

Analysis I *

Zhiyuan Bai

Compiled on May 12, 2023

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part IA course *Analysis I* in Lent 2020. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

0	Introduction	2
1	The Real and Complex Numbers	2
2	Convergence of Sequences and Series	3
2.1	Sequence	3
2.2	Series	5
2.3	Bolzano-Weierstrass	7
2.4	Bonus: Construct the Reals from the Rationals	9
3	Limits and Continuity	9
3.1	Limit of a function	9
3.2	Continuity	11
4	Differentiation	12
4.1	The Derivative	13
4.2	Mean Value Theorem	14
4.2.1	Inverse Function Theorem	15
4.2.2	L'Hôpital's Rule	16
4.3	Higher Derivatives and Taylor Series	17
4.4	Complex Differentiation	19
5	Special Functions	21
5.1	Power Series	21
5.2	The Exponentials, etc.	25

*Based on the lectures under the same name taught by Prof. J. Rasmussen in Lent 2020.

6	Integration	27
6.1	The Riemann Integral	27
6.2	Properties of the Integral	28
6.3	Fundamental Theorem of Calculus	29
6.4	Improper Integral	32

0 Introduction

Theorem 0.1. *Let $F, f : [a, b] \rightarrow \mathbb{R}$ are continuous and F differentiable with $F' = f$, we have*

$$\int_a^b f(t) dt = F(b) - F(a)$$

There is a reason why this theorem takes so long to prove. First of all obvious we need to actually define what we meant by all those terms in there and link them together. More subtly, this is a theorem about the real numbers. Suppose $F' = f = G'$, then F and G can be differed by a constant. So if we set

$$F(x) = \begin{cases} 1, & \text{if } x^2 > 2 \\ 0, & \text{otherwise} \end{cases}$$

and $G \equiv 0$. So F is differentiable at every rational number and those derivatives are 0, and so is G , but F is not constant. Hence the theorem is simply not true for rational numbers, so we must use properties of real numbers.

These are what real analysis (in a beginner's level) is going to be about.

1 The Real and Complex Numbers

Definition 1.1. The real numbers is a set \mathbb{R} such that

1. The rational numbers \mathbb{Q} is contained in \mathbb{R} .
2. \mathbb{R} is an ordered field.
3. It satisfies the least-upper-bound property.

\mathbb{R} is a field if it has two binary operations $+, \cdot$ such that $(\mathbb{R}, +), (\mathbb{R} \setminus \{0\}, \cdot)$ are abelian groups and $\forall a, b, c \in \mathbb{R}, (a + b)c = ac + bc$.

- Example 1.1.**
1. \mathbb{Q} is a field.
 2. $\mathbb{Z}/p\mathbb{Z}$ is a field for p prime and no otherwise.

An ordered field means a field such that there is a total order $<$ on \mathbb{R} such that $a > b \implies a + c > b + c$ and $a > 0, b > 0 \implies ab > 0$.

Example 1.2. \mathbb{Q} is an ordered field but $\mathbb{Z}/p\mathbb{Z}$ is not.

Definition 1.2. Let X be an ordered set and $A \subset X$. We say A has an upper bound in X if $\exists a \in X, \forall x \in A, x \leq a$. We call a an upper bound for A . If that is true, we say A is bounded above.

We say $a \in X$ is the least upper bound, or supremum of A if it is an upper bound and $a' \geq a$ whenever a' is an upper bound of A . So we write $a = \sup A$ since it is obvious that supremums are unique.

Definition 1.3. We say X has the LUBP if any nonempty subset of X that is bounded above has a supremum.

Note that \mathbb{Q} does not have LUBP by considering $\{x \in \mathbb{Q} : x^2 < 2\}$. These definitions can be easily extended to lower bounds and greatest lower bounds and greatest-lower-bound property. We write \inf or infimum for the greatest lower bound.

Proposition 1.1. 1. Given the rational numbers, we can construct the real numbers satisfying the axioms. 2. If X is an ordered field that satisfies the LUBP, then there is an isomorphism between ordered fields $\phi : X \rightarrow \mathbb{R}$.

Proof. Partially obvious. Some might be covered later. □

Proposition 1.2. 1. The natural numbers is not bounded above.

2. $\forall x \in \mathbb{R}, \exists N \in \mathbb{N}, N > x$.
3. $\forall x \in \mathbb{R}_{>0}, \exists N \in \mathbb{N}, 1/N < x$.
4. $\forall x, y \in \mathbb{R}, x < y \implies \exists q \in \mathbb{Q}, x < q < y$.
5. $\forall x \in \mathbb{R}_{>0}, \exists y \in \mathbb{R}, y^2 = x$.

Proof. All obvious. □

Definition 1.4. The complex numbers \mathbb{C} is

$$\left\{ z \in M_{2 \times 2}(\mathbb{R}) : z = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \right\}$$

which becomes a field under matrix addition and multiplication. Let $1 = I$ and $i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, then we have an obvious bijection $\mathbb{R}^2 \rightarrow \mathbb{C}$ with $(a, b) \mapsto a + bi$.

Definition 1.5. If $z = a + bi$, we set the modulus, or norm, by $|z| = \sqrt{\det z} = \sqrt{a^2 + b^2}$.

Note that $|zw| = |z||w|$.

Proposition 1.3. $\forall z, w \in \mathbb{C}, |z + w| \leq |z| + |w|$.

Proof. Trivial. □

2 Convergence of Sequences and Series

2.1 Sequence

Definition 2.1. Let X be a set, then a sequence in X is a function $f : \mathbb{N} \rightarrow X$.

We denote the sequence in the form $a_n = f(n)$.

- Example 2.1.** 1. The constant sequence $f \equiv a$ for some constant $a \in X$.
 2. The sequence $f(n) = \sqrt{n^2 + 17}$ is a sequence in positive real numbers.
 3. Flip a coin. The sequence

$$a_n = \begin{cases} 1, & \text{if the } n^{\text{th}} \text{ flip is head} \\ 0, & \text{if it is tail.} \end{cases}$$

Then a_n is a sequence in $\{0, 1\}$.

4. With last example, we can define $b_1 = 0, b_{n+1} = b_n + a_n$ makes a sequence b_n in \mathbb{Z} .

Definition 2.2. If (z_n) is a sequence of complex numbers, we say $z_n \rightarrow z$ (or z_n converges to z) as $n \rightarrow \infty$ if

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n > N, |z_n - z| < \epsilon$$

Otherwise, we say (z_n) diverges.

Note that N depends on ϵ .

- Example 2.2.**
1. The constant sequence $z_n = z$ converges to z .
 2. The sequence $z_n = 1/n$ converges to 0. Given $\epsilon > 0$ we can find N such that $N > 1/\epsilon$ since \mathbb{N} has no upper bound, then whenever $n > N$, we have $|z_n - z| = |1/n| < 1/N < \epsilon$.
 3. The sequence $z_n = n$ diverges.
 4. The sequence $z_n = (-1)^n$ diverges.

Proposition 2.1. $z_n \rightarrow z$ iff:

1. $z_n - z \rightarrow 0$.
2. $|z_n - z| \rightarrow 0$.
3. $\forall m \in \mathbb{N}, \exists N \in \mathbb{N}, \forall n > N, |z_n - z| < 1/m$.

