# Math Tripos Part IA: Analysis I

Michael Li

May 21, 2015

**Limits and convergence**
Sequences and series in R and 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 $\mathbb{R}$; Lagrangefis 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

# 1 The Real Number System

**Definition.** A *field* is a set $X$ with two binary operations $+$ and $\times$ and all the familiar properties (commutativity, associativity, identities, inverses [0 has no multiplicative inverse], distributivity) that are satisfied by addition and multiplication in $\mathbb{Q}$.

**Example.** $\mathbb{Q}$, $\mathbb{R}$, integers mod $p$, $\cdots$

**Definition.** An *ordered set* is a set $X$ with a transitive relation $<$ and $\forall x, y \in X$, exactly one of $x < y$, $x = y$ and $y < x$ is true. An *ordered field* is a field $\mathbb{F}$ with a relation $<$ that makes $F$ into an ordered set such that

(i) if $x, y, z \in \mathbb{F}$, $x < y$, then $x + z < y + z$

(ii) if $x, y, z \in \mathbb{F}$, $x < y$, $z > 0$, then $xz < yz$.

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

*Proof.* By trichotomy, $x < 0$, $x = 0$, or $x > 0$. If $x = 0$, then $x^2 = 0$, so $x^2 \geq 0$. If $x > 0$, then $x^2 > 0 \times x = 0$ by multiplication property. If $x < 0$, then $x - x < 0 - x$ by the additive property, so $0 < -x$. But then $x^2 = (-x)^2 > 0$. □

**Axiom** (Least upper bound Axiom, LUB)**.** Let $X$ be an ordered set and let $A \subset X$. 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.*

The upper bound $x$ is an *least upper bound* or *supremum*, denoted $\sup A$, if

(i) $\forall a \in A, a \leq x$

(ii) $\forall y < x, \exists a \in A$ such that $a > y$.

If $\sup A \in A$, then we call it $\max A$. Greatest lower bounds, or *infimum*, and lower bounds are defined similarly. It can be shown these two properties are equivalent.

**Example.** For $X = \mathbb{Q}$, $A = \{x; x^2 < 2\}$ has no supremum.

**Theorem.** An ordered field $F$ has the least upper bound property if every non-empty subset of $\mathbb{F}$ that is bounded above has a supremum.

**Lemma** (Archimedean property V1)**.** Let $\mathbb{F}$ be an ordered field with the least upper bound property. Then the set $\{1, 2, 3, 4, \cdots\}$ 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 $n \in \{1, 2, 3...\}$ such that $n > x - 1$. But then $n + 1 > x$. Contradiction. □

# 2 Convergence of sequences

**Definition** (Sequence)**.** A *sequence* is, formally, a function $a : \mathbb{N} \to \mathbb{R}$ (or $\mathbb{C}$). We write $a_n$ instead of $a(n)$ and instead of $a$, we write $(a_n)$, $(a_n)_1^\infty$ or $(a_n)_{n=1}^\infty$.

**Definition** (Convergence of sequence)**.** Let $(a_n)$ be a sequence and $\ell \in \mathbb{R}$. Then $a_n$ *converges to* $\ell$, *tends to* $\ell$, or $a_n \to \ell$ , if

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

**Remark.** Intuitively, from some point, the sequence is bounded somewhere around the limit.

**Lemma** (Archimedean property v2)**.** $1/n \to 0$.

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

**Note.** The red parts point to the *definition* of a sequence (proving *directly* from definition).

**Definition** (Bounded sequence)**.** A sequence $(a_n)$ is *bounded* if

$$\exists C \; \forall n : \; |a_n| \leq C.$$

A sequence is *eventually bounded* if

$$\exists C \; \exists N \; \forall n \geq N : \; |a_n| \leq C.$$

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

*Proof.* Let $C, N$ so $\forall n \geq N$ $|a_n| \leq C$. So $\forall n \in \mathbb{N}$, $|a_n| \leq \max\{|a_1|, \cdots, |a_{N-1}|, C\}$. □

**Remark.** Proving a sequence is eventually bounded is often easier than bounded.

## 2.1 Sums, products and quotients

**Theorem.**

(i) If $a_n \to a$ and $b_n \to b$, then $a_n + b_n \to a + b$.

(ii) Let $a_n \to a$ and $\lambda \in \mathbb{R}$. Then $\lambda a_n \to \lambda a$.

(iii) Let $(a_n)$ be bounded and $b_n \to 0$. Then $a_n b_n \to 0$.

(iv) Let $a_n \to a$ and $b_n \to b$. Then $a_n b_n \to ab$.

*Proof.*

(i) Let $\varepsilon > 0$. 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 $\forall n \geq N_2$, $|b_n - b| < \varepsilon/2$. Let $N = \max\{N_1, N_2\}$. By triangle inequality, when $n \geq N$,
$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon.$$

(ii) $\lambda = 0$ is trivial. Let $\varepsilon > 0, \lambda \neq 0$. $\exists N$ that $\forall n \geq N$, $|a_n - a| < \varepsilon/|\lambda|$. So $|\lambda a_n - \lambda a| < \epsilon$.

(iii) Let $C \neq 0$ so $\forall n : \; |a_n| \leq C, \varepsilon > 0$. $\exists N$ that $\forall n \geq N$, $|b_n| < \varepsilon/C$. Then $|a_n b_n| < \varepsilon$.

(iv) Let $c_n = a_n - a$ and $d_n = b_n - b$. Then $a_n b_n = (a + c_n)(b + d_n) = ab + ad_n + bc_n + c_n d_n$. $ad_n \to 0$ & $bc_n \to 0$ and since $c_n$ is bounded, $c_n d_n \to 0$. Thus, $a_n b_n \to ab$

□

*Alternative to* $(iv)$. $a_n b_n - ab = (a_n - a)b_n + (b_n - b)a$. 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 $a_n b_n \to ab$. □

**Note.** we no longer write "Let $\varepsilon > 0$" as we use the lemmas rather than the definitions.

**Lemma.** Let $(a_n)$ 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.* $a_n \to a$, so $\exists N$ that $\forall n \geq N$, $|a_n - a| \leq a/2$, so $\forall n \geq N$, $|a_n| \geq |a|/2$, $|1/(a_n a)| \leq 2/|a|^2$. So $1/(a_n a)$ is bounded. So $\frac{1}{a_n} - \frac{1}{a} = (a - a_n)/(aa_n) \to 0$. Result follows. □

**Corollary.** If $a_n \to a, b_n \to b$, and $\forall n$ $b_n \neq 0, b \neq 0$. 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 $|c_n - x| < \varepsilon$. □

**Example.** $1/2^n \to 0$. For every $n$, $n < 2^n$. So $0 < 1/2^n < 1/n$. Result follows.

**Example.** Let $k \in \mathbb{N}$ and let $\delta > 0$. Then

$$\frac{n^k}{(1+\delta)^n} \to 0.$$

By the binomial theorem,

$$(1+\delta)^n \geq \binom{n}{k+1}\delta^{k+1}.$$

Then if $n \geq 2k$,

$$\binom{n}{k+1} = \frac{n(n-1)\cdots(n-k)}{(k+1)!} \geq \frac{(n/2)^{k+1}}{(k+1)!}.$$

So for sufficiently large $n$,

$$\frac{n^k}{(1+\delta)^n} \leq \frac{n^k 2^{k+1}(k+1)!}{n^{k+1}\delta^{k+1}} = \frac{2^{k+1}(k+1)!}{\delta^{k+1}} \cdot \frac{1}{n} \to 0.$$

## 2.2 Monotone-sequences property, MSP

**Definition.** 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 similarly.

A sequence is *(strictly) monotone* if it is (strictly) increasing or decreasing.

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

**Lemma.** LUB $\Rightarrow$ MSP.

*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 LUB, it has a supremum $s$.

Let $\epsilon > 0$. Since $s = \sup\{a_n : n \in \mathbb{N}\}$, there exists an $N$ such that $a_N > s - \epsilon$. Then since $(a_n)$ is increasing, $\forall n \geq N$, we have $s - \epsilon < a_N \leq a_n \leq s$. So $|a_n - s| < \epsilon$. □

**Lemma.** MSP $\Rightarrow$ Archimedean property.

*Proof.* We prove version 2, ie. that $1/n \to 0$.

Since $1/n > 0$ and is decreasing, by MSP, it converges. Let $\delta$ be the limit. We must have $\delta \geq 0$, since if $\delta < 0$, then there would exist $n$ with $3\delta/2 < 1/n < \delta/2 < 0$. Contradiction.

