Part IA — Analysis I

Theorems with proof

Based on lectures by W. T. Gowers

Notes taken by Dexter Chua

Lent 2015

These notes are not endorsed by the lecturers, and I have modified them (often significantly) after lectures. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Limits and convergence

Sequences and series in $\mathbb{R}$ and $\mathbb{C}$. Sums, products and quotients. Absolute convergence; absolute convergence implies convergence. The Bolzano-Weierstrass theorem and applications (the General Principle of Convergence). Comparison and ratio tests, alternating series test. [6]

Continuity

Continuity of real- and complex-valued functions defined on subsets of $\mathbb{R}$ and $\mathbb{C}$. The intermediate value theorem. A continuous function on a closed bounded interval is bounded and attains its bounds. [3]

Differentiability

Differentiability of functions from $\mathbb{R}$ to $\mathbb{R}$. Derivative of sums and products. The chain rule. Derivative of the inverse function. Rolle's theorem; the mean value theorem. One-dimensional version of the inverse function theorem. Taylor's theorem from $\mathbb{R}$ to $\mathbb{R}$; Lagrange's form of the remainder. Complex differentiation. [5]

Power series

Complex power series and radius of convergence. Exponential, trigonometric and hyperbolic functions, and relations between them. *Direct proof of the differentiability of a power series within its circle of convergence*. [4]

Integration

Definition and basic properties of the Riemann integral. A non-integrable function. Integrability of monotonic functions. Integrability of piecewise-continuous functions. The fundamental theorem of calculus. Differentiation of indefinite integrals. Integration by parts. The integral form of the remainder in Taylor's theorem. Improper integrals. [6]

Contents

0 Introduction

1 The real number system

Lemma. 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 $x^2 = (-x)^2 > 0$. □

Lemma (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. □

2 Convergence of sequences

2.1 Definitions

Lemma (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 $|1/n - 0| < \varepsilon$. □

Lemma. Every eventually bounded sequence is bounded.

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

2.2 Sums, products and quotients

Lemma (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 $n \geq N$,

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

Lemma (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 $|\lambda a_n - \lambda a| < \varepsilon$. □

Lemma. 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 $|a_n b_n| < \varepsilon$. □

Lemma. 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. □

Lemma (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 $a_n b_n = (a + \varepsilon_n) b_n = a b_n + \varepsilon_n b_n$.

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 $a_n b_n \to ab$. □

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 $a_n b_n \to ab$. □

Lemma (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
$$\frac{1}{a_n} - \frac{1}{a} = \frac{a - a_n}{a a_n}.$$
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. □

Corollary. 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. □

Lemma (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 $|c_n - x| < \varepsilon$. □

2.3 Monotone-sequences property

Lemma. 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 $|a_n - s| < \varepsilon$. □

Lemma. 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. □

Lemma. 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 $\delta = 0$. □

Lemma. 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. □

Lemma. 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. □

Lemma (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 $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$.

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 $a \in \bigcap_{n=1}^{\infty} I_n$. □

Proposition. $\mathbb{R}$ 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. □

Theorem (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, $\bigcap_{n=1}^{\infty} [u_n, v_n] \neq \emptyset$. 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 $|a_{n_k} - x| < \varepsilon$. □

2.4 Cauchy sequences

Lemma. 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 $\forall p, q \geq N$, $|a_p - a_q| \leq |a_p - a| + |a - a_q| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Lemma. 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

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

Theorem (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. □

Lemma. 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. □

2.5 Limit supremum and infimum

Lemma. 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 $|a_m - a| < \varepsilon$. □

3 Convergence of infinite sums

3.1 Infinite sums

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

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

Lemma. 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. □

Lemma (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. □

3.2 Absolute convergence

Lemma. 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 $|S_p - S_q| < \varepsilon$. □

Theorem. 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. □

Theorem. 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. □

Lemma. 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 $\sum a_n = \sum a_{\pi(n)}$. □

3.3 Convergence tests

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

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

Lemma (Ratio test)**.** We have three versions:

(i) If $\exists c < 1$ such that
$$\frac{|a_{n+1}|}{|a_n|} \leq c,$$
for all n, then $\sum a_n$ converges.

(ii) 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.

(iii) 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.

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

(ii) 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 $\sum |a_k|$.

(iii) 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). □

Theorem (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

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

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

To prove the other way round, simply group as

$$\begin{aligned} &a_1 + (a_2 + a_3) + (a_4 + \cdots + a_7) + \cdots \\ \leq &a_1 + 2a_2 + 4a_4 + \cdots . \end{aligned}$$

□

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

3.4 Complex versions

Lemma (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

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

We now take the absolute value of everything to obtain

$$\begin{aligned}\left|\sum_{n=M}^{N} a_n z^n\right| &\leq \frac{1}{|z-1|}\left(a_N + a_M + \sum_{n=M}^{N-1}(a_n - a_{n+1})\right)\\ &= \frac{1}{|z-1|}\left(a_N + a_M + (a_M - a_{M+1}) + \cdots + (a_{N-1} - a_N)\right)\\ &= \frac{2a_M}{|z-1|} \to 0.\end{aligned}$$

So it is Cauchy. So it converges □

4 Continuous functions

4.1 Continuous functions

Lemma. 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)⇒(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)⇒(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 $f(a_n) \not\to f(a)$. □

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

(i) $f + g$ is continuous

(ii) fg is continuous

(iii) if g never vanishes, then f/g is continuous.

Proof.

(i) 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. □

Lemma. 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:

(i) 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.