Proposition 2.2. Suppose z_n and w_n are sequences in \mathbb{C} . If $z_n \rightarrow z, w_n \rightarrow w$, then

1. $z_n + w_n \rightarrow z + w$.
2. $z_n w_n \rightarrow zw$.
3. If $z \neq 0$, then $\exists N \in \mathbb{N}, \forall n > N, z_n \neq 0$ and for $n > N, 1/z_n \rightarrow 1/z$.
4. If $z_n = x_n + iy_n, z = x + iy$ with $x_n, x, y_n, y \in \mathbb{R}$, we have $z_n \rightarrow z \iff x_n \rightarrow x \wedge y_n \rightarrow y$.

Proof. Trivial. □

- Example 2.3.**
1. $1/(n^2) = (1/n)(1/n) \rightarrow 0 \cdot 0 = 0$.
 2. $1 \pm 1/n^2 \rightarrow 1 \pm 0 = 1$.
 3. $1/(1 \pm 1/n^2) \rightarrow 1/1 = 1$.
 4. $(n^2 - n)/(n^2 + n) = (1 + 1/n)^{-1}(1 - 1/n) \rightarrow 1 \cdot 1 = 1$.

Remark. 2 implies that if $z_n \rightarrow z$ then $cz_n \rightarrow cz$. Similarly 2,3 together show that if $z_n \rightarrow z, w_n \rightarrow w \neq 0$, then z_n/w_n is eventually well-defined and goes to z/w .

Corollary 2.3 (Uniqueness of Limit). If $z_n \rightarrow z$ and $z_n \rightarrow z'$, then $z = z'$

Proof. By 1 and 2, $0 \equiv z_n + (-1)z_n \rightarrow z + (-1)z' = z - z'$, but if $0 \rightarrow w \neq 0$, then $|w| > 0$, therefore there is some n such that $|w| = |w - 0| < |w|/2 < |w|$, contradiction. So $z = z'$. □

We sometimes write

$$\lim_{n \rightarrow \infty} z_n = z$$

if $z_n \rightarrow z$. We can write it so since limits are unique.

Proposition 2.4. Suppose $x_n \rightarrow x, y_n \rightarrow y$ are real sequences, then if $\forall n, x_n \geq y_n$, then $x \geq y$.

Note that it is not true if we replace \geq by $>$.

Proof. It suffices to prove the case when $y_n \equiv 0$.

If $x < 0$, then there is a natural number n such that $|x_n - x| < |x|$, but $|x_n - x| \geq |x|$, which is a contradiction.

For the general case, just consider the sequence $x_n - y_n$. □

Proposition 2.5 (Squeeze Rule). *If x_n, c_n, y_n are real sequences with $\forall n, x_n \geq c_n \geq y_n$ and $x_n \rightarrow c, y_n \rightarrow c$, then $c_n \rightarrow c$.*

Proof. Given $\epsilon > 0$, pick N_1 such that $\forall n > N_1, |x_n - c| < \epsilon$ and N_2 such that $\forall n > N_2, |y_n - c| < \epsilon$, so $\forall n > \max\{N_1, N_2\}$. we have

$$c + \epsilon > x_n \geq c_n \geq y_n > c - \epsilon$$

so $|c_n - c| < \epsilon$. □

Example 2.4. Since

$$\frac{n^2 - 1}{n^2} \leq \frac{n^2 + \sin n}{n^2} \leq \frac{n^2 + 1}{n^2}$$

we have $(n^2 + \sin n)/n^2 \rightarrow 1$.

Definition 2.3. Let x_n be a sequence in \mathbb{R} . We say it is monotone increasing if $x_{n+1} \geq x_n$ for all $n \in \mathbb{N}$.

Definition 2.4. A sequence x_n is bounded above if there is some constant $c \in \mathbb{R}$ such that $x_n \leq c$ for any n .

Theorem 2.6 (Monotone Sequence Theorem). *A monotone increasing sequence in \mathbb{R} that is bounded above converges.*

Note that this theorem is false if we replace \mathbb{R} by \mathbb{Q}

Proof. Let x_n be such a sequence and we let $c = \sup_{n \in \mathbb{N}} x_n$. c exists since x_n is bounded above. We shall show that $x_n \rightarrow c$.

If x_n does not tend to c , then there is some $\epsilon > 0$ such that $\forall N \in \mathbb{N}, \exists n > N, c - x_n = |x_n - c| \geq \epsilon$. However, this would mean that for any $N \in \mathbb{N}$, we can choose such an n , then it gives $x_N \leq x_n \leq c - \epsilon$, so c is not the least upper bound, contradiction. □

Definition 2.5. Let (x_n) be a sequence in \mathbb{R} , we say $x_n \rightarrow \infty$ if $\forall M \in \mathbb{R}, \exists N \in \mathbb{N}, \forall n > N, x_n > M$.

Note that a sequence that tends to infinity diverges.

Proposition 2.7. *If the sequence (x_n) is monotone increasing, then it either converges or tends to infinity.*

2.2 Series

Definition 2.6. We say the infinite series

$$\sum_{i=0}^{\infty} c_i$$

converges (diverges) if the sequence $s_n = \sum_{i=0}^n c_i$ converges (diverges).
 If s_n converges to s , we write

$$\sum_{i=0}^{\infty} c_i = s$$

Example 2.5. Consider the geometric series $\sum_{k=0}^{\infty} z^k$ where $|z| < 1$. So $s_n = (1 - z^{n+1})/(1 - z)$. Note that $z^{n+1} \rightarrow 0$ as $n \rightarrow \infty$, hence $s_n \rightarrow 1/(1 - z)$.

Lemma 2.8. *If the series sum $\sum_{i=0}^{\infty} c_i$ converges, then $c_i \rightarrow 0$ as $i \rightarrow \infty$.*

Note that the converse of the statement is false.

Proof. Trivial. □

Lemma 2.9. *Suppose $(a_n), (b_n)$ are sequences in \mathbb{C} ,*

1. *For any $K_1, K_2 \in \mathbb{N}$, $\sum_{i=K_1}^{\infty} a_n$ converges iff $\sum_{i=K_2}^{\infty} a_n$ converges.*
2. *Suppose $\sum_n a_n \rightarrow a$, $\sum_n b_n \rightarrow b$, then $\sum_n (a_n + \lambda b_n) \rightarrow a + \lambda b$.*

Proof. Obvious. □

Theorem 2.10 (Comparison Test). *Suppose $(a_n), (b_n)$ are sequences in \mathbb{R} , and $a_n \geq b_n \geq 0$ for all n , then if the series $\sum_i a_i$ converges, so does $\sum_i b_i$, and $\sum_i a_i \geq \sum_i b_i \geq 0$.*

Proof. Indubitable.¹ □

Definition 2.7. Let (c_n) be a complex sequence, we say $\sum_i c_i$ converges absolutely if $\sum_i |c_i|$ converges.

Lemma 2.11. *If $(c_n) \in \mathbb{R}$, then if $\sum_i |c_i|$ converges, then $\sum_i c_i$ converges.*

Proof. Just sum up $c_n + |c_n|$ and use Comparison Test. □

Theorem 2.12. *If $(c_n) \in \mathbb{C}$ converges absolutely, then it converges.*

Proof. Split real and imaginary part and use Comparison Test. □

Alternatively one can also use triangle inequality.

Theorem 2.13 (Strong Comparison Test). *Suppose $(c_n) \in \mathbb{C}$ and $(a_n) \in \mathbb{R}_+$ with $a_n \geq |c_n| \geq 0$. Then if $\sum_n a_n$ converges, then $\sum_n c_n$ converges absolutely.*

Proof. Analogous. □

Theorem 2.14 (Ratio Test). *Suppose $(c_n) \in \mathbb{C}$ such that*

$$[0, \infty] \ni r = \lim_{n \rightarrow \infty} \left| \frac{c_{n+1}}{c_n} \right|$$

exists, then if $r < 1$, $\sum_n c_n$ converges, and if $r > 1$, $\sum_n c_n$ diverges.

