with small $o(h)$ error.

Conversely, we have

Claim 6.5

If $f(x + h) = f(x) + h f'(x) + o(h)$, then f is differentiable at x with derivative $f'(x)$.

Proof.

$$\frac{f(x + h) - f(x)}{h} = f'(x) + \frac{o(h)}{h} \to f'(x). \qquad \square$$

We can take derivatives multiple times, and get multiple derivatives.

Definition 6.3 (Multiple derivatives)

This is defined recursively: f is $(n + 1)$-times differentiable if it is n-times differentiable and its nth derivative $f^{(n)}$ is differentiable. We write $f^{(n+1)}$ for the derivative of $f^{(n)}$, i.e. the $(n + 1)$th derivative of f.

Informally, we will say f is n-times differentiable if we can differentiate it n times, and the nth derivative is $f^{(n)}$.

We can prove the usual rules of differentiation using the small o-notation. It can also be proven by considering limits directly, but the notation will become a bit more daunting.

> **Lemma 6.6** (Sum and product rule)
>
> Let f, g be differentiable at x. Then $f + g$ and fg are differentiable at x, with
>
> $$(f + g)'(x) = f'(x) + g'(x)$$
> $$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

Proof.

$$
\begin{aligned}
(f + g)(x + h) &= f(x + h) + g(x + h) \\
&= f(x) + hf'(x) + o(h) + g(x) + hg'(x) + o(h) \\
&= (f + g)(x) + h(f'(x) + g'(x)) + o(h) \\
fg(x + h) &= f(x + h)g(x + h) \\
&= [f(x) + hf'(x) + o(h)][g(x) + hg'(x) + o(h)] \\
&= f(x)g(x) + h[f'(x)g(x) + f(x)g'(x)] \\
&\quad + \underbrace{o(h)[g(x) + f(x) + hf'(x) + hg'(x) + o(h)] + h^2 f'(x)g'(x)}_{\text{error term}}
\end{aligned}
$$

By limit theorems, the error term is $o(h)$. So we can write this as

$$= fg(x) + h(f'(x)g(x) + f(x)g'(x)) + o(h). \qquad \square$$

> **Lemma 6.7** (Chain rule)
>
> If f is differentiable at x and g is differentiable at $f(x)$, then $g \circ f$ is differentiable at x with derivative $g'(f(x))f'(x)$.

Proof. If one is sufficiently familiar with the small-o notation, then we can proceed as

$$g(f(x + h)) = g(f(x) + hf'(x) + o(h)) = g(f(x)) + hf'(x)g'(f(x)) + o(h).$$

If not, we can be a bit more explicit about the computations, and use $h\varepsilon(h)$ instead of $o(h)$:

$$
\begin{aligned}
(g \circ f)(x + h) &= g(f(x + h)) \\
&= g[f(x) + \underbrace{hf'(x) + h\varepsilon_1(h)}_{\text{the ``}h\text{'' term}}] \\
&= g(f(x)) + \big(fg'(x) + h\varepsilon_1(h)\big)g'(f(x)) \\
&\quad + \big(hf'(x) + h\varepsilon_1(h)\big)\varepsilon_2\big(hf'(x) + h\varepsilon_1(h)\big) \\
&= g \circ f(x) + hg'(f(x))f'(x) \\
&\quad + h\underbrace{\left[\varepsilon_1(h)g'(f(x)) + \big(f'(x) + \varepsilon_1(h)\big)\varepsilon_2\big(hf'(x) + h\varepsilon_1(h)\big)\right]}_{\text{error term}}.
\end{aligned}
$$

We want to show that the error term is $o(h)$, i.e. it divided by h tends to 0 as $h \to 0$.

But $\varepsilon_1(h)g'(f(x)) \to 0$, $f'(x) + \varepsilon_1(h)$ is bounded, and $\varepsilon_2(hf'(x) + h\varepsilon_1(h)) \to 0$ because $hf'(x) + h\varepsilon_1(h) \to 0$ and $\varepsilon_2(0) = 0$. So our error term is $o(h)$. $\qquad \square$

We usually don't write out the error terms so explicitly, and just use heuristics like $f(x + o(h)) = f(x) + o(h)$; $o(h) + o(h) = o(h)$; and $g(x) \cdot o(h) = o(h)$ for any (bounded) function g.

Example 6.8

1. Constant functions are differentiable with derivative 0.

2. $f(x) = \lambda x$ is differentiable with derivative λ.