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.** MSP $\Rightarrow$ LUB.

*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 $v_0 \geq v_1 \geq v_2 \geq \cdots$ and $v_n - u_n = \frac{v_0 - u_0}{2^n} \to 0$.

Note that here we used the Archimedean property since to prove $1/2^n \to 0$, we sandwich it with $1/n$. But to show $1/n \to 0$, we need the Archimedean property.

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.** 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 setting $\epsilon = a - x$, we can find $N$ such that $a_N > x$. □

**Remark.** This lemma can be used to streamline some of our arguments (MSP implies $\frac{1}{n} \to 0$)

**Lemma.** A sequence can have at most one limit.

*Proof.* Let $a_n \to x$ and $a_n \to y$. Let $\epsilon > 0$, and pick $N$ such that $\forall n \geq N$, $|a_n - x| < \frac{\epsilon}{2}$, $|a_n - y| < \frac{\epsilon}{2}$, then $|x - y| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Since $\epsilon$ was arbitrary, $x$ must equal $y$. □

**Definition.** Let $(a_n)$ be a sequence. A *subsequence* of $(a_n)$ is a sequence of the form $a_{n_1}, a_{n_2}, a_{n_3} \cdots$ with $n_1 < n_2 < n_3 < \cdots$.

**Lemma** (Nested Interval Property)**.** Let $\mathbb{F}$ be an ordered field with MSP. let $I_1 \supset I_2 \supset \cdots$ be closed bounded non-empty intervals. Then $\cap_{n=1}^{\infty} I_n \neq \emptyset$.

*Proof.* Let $I_n = [a_n, b_n]$ 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 $(a_n)$ is bounded above, by MSP it has a limit $a$. For each $n$, $a_n \leq a$, since if $a_n > a$, we would have $a_m \geq a_n, \ \forall m \geq n \Rightarrow a > a$. Also for each $n$, we have $\forall m \geq n$, $a_m \leq b_m \leq b_n$, so $a \leq b_n$. Thus, for all $n$, $a_n \leq a \leq b_n \Rightarrow a \in I_n$, so $a \in \cap_{n=1}^{\infty} I_n$. □

**Theorem** (Bolzano-Weierstrass Theorem)**.** Let $\mathbb{F}$ be an ordered field with MSP. Then every bounded sequence has a convergent subsequence.

*Proof.* Let $u_0/v_0$ be a lower/upper bound, respectively, for a sequence $(a_n)_1^{\infty}$. By repeated bisection, we can find a sequence of intervals $[u_0, v_0] \supset [u_1, v_1] \cdots$ such that $v_n - u_n = \frac{v_0 - u_0}{2^n}$ for each $n$, and such that each $[u_n, v_n]$ contains infinitely many terms of $(a_n)$.

By the nested-intervals property, let $x \in \cap_{n=1}^{\infty}[u_n, v_n] \neq \emptyset$. Now pick a subsequence $a_{n_1}, a_{n_2}, a_{n_3}. \cdots$ such that $a_{n_k} \in [u_k, v_k]$ for each $k$. We can do this since $[u_k, v_k]$ contains infinitely many $a_n$, so in particular contains $a_n$ with $n > n_{k-1}$.

Let $\epsilon > 0$. We can find $K$ that $\frac{v_0 - u_0}{2^K} < \epsilon$. This implies that $[u_K, v_K] \subset (x - \epsilon, x + \epsilon)$ since $x \in [u_K, v_K]$. Then $\forall k \geq K$, $a_{n_k} \in [u_k, v_k] \subset [u_K, v_K] \subset (x - \epsilon, x + \epsilon)$, so $|a_{n_k} - x| < \epsilon$. □

**Definition.** A sequence $(a_n)$ is *Cauchy* if $\forall \epsilon > 0$ there $\exists N$ such that $\forall p, q \geq N$, $|a_p - a_q| < \epsilon$.

**Remark.** This basically means the terms are going to be infinitely close to each other.

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

*Proof.* Let $a_n \to a$. Let $\epsilon > 0$, Pick $N$ such that $\forall n \geq N$, $|a_n - a| < \frac{\epsilon}{2}$. Then $\forall p, q \geq N$, $|a_p - a_q| \leq |a_p - a| + |a - a_q| < \epsilon$. □

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

*Proof.* Let $\epsilon > 0$. Pick $N$ such that $\forall p, q \geq N$, $|a_p - a_q| < \frac{\epsilon}{2}$. Then pick $K$ such that $n_K \geq N$, and $|a_{n-K} - a| < \frac{\epsilon}{2}$. Then $\forall n \geq N$, we have $|a_n - a| \leq |a_n - a_{n_K}| + |a_{n_K} - a| < \epsilon$. □

**Theorem** (The General Principle of Convergence)**.** Let $\mathbb{F}$ be an ordered field with MSP. Then every Cauchy sequence in $\mathbb{F}$ converges.

*Proof.* Let $(a_n)$ be Cauchy. It is eventually bounded, as $\exists N\ \forall n \geq N$, $|a_n - a_N| \leq 1$, so $(a_n)$ is bounded. By Bolzano-Weierstrass, it has a convergent subsequence, so $(a_n)$ converges. □

**Lemma.** Let $\mathbb{F}$ be an ordered field with the Archimedean property and that every Cauchy sequence converges. Then $\mathbb{F}$ satisfies the MSP.

*Proof.* We show that every increasing non-Cauchy sequence is not bounded above. Let $(a_n)$ be an increasing sequence. If $(a_n)$ is not Cauchy, then $\exists \epsilon > 0$ such that $\forall N$, $\exists p, q$ such that $|a_p - a_q| > \epsilon$. Since $(a_n)$ is increasing, if we set $q = n$, we know that $\exists \epsilon > 0$ such that $\forall N$, $\exists p > N$ such that $a_p \geq a_N + \epsilon$. Then we can construct a subsequence $a_{n_1}, a_{n_2}, \cdots$ such that every term is at least $\epsilon$ bigger than its previous one, so $a_{n_k} \geq a_{n_1} + (k-1)\epsilon$. By the Archimedean property, $a_{n_k}$ and hence $a_n$ is unbounded. □

**Definition.** An ordered field in which every Cauchy sequence converges is called *complete*.

**Definition.** Let $(a_n)$ be a bounded sequence. We define the following (same for infimum):

$$\lim_{n\to\infty} \sup a_n = \lim_{n\to\infty} \big(\sup_{m\geq n} a_m\big)$$

To show that this exists, set $b_n = \sup_{m\geq n} a_m$. Then the sequence $(b_n)$ is decreasing and bounded below by any lower bound for $(a_n)$ (since $b_n \geq a_n$). So it converges, by MSP.

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

(i) $a_n \to a$

(ii) $\lim\limits_{n\to\infty} \sup a_n = \lim\limits_{n\to\infty} \inf a_n = a$

*Proof.* If $a_n \to a$, then let $\epsilon \to 0$. Then $\exists n \forall m \geq n$, $a - \epsilon \leq a_m \leq a + \epsilon$. It follows that $a_\epsilon \leq \inf_{m\geq n} a_m \leq \sup_{m\geq n} a_m \leq a + \epsilon$. Since $\epsilon$ was arbitrary, it follows that:

$$\lim\inf a_n = \lim\sup a_n = a$$

Conversely, if $\lim\limits_{n\to\infty} \sup a_n = \lim\limits_{n\to\infty} \inf a_n = a$, then let $\epsilon > 0$. Then we can find $n$ such that $\inf_{m\geq n} a_m > a - \epsilon$ and $\sup_{m\geq n} a_m < a - \epsilon$. It follows that $\forall m \geq n$, $|a_m - a| < \epsilon$. □

# 3 Convergence of infinite series

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

**Lemma.** If $\sum\limits_{n=1}^{\infty}$ converges, then $a_n \to 0$.

*Proof.* Let $\sum_{n=1}^{\infty} = s$. Then $S_N \to s$, $S_{N-1} \to s$ and $a_N = S_N - S_{N-1} \to 0$. □

**Remark.** The converse is FALSE! (harmonic series)

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

*Proof.* The sequence $(S_N)$ is increasing, so the result follows from MSP. □

**Lemma** (Comparison Test)**.** Let $(a_n)$ be nonnegative sequences and suppose that $\exists C$ and $\exists N$ such that $\forall n \geq N$, we have $a_n \leq Cb_n$. If $\sum_{n=1}^{\infty} b_n$ 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 $T_R = \sum_{n=1}^{R} b_n$. Then we want the $S_R$ to be bounded above. $S_M - S_N = \sum_{N+1}^{M} a_n \leq C \sum_{N+1}^{M} b_n \leq C \sum_{n=N+1}^{\infty} b_n$, so $\forall M \geq N, S_M \leq S_N + C \sum_{n=N+1}^{\infty} b_n$. Since the $S_N$ are increasing, we have $\forall M, S_M \leq S_N + C \sum_{n=N+1}^{\infty} b_n$. □

**Example.**

(i) $\sum_1^{\infty} \frac{n}{2^n}$ converges.