If $r = 1$, then both ways are possible.

¹The author is short of such sort of adjectives.

Example 2.6.

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}$$

where z is a complex number. This series converges absolutely since

$$\lim_{n \rightarrow \infty} \frac{|z|^{n+1} n!}{|z|^n (n+1)!} = \lim_{n \rightarrow \infty} \frac{|z|}{n+1} \rightarrow 0 < 1$$

by Ratio Test. So we can really define $\exp(z)$.

Proof. If $r > 1$, then we can find η such that $1 < \eta < r$, and $\exists N \in \mathbb{N}, \forall n > N$ we have $|c_{n+1}| > \eta |c_n|$, therefore $|c_n| \rightarrow \infty$, hence the series diverges.

If $r < 1$, then we can find η such that $r < \eta < 1$ and $\exists N \in \mathbb{N}, \forall n > N, |c_{n+1}| \leq \eta |c_n|$, so a comparison test with the geometric series shows the convergence. \square

Example 2.7 (Non-examples). Consider the harmonic series $\sum_n 1/n$. The ratio test is useless here as $(n+1)/n \rightarrow 1$ as $n \rightarrow \infty$. The sequence of the terms converges as well, so we cannot get anything from there either. And in fact, it diverges. We can compare it with the sequence $1 + 1/2 + 1/4 + 1/4 + 1/8 + 1/8 + 1/8 + 1/8 + \dots$ to show that it is indeed actually diverges.

Proposition 2.15 (Cauchy Condensation Test). *Suppose $(a_n) \in \mathbb{R}$ and $a_n \geq a_{n+1} \geq 0$ for any n . Then $\sum_n a_n$ converges iff $\sum_k 2^k a_{2^k}$ converges.*

I'd like to think that we can replace 2 by any natural number $p \geq 2$.

Example 2.8. $\sum_n n^{-p}$ converges iff $\sum_k 2^k 2^{-kp} = \sum_k (2^{1-p})^k$ converges iff $p > 1$.

Proof. Comparison. \square

Definition 2.8. Suppose (x_n) is a sequence in some set X . A subsequence of x_n is a sequence of the form x_{n_k} where $(n_k)_{k \in \mathbb{N}}$ is strictly increasing.

Lemma 2.16. *Suppose c_n is a sequence in \mathbb{C} and c_{n_k} is a subsequence, then if $c_n \rightarrow c$, then $c_{n_k} \rightarrow c$*

Proof. Immediate. \square

The converse is obviously not true.

Lemma 2.17. *If $(a_n) \in \mathbb{R}$ is monotone and a subsequence of it converges $a_{n_k} \rightarrow a$, then $a_n \rightarrow a$.*

Proof. Easy. \square

2.3 Bolzano-Weierstrass

Definition 2.9. Let (c_n) be a sequence of complex numbers. We say (c_n) is bounded if $\exists M > 0, \forall n \in \mathbb{N}, |c_n| < M$.

Theorem 2.18 (Bolzano-Weierstrass). *Any bounded sequence has a converging subsequence.*

This is false in \mathbb{Q} as one will expect.

We shall prove this by “lion-hunting”. The philosophy is like this: Given an infinite sequence of lions in a rectangular zone. We divide the rectangle by halving each side, then one of the four parts shall contain infinitely many lions. Do this again on the region having infinitely many lions. And do this again and again and again so there is a sequence of rectangles $A_0 \subset A_1 \subset A_2 \subset \dots$ such that each A_n contains infinitely many lions and that the diameters of A_n decreases (exponentially) to 0 when $n \rightarrow \infty$. So choosing a lion in each A_n (such that its index is higher than the previous lion, possible since each A_n is infinite) gives a subsequence of converging lions.

Proof. Suffices to show the real case. Take $a_0 = -M, b_0 = M$, then $[a_0, b_0]$ contains infinitely many terms of the sequence. Once we have chosen a_n, b_n such that $[a_n, b_n]$ contains infinitely many terms of the sequence, at least one of $[a_n, (a_n + b_n)/2], [(a_n + b_n)/2, b_n]$ must contain infinitely many terms of the sequence. For the former case we set $a_{n+1} = a_n, b_{n+1} = (a_n + b_n)/2$ and for the latter $a_{n+1} = (a_n + b_n)/2, b_{n+1} = b_n$. Then $b_{n+1} - a_{n+1} = (b_n - a_n)/2 = 2M/2^{n+1} \rightarrow 0$ as $n \rightarrow \infty$. Also both b_n, a_n converges as monotone sequences. Then we can choose x_{n_k} inductively: Choose $x_{n_0} \in [a_0, b_0]$, then once we have chosen x_{n_k} , we can choose n_{k+1} by choosing one such that $x_{n_{k+1}} \in [a_{k+1}, b_{k+1}]$ and $n_{k+1} > n_k$, which is possible as $[a_{k+1}, b_{k+1}]$ contains infinitely many terms. So x_{n_k} is squeezed by a_k, b_k which converges to the same limit, hence this subsequence converges. \square

Definition 2.10. Let $(z_n) \in \mathbb{C}$ be a sequence in \mathbb{C} , we say z_n is Cauchy if $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n, m > N, |x_n - x_m| < \epsilon$.

Proposition 2.19. A complex sequence converges if and only if it is Cauchy.

Proof. The “only if” part is obvious. For the other direction, it is easy that Cauchy sequences are bounded and that if it has a subsequence that converges, then it converges to the same value. Then Bolzano-Weierstrass suffices. \square

Corollary 2.20. Let $(z_n) \in \mathbb{C}$, then $\sum_n z_n$ converges iff $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n > m > N,$

$$\left| \sum_{k=m+1}^n z_k \right| < \epsilon$$

Proof. Immediate. \square

Proposition 2.21 (Alternating Series Test). Suppose $(a_n) \in \mathbb{R}$ and $a_n \geq a_{n+1} \geq 0$ and $a_n \rightarrow 0$, then

$$\sum_{n=0}^{\infty} (-1)^n a_n$$

converges.

Example 2.9. So $\sum_n (-1)^n/n$ converges even if the harmonic series diverges.

Proof. Let s_n be the n^{th} partial sum, then $s_{2n+1} = s_{2n-1} + a_{2n} - a_{2n+1} \geq s_{2n-1}$ is monotone increasing and bounded above by a_0 , so it converges. Also s_{2n} converges as well and we can squeeze it. \square

2.4 Bonus: Construct the Reals from the Rationals

The definition of convergence and Cauchy sequence makes sense in \mathbb{Q} as well. The matters, there are sequences in \mathbb{Q} that “want to converge” (being Cauchy) but don’t, like 1, 1.4, 1.41, 1.414. In fact, every $x \in \mathbb{R}$ is a limit of a sequence in \mathbb{Q} (by e.g. decimal expansion). So we can define

$$\mathbb{R} = \{(x_n) \in \mathbb{Q} : (x_n) \text{ is Cauchy}\} / \sim$$

where \sim is an equivalence relation defined by $(x_n) \sim (y_n) \iff x_n - y_n \rightarrow 0$.

Proposition 2.22. *This is indeed the real number with the intuitive definition of operations and order.*

In particular $(x_n) \geq 0 = (0, 0, \dots)$ iff $\exists N \in \mathbb{N}, \forall n > N, x_n \geq 0$.