3. Using the product rule, we can show that x^n is differentiable with derivative nx^{n-1} by induction.

4. Hence all polynomials are differentiable.

Example 6.9

Let $f(x) = 1/x$. If $x \neq 0$, then

$$\frac{f(x+h) - f(x)}{h} = \frac{\frac{1}{x+h} - \frac{1}{x}}{h} = \frac{\left(\frac{-h}{x(x+h)}\right)}{h} = \frac{-1}{x(x+h)} \to \frac{-1}{x^2}$$

by limit theorems.

Lemma 6.10 (Quotient rule)

If f and g are differentiable at x, and $g(x) \neq 0$, then f/g is differentiable at x with derivative

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

Proof. First note that $1/g(x) = h(g(x))$ where $h(y) = 1/y$. So $1/g(x)$ is differentiable at x with derivative $\frac{-1}{g(x)^2} g'(x)$ by the chain rule.

By the product rule, f/g is differentiable at x with derivative

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

Lemma 6.11

If f is differentiable at x, then it is continuous at x.

Proof. As $y \to x$, $\dfrac{f(y) - f(x)}{y - x} \to f'(x)$. Since, $y - x \to 0$, $f(y) - f(x) \to 0$ by product theorem of limits. So $f(y) \to f(x)$. So f is continuous at x. $\qquad \square$

Theorem 6.12

Let $f : [a, b] \to [c, d]$ be differentiable on (a, b), continuous on $[a, b]$, and strictly increasing. Suppose that $f'(x)$ never vanishes. Suppose further that $f(a) = c$ and $f(b) = d$. Then f has an inverse g and for each $y \in (c, d)$, g is differentiable at y with derivative $1/f'(g(y))$.

In human language, this states that if f is invertible, then the derivative of f^{-1} is $1/f'$.

Note that the conditions will (almost) always require f to be differentiable on *open* interval (a, b), continuous on *closed* interval $[a, b]$. This is because it doesn't make sense to talk about differentiability at a or b since the definition of $f'(a)$ requires f to be defined on both sides of a.

Proof. g exists by an earlier theorem about inverses of continuous functions.

Let $y, y + k \in (c, d)$. Let $x = g(y)$, $x + h = g(y + k)$.

Since $g(y + k) = x + h$, we have $y + k = f(x + h)$. So $k = f(x + h) - y = f(x + h) - f(x)$. So

$$\frac{g(y + k) - g(y)}{k} = \frac{(x + h) - x}{f(x + h) - f(x)} = \left(\frac{f(x + h) - f(x)}{h}\right)^{-1}.$$

As $k \to 0$, since g is continuous, $g(y + k) \to g(y)$. So $h \to 0$. So

$$\frac{g(y + k) - g(y)}{k} \to [f'(x)]^{-1} = [f'(g(y))]^{-1}. \qquad \square$$

Example 6.13

Let $f(x) = x^{1/2}$ for $x > 0$. Then f is the inverse of $g(x) = x^2$. So

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

Similarly, we can show that the derivative of $x^{1/q}$ is $\frac{1}{q}x^{1/q - 1}$.

Then let's take $x^{p/q} = (x^{1/q})^p$. By the chain rule, its derivative is

$$p(x^{1/q})^{p-1} \cdot \frac{1}{q}x^{1/q - 1} = \frac{p}{q}x^{\frac{p-1}{q} + \frac{1}{q} - 1} = \frac{p}{q}x^{\frac{p}{q} - 1}.$$

---

§6.3 Differentiation theorems

Everything we've had so far is something we already know. It's just that now we can prove them rigorously. In this section, we will come up with genuinely new theorems, including but not limited to *Taylor's theorem*, which gives us Taylor's series.

Theorem 6.14 (Rolle's theorem)

Let f be continuous on a closed interval $[a, b]$ (with $a < b$) and differentiable on (a, b). Suppose that $f(a) = f(b)$. Then there exists $x \in (a, b)$ such that $f'(x) = 0$.

It is intuitively obvious: if you move up and down, and finally return to the same point, then you must have changed direction some time. Then $f'(x) = 0$ at that time.

Proof. If f is constant, then we're done.

Otherwise, there exists u such that $f(u) \neq f(a)$. wlog, $f(u) > f(a)$. Since f is continuous, it has a maximum, and since $f(u) > f(a) = f(b)$, the maximum is not attained at a or b.