(ii) 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 $|y - a| < \delta \Rightarrow |g(f(y)) - g(f(a))| < \varepsilon$. □

Theorem (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. □

Theorem (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:

(i) 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.

(ii) 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 $f(x) = 0$. □

Corollary. 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 $|z - y| < \delta \Rightarrow g(z) \in (x - \varepsilon, x + \varepsilon) \Rightarrow |g(z) - g(y)| < \varepsilon$. □

4.2 Continuous induction*

Proposition (Continuous induction v1)**.** Let $a < b$ and let $A \subseteq [a, b]$ have the following properties:

(i) $a \in A$

(ii) If $x \in A$ and $x \neq b$, then $\exists y \in A$ with $y > x$.

(iii) If $\forall \varepsilon > 0$, $\exists y \in A : y \in (x - \varepsilon, x]$, then $x \in A$.

Then $b \in A$.

Proof. Since $a \in A$, $A \neq \emptyset$. 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 $y > s$. □

Proposition (Continuous induction v2)**.** Let $A \subseteq [a, b]$ and suppose that

(i) $a \in A$

(ii) If $[a, x] \subseteq A$ and $x \neq b$, then there exists $y > x$ such that $[a, y] \subseteq A$.

(iii) 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 $A = [a, b]$. □

Theorem (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. □

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.

Proof. Let $f : [a, b]$ be continuous. Let $A = \{x : f \text{ is bounded on } [a, x]\}$. 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. □

Theorem (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. □

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 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 $\max |f(x_i)| + 1$. □

5 Differentiability

5.1 Limits

Proposition. 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.

5.2 Differentiation

Proposition.

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

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

Proof.

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

□

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

$$\begin{aligned} (f + g)'(x) &= f'(x) + g'(x) \\ (fg)'(x) &= f'(x)g(x) + f(x)g'(x) \end{aligned}$$

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

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

□

Lemma (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 $o(h)$:

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

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 $o(h)$. □

Lemma (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

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

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

Proof. As $y \to x$, $\frac{f(y) - f(x)}{y - x} \to f'(x)$. 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. □

Theorem. 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 $1/f'$.

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

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

5.3 Differentiation theorems

Theorem (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 $f'(x) = 0$.

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

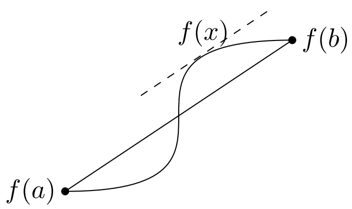
$$\frac{f(x+h) - f(x)}{h} \begin{cases} \leq 0 & h > 0 \\ \geq 0 & h < 0 \end{cases}$$

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 $f'(x) = 0$. $\square$

Corollary (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 $f(b)$.



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. $\square$

Theorem (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 $1/f'$.

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. $\square$

Theorem (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. $\square$

Corollary. 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 $f - g$. $\square$

Theorem (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 $x \in (a, a + h)$.

5.4 Complex differentiation

6 Complex power series

Lemma. 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. □

Lemma. 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 lim sup. 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. □

6.1 Exponential and trigonometric functions

Proposition. The derivative of e^z is e^z.

Proof.
$$\begin{aligned}\frac{e^{z+h} - e^z}{h} &= e^z\left(\frac{e^h - 1}{h}\right)\\ &= e^z\left(1 + \frac{h}{2!} + \frac{h^2}{3!} + \cdots\right)\end{aligned}$$

But
$$\left|\frac{h}{2!} + \frac{h^2}{3!} + \cdots\right| \leq \frac{|h|}{2} + \frac{|h|^2}{4} + \frac{|h|^3}{8} + \cdots = \frac{|h|/2}{1 - |h|/2} \to 0.$$
So
$$\frac{e^{z+h} - e^z}{h} \to e^z.$$
□

Theorem. 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
$$(a_0b_0) + (a_0b_1 + a_1b_1 + a_1b_0) + (a_0b_2 + a_1b_2 + a_2b_2 + a_2b_1 + a_2b_0) + \cdots \quad (*)$$
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:
$$a_0b_0 + (a_0b_1 + a_1b_0) + (a_0b_2 + a_1b_1 + a_2b_0) + \cdots$$
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
$$\sum_{n=0}^N c_n \to ST = \left(\sum_{n=0}^\infty a_n\right)\left(\sum_{n=0}^\infty b_n\right).$$
□

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

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

$$\begin{aligned}
e^z e^w &= \sum_{n=0}^{\infty}\left(1\cdot\frac{w^n}{n!}+\frac{z}{1!}\frac{w^{n-1}}{(n-1)!}+\frac{z^2}{2!}\frac{w^{n-2}}{(n-2)!}+\cdots+\frac{z^n}{n!}\cdot 1\right)\\
e^z e^w &= \sum_{n=0}^{\infty}\frac{1}{n!}\left(w^n+\binom{n}{1}zw^{n-1}+\binom{n}{2}z^2w^{n-2}+\cdots+\binom{n}{n}z^n\right)\\
&= \sum_{n=0}^{\infty}(z+w)^n \text{ by the binomial theorem}\\
&= e^{z+w}.
\end{aligned}$$

□

Proposition.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}z}\sin z &= \frac{ie^{iz}+ie^{-iz}}{2i} = \cos z\\
\frac{\mathrm{d}}{\mathrm{d}z}\cos z &= \frac{ie^{iz}-ie^{-iz}}{2} = -\sin z\\
\sin^2 z+\cos^2 z &= \frac{e^{2iz}+2+e^{-2iz}}{4}+\frac{e^{2iz}-2+e^{-2iz}}{-4} = 1.
\end{aligned}$$