Proof. Other properties are trivial, so we shall check the least-upper-bound property. Note that the least-upper-bound property, the monotone sequence theorem and the statement that every Cauchy sequence converges are equivalent. So we shall prove that every Cauchy sequence converges. Suppose $(x_{m,n}) = ((x_m)_n)$ is a Cauchy sequence in \mathbb{R} defined the way we made it. Then the sequence $(y_k) = x_{k,k}$ goes to the limit we want. \square

3 Limits and Continuity

If we want to differentiate, we need to make sense of a limit, so here goes.

3.1 Limit of a function

Definition 3.1. Suppose $a \in \mathbb{C}$ and $f : \mathbb{C} \setminus \{a\} \rightarrow \mathbb{C}$. We say $\lim_{z \rightarrow a} f(z) = c \in \mathbb{C}$ if

$$\forall \epsilon > 0, \exists \delta > 0, |z - a| < \delta \implies |f(z) - c| < \epsilon$$

Sometimes we write $f(z) \rightarrow c$ as $z \rightarrow a$.

We say $\lim_{z \rightarrow a} f(z)$ is undefined if there is no c with $\lim_{z \rightarrow a} f(z) = c$.

Proposition 3.1. *If $\lim_{z \rightarrow a} f(z) = c$ and $\lim_{z \rightarrow a} f(z) = d$, then $c = d$.*

Proof. Trivial. \square

Example 3.1. 1. If $f \equiv c$, then $\lim_{z \rightarrow a} f(z) = c$ for any a , just take for example $\delta = 1$.

2. If $f(z) = z$, then $\lim_{z \rightarrow a} f(z) = a$, as we can take $\delta = \epsilon$.

3. If $f(z) = 1_{\mathbb{C} \setminus \{0\}}$, then $\lim_{z \rightarrow 0} f(z) = 1$, taking for example $\delta = 1$.

4. If

$$f(z) = \begin{cases} 1, & \text{if } \operatorname{Re}(z) > 0 \\ 0, & \text{if } \operatorname{Re}(z) = 0 \\ -1, & \text{if } \operatorname{Re}(z) < 0 \end{cases}$$

then $\lim_{z \rightarrow 0} f(z)$ does not exist.

5. If $f(z) = z^2$, then $\lim_{z \rightarrow a} f(z) = a^2 = f(a)$ by choosing $\delta = \min\{\epsilon/(2|a| + 1), 1\}$.

Proposition 3.2. $\lim_{z \rightarrow a} f(z) = c$ if and only if $\forall (z_n) \in \mathbb{C}$ with $z_n \rightarrow a$, we have $f(z_n) \rightarrow c$

Proof. If the limit of the function is c , given $\epsilon > 0$, we can find δ such that $|z - a| < \delta \implies |f(z) - c| < \epsilon$. Given a sequence $z_n \rightarrow a$, we can find $N \in \mathbb{N}$ such that $\forall n > N, |z_n - a| < \delta$, so using this $N, \forall n > N, |z_n - a| < \delta \implies |f(z_n) - c| < \epsilon$, so $f(z_n) \rightarrow c$.

Conversely, if c is not the limit of that function, then $\exists \epsilon > 0, \forall \delta > 0, \exists z \in \mathbb{C}, |z - a| < \delta, |f(z) - c| > \epsilon$. So we choose such z_n for $\delta = 1/n$, then $z_n \rightarrow a$ but $|f(z_n) - c| > \epsilon$ for any n , hence $f(z_n)$ does not tend to c . \square

Proposition 3.3. Let $f, g : \mathbb{C} \setminus \{a\} \rightarrow \mathbb{C}$, if $\lim_{z \rightarrow a} f(z) = c$ and $\lim_{z \rightarrow a} g(z) = d$, then

1. $\lim_{z \rightarrow a} (f(z) + g(z)) = c + d$.
2. $\lim_{z \rightarrow a} f(z)g(z) = cd$.
3. If $c \neq 0$, then $\lim_{z \rightarrow a} 1/f(z) = 1/c$.

Proof. Using the preceding proposition and laws of limits for sequences. \square

Proposition 3.4. Suppose $f, g : \mathbb{C} \setminus \{a\} \rightarrow \mathbb{R}$ such that

$$\lim_{z \rightarrow a} f(z) = c, \lim_{z \rightarrow a} g(z) = d$$

then,

1. If $\forall z, f(z) \geq g(z)$, then $c \geq d$.
2. (Squeeze Rule) Let $h : \mathbb{C} \setminus \{a\}$ be such that $f \geq h \geq g$. Then if $c = d$, we have $\lim_{z \rightarrow a} h(z) = c$.

Proof. Immediate. \square

What if $f : A \rightarrow \mathbb{C}$ where A is a subset of \mathbb{C} ? Does it still makes sense? Does it change the limit? The answer to the second question is obviously no. But we can make sense of the limit.

Definition 3.2. Suppose A is a subset of \mathbb{C} . We say $a \in \mathbb{C}$ is a limit point of A , if $\forall \delta > 0, \exists z \in A, 0 < |z - a| < \delta$.

Example 3.2. Any $a \in [0, \infty)$ is a limit point of $(0, \infty)$, and \mathbb{Z} has no limit point.

Definition 3.3. If $f : A \rightarrow \mathbb{C}$ and $a \in \mathbb{C}$ is a limit point of A , we say $\lim_{z \rightarrow a} f(z) = c$ if $\forall \epsilon > 0, \exists \delta > 0, z \in A, 0 < |z - a| < \delta \implies |f(z) - c| < \epsilon$.

Equivalently, $f(z_n) \rightarrow c$ for any sequence $(z_n) \in A$ such that $z_n \rightarrow a$. Also, it is also easy to show that the limit is unique.

Definition 3.4. For $f : \mathbb{R} \rightarrow \mathbb{R}$, we define

$$\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a} f|_{(a, \infty)}(x), \lim_{x \rightarrow a^-} f(x) = \lim_{x \rightarrow a} f|_{(-\infty, a)}(x)$$

3.2 Continuity

Definition 3.5. Suppose $A \subset \mathbb{C}$, and $f : A \rightarrow \mathbb{C}$. We say f is continuous at $a \in A$ if

$$\forall \epsilon > 0, \exists \delta > 0, z \in A, |z - a| < \delta \implies |f(z) - f(a)| < \epsilon$$

Remark. $|f(a) - f(a)| = 0 < \epsilon$, so this condition is equivalent to say $|f(z) - f(a)| < \epsilon$ whenever $0 < |z - a| < \delta$.

Proposition 3.5. Suppose $f : A \rightarrow \mathbb{C}$ and $a \in A$, then the following conditions are equivalent:

1. f is continuous at a .
2. Either a is not a limit point of A or $\lim_{z \rightarrow a} f(z) = f(a)$.
3. $f(z_n) \rightarrow f(a)$ whenever $z_n \rightarrow a$ is a sequence in A .

Proof. Verify from definition. □

Example 3.3. 1. Constant functions are continuous at every point.

Definition 3.6. We say $f : A \rightarrow \mathbb{C}$ is continuous if f is continuous at every $a \in A$.

Definition 3.7. 2. The identity function is continuous.

3. (non-example) $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f = 1_{\mathbb{Q}}$ is not continuous at any point.
4. Consider $f : \mathbb{R} \rightarrow \mathbb{R}$ by

$$f(x) = \begin{cases} 1/q, & \text{if } x \text{ is rational and } q = \min\{q \in \mathbb{N} : \exists p \in \mathbb{Z}, x = p/q\} \\ 0, & \text{otherwise} \end{cases}$$

Then f is not continuous at any $x \in \mathbb{Q} \setminus \{0\}$ but it is continuous at $a \notin \mathbb{Q}$.