Suppose maximum is attained at $x \in (a, b)$. Then for any $h \neq 0$, we have

$$\frac{f(x + h) - f(x)}{h} \begin{cases} \leq 0 & h > 0 \\ \geq 0 & h < 0 \end{cases}$$

since $f(x + h) - f(x) \leq 0$ by maximality of $f(x)$. By considering both sides as we take the limit $h \to 0$, we know that $f'(x) \leq 0$ and $f'(x) \geq 0$. So $f'(x) = 0$. $\qquad \square$

Corollary 6.15 (Mean value theorem)

Let f be continuous on $[a, b]$ ($a < b$), and differentiable on (a, b). Then there exists $x \in (a, b)$ such that

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

Note that $\frac{f(b) - f(a)}{b - a}$ is the slope of the line joining $f(a)$ and $f(b)$.

The mean value theorem is sometimes described as "rotate your head and apply Rolle's". However, if we actually rotate it, we might end up with a non-function. What we actually want is a shear.

Proof. Let

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

Then

$$g(b) - g(a) = f(b) - f(a) - \frac{f(b) - f(a)}{b - a}(b - a) = 0.$$

So by Rolle's theorem, we can find $x \in (a, b)$ such that $g'(x) = 0$. So

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

as required.   □

We've always assumed that if a function has a positive derivative everywhere, then the function is increasing. However, it turns out that this is really hard to prove directly. It does, however, follow quite immediately from the mean value theorem.

Example 6.16

Suppose $f'(x) > 0$ for every $x \in (a, b)$. Then for u, v in $[a, b]$, we can find $w \in (u, v)$ such that

$$\frac{f(v) - f(u)}{v - u} = f'(w) > 0.$$

It follows that $f(v) > f(u)$. So f is strictly increasing.
    Similarly, if $f'(x) \geq 2$ for every x and $f(0) = 0$, then $f(1) \geq 2$, or else we can find $x \in (0, 1)$ such that

$$2 \leq f'(x) = \frac{f(1) - f(0)}{1 - 0} = f(1).$$

Theorem 6.17 (Local version of inverse function theorem)

Let f be a function with continuous derivative on (a, b).
    Let $x \in (a, b)$ and suppose that $f'(x) \neq 0$. Then there is an open interval (u, v) containing x on which f is invertible (as a function from (u, v) to $f((u, v))$). Moreover, if g is the inverse, then $g'(f(z)) = \frac{1}{f'(z)}$ for every $z \in (u, v)$.
    This says that if f has a non-zero derivative, then it has an inverse locally and the derivative of the inverse is $1/f'$.

Note that this not only requires f to be differentiable, but the derivative itself also has to be continuous.

Proof. wlog, $f'(x) > 0$. By the continuity, of f', we can find $\delta > 0$ such that $f'(z) > 0$ for every $z \in (x - \delta, x + \delta)$. By the mean value theorem, f is strictly increasing on $(x - \delta, x + \delta)$, hence injective. Also, f is continuous on $(x - \delta, x + \delta)$ by differentiability.
    Then done by the inverse function theorem.   □

Finally, we are going to prove Taylor's theorem. To do so, we will first need some lemmas.

Theorem 6.18 (Higher-order Rolle's theorem)

Let f be continuous on $[a, b]$ ($a < b$) and n-times differentiable on an open interval containing $[a, b]$. Suppose that

$$f(a) = f'(a) = f^{(2)}(a) = \cdots = f^{(n-1)}(a) = f(b) = 0.$$

Then $\exists x \in (a, b)$ such that $f^{(n)}(x) = 0$.

Proof. Induct on n. The $n = 0$ base case is just Rolle's theorem.

Suppose we have $k < n$ and $x_k \in (a, b)$ such that $f^{(k)}(x_k) = 0$. Since $f^{(k)}(a) = 0$, we can find $x_{k+1} \in (a, x_k)$ such that $f^{(k+1)}(x_{k+1}) = 0$ by Rolle's theorem.

So the result follows by induction. $\qquad\square$

Corollary 6.19

Suppose that f and g are both differentiable on an open interval containing $[a, b]$ and that $f^{(k)}(a) = g^{(k)}(a)$ for $k = 0, 1, \cdots, n - 1$, and also $f(b) = g(b)$. Then there exists $x \in (a, b)$ such that $f^{(n)}(x) = g^{(n)}(x)$.

Proof. Apply generalised Rolle's to $f - g$. $\qquad\square$

Now we shall show that for any f, we can find a polynomial p of degree at most n that satisfies the conditions for g, i.e. a p such that $p^{(k)}(a) = f^{(k)}(a)$ for $k = 0, 1, \cdots, n - 1$ and $p(b) = f(b)$.