*Proof.* If $n \geq 4$ then $n \leq 2^n$. That's because $4 = 2^{4/2}$ and for $n \geq 4, \frac{n+1}{n} < \sqrt{2}$. Hence, by comparison test, it is enough if $\sum 2^{-\frac{n}{2}}$ converges, and it does. □

(ii) $\sum_{n=1}^{\infty} \frac{n+5}{n^3 - \frac{7n^2}{2}}$ converges if $\sum \frac{1}{n^2}$ converges.

*Proof.* $n^3 - \frac{7n^2}{2} = n^2(n - \frac{7}{2})$. So if $n \geq 8$, then $n^3 - \frac{7n^2}{2} \geq \frac{n^3}{2}$. also $n + 5 \leq 2n$. So $\frac{n+5}{n^3 - \frac{7n^2}{2}} \leq \frac{4}{n^2}$. SO true by comparison test. □

(iii) If $\alpha > 1$, then $\sum \frac{1}{n^\alpha}$ converges.

*Proof.* Let $S_N = \sum_{n=1}^{N} \frac{1}{n^\alpha}$. Then $S_{2^n} - S_{2^n-1} = \frac{1}{(2^{n-1}+1)^\alpha} + \cdots + \frac{1}{(2^n)^\alpha} \leq \frac{2^{n-1}}{(2^{n-1})^\alpha} = (2^{1-\alpha})^{n-1}$. But $2^{1-\alpha} < 1$. So $S_{2^n}$ is bounded above by the GP $1, 2^{1-\alpha}, (2^{1-\alpha})^2, \cdots$. □

## 3.1 Absolute convergence

**Definition.** A series $\sum a_n$ converges *absolutely* if the series $\sum |a_n|$ converges.

**Lemma.** If $\sum a_n$ converges absolutely, then $a_n$ 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 shall show that the sequence $(S_N)$ is Cauchy. If $p > q$, then $|S_p - S_q|$=$\sum_{n=q+1}^{p} a_n \leq \sum_{q+1}^{p} |a_n| = T_p - T_q$. But the sequence $(T_p)$ converges, so $\forall \epsilon > 0$ we can find $N$ such that $\forall p > q \geq N$, $T_p - T_q < \epsilon$, which implies that $|S_p - S_q| < \epsilon$. □

**Definition.** A series $\sum a_n$ *converges unconditionally* if $\sum_{n=1}^{\infty} a_{\pi(n)}$ converges for every bijection $\pi : \mathbb{N} \to \mathbb{N}$.

**Theorem.** If $\sum a_n$ converges absolutely then it converges unconditionally.

*Proof.* Let $S_N = \sum_{n=1}^{N} a_{\pi(n)}$. Then if $p > q$, $|S_p - S_q|$=$|\sum_{n=q+1}^{p} a_{\pi(n)}| \leq \sum_{n=q+1}^{p} |a_{\pi(n)}|$. Let $\epsilon > 0$. Pick $M$ such that $\sum_{n=M+1}^{\infty} |a_n| < \epsilon$. pick $N$ large enough that $\{1, 2, \cdots, M\} \subset \{\pi(1), \cdots, \pi(N)\}$. Then if $n > N$ we have $\pi(n) > M$. Therefore if $p > q \geq N$, then $|S_p - S_q| \leq \sum_{n=q+1}^{p} |a_{\pi(n)}| \leq \sum_{n=m+1}^{\infty} |a_n| < \epsilon$. Therefore the sequence $(S_N)$ is Cauchy. □

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

*Proof.* Suppose that $\sum |a_n| = \infty$. Let $(b_n)$ be the subsequence of non-negative terms of $(a_n)$ and let $(c_n)$ be the subsequence of negative terms. Then $\sum b_n$ and $\sum c_n$ cannot both converge, or it is easy to prove that $\sum |a_n|$ converges. Without loss of generality, $\sum b_n = \infty$. Now construct a sequence $0 = n_0 < n_1 < n_2 < \cdots$ such that for all $k$, $b_{n_{k-1}+1} + \cdots + b_{n_k} + c_k \geq 1$. Then we arrange the rearrangement in this form $(b_1, b_2, \cdots, b_{n_1}, c_1, b_{n_1+1} \cdots)$ and it diverges as it sums up to at least $k$ to $c_k$, so the partial sums tend to infinity. □

**Lemma** (The Alternating Series 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.

*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 increasing. $S_{2n+1} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2n} - a_{2n+1}) \leq a_1$ and $(S_{2n+1})$ is decreasing. Also, $S_{2n+1} - S_{2n} = a_{2n+1} \geq 0$. From MSP we have $(S_{2n})$ and $(S_{2n+1})$ converges and since $S_{2n+1} - S_{2n} = a_{2n+1} \to 0$, they tend to the same limit. □

**Lemma** (Ratio Test)**.**

(i) If then exists $\rho \in (-1, 1)$ such that $\frac{a_{n+1}}{a_n} \to \rho$, then $\sum a_n$ converges.

(ii) If there exists $N$ and $0 \leq c < 1$ such that $|\frac{a_{n+1}}{a_n}| \leq c$ for every $n \geq N$, then$\sum a_n$ converges.

*Proof.*

(i) If $\frac{a_{n+1}}{a_n} \to \rho$, so setting $\epsilon = \frac{1-|\rho|}{2}$, there exists $N$ such that $\forall n \geq N$, $|\frac{a_{n+1}}{a_n} \leq \frac{1+|\rho|}{2} < 1$. Now the result follows from the proof below.

(ii) For all $k \geq 0$ we have $|a_{N+k} \leq c^k|a_N|$, so $\sum_{k=0}^{\infty}|a_{N+k}|$ converges, and thus so does $\sum_{n=1}^{\infty}|a_n|$.

□

**Theorem** (Condensation test)**.** Let $(a_n)$ be decreasing and non-negative. Then $\sum_{n=1}^{\infty} a_n < \infty$ iff

$$\sum_{k=1}^{\infty} 2^k a_{2^k} < \infty.$$

*Proof.*

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

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

To prove the other way round, simply group as

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

□

**Example.** 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 $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

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

## 3.2 Complex Versions

Most definitions in the course so far carry over unchanged. The two exceptions are least upper bounds, and monotone sequences. There is still convergence and Cauchy sequences as we can separate the real and complex parts.

The Bolzano-Weierstrass Theorem still holds. If $(z_n)$ is bounded, $z_n = x_n + iy_n$, then $(x_n), (y_n)$ are bounded. Find a subsequence $(x_{n_k})$ converges. Then find a subsequence $(y_{n_{k_j}})_{j=1}^{\infty}$ of $(y_{n_k})$ that converges. Then the sequence $(z_{n_{k_j}})_{j=1}^{\infty}$ is a convergent subsequence of $(z_n)$. The proof that absolutely convergent sequences converge still works. so ratio test still works.

#### 3.2.1 Partial Summation

Let $(a_n), (b_n)$ be two sequences, and for each $n$ let $S_n$ be the partial sum $\sum_{k=1}^{n} a_k$. Then

$$\begin{aligned}\sum_{n=M+1}^{N} a_n b_n &= \sum_{n=M+1}^{N} (S_n - S_{n-1}) b_n \\ &= \sum_{n=M+1}^{N} S_n b_n - \sum_{n=M}^{N-1} S_n b_{n+1} \\ &= S_N b_N - S_M b_{M+1} + \sum_{n=M+1}^{N-1} S_n (b_n - b_{n+1})\end{aligned}$$

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

*Proof.* Let $S_n = \sum_{k=1}^{n} z^k = \frac{z - z^{n+1}}{1-z}$. Then $|S_n| \leq \frac{2}{|1-z|}$ for every $n$. By partial summation:

$$\sum_{n=M+1}^{N} z^n a_n = S_N a_N - S_M a_{M+1} + \sum_{n=M+1}^{N-1} S_n (a_n - a_{n+1})$$

$$\begin{aligned}\left| \sum_{n=M+1}^{N} z^n a_n \right| &\leq \frac{2}{|1-z|} \left( a_N + a_{M+1} + \sum_{n=M+1}^{N-1} (a_n - a_{n+1}) \right) \\ &= \frac{2}{|1-z|} (a_N + a_{M+1} + a_{M+1} - a_N) = \frac{4a_{M+1}}{|1-z|}\end{aligned}$$

So if $T_N = \sum_{n=1}^{N} a_n z^n$, then we have that $|T_N - T_M| \leq \frac{4|a_{N+1}|}{|a-z|}$ which tends to 0 as $M \to \infty$. It follows that $(T_N)$ is Cauchy, so $T_N$ converges. □

# 4 Continuous functions

## 4.1 Continuous Induction*

**Theorem** (Version 1)**.** let $a < b$ and let $A \subset [a, b]$ satisfy:

(i) $a \in A$

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

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

Then $b \in A$.