5. $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ by $x \mapsto \sqrt{x}$ is continuous.

Proposition 3.6. Suppose $f, g : A \rightarrow \mathbb{C}$ for some $A \subset \mathbb{C}$ and they are both continuous at $a \in A$. Then $f + g, fg$ are continuous at a . And if $f(a) \neq 0$, then $1/f$ is continuous at a .

Proof. Trivial. □

Example 3.4. Everywhere in \mathbb{C} we have any constant function is continuous, the identity is continuous, hence any linear function is continuous, therefore any polynomial is continuous. Also any rational function is continuous at anywhere it is defined.

The function \exp is continuous, but we will not prove it now.

Proposition 3.7. Let $A, B, C \subset \mathbb{C}$ and $f : A \rightarrow B, g : B \rightarrow C$ such that f is continuous at a and g continuous at $f(a)$, then $g \circ f$ is continuous at a .

Proof. Straightforward. □

Example 3.5. $g : (0, \infty) \rightarrow (0, \infty), g(x) = \sqrt{x}$ and $f : (0, \infty) \rightarrow (0, \infty), f(x) = x^2 + 1$, so $x \mapsto (g \circ f)(x) = \sqrt{x^2 + 1}$ is continuous.

Lemma 3.8. Suppose (x_n) is a sequence in $[a, b]$ and $x_n \rightarrow x \in \mathbb{R}$, then $x \in [a, b]$.

Note that this is false if we replace $[a, b]$ by (a, b) .

Proof. Obvious. □

Theorem 3.9 (Intermediate Value Theorem). *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, and $f(a) \leq 0, f(b) \geq 0$, then there is some $c \in [a, b]$ such that $f(c) = 0$.*

This is false for \mathbb{Q} .

Remark. Note that it does not matter if it is $f(a) \leq 0, f(b) \geq 0$ or $f(b) \leq 0, f(a) \geq 0$. Also it suffices to show the case where the inequalities are both strict, so we will assume that in the proof.

Proof. For $n \in \mathbb{N}$, we define $Q_n = \{a + k(b - a)/n : 0 \leq k \leq n\}$ and $P_n \subset Q_n$ such that $\forall x \in P_n, f(x) \leq 0$, so P_n is never empty. Define $x_n = \max P_n$, then $a \leq x_n \leq b - (b - a)/n$ and $f(x_n) \leq 0$. We let $y_n = x_n + (b - a)/n = \min Q_n \setminus P_n \subset [a, b]$, then by maximality of x_n we have $f(y_n) > 0$. By Bolzano-Weierstrass Theorem, there is a subsequence x_{n_k} that converges to some $c \in [a, b]$ by the preceding lemma, so by continuity of f , $f(c) = \lim_{n \rightarrow \infty} f(x_n) \leq 0$. Therefore $y_{n_k} \rightarrow c$ as well since $(b - a)/n_k \rightarrow 0$, but then $f(c) = \lim_{n \rightarrow \infty} f(y_n) \geq 0$. So $f(c) = 0$. □

Corollary 3.10. *An odd degree polynomial has a root in \mathbb{R} .*

Proof. Immediate. □

Corollary 3.11. *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, and $f(a) = x, f(b) = y$, then for all z in between x, y there is some $c \in [a, b]$ such that $f(c) = z$.*

Proof. $g : x \mapsto x - z$ is continuous. Then apply the preceding theorem. □

Theorem 3.12. *Continuous function on bounded closed interval is bounded and attains its bounds.*

Again it is not true in general for open intervals.

Proof. Suffice to show this for upper bounds. Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous. Let $A = f([a, b])$. Suppose A is not bounded above, then $\forall n \in \mathbb{N}, \exists x_n \in [a, b], f(x_n) > n$. Then choose a convergent subsequence (x_{n_k}) of (x_n) , then $x_{n_k} \rightarrow x \in [a, b]$, but by continuity $f(x) = \lim_{k \rightarrow \infty} f(x_k) \geq \lim_{k \rightarrow \infty} n_k \rightarrow \infty$, contradiction.

Now that A is bounded above, we choose $M = \sup A$, then $\forall n \in \mathbb{N}$, we can choose $x_n \in [a, b]$ such that $M - 1/n < x_n \leq M$. Choose a convergent subsequence $x_{n_k} \rightarrow x \in [a, b]$, then continuity tells us that $f(x) = \lim_{k \rightarrow \infty} f(x_{n_k})$ which is bounded by $M - 1/n_k, M$, hence $f(x) = M$ by Squeeze rule. □

4 Differentiation

Until further notice, we will just study differentiation of a function $f : I \rightarrow \mathbb{R}$ where I is an interval.

4.1 The Derivative

Definition 4.1. If $a \in I$ and we have

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

exists and equals $c \in \mathbb{R}$, we say f is differentiable at a and the derivative is $f'(a) = c$. Moreover, if f is differentiable at all $a \in I$, we say f is differentiable.

Geometrically $f'(a)$ is the slope of the tangent line of f at a .

Another (formal) way to put this is to observe that the definition is essentially $\forall \epsilon > 0, \exists \delta > 0$,

$$|x - a| < \delta \implies \left| \frac{f(x) - f(a)}{x - a} - c \right| < \epsilon \implies |f(x) - f(a) - c(x - a)| < \epsilon|x - a|$$

Another way we can write is

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}$$

where $h \in I - a$.

Example 4.1. 1. The constant function is differentiable everywhere and has zero derivative everywhere.

2. The identity function is differentiable everywhere and has derivative 1.

3. $f : (0, \infty) \rightarrow \mathbb{R}$ by $x \mapsto 1/x$ is differentiable and has derivative $f'(a) = -a^{-2}$.

4. (non-example) The function $x \mapsto |x|$ is not differentiable at 0.

Lemma 4.1. Suppose f is differentiable at a with derivative $f'(a)$ iff

$$f(a + h) = f(a) + f'(a)h + \alpha(h)h$$

where $\alpha(0) = 0$ and α is continuous at 0.

Note that given this formula, then $\alpha(h)$ is completely determined by f and $f'(a)$.

Proof. Obvious. □

Corollary 4.2. f is differentiable at a implies f is continuous at a .

Proof. Follows directly. □

Proposition 4.3. If $f, g : I \rightarrow \mathbb{R}$ are differentiable at some $a \in I$, so are $f + g, fg$ with derivatives $f'(a) + g'(a), f'(a)g(a) + f(a)g'(a)$ respectively.

Proof. Trivial. □

Theorem 4.4 (Chain Rule). Suppose $f : I_1 \rightarrow I_2$ is differentiable at $a \in I_1$ and $g : I_2 \rightarrow \mathbb{R}$ is differentiable at $f(a)$, then $(g \circ f)'(a) = g'(f(a))f'(a)$.

Proof. Let $b = f(a)$, then by our conditions,

$$f(a+h) = f(a) + f'(a)h + \alpha(h)h, g(b+k) = g(b) + g'(b)k + \beta(k)k$$

And $\alpha(0) = \beta(0) = 0$ and they are both continuous at 0. Take $k = f(a+h) - f(a) = f'(a)h + \alpha(h)h$, so

$$\begin{aligned} (g \circ f)(a) &= g(b+k) \\ &= g(b) + g'(b)k + \beta(k)k \\ &= g(b) + g'(b)(f'(a)h + \alpha(h)h) + \beta(f'(a)h + \alpha(h)h)(f'(a)h + \alpha(h)h) \\ &= (g \circ f)(a) + g'(f(a))f'(a)h + \phi(h)h \end{aligned}$$

Where $\phi(h) = g'(b)\alpha(h) + \beta(f'(a)h + \alpha(h)h)(f'(a) + \alpha(h))$ which is 0 at 0 and is continuous at 0. The chain rule follows. \square

Corollary 4.5. Suppose $f, g : I \rightarrow \mathbb{R}$ are differentiable at a and $g(a) \neq 0$, then f/g is differentiable at a and $(f/g)' = (f'g - fg')/g^2$.