A useful ingredient is the observation that if

$$Q_k(x) = \frac{(x - a)^k}{k!},$$

then

$$Q_k^{(j)}(a) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

Therefore, if

$$Q(x) = \sum_{k=0}^{n-1} f^{(k)}(a) Q_k(x),$$

then

$$Q^{(j)}(a) = f^{(j)}(a)$$

for $j = 0, 1, \cdots, n - 1$. To get $p(b) = f(b)$, we use our nth degree polynomial term:

$$p(x) = Q(x) + \frac{(x - a)^n}{(b - a)^n}\big(f(b) - Q(b)\big).$$

Then our final term does not mess up our first $n - 1$ derivatives, and gives $p(b) = f(b)$.

By the previous corollary, we can find $x \in (a, b)$ such that

$$f^{(n)}(x) = p^{(n)}(x).$$

That is,

$$f^{(n)}(x) = \frac{n!}{(b - a)^n}\big(f(b) - Q(b)\big).$$

Therefore

$$f(b) = Q(b) + \frac{(b - a)^n}{n!} f^{(n)}(x).$$

Alternatively,

$$f(b) = f(a) + (b - a)f'(a) + \cdots + \frac{(b - a)^{n-1}}{(n - 1)!} f^{(n-1)}(a) + \frac{(b - a)^n}{n!} f^{(n)}(x).$$

Setting $b = a + h$, we can rewrite this as

Theorem 6.20 (Taylor's theorem with the Lagrange form of remainder)

$$f(a + h) = \underbrace{f(a) + hf'(a) + \cdots + \frac{h^{n-1}}{(n - 1)!} f^{(n-1)}(a)}_{(n-1)\text{-degree approximation to } f \text{ near } a} + \underbrace{\frac{h^n}{n!} f^{(n)}(x)}_{\text{error term}}.$$

for some $x \in (a, a + h)$.

Strictly speaking, we only proved it for the case when $h > 0$, but we can easily show it holds for $h < 0$ too by considering $g(x) = f(-x)$.

Note that the remainder term is *not* necessarily small, but this often gives us the best $(n-1)$-degree approximation to f near a. For example, if $f^{(n)}$ is bounded by C near a, then

$$\left| \frac{h^n}{n!} f^{(n)}(x) \right| \leq \frac{C}{n!} |h|^n = o(h^{n-1}).$$

Example 6.21

Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $f(0) = 1$ and $f'(x) = f(x)$ for every x (intuitively, we know it is e^x, but that thing doesn't exist!). Then for every x, we have

$$f(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

While it seems like we can prove this works by differentiating it and see that $f'(x) = f(x)$, the sum rule only applies for finite sums. We don't know we can differentiate a sum term by term. So we have to use Taylor's theorem.

Since $f'(x) = f(x)$, it follows that all derivatives exist. By Taylor's theorem,

$$f(x) = f(0) + f'(0)x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1} + \frac{f^{(n)}(u)}{n!}x^n.$$

for some u between 0 and x. This equals to

$$f(x) = \sum_{k=0}^{n-1} \frac{x^k}{k!} + \frac{f^{(n)}(u)}{n!}x^n.$$

We must show that the remainder term $\frac{f^{(n)}(u)}{n!}x^n \to 0$ as $n \to \infty$. Note here that x is fixed, but u can depend on n.

But we know that $f^{(n)}(u) = f(u)$, but since f is differentiable, it is continuous, and is bounded on $[0, x]$. Suppose $|f(u)| \leq C$ on $[0, x]$. Then

$$\left| \frac{f^{(n)(u)}}{n!}x^n \right| \leq \frac{C}{n!} |x|^n \to 0$$

from limit theorems. So it follows that

$$f(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

---

§6.4 Complex differentiation

Definition 6.4 (Complex differentiability)

Let $f : \mathbb{C} \to \mathbb{C}$. Then f is differentiable at z with derivative $f'(z)$ if

$$\lim_{h \to 0} \frac{f(z+h) - f(z)}{h} \text{ exists and equals } f'(z).$$

Equivalently,

$$f(z+h) = f(z) + hf'(z) + o(h).$$

This is exactly the same definition with real differentiation, but has very different properties!

All the usual rules — chain rule, product rule etc. also apply (with the same proofs). Also the derivatives of polynomials are what you expect. However, there are some more interesting cases.

> **Example 6.22**
>
> $f(z) = \bar{z}$ is *not* differentiable.
> $$\frac{\overline{z+h} - \bar{z}}{h} = \frac{\bar{h}}{h} = \begin{cases} 1 & h \text{ is real} \\ -1 & h \text{ is purely imaginary} \end{cases}$$

If this seems weird, this is because we often think of $\mathbb{C}$ as $\mathbb{R}^2$, but they are not the same. For example, reflection is a linear map in $\mathbb{R}^2$, but not in $\mathbb{C}$. A linear map in $\mathbb{C}$ is something in the form $x \mapsto bx$, which can only be a dilation or rotation, not reflections or other weird things.