*Proof.* Since $a \in A$, $A \neq \emptyset$. Let $s = \sup A$. Then $\forall \epsilon > 0$, there exists $y \in A$, $y > s - \epsilon$. By $(iii)$, $s \in A$. If $s \neq b$, then by (ii) we can find $y \in A$ such that $y > s$. □

**Theorem** (Version 2)**.** Let $A \subset [a, b]$ and suppose that:

(i) $a \in A$

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

(iii) If $[a, x) \subset A$, then $[a, x] \in A$.

*Version 1 → Version 2.* let $A' = \{x \in [a, b] : [a, x] \subset A\}$. Then $a \in A'$. If $x \in A', x \neq b$, then $[a, x] \subset A$, so $\exists y > x$ such that $[a, y] \subset A$, so $\exists y > x$ such that $y \in A'$. If $\forall \epsilon > 0$, $\exists y \in (x - \epsilon, x]$ such that $[a, y] \in A$, then $[a, x) \subset A$, so by $(iii)$, $[a, x] \in A$, so $x \in A'$. Thus $A'$ satisfies conditions of version 1, and it is a subset of $A$, so version 2 is proved. □

## 4.2 Continuous functions

**Definition** (Continuous function)**.** Let $A \subseteq \mathbb{R}$, $a \in A$, and $f : A \to \mathbb{R}$. Then $f$ is *continuous at* $a$ if

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

$f$ is *continuous* if it is continuous at each every $a \in A$. That is

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

**Remark.** Intuitively, this means that we can approximate $f(a)$ as closely as we want, as long we take our $y$ to be within a certain $\sigma$ of $a$.

**Example.**

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

**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 $\forall n \geq N$, $|f(a_n) - f(a)| < \varepsilon$. It is therefore 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 $a_n \to a$ , but $f(a_n) \not\to f(a)$.

□

**Example.** 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 $a$.

**Lemma.** Let $A \subseteq \mathbb{R}$ and $f, g : A \to \mathbb{R}$ be continuous functions. Then $f + g$, $fg$ and $f/g$ (if $g$ never vanishes) are continuous.

*Proof.* 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

$$f(a_n) + g(a_n) \to f(a) + g(a) = (f + g)(a).$$

Proof for others are similar. □

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.** Let $f : A \to B$ ad $g : B \to \mathbb{R}$. 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 let $\epsilon \to 0$. Since $g$ is continuous at $f(a)$, there exists $\eta > 0$, such that $\forall z \in B$, $|z - f(a)| < \eta \Rightarrow |g(z) - g(f(a))| < \epsilon$. Since $f$ is continuous at $a$, $\exists \sigma > 0$ such that $\forall \eta \in A$, $|y - a| < \sigma \Rightarrow |f(y) - f(a)| < \eta$. Therefore, $|y - a| < \sigma \Rightarrow |g(f(y)) - g(f(a))| < \epsilon$.

□

**Definition.** Let $A \subset \mathbb{R}$. A *cover* of A by open intervals is a set $\{I_\gamma : \gamma \in \Gamma\}$ where each $I_\gamma$ is an open interval and $A \subset \cup_{\gamma\in\Gamma} I_\gamma$. A *finite subset* is a set $\{I_{\gamma_1}, \cdots, I_{\gamma_n}\}$, $\gamma_i \in \Gamma$ such that $A \subset \cup_{i=1}^n I_{\gamma_i}$,

**Theorem** (Heine-Borel)**.** Every cover of a closed interval $[a, b]$ by open intervals has a finite sub cover.

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

Now suppose that $\forall\epsilon > 0, \exists y \in (x - \epsilon, x]$ for some $\epsilon > 0$. Pick $y \in A$, $y \in (x - \epsilon, x]$. Now combine $I_y$ with a finite subset of $[a, y]$ to get a finite subset of $[a, x]$. 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 its bounds.

*Proof.* If $f$ is not bounded above, then for each $n$, we can find $x_n$ in $[a, b]$ such that $f(x_n) \geq n$. By Bolzano-Weierstrass, 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. Similarly, one prove it is bounded below.

Now let $C = \sup\{f(x) : x \in [a, b]\}$. Then for every $n$ we can find $x_n$ such that $f(x_n) \geq C - \frac{1}{n}$. 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 $f(x) = C$. Similarly, we prove the infimum. □

*Continuous Induction.* 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 $\exists\sigma > 0$ such that $|y - x| < \sigma \Rightarrow |f(y) - f(x)| < 1$. So $\exists y > x$ such that $f$ is bounded in $[x, y]$, which implies that $y \in A$.

Now suppose that $\forall\epsilon > 0, \exists y \in (x - \epsilon, x]$ such that $y \in A$. Again we can find $\sigma > 0$ such that $f$ is bounded on $(x - \sigma, x + \sigma)$ and in particular on $(x - \sigma, x]$. Pick $y$ such that $f$ is bounded on $[a, y]$ and $y > x - \sigma$. Then $f$ is bounded on $[a, x]$, so $x \in A$. Done by continuous induction. □

*Henie-Borel.* Let $f : [a, b] \to \mathbb{R}$ be continuous. Then by continuity, $\forall x \in [a, b]$, $\exists \sigma_x > 0$ such that $\forall y, y - x < \sigma_x \Rightarrow |f(y) - f(x)| < 1$. Let $\Gamma = [a, b]$ and for each $x \in \Gamma$ let $I_x = \{x - \sigma_x, x + \sigma_x\}$. By Henie-Borel we have a finite subset that covers $[a, b]$. But $f$ is bounded in each interval $(x_i - \sigma_x, x_i + \sigma_x)$, so it is bounded on $[a, b]$. □

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

*Proof.* Let $A = \{x : f(x) < 0\}$ and let $s = \sup A$. We shall show that $f(s) = 0$. If $f(s) < 0$, then we can find $\delta > 0$ (setting $\epsilon = \frac{|f(s)|}{2}$) such that $\forall y, |y - s| < \sigma \Rightarrow f(y) < \frac{f(s)}{2} < 0$. In particular, $s + \frac{\epsilon}{2} \in A$, so $s$ isn't an upper bound.

If $f(s) > 0$, then we can find $\sigma > 0$ such that $\forall y, |y - s| < \sigma \Rightarrow f(y) > \frac{f(s)}{2}$. So $s$ is an upper bound for A, and $s - \frac{\sigma}{2}$ is a smaller upper bound.

□

*Repeated Bisection.* 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 $f(a_n) < 0$, and $f(b_n) \geq 0$. Then by nested interval property, the intersection of all of these intervals is non-empty. We can find $x$ in this intersection. Moreover, since $a_n \to x$ and $b_n \to x$. Since $f(a_n) < 0$ and $f(b_n) \geq 0$ for all n, $\lim f(a_n) \leq 0$ and $\lim f(b_n) \geq 0$, so $f(x) = 0$. □

*Continuous Induction.* Assume $f$ is continuous, $f(a) < 0$, $f(b) > 0$. Let $A = \{x : f(x) < 0\}$. Then $a \in A$. If $x \in A$, then $f(x) < 0$ and by continuity we can find $\sigma > 0$ such that $|y - x| < \sigma \Rightarrow f(y) < 0$ so if $x \neq b$ then we can find $y \in A$ such that $y > x$. If $x \neq A$, then $f(x) > 0$ (if we assume that $f$ is never zero). Then by continuity $\exists \sigma > 0$ such that $|y - x| < \sigma \Rightarrow f(y) > 0$, so $y \not\subset A$. Therefore, $A \cap (x - \sigma, x] = \emptyset$. Hence, by continuous induction, $b \in A$. □

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

*Proof.* Since $f$ is strictly increasing, it is an injection. Now let $y \in (c, d)$. By IVT, $\exists x \in (a, b)$ 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 $\epsilon > 0$. Let $x = g(y)$, so $f(x) = y$. Let $u = f(x - \epsilon), v = f(x + \epsilon)$. Then $u < y < v$, so we can find $\sigma > 0$ such that $(y - \sigma, y + \sigma) \subset (u, v)$. Then $|z - y| < \sigma \Rightarrow g(z) \in (x - \epsilon, x + \epsilon) \to |g(z) - g(y)| < \epsilon$. □

# 5 Differentiability

## 5.1 Limits

**Definition.** Let $A \subset \mathbb{R}$ and let $f : A \to R$. We say $\lim\limits_{x \to a} f(x) = l$ if $\forall \epsilon > 0$, $\exists \sigma > 0$ such that $\forall x \in A, x \neq a\ |x - a| < \sigma \Rightarrow |f(x) - l| < \epsilon$. We don't care what happens at $x = a$.