Proof. Immediate from the chain rule. \square

Corollary 4.6. The polynomial $p(x) = a_0 + a_1x + \dots + a_nx^n$ is differentiable at all $x \in \mathbb{R}$, and if p, q are polynomials, p/q is differentiable at a when $q(a) \neq 0$.

Proof. Trivial. \square

4.2 Mean Value Theorem

Suppose we have some $f : I \rightarrow \mathbb{R}$.

Definition 4.2. We say some $c \in I$ is a global maximum of f if $\forall x \in I, f(x) \leq f(c)$ and a global minimum if $\forall x \in I, f(x) \geq f(c)$.

We know that if I is a closed bounded interval, then f has global maxima and minima.

Definition 4.3. Some $c \in I$ is called a local maximum of f if $\exists \epsilon > 0, \forall x \in (c - \epsilon, c + \epsilon) \cap I, f(x) \leq f(c)$. Similarly it is a local minimum if $\exists \epsilon > 0, \forall x \in (c - \epsilon, c + \epsilon) \cap I, f(x) \geq f(c)$.

Definition 4.4. $c \in I$ is called an interior point of I if $\exists \epsilon > 0, (c - \epsilon, c + \epsilon) \subset I$.

Proposition 4.7. Suppose $f : I \rightarrow \mathbb{R}$ and c is an interior point of I where f attains local maximum or local minimum at c and f is differentiable at c , then $f'(c) = 0$.

Proof. Suffice to prove the local maximum case. Since c is an interior point of I , we can choose $\epsilon' > 0, (c - \epsilon', c + \epsilon') \subset I$. As c is a local maximum, then there is some ϵ such that $0 < \epsilon \leq \epsilon'$ and that $\forall x \in (c - \epsilon, c + \epsilon), f(x) \leq f(c)$. We can easily find sequences $(x_n^-) \in (c - \epsilon, c), (x_n^+) \in (c, c + \epsilon)$ both converging to c . Let $g : (c - \epsilon, c + \epsilon) \setminus \{c\} \rightarrow \mathbb{R}$ defined by $g(x) = (f(x) - f(c))/(x - c)$, then

$$\begin{aligned} 0 &\leq \lim_{n \rightarrow \infty} \frac{f(x_n^-) - f(c)}{x_n^- - c} \\ &= \lim_{x \rightarrow c^-} g(x) = \lim_{x \rightarrow c} g(x) = \lim_{x \rightarrow c^+} g(x) \\ &= \lim_{n \rightarrow \infty} \frac{f(x_n^+) - f(c)}{x_n^+ - c} \leq 0 \end{aligned}$$

Hence $f'(c) = \lim_{x \rightarrow c} g(x) = 0$. \square

Theorem 4.8 (Rolle's Theorem). *Suppose $f : [a, b] \rightarrow \mathbb{R}$ is continuous and is differentiable in (a, b) and $f(a) = f(b) = 0$, then there is some $c \in (a, b)$, $f'(c) = 0$.*

Proof. It is obvious if $f \equiv 0$, otherwise WLOG f attains some positive value. Then let c be a global maximum of f , then c is a local maximum and an interior point of f (since f is zero on endpoints but we have assumed that f attains some positive value), then $c \in (a, b)$ and we have $f'(c) = 0$ by the preceding proposition. \square

Theorem 4.9 (Mean Value Theorem). *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) , then there is some $c \in (a, b)$ with $f'(c) = (f(b) - f(a))/(b - a)$.*

Proof. Let

$$p(x) = \frac{1}{b-a}((x-a)f(b) - (x-b)f(a))$$

which although looks dreadful is just a linear polynomial. Note that $p(a) = f(a)$, $p(b) = f(b)$, so it is just a line joining the endpoints. Then $g(x) = f(x) - p(x)$ is continuous on $[a, b]$ and differentiable on (a, b) , so by Rolle's Theorem, there is some $c \in (a, b)$ such that

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b-a} \implies f'(c) = \frac{f(b) - f(a)}{b-a}$$

As desired. \square

Proposition 4.10. *Suppose $f : [a, b] \rightarrow \mathbb{R}$ is continuous and differentiable in (a, b) .*

1. $f' \equiv 0 \iff f$ is constant.
2. $f' \geq 0 \iff f$ is increasing.
3. $f' > 0 \implies f$ is strictly increasing.

Proof. Immediate from Mean Value Theorem. \square

Note that the converse of the third statement is false in general. E.g, $f(x) = x^3$.

Note that all these depends on the least upper bound property since we used the Mean Value Theorem which depends on Rolle's Theorem which depends on maximum value theorem which depends on Bolzano-Weierstrass Theorem which depends on the Monotone Sequence Theorem which depends on the least-upper-bound property.

4.2.1 Inverse Function Theorem

Lemma 4.11. *If $f : [a, b] \rightarrow \mathbb{R}$ satisfies $f'(x) > 0$ for all $x \in [a, b]$, then $f : [a, b] \rightarrow [f(a), f(b)]$ is a bijection.*

Proof. Injectivity is obvious by Mean Value Theorem. For surjectivity we can just exploit Intermediate Value Theorem. \square

Lemma 4.12. *Let I be an interval. Suppose $f : I \rightarrow f(I)$ is a continuous bijection, then $f^{-1} : f(I) \rightarrow I$ is also continuous.*

Proof. Just check definition. \square

Theorem 4.13 (Inverse Function Theorem). *Suppose $I \subset \mathbb{R}$ is an interval and $f'(x) > 0$ for any $x \in I$, then the continuous inverse (exists by the preceding lemmas) $f^{-1} : f(I) \rightarrow I$ is differentiable and $(f^{-1})'(y) = 1/f'(f^{-1}(y))$.*

Example 4.2. 1. $f : x \mapsto x^n$ on $(0, \infty)$ satisfies the conditions, so its inverse $f^{-1} : y \mapsto y^{1/n}$ has $(f^{-1})'(y) = (1/n)y^{1/n-1}$.

2. Consider $\tan : (-\pi/2, \pi/2) \rightarrow \mathbb{R}$, it shall be invertible with differentiable inverse \tan^{-1} , then we can calculate, by the formula, that $(\tan^{-1})'(y) = 1/(y^2 + 1)$

Proof assuming differentiability. Assuming f^{-1} is differentiable, then we know that $f(f^{-1}(y)) = y$, then by the chain rule, we get $f'(f^{-1}(y))(f^{-1})'(y) = 1$, rearrange to give the formula. \square

Actual proof. Fix $b \in f(I)$ and let $a = f^{-1}(b)$. Suppose (y_n) is a sequence in $f(I) \setminus \{b\}$ with $y_n \rightarrow b$, and let $x_n = f^{-1}(y_n)$. We know that f is differentiable hence continuous, so f^{-1} is continuous by the preceding lemma. Therefore $x_n = f^{-1}(y_n) \rightarrow f^{-1}(b) = a$. Since we know that

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a)$$

We have

$$\frac{f(x_n) - f(a)}{x_n - a} \rightarrow f'(a)$$

By hypothesis, $f'(a) \neq 0$, so

$$\frac{f^{-1}(y_n) - f^{-1}(b)}{y_n - b} = \frac{x_n - a}{f(x_n) - f(a)} \rightarrow \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}$$

which is true for every sequence $y_n \rightarrow b$, so the theorem is proved. \square

4.2.2 L'Hôpital's Rule

Proposition 4.14 (L'Hôpital's Rule). *Suppose $f, g : I \rightarrow \mathbb{R}$ are differentiable with $f(a) = g(a) = 0$, and $\exists r > 0, g(x) \neq 0 \neq g'(x)$ for $0 < |x - a| < r$. Then if $\lim_{x \rightarrow a} f'(x)/g'(x)$ exists and equals k , then*

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = k$$

Proof. Suppose $x \in I$ and $x > a$. Consider $h(t) = f(t)g(x) - f(x)g(t), t \in [a, x]$. Since f, g are differentiable and we have $h(a) = h(x) = 0$, so by Rolle's Theorem, there is some $c \in (a, x)$ with $f'(c)g(x) - f(x)g'(c) = h'(c) = 0$. Therefore

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

Same for $x < a$.