> **Example 6.23**
>
> $f(z) = |z|$ is also not differentiable. If it were, then $|z|^2$ would be as well (by the product rule). So would $\frac{|z|^2}{z} = \bar{z}$ when $z \neq 0$ by the quotient rule. At $z = 0$, it is certainly not differentiable, since it is not even differentiable on $\mathbb{R}$.

§7 Complex power series

Before we move on to integration, we first have a look at complex power series. This will allow us to define the familiar exponential and trigonometric functions.

> **Definition 7.1** (Complex power series)
>
> A *complex power series* is a series of the form
> $$\sum_{n=0}^{\infty} a_n z^n.$$
> when $z \in \mathbb{C}$ and $a_n \in \mathbb{C}$ for all n. When it converges, it is a function of z.

When considering complex power series, a very important concept is the *radius of convergence*. To make sense of this concept, we first need the following lemma:

> **Lemma 7.1**
>
> Suppose that $\sum a_n z^n$ converges and $|w| < |z|$, then $\sum a_n w^n$ converges (absolutely).

Proof. We know that
$$|a_n w^n| = |a_n z^n| \cdot \left|\frac{w}{z}\right|^n.$$

Since $\sum a_n z^n$ converges, the terms $a_n z^n$ are bounded. So pick C such that
$$|a_n z^n| \leq C$$

for every n. Then
$$0 \leq \sum_{n=0}^{\infty} |a_n w^n| \leq \sum_{n=0}^{\infty} C \left|\frac{w}{z}\right|^n,$$

which converges (geometric series). So by the comparison test, $\sum a_n w^n$ converges absolutely. $\square$

It follows that if $\sum a_n z^n$ does not converge and $|w| > |z|$, then $\sum a_n w^n$ does not converge.

Now let $R = \sup\{|z| : \sum a_n z^n \text{ converges }\}$ (R may be infinite). If $|z| < R$, then we can find z_0 with $|z_0| \in (|z|, R]$ such that $\sum_n^{\infty} a_n z_0^n$ converges. So by lemma above, $\sum a_n z^n$ converges. If $|z| > R$, then $\sum a_n z^n$ diverges by definition of R.

> **Definition 7.2** (Radius of convergence)
>
> The *radius of convergence* of a power series $\sum a_n z^n$ is
>
> $$R = \sup \left\{ |z| : \sum a_n z^n \text{ converges} \right\}.$$
>
> $\{z : |z| < R\}$ is called the *circle of convergence*.[a]
>
> If $|z| < R$, then $\sum a_n z^n$ converges. If $|z| > R$, then $\sum a_n z^n$ diverges. When $|z| = R$, the series can converge at some points and not the others.
>
> ―――――――――――
>
> [a]Note to pedants: yes it is a disc, not a circle

> **Example 7.2**
>
> $\sum\limits_{n=0}^{\infty} z^n$ has radius of convergence of 1. When $|z| = 1$, it diverges (since the terms do not tend to 0).

> **Example 7.3**
>
> $\sum\limits_{n=0}^{\infty} \dfrac{z^n}{n}$ has radius of convergence 1, since the ratio of $(n+1)$th term to nth is
>
> $$\frac{z^{n+1}/(n+1)}{z^n/n} = z \cdot \frac{n}{n+1} \to z.$$
>
> So if $|z| < 1$, then the series converges by the ratio test. If $|z| > 1$, then eventually the terms are increasing in modulus.
>
> If $z = 1$, then it diverges (harmonic series). If $|z| = 1$ and $z \neq 1$, it converges by Abel's test.

> **Example 7.4**
>
> The series $\sum\limits_{n=1}^{\infty} \dfrac{z^n}{n^2}$ converges for $|z| \leq 1$ and diverges for $|z| > 1$.

As evidenced by the above examples, the ratio test can be used to find the radius of convergence. We also have an alternative test based on the nth root.

> **Lemma 7.5**
>
> The radius of convergence of a power series $\sum a_n z^n$ is
>
> $$R = \frac{1}{\limsup \sqrt[n]{|a_n|}}.$$
>
> Often $\sqrt[n]{|a_n|}$ converges, so we only have to find the limit.

Proof. Suppose $|z| < 1/\limsup \sqrt[n]{|a_n|}$. Then $|z| \limsup \sqrt[n]{|a_n|} < 1$. Therefore there exists N and $\varepsilon > 0$ such that

$$\sup_{n \geq N} |z| \sqrt[n]{|a_n|} \leq 1 - \varepsilon$$

by the definition of $\limsup$. Therefore

$$|a_n z^n| \leq (1 - \varepsilon)^n$$

for every $n \geq N$, which implies (by comparison with geometric series) that $\sum a_n z^n$ converges absolutely.