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

## 5.2 Differentiation

**Definition.** We say f is *differentiable at $a$ with derivative $\lambda$* if:

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

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

**Proposition.**

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

*Proof.*

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

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

$$f(x+h) = f(x) + hf'(x) + h\epsilon(h),$$

□

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

*Proof.*

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

□

**Definition** (Multiple derivatives)**.** This is defined recursively: $f$ is $(n+1)$-times differentiable if it is $n$-times differentiable and its $n$-th derivative $f^{(n)}$ is differentiable. We write $f^{(n+1)}$ for the derivative of $f^{(n)}$, ie. then $(n+1)$th derivative of $f$.

Informally, we will say $f$ is $n$-times differentiable if we can differentiate it $n$ times, and the $n$th derivative is $f^{(n+1)}$.

**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.* Both proofs are trivial and use the small $o$ approximation above. □

**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.* In this proof, we use $h\epsilon(h)$ instead of $o(h)$.

$$\begin{aligned}(g\circ f)(x+h) &= g(f(x+h))\\ &= g[f(x) + \underbrace{hf'(x) + h\epsilon_1(h)}_{\text{the ``}h\text{'' term}}]\end{aligned}$$

where $\epsilon_1(h) \to 0$ and $\epsilon_1(0) = 0$

$$\begin{aligned}&= g(f(x)) + \big(fg'(x) + h\epsilon_1(h)\big)g'(f(x))\\ &+ \big(hf'(x) + h\epsilon_1(h)\big)\epsilon_2(hf'(x) + h\epsilon_1(h))\\ &= g\circ f(x) + hg'(f(x))f'(x)\\ &\underbrace{+\, h\Big[\epsilon_1(h)g'(f(x)) + \big(f'(x) + \epsilon_1(h)\big)\epsilon_2\big(hf'(x) + h\epsilon_1(h)\big)\Big]}_{\text{error term}}.\end{aligned}$$

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

But $\epsilon_1(h)g'(f(x)) \to 0$, $f'(x) + \epsilon_1(h)$ is bounded, and $\epsilon_2(hf'(x) + h\epsilon_1(h)) \to 0$ because $hf'(x) + h\epsilon_1(h) \to 0$ and $\epsilon_2(0) = 0$. So our error term is $o(h)$. □

**Example.** 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** (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.* Use product rule on $f$ and $1/g$. □

**Lemma.** 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 $x$. □

**Theorem.** Let $f : [a, b] \to [c, d]$ be a 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 $1/f'(g(y))$.

Decoded, this means 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 $x = g(y)$, $x+h = g(y+k)$. Then $f(x) = y$ and $f(x+h) = y+k$. We want to write $g(y+k)$ as $g(y) + kg'(y) + o(k)$. We have

$$g(y+k) = x + h = g(y) + h. \qquad (*)$$

Also

$$y + k = f(x+h) = f(x) + hf'(x) + h\epsilon(h) = y + hf'(x) + h\epsilon(h).$$

So $k = hf'(x) + h\epsilon(h)$. Therefore

$$h = \frac{k}{f'(x) + \epsilon(h)}.$$

As $k \to 0$, so does $h$ (by continuity of $g$), and $\epsilon(0) = 0$. So

$$h = \frac{k}{f'(x) + \epsilon_1(k)} = \frac{k}{f'(x)} - \frac{k\epsilon_1(k)}{f'(x)\big(f'(x) + \epsilon_1(k)\big)} = \frac{k}{f'(x)} + k\epsilon_2(k).$$

Since $f'(x) = f'(g(y))$, substituting into $(*)$, we obtain

$$g(y+k) = g(y) + k\frac{1}{f'(g(y))} + \epsilon_2(k).$$

□

## 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 $f(x)$. As we take the limit as $h \to 0$, we know that $f'(x) \leq 0$ and $f'(x) \geq 0$. So $f'(x) = 0$. □

**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

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

*Proof.* Let

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

Then apply Rolle's theorem on $g(x)$ to get result.

□

**Theorem** (1D form 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. □

**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

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

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

*Proof.* Induct on $n$ using Rolle's theorem on $n = k$ case. □

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

*Proof.* Apply Theorem above to $f$-$g$. □

Now we show we can find a polynomial of degree $\leq n$ that satisfies the condition for $g$.

*Proof.* 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 $Q(x) = \sum_{i=0}^{n-1} f^{(k)}(a)Q_k(x)$, then $Q^{(j)}(a) = f^{(j)}(a)$ for $n = 0, 1, \cdots, n-1$, then

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

By corollary 10, we can find $x \in (a, b)$ such that $f^{(n)}(x) = P^{(n)}(x)$. That is:

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

Therefore:

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

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

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

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

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:

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

□

**Remark.** The result, known as *Taylor's theorem*, holds for negative $h$ too-just consider $g(x) = f(-x)$.

**Example.** Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $f(0) = 1$ and $f'(x) = f(x)$ for every $x$. Then for every $x$ we have $f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

*Proof.* By Taylor's theorem,

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

for some $u$ between 0 and $x$. Now we show the remainder term tends to 0 as $n \to \infty$. Now $f^{(n)}(u) = f(u)$. Since $f$ is differentiable, it is continuous, and therefore bounded on $[0, x]$. Suppose that $|f(u)| \leq C$ on $[0, x]$. Then $|\frac{f^{(n)}(u)}{n!}x^n| \leq \frac{C}{n!}|x|^n$ tends to 0 as the ratio between successive terms tend to 0. □

### 5.3.1 Complex Differentiation

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

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

exists and equals $f'(z)$.

All the usual rules and theorems apply. The derivatives of polynomials are also what you expect. However, it is very hard to have a function that is complex differentiable (much harder than real differentiable) due to the added "dimension".

## 5.4 Complex Power Series

**Definition** (Informal)**.** A complex power series is a series of form $\sum?\infty_{n=0} a_n z^n$ when $z \in \mathbb{C}$ and $a_n \in \mathbb{C}$ for each $n$. When it converges, it is a function of $z$.

**Lemma.** Suppose $\sum_{n=0}^{\infty} a_n z^n$ converges and $|w| < |z|$, Then $\sum_{n=0}^{\infty} a_n w^n$ converges absolutely.

*Proof.* $|a_n w^n| = |a_n z^n| * |\frac{w}{z}|^n$. Since $\sum_{n=0}^{\infty} a_n z^n$ converges, the terms $a_n z^n$ are bounded, so pick $C$ such that $|a_n z^n| \leq C$ for every $n$. Then $\sum_{n=0}^{\infty} |a_n w^n| \leq \sum_{n=0}^{\infty} C|\frac{w}{n}|^n$, which converges since it is a geometric progression. □

**Remark.** It follows that if $\sum_{n=0}^{\infty} a_n z^n$ does not converge and $|w| > |z|$, then $\sum_{n=0}^{\infty} a_n w^n$ does not converge.

Now let $R = \sup\{|z| : \sum_{n=0}^{\infty} a_n z^n \text{ converges}\}$ ($R$ may be infinite). If $|z| < R$, then we can find $z_0$ with $|z_0| \in (|z|, R]$ such that $\sum_{n=0}^{\infty} a_n z_0^n$ converges. So by lemma above $\sum_{n=0}^{\infty} a_n z^n$ converges. If $|z| > R$, the infinite sum of the series diverges from definition.

**Definition.** $R$ is called the *radius of convergence.* $\{z : |z| < r\}$ is called the *circle of convergence.* On the set $\{z : |z| = R\}$, the series can converges at some points and not at others.

**Lemma.** The radius of convergence of the power series $\sum_{n=0}^{\infty} a_n z^n$ is $\frac{1}{\lim\limits_{n\to\infty} \sup |a_n|^{\frac{1}{n}}}$.