Proposition.

$$\begin{aligned}
\cos(z+w) &= \cos z\cos w-\sin z\sin w\\
\sin(z+w) &= \sin z\cos w+\cos z\sin w
\end{aligned}$$

Proof.

$$\begin{aligned}
\cos z\cos w-\sin z\sin w &= \frac{(e^{iz}+e^{-iz})(e^{iw}+e^{-iw})}{4}+\frac{(e^{iz}-e^{-iz})(e^{iw}-e^{-iw})}{4}\\
&= \frac{e^{i(z+w)}+e^{-i(z+w)}}{2}\\
&= \cos(z+w).
\end{aligned}$$

□

Differentiating both sides wrt z gives

$$-\sin z\cos w-\cos z\sin w = -\sin(z+w).$$

So

$$\sin(z+w) = \sin z\cos w+\cos z\sin w.$$

Proposition.

$$\begin{aligned}
\cos\left(z+\frac{\pi}{2}\right) &= -\sin z\\
\sin\left(z+\frac{\pi}{2}\right) &= \cos z\\
\cos(z+\pi) &= -\cos z\\
\sin(z+\pi) &= -\sin z\\
\cos(z+2\pi) &= \cos z\\
\sin(z+2\pi) &= \sin z
\end{aligned}$$

Proof.

$$\begin{aligned}\cos\left(z+\frac{\pi}{2}\right) &= \cos z\cos\frac{\pi}{2} - \sin z\sin\frac{\pi}{2}\\ &= -\sin z\sin\frac{\pi}{2}\\ &= -\sin z\end{aligned}$$

and similarly for others. □

6.2 Differentiating power series

Lemma. Let a and b be complex numbers. Then

$$b^n - a^n - n(b-a)a^{n-1} = (b-a)^2(b^{n-2} + 2ab^{n-3} + 3a^2b^{n-4} + \cdots + (n-1)a^{n-2}).$$

Proof. If $b = a$, we are done. Otherwise,

$$\frac{b^n - a^n}{b-a} = b^{n-1} + ab^{n-2} + a^2b^{n-3} + \cdots + a^{n-1}.$$

Differentiate both sides with respect to a. Then

$$\frac{-na^{n-1}(b-a) + b^n - a^n}{(b-a)^2} = b^{n-2} + 2ab^{n-3} + \cdots + (n-1)a^{n-2}.$$

Rearranging gives the result.

Alternatively, we can do

$$b^n - a^n = (b-a)(b^{n-1} + ab^{n-2} + \cdots + a^{n-1}).$$

Subtract $n(b-a)a^{n-1}$ to obtain

$$(b-a)[b^{n-1} - a^{n-1} + a(b^{n-2} - a^{n-2}) + a^2(b^{n-3} - a^{n-3}) + \cdots]$$

and simplify. □

Lemma. Let $a_n z^n$ have radius of convergence R, and let $|z| < R$. Then $\sum na_n z^{n-1}$ converges (absolutely).

Proof. Pick r such that $|z| < r < R$. Then $\sum |a_n| r^n$ converges, so the terms $|a_n|r^n$ are bounded above by, say, C. Now

$$\sum n|a_n z^{n-1}| = \sum n|a_n|r^{n-1}\left(\frac{|z|}{r}\right)^{n-1} \leq \frac{C}{r}\sum n\left(\frac{|z|}{r}\right)^{n-1}$$

The series $\sum n\left(\frac{|z|}{r}\right)^{n-1}$ converges, by the ratio test. So $\sum n|a_n z^{n-1}|$ converges, by the comparison test. □

Corollary. Under the same conditions,

$$\sum_{n=2}^{\infty}\binom{n}{2}a_n z^{n-2}$$

converges absolutely.

Proof. Apply Lemma above again and divide by 2. □

Theorem. Let $\sum a_n z^n$ be a power series with radius of convergence R. For $|z| < R$, let

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \text{ and } g(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}.$$

Then f is differentiable with derivative g.

Proof. We want $f(z+h) - f(z) - hg(z)$ to be $o(h)$. We have

$$f(z+h) - f(z) - hg(z) = \sum_{n=2}^{\infty} a_n((z+h)^n - z^n - hnz^n).$$

We started summing from $n = 2$ since the $n = 0$ and $n = 1$ terms are 0. Using our first lemma, we are left with

$$h^2 \sum_{n=2}^{\infty} a_n\big((z+h)^{n-2} + 2z(z+h)^{n-3} + \cdots + (n-1)z^{n-2}\big)$$

We want the huge infinite series to be bounded, and then the whole thing is a bounded thing times h^2, which is definitely $o(h)$.

Pick r such that $|z| < r < R$. If h is small enough that $|z+h| \leq r$, then the last infinite series is bounded above (in modulus) by

$$\sum_{n=2}^{\infty} |a_n|(r^{n-2} + 2r^{n-2} + \cdots + (n-1)r^{n-2}) = \sum_{n=2}^{\infty} |a_n|\binom{n}{2} r^{n-2},$$

which is bounded. So done. □

6.3 Hyperbolic trigonometric functions

Proposition.

$$\frac{\mathrm{d}}{\mathrm{d}z}\cosh z = \sinh z$$
$$\frac{\mathrm{d}}{\mathrm{d}z}\sinh z = \cosh z$$

Proposition.

$$\cosh iz = \cos z$$
$$\sinh iz = i \sin z$$

Proposition.