On the other hand, if $|z| \limsup \sqrt[n]{|a_n|} > 1$, it follows that $|z| \sqrt[n]{|a_n|} \geq 1$ for infinitely many n. Therefore $|a_n z^n| \geq 1$ for infinitely many n. So $\sum a_n z^n$ does not converge. $\qquad\square$

> **Example 7.6**
>
> The radius of convergence of $\dfrac{z^n}{2^n}$ is 2 because $\sqrt[n]{|a_n|} = \frac{1}{2}$ for every n. So $\limsup \sqrt[n]{|a_n|} = \frac{1}{2}$. So $1/\limsup \sqrt[n]{|a_n|} = 2$.

But often it is easier to find the radius convergence from elementary methods such as the ratio test, e.g. for $\sum n^2 z^n$.

§7.1 Exponential and trigonometric functions

> **Definition 7.3** (Exponential function)
> The *exponential function* is
> $$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$
> By the ratio test, this converges on all of $\mathbb{C}$.

A fundamental property of this function is that

$$e^{z+w} = e^z e^w.$$

Once we have this property, we can say that

> **Claim 7.7**
> The derivative of e^z is e^z.

Proof.

$$\frac{e^{z+h} - e^z}{h} = e^z \left(\frac{e^h - 1}{h} \right)$$
$$= e^z \left(1 + \frac{h}{2!} + \frac{h^2}{3!} + \cdots \right)$$

But

$$\left| \frac{h}{2!} + \frac{h^2}{3!} + \cdots \right| \le \frac{|h|}{2} + \frac{|h|^2}{4} + \frac{|h|^3}{8} + \cdots = \frac{|h|/2}{1 - |h|/2} \to 0.$$

So

$$\frac{e^{z+h} - e^z}{h} \to e^z. \qquad \square$$

But we must still prove that $e^{z+w} = e^z e^w$.

Consider two sequences $(a_n), (b_n)$. Their *convolution* is the sequence (c_n) defined by

$$c_n = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0.$$

The relevance of this is that if you take

$$\left(\sum_{n=0}^{N} a_n z^n \right) \left(\sum_{n=0}^{N} b_n z^n \right) \text{ and } \sum_{n=0}^{N} c_n z^n,$$

and equate coefficients of z^n, you get

$$c_n = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0.$$

> **Theorem 7.8**
>
> Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ be two absolutely convergent series, and let (c_n) be the convolution of the sequences (a_n) and (b_n). Then $\sum_{n=0}^{\infty} c_n$ converges (absolutely), and
>
> $$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n \right) \left(\sum_{n=0}^{\infty} b_n \right).$$

Proof. We first show that a rearrangement of $\sum c_n$ converges absolutely. Hence it converges unconditionally, and we can rearrange it back to $\sum c_n$.

Consider the series

$$(a_0 b_0) + (a_0 b_1 + a_1 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_2 + a_2 b_2 + a_2 b_1 + a_2 b_0) + \cdots \qquad (*)$$

Let

$$S_N = \sum_{n=0}^{N} a_n, \quad T_N = \sum_{n=0}^{N} b_n, \quad U_N = \sum_{n=0}^{N} |a_n|, \quad V_N = \sum_{n=0}^{N} |b_n|.$$

Also let $S_N \to S, T_N \to T, U_N \to U, V_N \to V$ (these exist since $\sum a_n$ and $\sum b_n$ converge absolutely).

If we take the modulus of the terms of $(*)$, and consider the first $(N+1)^2$ terms (i.e. the first $N+1$ brackets), the sum is $U_N V_N$. Hence the series converges absolutely to UV. Hence $(*)$ converges.

The partial sum up to $(N+1)^2$ of the series $(*)$ itself is $S_N T_N$, which converges to ST. So the whole series converges to ST.