*Proof.* If $|z| < \frac{1}{\limsup |a_n|^{\frac{1}{n}}}$, then $|z| \limsup |a_n|^{\frac{1}{n}} < 1$. Therefore, there exists $N$ and $\epsilon > 0$ such that

$$\sup_{n\geq N} |a_n|^{\frac{1}{n}} |z| \leq 1 - \epsilon$$

Therefore, $|a_n z^n| \leq (1-\epsilon)^n$ for every $n \geq N$, which implies (by comparison test) that $\sum a_n z^n$ converges absolutely.

if $|z| > \frac{1}{\limsup |a_n|^{\frac{1}{n}}}$, then $|z| \limsup |a_n|^{\frac{1}{n}} > 1$, so $|z||a_n|^{\frac{1}{n}} \geq 1$ for infinitely many $n$. Therefore, $|a_n z^n| \geq 1$ for infinitely many $n$, so the series doesn't converge. □

### 5.4.1 Exponential and trignometric functions

We have $e^z$ to be $\sum_{n=0}^{\infty} \frac{z^n}{n!}$. By the ratio test, this converges on all of $\mathbb{C}$. A fundamental property of this function is:

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

once we have that property, then we can say that:

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

But $|\frac{h}{2!} + \frac{h^2}{3!} + \cdots| \leq \frac{|h|}{2} + \frac{|h|^2}{4} + \cdots = \frac{\frac{|h|}{2}}{1-\frac{|h|}{2}} \to 0$ Therefore, the derivative of $e^z$ is $e^z$.

**Definition.** Consider two sequences, $(a_n)$, $(b_n)$, their *convolution* is the sequence $(c_n)$, defined by $c_n = a_0 b_n + \cdots + a_n b_0$.

**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 $a_n$ and $b_n$. Then $\sum_{n=\infty}^{\infty} c_n$ converges absolutely, and $\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} a_n \sum_{n=0}^{\infty} b_n$.

*Proof.* Consider the series:

$$a_0 b_0 + a_0 b_1 + a_1 b_1 + a_1 b_0 + a_0 b_2 + a_1 b_1 + a_0 b_2 + \cdots$$

Let $S_N = \sum^N N = 0 a_n$, $T_N = \sum_{n=0}^{N} b_n$, $U_n = \sum_{n-0}^{N} |a_n|$, $V_n = \sum_{n=0}^{N} |b_n|$m and $S_n \to S$, $T_n \to T$, $U_n \to U$, $V_n \to V$ (these limits exists as $\sum a_n$ and $\sum b_n$ converges absolutely).

Then the sum of the first $(N+1)^2$ moduli of the terms of the series is $U_N V_N$. Hence, the series converges absolutely (as all partial sums are bounded above by $UV$.

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 look at the rearrangement $a_0 b_0 + a_0 b_1 + a_1 b_0 + \cdots$ Then this converges to $ST$ as well, and the partial sum up to $N$ is sum of $c_n$ up to $N$. So $\sum_{n=0}^{\infty} c_n \to ST = \sum_{n=0}^{\infty} a_n \sum_{n=0}^{\infty} b_n$. The absolute convergences follow easily by applying this to $|a_n|$ and $|b_n|$. □

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

*Proof.* By theorem above, we have:

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

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

$$= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} = e^{z+w}$$

by the binomial theorem. □

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

**Definition.**

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \cos z = \frac{e^{iz} + e^{-iz}}{2}$$

Then, we can use the definition of $e^z$ to prove some of the basic properties of trig functions:

$$\frac{d(\sin z)}{dz} = \frac{ie^{iz} + ie^{iz}}{2i} = \cos z$$

$$\frac{d(\cos z)}{dz} = \frac{ie^{iz} - ie^{-iz}}{2} = -\sin z$$

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

$$\cos^2 z + \sin^2 z = 1$$

The last line states that if $\cos z$ and $\sin z$ are real, then their modulus is at most 1. Since $\frac{d}{dx} \sin x = \cos x \leq 1$ and $\sin 0 = 0$, we have $\sin x \leq x$ by the mean value theorem. Similarly we have $\cos x \geq 1 - \frac{x^2}{2}$ when $x \geq 0$. Continuing in this way, it will be $\geq \sin z, \cos z$ if you stop at a positive term and $\leq$ if you stop at a negative one.

Since $\cos 2 < 0$, we deduce from intermediate value theorem that there exists $x \in (0, 2)$ such that $\cos x = 0$.

**Definition.** Define the smallest such $x$ to be $\frac{\pi}{2}$.

Then $\cos(z + \frac{\pi}{2}) = -\sin z \sin\frac{\pi}{2}$. $\sin\frac{\pi}{2} = \pm 1$. Since $\cos x \geq 0$ for $[0, \frac{\pi}{2}]$, $\sin\frac{\pi}{2} \geq 0$ by MVT, so $\sin\frac{\pi}{2} = 1$.

### 5.5 Differentiating power series

We shall show that inside the circle of convergence, the derivative of $\sum_{n=0}^{\infty} a_z^n$ is given by the obvious formula $\sum_{n=1}^{\infty} na_n z^{n-1}$.

**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. □

**Lemma.** Let $a_n z^n$ have radius of convergence $R$, and let $|z| < R$. Then $\sum na_n z^{n-1}$, $\sum_{n=2}^{\infty}\binom{n}{2}a_n z^{n-2}, \cdots$ 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

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

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. Apply first case again for successive cases. □

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

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

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

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

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 the remainder term is of $O(h^2) = o(h)$. □

### 5.6 Hyperbolic trigonometric functions

**Definition** (Hyperbolic sine and cosine)**.** We define

$$\cosh z = \frac{e^z + e^{-z}}{2} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \frac{z^6}{6!} + \cdots$$
$$\sinh z = \frac{e^z - e^{-z}}{2} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \frac{z^7}{7!} + \cdots$$

Either from the definition or from differentiating the power series, we get

(i) $\frac{\mathrm{d}}{\mathrm{d}z} \cosh z = \sinh z$

(ii) $\frac{\mathrm{d}}{\mathrm{d}z} \sinh z = \cosh z$

(iii) $\cosh iz = \cos z$

(iv) $\sinh iz = i \sin z$

(v) $\cosh^2 z - \sinh^2 z = 1$

# 6 The Riemann Integral

## 6.1 Integration

**Definition** (Dissections)**.** Let $[a, b]$ be a closed interval. A *dissection* of $[a, b]$ is sequence $a = x_0 < x_1 < x_2 < \cdots < x_n = b$.

**Definition** (Upper and lower sums)**.** Given a dissection $\mathcal{D}$, the *upper sum* and *lower sum* are defined by the formulae

$$S_{\mathcal{D}}(f) = \sum_{i=1}^{n} (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} f(x)$$
$$s_{\mathcal{D}}(f) = \sum_{i=1}^{n} (x_i - x_{i-1}) \inf_{x \in [x_{i-1}, x_i]} f(x)$$

Sometimes we use the shorthand

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

**Definition** (Refining dissections)**.** If $\mathcal{D}_1$ and $\mathcal{D}_2$ are dissections of $[a, b]$, we say that $\mathcal{D}_2$ refines $\mathcal{D}_1$ if every point of $\mathcal{D}_1$ is a point of $\mathcal{D}_2$.

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

$$S_{\mathcal{D}_2} f \leq S_{\mathcal{D}_1} f \text{ and } s_{\mathcal{D}_2 f} \geq s_{\mathcal{D}_1} f.$$

*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

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

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

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

**Definition** (Least common refinement)**.** If $\mathcal{D}_1$ and $\mathcal{D}_2$ be dissections of $[a, b]$. Then the least common refinement of $\mathcal{D}_1$ and $\mathcal{D}_2$ is the dissection made out of the points of $\mathcal{D}_1$ and $\mathcal{D}_2$.

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

$$S_{\mathcal{D}_1} f \geq s_{\mathcal{D}_2} f.$$

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

$$S_{\mathcal{D}_1} f \geq S_{\mathcal{D}} f \geq s_{\mathcal{D}} \geq s_{\mathcal{D}_2} f.$$

□

**Definition** (Upper, lower, and Riemann integral)**.** The *upper integral* is

$$\overline{\int_a^b} f(x)\ \mathrm{d}x = \inf_{\mathcal{D}} S_{\mathcal{D}} f.$$

The *lower integral* is

$$\underline{\int_a^b} f(x)\ \mathrm{d}x = \sup_{\mathcal{D}} s_{\mathcal{D}} f.$$

If these are equal, then we call their common value the *Riemann integral* of $f$, and is denoted $\int_a^b f(x)\ \mathrm{d}x$.