$$\cosh^2 z - \sinh^2 z = 1,$$

7 The Riemann Integral

7.1 Riemann Integral

Lemma. If $\mathcal{D}_2$ refines $\mathcal{D}_1$, then

$$U_{\mathcal{D}_2} f \leq U_{\mathcal{D}_1} f \text{ and } L_{\mathcal{D}_2 f} \geq L_{\mathcal{D}_1} f.$$

Proof. Let $\mathcal{D}$ be $x_0 < x_1 < \cdots < x_n$. Let $\mathcal{D}_2$ be obtained from $\mathcal{D}_1$ by the addition of one point z. If $z \in (x_{i-1}, x_i)$, then

$$\begin{aligned} U_{\mathcal{D}_2} f - U_{\mathcal{D}_1} f = & \left[(z - x_{i-1}) \sup_{x \in [x_{i-1}, z]} f(x)\right] \\ & + \left[(x_i - z) \sup_{x \in [z, x_i]} f(x)\right] - (x_i - x_{i-1}) M_i. \end{aligned}$$

But $\sup_{x \in [x_{i-1}, z]} f(x)$ and $\sup_{x \in [z, x_i]} f(x)$ are both at most M_i. So this is at most $M_i(z - x_{i-1} + x_i - z - (x_i - x_{i-1})) = 0$. So

$$U_{\mathcal{D}_2} f \leq U_{\mathcal{D}_1} f.$$

By induction, the result is true whenever $\mathcal{D}_2$ refines $\mathcal{D}_1$.

A very similar argument shows that $L_{\mathcal{D}_2 f} \geq L_{\mathcal{D}_1} f$. □

Corollary. Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be two dissections of $[a, b]$. Then

$$U_{\mathcal{D}_1} f \geq L_{\mathcal{D}_2} f.$$

Proof. Let $\mathcal{D}$ be the least common refinement (or indeed any common refinement). Then by lemma above (and by definition),

$$U_{\mathcal{D}_1} f \geq U_{\mathcal{D}} f \geq L_{\mathcal{D}} f \geq L_{\mathcal{D}_2} f.$$ □

Proposition (Riemann's integrability criterion)**.** This is sometimes known as Cauchy's integrability criterion.

Let $f : [a, b] \to \mathbb{R}$. Then f is Riemann integrable if and only if for every $\varepsilon > 0$, there exists a dissection $\mathcal{D}$ such that

$$U_{\mathcal{D}} - L_{\mathcal{D}} < \varepsilon.$$

Proof. ($\Rightarrow$) Suppose that f is integrable. Then (by definition of Riemann integrability), there exist $\mathcal{D}_1$ and $\mathcal{D}_2$ such that

$$U_{\mathcal{D}_1} < \int_a^b f(x)\,\mathrm{d}x + \frac{\varepsilon}{2},$$

and

$$L_{\mathcal{D}_2} > \int_a^b f(x)\,\mathrm{d}x - \frac{\varepsilon}{2}.$$

Let $\mathcal{D}$ be a common refinement of $\mathcal{D}_1$ and $\mathcal{D}_2$. Then

$$U_{\mathcal{D}} f - L_{\mathcal{D}} f \leq U_{\mathcal{D}_1} f - L_{\mathcal{D}_2} f < \varepsilon.$$

($\Leftarrow$) Conversely, if there exists $\mathcal{D}$ such that

$$U_{\mathcal{D}} f - L_{\mathcal{D}} f < \varepsilon,$$

then

$$\inf U_{\mathcal{D}} f - \sup L_{\mathcal{D}} f < \varepsilon,$$

which is, by definition, that

$$\overline{\int_a^b} f(x)\, \mathrm{d}x - \underline{\int_a^b} f(x)\, \mathrm{d}x < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this gives us that

$$\overline{\int_a^b} f(x)\, \mathrm{d}x = \underline{\int_a^b} f(x)\, \mathrm{d}x.$$

So f is integrable. □

Proposition. Let $f : [a, b] \to \mathbb{R}$ be integrable, and $\lambda \geq 0$. Then λf is integrable, and

$$\int_a^b \lambda f(x)\, \mathrm{d}x = \lambda \int_a^b f(x)\, \mathrm{d}x.$$

Proof. Let $\mathcal{D}$ be a dissection of $[a, b]$. Since

$$\sup_{x \in [x_{i-1}, x_i]} \lambda f(x) = \lambda \sup_{x \in [x_{i-1}, x_i]} f(x),$$

and similarly for inf, we have

$$U_{\mathcal{D}}(\lambda f) = \lambda U_{\mathcal{D}} f$$
$$L_{\mathcal{D}}(\lambda f) = \lambda L_{\mathcal{D}} f.$$

So if we choose $\mathcal{D}$ such that $U_{\mathcal{D}} f - L_{\mathcal{D}} f < \varepsilon/\lambda$, then $U_{\mathcal{D}}(\lambda f) - L_{\mathcal{D}}(\lambda f) < \varepsilon$. So the result follows from Riemann's integrability criterion. □

Proposition. Let $f : [a, b] \to \mathbb{R}$ be integrable. Then $-f$ is integrable, and

$$\int_a^b -f(x)\, \mathrm{d}x = -\int_a^b f(x)\, \mathrm{d}x.$$

Proof. Let $\mathcal{D}$ be a dissection. Then

$$\sup_{x \in [x_{i-1}, x_i]} -f(x) = -\inf_{x \in [x_{i-1}, x_i]} f(x)$$
$$\inf_{x \in [x_{i-1}, x_i]} -f(x) = -\sup_{x \in [x_{i-1}, x_i]} f(x).$$

Therefore

$$U_{\mathcal{D}}(-f) = \sum_{i=1}^{n} (x_i - x_{i-1})(-m_i) = -L_{\mathcal{D}}(f).$$