Since it converges absolutely, it converges unconditionally. Now consider a rearrangement:

$$a_0 b_0 + (a_0 b_1 + a_1 b_0) + (a_0 b_2 + a_1 b_1 + a_2 b_0) + \cdots$$

Then this converges to ST as well. But the partial sum of the first $1 + 2 + \cdots + N$ terms is $c_0 + c_1 + \cdots + c_N$. So

$$\sum_{n=0}^{N} c_n \to ST = \left(\sum_{n=0}^{\infty} a_n \right) \left(\sum_{n=0}^{\infty} b_n \right). \qquad \square$$

> **Corollary 7.9**
>
> $$e^z e^w = e^{z+w}.$$

Proof. By theorem above (and definition of e^z),

$$e^z e^w = \sum_{n=0}^{\infty} \left(1 \cdot \frac{w^n}{n!} + \frac{z}{1!} \frac{w^{n-1}}{(n-1)!} + \frac{z^2}{2!} \frac{w^{n-2}}{(n-2)!} + \cdots + \frac{z^n}{n!} \cdot 1 \right)$$

$$e^z e^w = \sum_{n=0}^{\infty} \frac{1}{n!} \left(w^n + \binom{n}{1} z w^{n-1} + \binom{n}{2} z^2 w^{n-2} + \cdots + \binom{n}{n} z^n \right)$$

$$= \sum_{n=0}^{\infty} (z + w)^n \text{ by the binomial theorem}$$

$$= e^{z+w}. \qquad \square$$

Note that if (c_n) is the convolution of (a_n) and (b_n), then the convolution of $(a_n z^n)$ and $(b_n z^n)$ is $(c_n z^n)$. Therefore if both $\sum a_n z^n$ and $\sum b_n z^n$ converge absolutely, then their product is $\sum c_n z^n$.

Note that we have now completed the proof that the derivative of e^z is e^z.

Now we define $\sin z$ and $\cos z$:

Definition 7.4 (Sine and cosine)

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots$$

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \cdots .$$

We now prove certain basic properties of $\sin$ and $\cos$, using known properties of e^z.

It follows that if x is real, then $|\cos x|$ and $|\sin x|$ are at most 1.

List of Appendices

§A Heading on level 1 (section)

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

§A.1 Heading on level 2 (subsection)

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

§A.1.1 Heading on level 3 (subsubsection)

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Heading on level 4 (paragraph)   Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

§B Lists

§B.1 Example for list (itemize)

- First itemtext
- Second itemtext
- Last itemtext
- First itemtext
- Second itemtext

§B.1.1 Example for list (4*itemize)

- First itemtext
  - First itemtext
    - First itemtext
      - First itemtext
      - Second itemtext
    - Last itemtext
  - First itemtext
- Second itemtext

§B.2 Example for list (enumerate)

1. First itemtext
2. Second itemtext
3. Last itemtext
4. First itemtext
5. Second itemtext

§B.2.1 Example for list (4*enumerate)

1. First itemtext
   a) First itemtext
      i. First itemtext
         A. First itemtext
         B. Second itemtext
      ii. Last itemtext
   b) First itemtext
2. Second itemtext

§B.3 Example for list (description)

First itemtext

Second itemtext

Last itemtext

First itemtext

Second itemtext

§B.3.1 Example for list (4*description)

First itemtext

> **First** itemtext
>
>> **First** itemtext
>>
>>> **First** itemtext
>>>
>>> **Second** itemtext
>>
>> **Last** itemtext
>
> **First** itemtext

Second itemtext

§C Sample Environments

Here is an example of the types of environments that can be used, as well as a showcase of some of the macros to aid in typesetting mathematics.

§C.1 Macros

Notation C.1 (Limits)

Let f be a real-valued function. We say that the *limit* of $f(x)$ as x approaches a is L:

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

In notation C.1 we use the macro `\Lim{x\to a}f(x)=L` as a short cut for `\displaystyle\lim_{x\to a}f(x)=L`.

Remark C.1

In this example, as the limit is in display mode (due being wrapped in `\[ ...\]`), the `\displaystyle` aspect of the macro is not used. However, if we were in inline mode, such as now, the command forces `\lim` to show as if it were in display mode. For instance, this is `\lim` as an inline function: $\lim_{x \to \infty}$ and here it is again using the macro which forces the display mode: $\lim_{x \to \infty}$. The product is that the limit notation is now under the actual notation, leading to neater and better looking mathematics.

Similar to the limit macro, there is also the absolute value macro i.e `\abs{x}` which shows $|x|$. There are others but for now, these two seem the most helpful for introductory mathematics. I will update and add more examples as I see fit. To see all the macros, I highly suggest you take a look at the well documented[1] preamble.

§C.2 Theorem Styles

The theorem style for Theorems, Lemmas and propositions are all the same:

Theorem C.2 (Intermediate Value Theorem)

If f is a continuous, real-valued function defined on the interval $[a, b]$, with $f(a) \neq f(b)$, and y is a real number between $f(a)$ and $f(b)$, then there exists some $c \in (a, b)$ such that $f(c) = y$

Similarly for Corollary, the environment is the same for the claim environment.