If this exists, we say $f$ is *Riemann integrable.*

**Note.** If $f$ is unbounded in $[a, b]$, then for any dissection $\mathcal{D}$, there must be some $i$ such that $f$ is unbounded on $[x_{i-1}, x_i]$. So $M_i = \pm\infty$. So $S_{\mathcal{D}} f = \pm\infty$. So unbounded functions are not Riemann integrable.

**Example.** Define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}.$$

Let $x_0 < x_1 < \cdots, x_n$ be a dissection. Then for every $i$, we have $m_i = 0$ (since there is an irrational in every interval), and $M_i = 1$ (since there is a rational in every interval). So

$$S_{\mathcal{D}} f = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) = \sum_{i=1}^{n} (x_i - x_{i-1}) = 1.$$

Similarly, $s_{\mathcal{D}} f = 0$. Since $\mathcal{D}$ was arbitrary, we have

$$\overline{\int_0^1} f(x)\ \mathrm{d}x = 1, \quad \underline{\int_0^1} f(x)\ \mathrm{d}x = 0.$$

So $f$ is *not* Riemann integrable.

**Proposition** (Riemann's Integrability criterion)**.** Let $f : [a,b] \to \mathbb{R}$. Then $f$ is Riemann integrable if and only if for every $\epsilon > 0$ then exists a dissection $D$ such that:

$$S_A f - s_A f < \epsilon$$

*Proof.* Suppose that $f$ is integrable. Then there exists $D_1$ and $D_2$ such that $S_{D_1} f < \int_a^b f(x)dx + \frac{\epsilon}{2}$ and $S_{D_2} f > \int_a^b f(x)dx - \frac{\epsilon}{2}$. Let $D$ be a common refinement of $D_1$ and $D_2$. Then $S_D f - s_D f < S_{D_1} f - s_{D_2} f < \epsilon$. Conversely, if there exists $D$ such that $S_D f - s_D f < \epsilon$, then $\inf S_D f - \sup s_D f < \epsilon$, which implies that $\overline{\int_a^b} f(x)\,dx - \underline{\int_a^b} f(x)\,dx < \epsilon$.

Since $\epsilon > 0$ is arbitrary, this gives us:

$\overline{\int_a^b} f(x)\,dx = \underline{\int_a^b} f(x)\,dx < \epsilon$.

So $f$ is integrable. □

**Proposition.** Let $f : [a,b] \to \mathbb{R}$ be integrable and let $\lambda > 0$. Then $\lambda f$ is integrable and $\int_a^b \lambda f(x)\mathrm{d}x = \lambda \int_a^b f(x)\mathrm{d}x$.

*Proof.* Let $D$ be a dissection of $[a,b]$. Since $\sup \lambda f(x) = \lambda \sup f(x)$ in the range $x \in [x_{i-1}, x_i]$ and similarly for $\inf$, so:

$$S_D(\lambda f) = \lambda S_D f$$

$$s_D(\lambda f) = \lambda s_D f$$

So if we choose $D$ such that $S_D f - s_D f < \frac{\epsilon}{\lambda}$ then $S_D(\lambda f) - s_D(\lambda f) < \epsilon$, So the result follows from Riemann's integrability criterion. □

**Proposition.** Let $f : [a,b] \to \mathbb{R}$ be integrable. Then $-\lambda f$, for $\lambda > 0$ is integrable and $\int_a^b -\lambda f(x)\mathrm{d}x = -\lambda \int_a^b f(x)\mathrm{d}x$.

*Proof.* Let $D$ be a dissection. Then $\sup -\lambda f(x) = -\lambda \sup f(x)$ and similarly for $\inf$. Therefore, $S_D(-\lambda f) = \sum_{i=1}^n (x_i - x_{i-1})(-\lambda m_i) = -\lambda s_D(f)$. and similarly $s_D(-f) = -S_D(f)$. So $S_D(-f) - s_D(-f) = S_D f - s_D f$. Result follows from Riemann integrability criterion. □

**Proposition.** Let $f, g$ be integrable on $[a,b]$. 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 $D_1$ be a dissection such that $S_{D_1} f - S_{D_1} f < \epsilon$ and let $D_2$ be a dissection such that $S_{D_2} g - S_{D_2} g < \epsilon$. Let $D$ be a common refinement. Then:

$$S_D(f+g) = \sum_{i=1}^n (X_i - X_{i-1}) \sup_{x \in [x_{i-1}, x_i]} (f(x) + g(x))$$

$$= \sum_{i=1}^n (X_i - X_{i-1}) \sup_{u \in [x_{i-1}, x_i]} f(x) + \sup_{v \in [x_{i-1}, x_i]} g(x)$$

$$= S_D f + S_D g \leq S_{D_1} f + S_{D_2} g$$

Therefore $\overline{\int_a^b} (f(x) + g(x))\mathrm{d}x \leq \overline{\int_a^b} f(x)\mathrm{d}x + \overline{\int_a^b} g(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x$. This is similarly true for the lower integral. So the result follows. □

**Proposition.** Let $f, g$ be integrable and $f(x) \leq g(x)$ for every $x$ then $\int_a^b f(x)\mathrm{d}x \leq \int_a^b g(a)\mathrm{d}x$.

*Proof.* This immediately follows 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

$$S_{\mathcal{D}}(|f|) - s_{\mathcal{D}}(|f|) \leq S_{\mathcal{D}}(f) - Ss_{\mathcal{D}}(f).$$

So the result follows from Riemann's integrability criterion.

□

**Remark.** Combining propositions above, we see if $|f(x) - g(x)| \leq C$ for every $x \in [a, b]$ then:

$$|\int_a^b f(x)\mathrm{d}x - \int_a^b g(x)\mathrm{d}x| < C(b - a)$$

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

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

*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

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

and

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

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

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

and

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

Since $S_{\mathcal{D}}(f) - s_{\mathcal{D}}(f) < 2\varepsilon$, and both $S_{\mathcal{D}_2}(f) - s_{\mathcal{D}_2}(f)$ and $S_{\mathcal{D}_1}(f) - s_{\mathcal{D}_1}(f)$ are greater than or equal to 0, we have $S_{\mathcal{D}_1}(f) - s_{\mathcal{D}_1}(f)$ and $S_{\mathcal{D}_2}(f) - s_{\mathcal{D}_2}(f)$ are less that $2\varepsilon$. Since $\varepsilon$

is arbitrary, it follows that the restrictions of $f$ to $[a, b]$ and $[b, c]$ are both Riemann integrable. Furthermore,

$$\int_a^b f(x)\ \mathrm{d}x + \int_b^c f(x)\ \mathrm{d}x \le S_{\mathcal{D}_1}(f) + S_{\mathcal{D}_2}(f) = S_{\mathcal{D}'}(f) \le S_{\mathcal{D}}(f) \le \int_a^c f(x)\ \mathrm{d}x + \varepsilon.$$

The other direction follows similarly. Since $\varepsilon$ is arbitrary, it follows that

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

□

**Definition.** Let $A \subset \mathbb{R}$ and let $f : A \to R$. Then $f$ is *uniformly continuous* if $\forall \epsilon > 0, \exists \sigma > 0, \forall x, y, |x - y| < \sigma \Rightarrow |f(x) - f(y)| < \epsilon$.

**Theorem.** 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

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x, y :\ |x - y| \le \delta \text{ and } |f(x) - f(y)| \ge \varepsilon.$$

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

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

Then by Bolzano-Weierstrass theorem, we can find a subsequence $(x_{n_k})$ converging to some $x$. Since $|x_{n_k} - y_{n_k}| \le \frac{1}{n_k}$, $y_{n_k} \to x$ as well. But $|f(x_{n_k}) - f(y_{n_k})| \ge \varepsilon$ 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.** Every continuous function $f$ on a closed bounded interval $[a, b]$ is Riemann integrable.

*Proof.* Let $\varepsilon > 0$. We want to construct a dissection of $[a, b]$ such that the difference between the supremum and infimum of each interval is small, so that the difference between $s_{\mathcal{D}}$ and $S_{\mathcal{D}}$ is small. So we want $f$ to vary very little within each interval. We can do this by uniform continuity.