Given $\epsilon > 0$, choose $\delta > 0$ such that

$$0 < |c - a| < \delta \implies \left| \frac{f'(c)}{g'(c)} - k \right| < \epsilon$$

So when $0 < |x - a| < \delta$, one can find c such that $0 < |c - a| < |x - a| < \delta$ and $f'(c)/g'(c) = f(x)/g(x)$, which means that

$$\left| \frac{f(x)}{g(x)} - k \right| = \left| \frac{f'(c)}{g'(c)} - k \right| < \epsilon$$

which is as we wanted. \square

4.3 Higher Derivatives and Taylor Series

Definition 4.5. We say (inductively) that $f : I \rightarrow \mathbb{R}$ is k times differentiable with k^{th} derivative $f^{(k)}$ if f is $(k - 1)$ times differentiable with differentiable $(k - 1)^{\text{th}}$ derivative $f^{(k-1)}$ and $f^{(k-1)'} = f^{(k)}$.

Example 4.3. A polynomial is k times differentiable for any $k \in \mathbb{N}$.

Definition 4.6. We say f is C^k if f is k times differentiable and $f^{(k)}$ is continuous.

Example 4.4 (Non-example). Consider

$$f(x) = \begin{cases} x^2 \sin(1/x), & \text{for } x \neq 0 \\ 0, & \text{for } x = 0 \end{cases}$$

So one can verify that f is differentiable but the derivative, namely,

$$f'(x) = \begin{cases} 2x \sin(1/x) - \cos(1/x), & \text{for } x \neq 0 \\ 0, & \text{for } x = 0 \end{cases}$$

is not continuous at 0.

Now, given $f : I \rightarrow \mathbb{R}$, we want a degree k polynomial that best approximates f . The answer depends on the meaning of “best”. We can either calculate the total squared difference (we have not defined it and will not care about it), but on the other hand, the theory of differentiation hints that

$$f(x) = f(a) + f'(a)(x - a) = P_1(x)$$

is the best linear approximation of f around a . Why “around a ”? say for example we want to approximate $\sin x$ over a large interval, then the best approximation intuitively might be a line with slope very close to 0, which is obviously not the tangent at, e.g., 0. And we are only interested in the best approximation close to a .

Now back to our best linear approximation. The reason we say it is the “best” in the sense that $f(x) - P_1(h) = \alpha(h)h$ with $\alpha(0) = 0$ and α is continuous at 0. We can also easily know that such a P_1 is unique so as for such an α to exist by working out their coefficients by substitution and differentiating (with appropriate comments on differentiability, of course). More generally, for a sufficiently smooth f , we can try to approximate $f(a + h) = P_k(a + h) + h^k \alpha(h)$ where $\alpha(0) = 0$ and α is continuous at 0 and P_k is a polynomial. Using the same trick, such an P_k is unique. In particular, we obtain $P_k^{(r)} = f^{(r)}$ for any $r \leq k$.

Lemma 4.15. *If $a, c_0, c_1, \dots, c_k \in \mathbb{R}$, then there is a unique degree k polynomial P with $P^{(i)}(a) = c_i$.*

Proof. Write out $P(x) = a_0 + a_1(x - a) + \dots + a_k(x - a)^k$ and differentiating gives $i!a_i = c_i = P^{(i)}(a)$, so the polynomial is uniquely determined. \square

Definition 4.7. Given $f : I \rightarrow \mathbb{R}$ which is k times differentiable and a is an interior point of I , then the k^{th} Taylor polynomial of f centered at a is

$$P_k(x) = \sum_{i=0}^k \frac{f^{(i)}(a)}{i!} (x - a)^i$$

which is chosen such that $P_k^{(i)}(a) = f^{(i)}(a)$.

Remark. $P_k(x)$ is the unique polynomial with degree at most k having the property $P_k^{(i)}(a) = f^{(i)}(a)$ by the preceding lemma.

Theorem 4.16 (Taylor's Theorem with Remainder). *Suppose $f : I \rightarrow \mathbb{R}$ is $(k + 1)$ times differentiable, and $[a, x] \subset I$, then*

$$f(x) = \left(\sum_{i=0}^k \frac{f^{(i)}(a)}{i!} (x - a)^i \right) + \frac{f^{(k+1)}(c)}{(k+1)!} (x - a)^{k+1}$$

for some $c \in (a, x)$.

For $k = 0$, this reduces to the mean value theorem.

Proof. Fix $x > a$ and define

$$g(t) = f(t) + \sum_{i=1}^k \frac{f^{(i)}(t)}{i!} (x - t)^i + \alpha \frac{(x - t)^{k+1}}{(k+1)!}$$

where α is chosen such that $g(a) = f(x)$ (which exists as a solution to a nondegenerate linear equation). Our goal is to find $c \in (a, x)$ such that $f^{(k+1)}(c) = \alpha$. We have, by definition,

$$f(x) = g(a) = f(a) + \sum_{i=1}^k \frac{f^{(i)}(a)}{i!} (x - a)^i + \alpha \frac{(x - a)^{k+1}}{(k+1)!}$$

Now f is $k + 1$ times differentiable so g is differentiable and $g(a) = f(x) = g(x)$, so by Rolle's Theorem, there is some $c \in (a, x)$ such that $g'(c) = 0$. But we have

$$g'(t) = f'(t) + \sum_{i=1}^k \left(\frac{f^{(i+1)}(t)}{i!} (x - t)^i - \frac{f^{(i)}(t)}{(i-1)!} (x - t)^{i-1} \right) - \alpha \frac{(x - t)^k}{k!}$$

The first part of this expression is telescoping, so we can simplify it to

$$g'(t) = (f^{(k+1)}(t) - \alpha) \frac{(x - t)^k}{k!}$$

Now plugging in $t = c$ shows $f^{(k+1)}(c) = \alpha$, which completes the proof. \square

If f is infinitely differentiable, one can then form the Taylor series

$$\sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!} (x-a)^i$$

as the formal limit of the Taylor polynomials $P_k(x)$, which sometimes does converge and equals f . If we understand the derivatives of f well-enough and that those derivatives do behave nicely, then it is sometimes possible to use Taylor's Theorem to prove that

$$p(x) = \lim_{n \rightarrow \infty} P_n(x) = f(x)$$

Example 4.5. Take $f(x) = \cos x$ (which, albeit hasn't been properly defined, we shall assume standard facts about), then the Taylor series about 0 is

$$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

The series always converges by the ratio test. Must the series converge to $\cos x$ then? The $(2k)^{\text{th}}$ partial sum of the series is exactly $P_{2k}(x)$ of \cos about 0, then there is $c \in (0, x)$ with

$$|f(x) - P_{2k}(x)| = \left| \frac{f^{(2k+1)}(c)}{(2k+1)!} x^{2k+1} \right| \leq \frac{|x|^{2k+1}}{(2k+1)!} \rightarrow 0$$

as $k \rightarrow \infty$, so the series does converge to $\cos x$.