Similarly,

$$L_{\mathcal{D}}(-f) = -U_{\mathcal{D}} f.$$

So
$$U_{\mathcal{D}}(-f) - L_{\mathcal{D}}(-f) = U_{\mathcal{D}}f - L_{\mathcal{D}}f.$$
Hence if f is integrable, then $-f$ is integrable by the Riemann integrability criterion. □

Proposition. Let $f, g : [a, b] \to \mathbb{R}$ be integrable. Then $f + g$ is integrable, and
$$\int_a^b (f(x) + g(x))\ \mathrm{d}x = \int_a^b f(x)\ \mathrm{d}x + \int_a^b g(x)\ \mathrm{d}x.$$

Proof. Let $\mathcal{D}$ be a dissection. Then
$$\begin{aligned} U_{\mathcal{D}}(f+g) &= \sum_{i=1}^n (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} (f(x) + g(x)) \\ &\leq \sum_{i=1}^n (x_i - x_{i-1}) \left(\sup_{u \in [x_{i-1}, x_i]} f(u) + \sup_{v \in [x_{i-1}, x_i]} g(v) \right) \\ &= U_{\mathcal{D}}f + U_{\mathcal{D}}g \end{aligned}$$
Therefore,
$$\overline{\int_a^b} (f(x) + g(x))\ \mathrm{d}x \leq \overline{\int_a^b} f(x)\ \mathrm{d}x + \overline{\int_a^b} g(x)\ \mathrm{d}x = \int_a^b f(x)\ \mathrm{d}x + \int_a^b g(x)\ \mathrm{d}x.$$
Similarly,
$$\underline{\int_a^b} (f(x) + g(x))\ \mathrm{d}x \geq \int_a^b f(x)\ \mathrm{d}x + \int_a^b g(x)\ \mathrm{d}x.$$
So the upper and lower integrals are equal, and the result follows. □

Proposition. Let $f, g : [a, b] \to \mathbb{R}$ be integrable, and suppose that $f(x) \leq g(x)$ for every x. Then
$$\int_a^b f(x)\ \mathrm{d}x \leq \int_a^b g(x)\ \mathrm{d}x.$$

Proof. Follows immediately from the definition. □

Proposition. Let $f : [a, b] \to \mathbb{R}$ be integrable. Then $|f|$ is integrable.

Proof. Note that we can write
$$\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) = \sup_{u, v \in [x_{i-1}, x_i]} |f(u) - f(v)|.$$
Similarly,
$$\sup_{x \in [x_{i-1}, x_i]} |f(x)| - \inf_{x \in [x_{i-1}, x_i]} |f(x)| = \sup_{u, v \in [x_{i-1}, x_i]} ||f(u)| - |f(v)||.$$
For any pair of real numbers, x, y, we have that $||x| - |y|| \leq |x - y|$ by the triangle inequality. Then for any interval $u, v \in [x_{i-1}, x_i]$, we have
$$||f(u)| - |f(v)|| \leq |f(u) - f(v)|.$$

Hence we have

$$\sup_{x\in[x_{i-1},x_i]} |f(x)| - \inf_{x\in[x_{i-1},x_i]} |f(x)| \leq \sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x).$$

So for any dissection $\mathcal{D}$, we have

$$U_{\mathcal{D}}(|f|) - L_{\mathcal{D}}(|f|) \leq U_{\mathcal{D}}(f) - L_{\mathcal{D}}(f).$$

So the result follows from Riemann's integrability criterion. □

Proposition (Additivity property)**.** Let $f : [a, c] \to \mathbb{R}$ be integrable, and let $b \in (a, c)$. Then the restrictions of f to $[a, b]$ and $[b, c]$ are Riemann integrable, and

$$\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x$$

Similarly, if f is integrable on $[a, b]$ and $[b, c]$, then it is integrable on $[a, c]$ and the above equation also holds.

Proof. Let $\varepsilon > 0$, and let $a = x_0 < x_1 < \cdots < x_n = c$ be a dissection of $\mathcal{D}$ of $[a, c]$ such that

$$U_{\mathcal{D}}(f) \leq \int_a^c f(x)\,\mathrm{d}x + \varepsilon,$$

and

$$L_{\mathcal{D}}(f) \geq \int_a^c f(x)\,\mathrm{d}x - \varepsilon.$$

Let $\mathcal{D}'$ be the dissection made of $\mathcal{D}$ plus the point b. Let $\mathcal{D}_1$ be the dissection of $[a, b]$ made of points of $\mathcal{D}'$ from a to b, and D_2 be the dissection of $[b, c]$ made of points of $\mathcal{D}'$ from b to c. Then

$$U_{\mathcal{D}_1}(f) + U_{\mathcal{D}_2}(f) = U_{\mathcal{D}'}(f) \leq U_{\mathcal{D}}(f),$$

and

$$L_{\mathcal{D}_1}(f) + L_{\mathcal{D}_2}(f) = L_{\mathcal{D}'}(f) \geq L_{\mathcal{D}}(f).$$

Since $U_{\mathcal{D}}(f) - L_{\mathcal{D}}(f) < 2\varepsilon$, and both $U_{\mathcal{D}_2}(f) - L_{\mathcal{D}_2}(f)$ and $U_{\mathcal{D}_1}(f) - L_{\mathcal{D}_1}(f)$ are non-negative, we have $U_{\mathcal{D}_1}(f) - L_{\mathcal{D}_1}(f)$ and $U_{\mathcal{D}_2}(f) - L_{\mathcal{D}_2}(f)$ are less than 2ε. Since ε is arbitrary, it follows that the restrictions of f to $[a, b]$ and $[b, c]$ are both Riemann integrable. Furthermore,