Let $\varepsilon > 0$. Since $f$ is continuous, it is uniformly continuous. So we can find $\delta > 0$ such that

$$|f(y) - f(x)| < \frac{\varepsilon}{b - a}$$

whenever $|y - x| < \delta$. Let $\mathcal{D}$ be a dissection of $[a, b]$ of mesh less than $\delta$ (ie. the width of each interval is less than $\delta$). Then

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

Since for each $i$, we have $x_i - x_{i-1} < \delta$, we have $M_i - m_i < \frac{\varepsilon}{b-a}$ for each $i$. So

$$S_{\mathcal{D}} f - s_{\mathcal{D}} f < \frac{\varepsilon}{b - a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \varepsilon.$$

So we are done by Riemann's integrability criterion. □

**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 that $\frac{\varepsilon}{f(b)-f(a)}$. Then

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

□

**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.

The idea of the proof is to integrate from a point $x_1$ very near $a$ up to a point $x_{n-1}$ very close to $b$. Since $f$ is bounded, the regions $[a, x_1]$ and $[x_{n-1}, b]$ is small enough to not cause trouble.

*Proof.* Let $\varepsilon > 0$. Suppose that $|f(x)| \le C$ 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

$$S_{\mathcal{D}'}f - s_{\mathcal{D}'}f < \frac{\varepsilon}{2}.$$

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

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

So done by Riemann integrability criterion. □

**Example.**

$f(x) = \begin{cases} \sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$ defined on $[-1, 1]$ is integrable.

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

*Proof.* Partition $[a, b]$ with intervals $I_1, \cdots, I_k$ on each of which $f$ is bounded and continuous. By lemma 13, if $I_j$ is one of these intervals and has end points $x_{j-1}$, $_j$, then $f$ is integrable on $[x_{j-1}, x_j]$. But then by the additivity property of integrals, we have $f$ is integrable on $[a, b]$. □

**Lemma.** Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable and for each $n$ let $D_n$ 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 $S_{D_n}f \to \int_a^b f(x)\mathrm{d}x$ and $s_{D_n}f \to \int_a^b f(x)\mathrm{d}x$.

*Proof.* Let $\epsilon > 0$. Since $f$ is integrable, there is a dissection $D$ such that $S_D f - \int_a^b f(x)\mathrm{d}x < \frac{\epsilon}{2}$. Let $D$ be $u_0 < u_1 < \cdots < u_m$. If $n$ is such that

$$\frac{m(b-a)}{n}\cdot 2C < \frac{\epsilon}{2}$$

(where $|f(x)| \le C$ for all $x$), then we shall show that $S_{D_n}f - S_D f < \frac{\epsilon}{2}$.

Let $D'$ be the least common refinement of $D_n$ and $D$. Then $S_{D'}f \le S_D f$. Also the sums $S_{D_n}f$ and $S_{D'}f$ are the same, except that at most $m$ of the subintervals $[x_{i-1}, x_i]$ are subdivided in $D'$. For each interval that gets chopped up, the upper sum decreases by at most $\frac{b-a}{n} \cdot 2C$. Therefore

$$S_{D_n} f - S_{D'} f \leq \frac{b-a}{n} \cdot 2C \cdot m \leq \frac{\epsilon}{2}$$

. Therefore , $S_{D_n} f - S_D f < \frac{\epsilon}{2}$, so $S_{D_n} f - \int_a^b f(x) \mathrm{d}x < \epsilon$. That is true whenever $n > \frac{4C(b-a)m}{\epsilon}$. Note also that $S_{D_n} f \geq \int_a^b f(x) \mathrm{d}x$. So we have $S_{D_n} f \to \int_a^b f(x) \mathrm{d}x$. The other side is similar. □

**Definition.** If $b > a$, then we define $\int_b^a f(x) \mathrm{d}x$ to be $-\int_a^b f(x) \mathrm{d}x$.

**Theorem** (Fundamental theorem of calculus, Part 1)**.** Let $f : [a, b] \to \mathbb{R}$ be continuous, and for $x \in [a, b]$ define $F(x)$ to be $\int_a^x f(t) \mathrm{d}t$. 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 $\epsilon > 0$. Since $f$ is continuous at $z$, then exists $\sigma > 0$ such that $|y - x| < \sigma$ implies that $|f(y) - f(x)| < \epsilon$. If $|h| < \sigma$, then

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

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

□

**Corollary.** If $f$ is continuously differentiable at $[a, b]$, 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 $g'(x) = f'(x) = \frac{d}{dx}(f(x) - f(a))$, with $g(0) = 0 = f(a) - f(a)$. So by the mean value theorem, $g(x) = (f(x) - f(a))$ 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 with derivative $f$ and suppose that $f$ is integrable. Then $\int_a^b f(t) \mathrm{d}t = F(b) - F(a)$.

*Proof.* Let $D$ be a dissection $x_0 < x_1 < \cdots < x_n$. Then

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

For each $i$, there exists $u_1 \in (x_{i-1}, x_i)$ such that $F(x_i) - F(x_{i-1}) = (x_i - x_{i-1}) f(u_i)$ by the mean value theorem. So

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

Therefore, $s_A f \leq F(b) - F(a) \leq S_D f$. □

Since $f$ is integrable and $D$ was arbitrary, it follows that

$$F(b) - F(a) = \int_a^b f(t) \mathrm{d}t$$

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

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

$$\left|\sum_{i=1}^{n}(x_i - x_{i-1})(f(v_i)g(v_i) - f(u_i)g(u_i))\right|$$

$$= \sum_{i=1}^{n}(x_i - x_{i-1})(f(v_i)(g(v_i) - g(u_i)) - (f(v_i) - f(u_i))g(u_i))$$

$$\leq \sum_{i=1}^{n}(x_i - x_{i-1})(C(L_i - l_i) + (M_i - m_i)C) = C(S_D g - s_D g + S_D f - s_D f)$$

If follows since $u_i$ and $v_i$ are arbitrary, that:

$$S_D fg - s_D fg \leq C(S_D g - s_D g + S_D f - s_D f)$$

Result follows. □

**Theorem** (Integration by parts)**.** Let $f, g : [a,b] \to \mathbb{R}$ be such that $f'$ and $g'$ exist and are integrable. 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.* Note that $f(x)g'(x) + f'(x)g(x)$ is integrable, by proposition 19. Also $\int_a^b (f(x)g'(x) + f'(x)g(x))\mathrm{d}x = \int_a^b (fg)'(x)\mathrm{d}x = f(b)g(b) - f(a)g(a)$, by FTC part 2. □

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

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

*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 fundamental theorem of calculus part 2.

Now observe that:

$$\int_a^b \frac{(b-t)^n}{n!}f^{(n+1)}(t)\mathrm{d}t = [\frac{-f^{(n+1)}(t)(b-t)^{n+1}}{(n+1)!}]_a^b + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\mathrm{d}t$$

$$= f^{(n+1)}(a)\frac{(b-a)^{n+1}}{(n+1)!} + \int_a^b \frac{(b-t)^{n+1}}{(n+1)!}f^{(n+2)}(t)\mathrm{d}t$$

Therefore, if the result is true for n, then it is true for $n+1$. □

**Theorem** (Integration by Substitution)**.** Let $f : [a,b] \to \mathbb{R}$ be continuous. Let $g : [u,v] \to \mathbb{R}$ be continuous. Let $g : [u,v] \to \mathbb{R}$ 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 FTC part 1, $f$ has an antiderivative $F$ defined on $g([u,v])$. Then:

$$\int_u^v f(g(t))g'(t)\mathrm{d}t = \int_u^v F'(g(t))g'(t)\mathrm{d}t$$

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

□

## 6.2 Improper integrals

**Definition** (Improper integral)**.** Suppose that we have a function $f : [a, b] \to \mathbb{R}$ such that, for every $\varepsilon > 0$, $f$ is integrable on $[a + \varepsilon, b]$ and $\lim_{\varepsilon \to 0} \int_{a+\varepsilon}^{b} f(x)\ \mathrm{d}x$ exists. Then we define the improper integral

$$\int_a^b f(x)\ \mathrm{d}x \text{ to be } \lim_{\varepsilon \to 0} \int_{a+\varepsilon}^{b} f(x)\ \mathrm{d}x.$$

We can do similarly for integral to infinity and $[a, b - \epsilon]$.

**Example.**

$$\int_\varepsilon^1 x^{-1/2}\ \mathrm{d}x = \left[2x^{-1/2}\right]_\varepsilon^1 = 2 - 2\varepsilon^{1/2} \to 2.$$

So

$$\int_0^1 x^{-1/2}\ \mathrm{d}x = 2,$$

**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

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

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

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

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. ☐

**Example.** Since $\int_1^x \frac{1}{t^2}\ \mathrm{d}t < \infty$, it follows that $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.