Corollary C.3 (The Boundedness Theorem)

Suppose f is continuous on a closed interval $[a, b]$. Then there exists some value $M \in \mathbb{R}$ such that $f(c) \leq M$ for all $c \in [a, b]$. (That is, M is an upper bound for f on the interval $[a, b]$.)

Example C.4 (extreme Value Theorem)

Suppose that f is continuous on a closed interval $[a, b]$. Show that f attains a maximum on $[a, b]$; that is, there exists a real number $c \in [a, b]$ such that $f(c) \geq f(x)$ for any $x \in [a, b]$

Proof. By corollary C.3, we know that $f([a, b])$ has an upper bound, and thus, by the completeness property of $\mathbb{R}$, it has a least upper bound. Call this least upper bound M. We need to show that M is the image of some point $c \in [a, b]$. We will prove this by contradiction, so assume that M is not in the image of $f([a, b])$. Then the function $g(x) = M - f(x)$ is continuous and strictly positive on all of $[a, b]$. But more than that: because M) is the least upper bound, we know that $g^{-1}((0, \varepsilon)) \neq 0$ for all $\varepsilon > 0$, because otherwise $M - \varepsilon$ would be an upper bound for $f[a, b]$.

---

[1] If I do say so myself

Since $g(x) \neq 0$ on $[a, b]$, the function $h(x) - \frac{1}{g(x)}$ is a continuous function on $[a, b]$. However, $h(x)$ is unbounded, since for all $\varepsilon > 0$, there is some x such that $g(x) < \varepsilon$, and hence $\frac{1}{\varepsilon}$ is not an upper bound for h. But the unboundedness of h contradicts the Boundedness Theorem of corollary C.3.

Therefore, M must be in $f[a, b]$, as desired. $\qquad\square$

Remark C.5

Naturally, the Boundedness Theorem and the Extreme Value Theorem work just as well for lower bounds and minimums - we can simply replace f with $-f$, since a lower bound for f is -1 times an upper bound of $-f$, and vice versa.

---

§C.2.1 Exercise environment

Here is an example of the exercise environment making use of the solution, claim and subproof environments, as well as other environments for questions and problems.

Exercise C.6 (Weighted AM-GM).

Show that for all nonnegative reals $a_1, \ldots, a_n$ and nonnegative reals $\lambda_1, \ldots, \lambda_n$ such that $\displaystyle\sum_{i=1}^{n} \lambda_i = 1$, then

$$\sum_{i=1}^{n} \lambda_i a_i \geq \prod_{i=1}^{n} a_i^{\lambda_i}$$

With equality if and only if $a_i = a_j$ for all i, j such that $\lambda_i, \lambda_j \neq 0$.

Source: Proofs of AM-GM, AoPS

Solution. We first note that we may disregard any a_i for which $\lambda_i = 0$, as they contribute to neither side of the desired inequality. We also note that if $a_i = 0$ and $\lambda_i \neq 0$, for some i, then the right-hand side of the inequality is 0 and the left hand of the inequality is greater or equal to zero, with equality if and only if $a_j = 0 = a_i$ whenever $\lambda_j \neq 0$. Thus we may henceforth assume that all a_i and λ_i are *strictly positive*.

Claim C.7 (Jensen's Inequality)

Let F be a concave function[a] of one real variable. Let $x_1, \ldots, x_n \in \mathbb{R}_n$ and let $a_i, \ldots, a_n \geq 0$ satisfy $a_1 + \cdots + a_n = 1$. Then:

$$F(a_1 x_1 + \cdots + a_n x_n) \geq a_1 F(x_1) + \cdots + a_n F(x_n)$$

---
[a]Should F be a convex function we have $F(a_1 x_1 + \cdots + a_n x_n) \leq a_1 F(x_1) + \cdots + a_n F(x_n)$

Proof. There are many ways to prove the inequality, however, it simply follows from doing induction on n. This is only here to show the subproof Marco, not to teach induction. $\qquad\blacksquare$

We note that the function $x \mapsto \ln x$ is strictly concave. Then by Jensens Inequality (claim C.7),

$$\ln \sum_{i=1}^{n} \lambda_i a_i \geq \sum_{i=1}^{n} \lambda_i \ln a_i = \ln \prod_{i=1}^{n} a_i^{\lambda_i}$$

with equality of and only if all the a_i are equal. Since $x \mapsto \ln x$ is a strictly increasing function, it then follows that

$$\sum_{i=1}^{n} \lambda_i a_i \geq \prod_{i=1}^{n} a_i^{\lambda_i}$$

with equality if and only if all the a_i are equal, as desired.

$\qquad\square$