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x \leq U_{\mathcal{D}_1}(f) + U_{\mathcal{D}_2}(f) = U_{\mathcal{D}'}(f) \leq U_{\mathcal{D}}(f)\\ \leq \int_a^c f(x)\,\mathrm{d}x + \varepsilon.\end{aligned}$$

Similarly,

$$\begin{aligned}\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x \geq L_{\mathcal{D}_1}(f) + L_{\mathcal{D}_2}(f) = L_{\mathcal{D}'}(f) \geq L_{\mathcal{D}}(f)\\ \geq \int_a^c f(x)\,\mathrm{d}x - \varepsilon.\end{aligned}$$

Since ε is arbitrary, it follows that

$$\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x.$$

The other direction is left as an (easy) exercise. □

Proposition. Let $f, g : [a, b] \to \mathbb{R}$ be integrable. Then fg is integrable.

Proof. Let C be such that $|f(x)|, |g(x)| \le C$ for every $x \in [a, b]$. Write L_i and ℓ_i for the sup and inf of g in $[x_{i-1}, x_i]$. Now let $\mathcal{D}$ be a dissection, and for each i, let u_i and v_i be two points in $[x_{i-1}, x_i]$.

We will pretend that u_i and v_i are the minimum and maximum when we write the proof, but we cannot assert that they are, since fg need not have maxima and minima. We will then note that since our results hold for arbitrary u_i and v_i, it must hold when fg is at its supremum and infimum.

We find what we pretend is the difference between the upper and lower sum:

$$\begin{aligned}
&\left|\sum_{i=1}^n \big(x_i - x_{i-1}\big)\big(f(v_i)g(v_i) - f(u_i)g(u_i)\big)\right| \\
&= \left|\sum_{i=1}^n (x_i - x_{i-1})\big(f(v_i)(g(v_i) - g(u_i)) + (f(v_i) - f(u_i))g(u_i)\big)\right| \\
&\le \sum_{i=1}^n \big(C(L_i - \ell_i) + (M_i - m_i)C\big) \\
&= C(U_{\mathcal{D}}g - L_{\mathcal{D}}g + U_{\mathcal{D}}f - L_{\mathcal{D}}f).
\end{aligned}$$

Since u_i and v_i are arbitrary, it follows that

$$U_{\mathcal{D}}(fg) - L_{\mathcal{D}}(fg) \le C(U_{\mathcal{D}}f - L_{\mathcal{D}}f + U_{\mathcal{D}}g - L_{\mathcal{D}}g).$$

Since C is fixed, and we can get $U_{\mathcal{D}}f - L_{\mathcal{D}}f$ and $U_{\mathcal{D}}g - L_{\mathcal{D}}g$ arbitrary small (since f and g are integrable), we can get $U_{\mathcal{D}}(fg) - L_{\mathcal{D}}(fg)$ arbitrarily small. So the result follows. □

Theorem. Every continuous function f on a closed bounded interval $[a, b]$ is Riemann integrable.

Proof. wlog assume $[a, b] = [0, 1]$.

Suppose the contrary. Let f be non-integrable. This means that there exists some ε such that for every dissection $\mathcal{D}$, $U_{\mathcal{D}} - L_{\mathcal{D}} > \varepsilon$. In particular, for every n, let $\mathcal{D}_n$ be the dissection $0, \frac{1}{n}, \frac{2}{n}, \cdots, \frac{n}{n}$.

Since $U_{\mathcal{D}_n} - L_{\mathcal{D}_n} > \varepsilon$, there exists some interval $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ in which $\sup f - \inf f > \varepsilon$. Suppose the supremum and infimum are attained at x_n and y_n respectively. Then we have $|x_n - y_n| < \frac{1}{n}$ and $f(x_n) - f(y_n) > \varepsilon$.

By Bolzano Weierstrass, (x_n) has a convergent subsequence, say (x_{n_i}). Say $x_{n_i} \to x$. Since $|x_n - y_n| < \frac{1}{n} \to 0$, we must have $y_{n_i} \to x$. By continuity, we must have $f(x_{n_i}) \to f(x)$ and $f(y_{n_i}) \to f(x)$, but $f(x_{n_i})$ and $f(y_{n_i})$ are always apart by ε. Contradiction. □

Theorem (non-examinable)**.** Let $a < b$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is uniformly continuous.

Proof. Suppose that f is not uniformly continuous. Then

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

Therefore, we can find sequences $(x_n), (y_n)$ such that for every n, we have

$$|x_n - y_n| \leq \frac{1}{n} \text{ and } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Then by Bolzano-Weierstrass theorem, we can find a subsequence (x_{n_k}) converging to some x. Since $|x_{n_k} - y_{n_k}| \leq \frac{1}{n_k}$, $y_{n_k} \to x$ as well. But $|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon$ for every k. So $f(x_{n_k})$ and $f(y_{n_k})$ cannot both converge to the same limit. So f is not continuous at x. □

Theorem. Let $f : [a, b] \to \mathbb{R}$ be monotone. Then f is Riemann integrable.

Proof. let $\varepsilon > 0$. Let $\mathcal{D}$ be a dissection of mesh less than $\frac{\varepsilon}{f(b)-f(a)}$. Then

$$\begin{aligned} U_{\mathcal{D}}f - L_{\mathcal{D}}f &= \sum_{i=1}^{n}(x_i - x_{i-1})(f(x_i) - f(x_{i-1})) \\ &\leq \frac{\varepsilon}{f(b) - f(a)} \sum_{i=1}^{n}(f(x_i) - f(x_{i-1})) \\ &= \varepsilon. \end{aligned}$$

□

Lemma. Let $a < b$ and let f be a bounded function from $[a, b] \to \mathbb{R}$ that is continuous on (a, b). Then f is integrable.

Proof. Let $\varepsilon > 0$. Suppose that $|f(x)| \leq C$ for every $x \in [a, b]$. Let $x_0 = a$ and pick x_1 such that $x_1 - x_0 < \frac{\varepsilon}{8C}$. Also choose z between x_1 and b such that $b - z < \frac{\varepsilon}{8C}$.

Then f is continuous $[x_1, z]$. Therefore it is integrable on $[x_1, z]$. So we can find a dissection $\mathcal{D}'$ with points $x_1 < x_2 < \cdots < x_{n-1} = z$ such that

$$U_{\mathcal{D}'}f - L_{\mathcal{D}'}f < \frac{\varepsilon}{2}.$$

Let $\mathcal{D}$ be the dissection $a = x_0 < x_1 < \cdots < x_n = b$. Then

$$U_{\mathcal{D}}f - L_{\mathcal{D}}f < \frac{\varepsilon}{8C} \cdot 2C + \frac{\varepsilon}{2} + \frac{\varepsilon}{8C} \cdot 2C = \varepsilon.$$

So done by Riemann integrability criterion. □

Corollary. Every piecewise continuous and bounded function on $[a, b]$ is integrable.

Proof. Partition $[a, b]$ into intervals $I_1, \cdots, I_k$, on each of which f is (bounded and) continuous. Hence for every I_j with end points x_{j-1}, x_j, f is integrable on $[x_{j-1}, x_j]$ (which may not equal I_j, e.g. I_j could be $[x_{j-1}, x_j)$). But then by the additivity property of integration, we get that f is integrable on $[a, b]$ □

Lemma. Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable, and for each n, let $\mathcal{D}_n$ be the dissection $a = x_0 < x_1 < \cdots < x_n = b$, where $x_i = a + \frac{i(b-a)}{n}$ for each i. Then

$$U_{\mathcal{D}_n} f \to \int_a^b f(x)\,\mathrm{d}x$$

and

$$L_{\mathcal{D}_n} f \to \int_a^b f(x)\,\mathrm{d}x.$$

Proof. Let $\varepsilon > 0$. We need to find an N. The only thing we know is that f is Riemann integrable, so we use it:

Since f is integrable, there is a dissection $\mathcal{D}$, say $u_0 < u_1 < \cdots < u_m$, such that

$$U_{\mathcal{D}} f - \int_a^b f(x)\,\mathrm{d}x < \frac{\varepsilon}{2}.$$

We also know that f is bounded. Let C be such that $|f(x)| \leq C$.

For any n, let $\mathcal{D}'$ be the least common refinement of $\mathcal{D}_n$ and $\mathcal{D}$. Then

$$U_{\mathcal{D}'} f \leq U_{\mathcal{D}} f.$$

Also, the sums $U_{\mathcal{D}_n} f$ and $U_{\mathcal{D}'} f$ are the same, except that at most m of the subintervals $[x_{i-1}, x_i]$ are subdivided in $\mathcal{D}'$.

For each interval that gets chopped up, the upper sum decreases by at most $\frac{b-a}{n} \cdot 2C$. Therefore

$$U_{\mathcal{D}_n} f - U_{\mathcal{D}'} f \leq \frac{b-a}{n} 2C \cdot m.$$

Pick n such that $2Cm(b-a)/n < \frac{\varepsilon}{2}$. Then

$$U_{\mathcal{D}_n} f - U_{\mathcal{D}} f < \frac{\varepsilon}{2}.$$

So

$$U_{\mathcal{D}_n} f - \int_a^b f(x)\,\mathrm{d}x < \varepsilon.$$

This is true whenever $n > \frac{4C(b-a)m}{\varepsilon}$. Since we also have $U_{\mathcal{D}_n} f \geq \int_a^b f(x)\,\mathrm{d}x$, therefore

$$U_{\mathcal{D}_n} f \to \int_a^b f(x)\,\mathrm{d}x.$$

The proof for lower sums is similar. □

Theorem (Fundamental theorem of calculus, part 1)**.** Let $f : [a, b] \to \mathbb{R}$ be continuous, and for $x \in [a, b]$, define

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

Then F is differentiable and $F'(x) = f(x)$ for every x.

Proof.

$$\frac{F(x+h)-F(x)}{h} = \frac{1}{h}\int_x^{x+h} f(t)\ \mathrm{d}t$$

Let $\varepsilon > 0$. Since f is continuous, at x, then there exists δ such that $|y-x| < \delta$ implies $|f(y) - f(x)| < \varepsilon$.

If $|h| < \delta$, then

$$\begin{aligned}\left|\frac{1}{h}\int_x^{x+h} f(t)\ \mathrm{d}t - f(x)\right| &= \left|\frac{1}{h}\int_x^{x+h} (f(t) - f(x))\ \mathrm{d}t\right| \\ &\leq \frac{1}{|h|}\left|\int_x^{x+h} |f(t) - f(x)|\ \mathrm{d}t\right| \\ &\leq \frac{\varepsilon |h|}{|h|} \\ &= \varepsilon.\end{aligned}$$

□

Corollary. If f is continuously differentiable on $[a, b]$, then

$$\int_a^b f'(t)\ \mathrm{d}t = f(b) - f(a).$$

Proof. Let

$$g(x) = \int_a^x f'(t)\ \mathrm{d}t.$$

Then

$$g'(x) = f'(x) = \frac{\mathrm{d}}{\mathrm{d}x}(f(x) - f(a)).$$

Since $g'(x) - f'(x) = 0$, $g(x) - f(x)$ must be a constant function by the mean value theorem. We also know that

$$g(a) = 0 = f(a) - f(a)$$

So we must have $g(x) = f(x) - f(a)$ for every x, and in particular, for $x = b$. □

Theorem (Fundamental theorem of calculus, part 2)**.** Let $f : [a, b] \to \mathbb{R}$ be a differentiable function, and suppose that f' is integrable. Then

$$\int_a^b f'(t)\ \mathrm{d}t = f(b) - f(a).$$

Proof. Let $\mathcal{D}$ be a dissection $x_0 < x_1 < \cdots < x_n$. We want to make use of this dissection. So write

$$f(b) - f(a) = \sum_{i=1}^n (f(x_i) - f(x_{i-1})).$$

For each i, there exists $u_i \in (x_{i-1}, x_i)$ such that $f(x_i) - f(x_{i-1j}) = (x_i - x_{i-1})f'(u_i)$ by the mean value theorem. So

$$f(b) - f(a) = \sum_{i=1}^n (x_i - x_{i-1}) f'(u_i).$$

We know that $f'(u_i)$ is somewhere between $\sup_{x\in[x_i,x_{i-1}]} f'(x)$ and $\inf_{x\in[x_i,x_{i-1}]} f'(x)$ by definition. Therefore

$$L_{\mathcal{D}}f' \leq f(b) - f(a) \leq U_{\mathcal{D}}f'.$$

Since f' is integrable and $\mathcal{D}$ was arbitrary, $L_{\mathcal{D}}f'$ and $U_{\mathcal{D}}f'$ can both get arbitrarily close to $\int_a^b f'(t)\ \mathrm{d}t$. So

$$f(b) - f(a) = \int_a^b f'(t)\ \mathrm{d}t. \qquad \square$$

Theorem (Integration by parts)**.** Let $f, g : [a, b] \to \mathbb{R}$ be integrable such that everything below exists. Then

$$\int_a^b f(x)g'(x)\ \mathrm{d}x = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\ \mathrm{d}x.$$

Proof. By the fundamental theorem of calculus,

$$\int_a^b (f(x)g'(x) + f'(x)g(x))\ \mathrm{d}x = \int_a^b (fg)'(x)\ \mathrm{d}x = f(b)g(b) - f(a)g(a).$$

The result follows after rearrangement. $\square$

Theorem (Taylor's theorem with the integral form of the remainder)**.** Let f be $n+1$ times differentiable on $[a, b]$ with with $f^{(n+1)}$ continuous. Then

$$\begin{aligned} f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2!}f^{(2)}(a) + \cdots \\ + \frac{(b-a)^n}{n!}f^{(n)}(a) + \int_a^b \frac{(b-t)^n}{n!}f^{(n+1)}(t)\ \mathrm{d}t. \end{aligned}$$

Proof. Induction on n.

When $n = 0$, the theorem says

$$f(b) - f(a) = \int_a^b f'(t)\ \mathrm{d}t.$$

which is true by the fundamental theorem of calculus.

Now observe that

$$\begin{aligned} \int_a^b \frac{(b-t)^n}{n!}f^{(n+1)}(t)\ \mathrm{d}t &= \left[\frac{-(b-t)^{n+1}}{(n+1)!}f^{(n+1)}(t)\right]_a^b \\ &\quad + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!}f^{(n+1)}(t)\ \mathrm{d}t \\ &= \frac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(a) + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\ \mathrm{d}t. \end{aligned}$$

So the result follows by induction. $\square$

Theorem (Integration by substitution)**.** Let $f : [a, b] \to \mathbb{R}$ be continuous. Let $g : [u, v] \to \mathbb{R}$ be continuously differentiable, and suppose that $g(u) = a, g(v) = b$, and f is defined everywhere on $g([u, v])$ (and still continuous). Then

$$\int_a^b f(x)\ \mathrm{d}x = \int_u^v f(g(t))g'(t)\ \mathrm{d}t.$$

Proof. By the fundamental theorem of calculus, f has an anti-derivative F defined on $g([u, v])$. Then

$$\begin{aligned}\int_u^v f(g(t))g'(t)\ \mathrm{d}t &= \int_u^v F'(g(t))g'(t)\ \mathrm{d}t \\ &= \int_u^v (F \circ g)'(t)\ \mathrm{d}t \\ &= F \circ g(v) - F \circ g(u) \\ &= F(b) - F(a) \\ &= \int_a^b f(x)\ \mathrm{d}x.\end{aligned}$$

□

7.2 Improper integrals

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

Proof. We have

$$\int_n^{n+1} f(x)\ \mathrm{d}x \leq f(n) \leq \int_{n-1}^n f(x)\ \mathrm{d}x,$$

since f is decreasing (the right hand inequality is valid only for $n \geq 2$). It follows that

$$\int_1^{N+1} f(x)\ \mathrm{d}x \leq \sum_{n=1}^N f(n) \leq \int_1^N f(x)\ \mathrm{d}x + f(1)$$

So if the integral exists, then $\sum f(n)$ is increasing and bounded above by $\int_1^\infty f(x)\ \mathrm{d}x$, so converges.

If the integral does not exist, then $\int_1^N f(x)\ \mathrm{d}x$ is unbounded. Then $\sum_{n=1}^N f(n)$ is unbounded, hence does not converge. □