Definition 4.8. We say $f : I \rightarrow \mathbb{R}$ is C^∞ (or smooth) if f is C^k for all $k \in \mathbb{N}$.

So for a smooth $f : I \rightarrow \mathbb{R}$, one can form the Taylor series (formally)

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

But even if the series converges for all $x \in I$, it may not converge to $f(x)$ for $x \neq a$. Philosophically, one cannot fully expect that the behaviour of f near a control its global behaviour. ² Indeed, for $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ smooth and $a < b$, it is possible to construct some smooth $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f|_{<a} = f_1|_{<a}$ and $f|_{>b} = f_2|_{>b}$. This will be proven in example sheet.

4.4 Complex Differentiation

Definition 4.9. If $c \in \mathbb{C}$, $r \in [0, \infty)$, we define the open ball of radius r centered at c by $B_r(c) = \{z \in \mathbb{C} : |z - c| < r\}$.

Definition 4.10. A subset $\Omega \subset \mathbb{C}$ is open if $\forall c \in \Omega, \exists \epsilon > 0, B_\epsilon(c) \subset \Omega$.

Example 4.6. 1. \mathbb{C}, \emptyset are open.

2. Open balls are open.

3. (non-example) $\{z \in \mathbb{C} : |z - a| \leq r\}$ is not open, but its complement is.

²For complex-valued functions, however, the local behaviour of a nice enough function f can be controlled by its local properties.

Definition 4.11. Suppose $\Omega \subset \mathbb{C}$ is open and $c \in \Omega$. For a function $f : \Omega \rightarrow \mathbb{C}$, we say f is complex differentiable at c if

$$\lim_{z \rightarrow c} \frac{f(z) - f(c)}{z - c}$$

exists. In case it does exist, we write it as $f'(c)$.

All the rules in real differentiation still applies, but we don't really have mean value theorem. We can write $f(x + iy) = u(x, y) + iv(x, y)$ where $u, v : \tilde{\Omega} \rightarrow \mathbb{R}$ where $\tilde{\Omega} = \{(x, y) \in \mathbb{R}^2 : x + iy \in \Omega\}$.

Definition 4.12. Say $\tilde{\Omega}$ is open if Ω is. Then the partial derivatives of $F : \tilde{\Omega} \rightarrow \mathbb{R}$ at $(a, b) \in \tilde{\Omega}$ are

$$\begin{aligned} \left. \frac{\partial F}{\partial x} \right|_{(a,b)} &= \lim_{h \rightarrow 0} \frac{F(a+h, b) - F(a, b)}{h} \\ \left. \frac{\partial F}{\partial y} \right|_{(a,b)} &= \lim_{h \rightarrow 0} \frac{F(a, b+h) - F(a, b)}{h} \end{aligned}$$

Provided that they exist.

Alternatively we can define $g_b = F(x, b)$ and so

$$\left. \frac{\partial F}{\partial x} \right|_{(a,b)} = g'_b(a)$$

Similar for the other component.

Proposition 4.17. If $f : \Omega \rightarrow \mathbb{C}$ is complex differentiable at $c = a + ib$ with derivative $f'(c) = \alpha + i\beta$, then the partial derivatives of u, v at (a, b) exist and satisfy

$$\begin{aligned} \left. \frac{\partial u}{\partial x} \right|_{(a,b)} &= \left. \frac{\partial v}{\partial y} \right|_{(a,b)} = \alpha \\ - \left. \frac{\partial u}{\partial y} \right|_{(a,b)} &= \left. \frac{\partial v}{\partial x} \right|_{(a,b)} = \beta \end{aligned}$$

Proof. For any sequence $h_n \rightarrow 0$ in $\mathbb{R} \setminus \{0\}$, then it is also a sequence in $\mathbb{C} \setminus \{0\}$, so since f is differentiable at c ,

$$\begin{aligned} \alpha + i\beta &= f'(c) \\ &= \lim_{h \rightarrow 0} \frac{f(c+h) - f(c)}{h} \\ &= \lim_{n \rightarrow \infty} \frac{u(a+h_n, b) - u(a, b)}{h_n} + i \lim_{n \rightarrow \infty} \frac{v(a+h_n, b) - v(a, b)}{h_n} \end{aligned}$$

So $\partial u / \partial x$ and $\partial v / \partial x$ exist and are α, β respectively. We can get the other half of the equations by replacing h_n by ih_n and casting exactly the same argument. \square

Remark. If $f : \Omega \rightarrow \mathbb{C}$ is complex differentiable, then (assuming C^1 partial derivatives and symmetry of partial derivatives)

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

which is the Laplace Equation.

Proposition 4.18. Suppose $F : B_r((a, b)) \rightarrow \mathbb{R}$ has partial derivatives with

$$\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = 0$$

then F is constant.

Proof. For $(a', b') \in B_r((a, b))$, we consider the functions $f_b(x) = F(x, b)$ and $f_a(x) = F(a', x)$ on the reals. Then they are differentiable and using mean value theorem on f_b wrt a, a' shows $f(a, b) = f(a', b)$, and using it again on f_a wrt b, b' shows $f(a', b) = f(a', b')$, so $f(a, b) = f(a', b')$. \square

Corollary 4.19. If $f : B_r(c) \rightarrow \mathbb{C}$ has derivative zero, then f is constant.

Proof. Immediate. \square

Remark. Complex differentiable functions $f : \mathbb{C} \rightarrow \mathbb{C}$ behave very nicely. Firstly, if f is complex differentiable, so is f' , consequently f is C^∞ . Even better, its Taylor series always converges to itself. Secondly, if f is bounded then f is constant.

5 Special Functions

5.1 Power Series

Recall that if $\sum_n c_n$ converges then $c_n \rightarrow 0$. We also have a strong convergence test where if $|c_n| \leq a_n$ for all n and $\sum_n a_n$ converges, then $\sum_n c_n$ converges. We used it to prove the ratio test: suppose $\lim_{n \rightarrow \infty} |c_{n+1}/c_n|$ exists and equals α , then $\alpha < 1$ implies $\sum_n c_n$ converges and $\alpha > 1$ implies it diverges.

Definition 5.1. A power series centered at $a \in \mathbb{C}$ is a series that can be written in the form

$$\sum_{n=0}^{\infty} c_n (z - a)^n, c_n \in \mathbb{C}$$

Example 5.1. The series

$$\sum_{n=0}^{\infty} (z - 1)^n, \sum_{n=0}^{\infty} \frac{z^n}{n!}, \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

are power series. In general, any Taylor series around a is a power series centered at a .

For a power series and each $z \in \mathbb{C}$, we can get a separate series. Then we want to ask for which $z \in \mathbb{C}$ does the series converge.

Example 5.2. $\sum_n (z - 1)^n$ converges in $|z - 1| < 1$ and diverges otherwise. On the other hand, $\sum_n z^n/n!$, $\sum_n (-1)^n z^{2n}/(2n)!$ converge absolutely. But $\sum_n n!z^n$ converges iff $z = 0$.

Note that suppose $|c_{n+1}/c_n| \rightarrow \alpha$, then

$$\lim_{n \rightarrow \infty} \left| \frac{c_{n+1}(z - a)^{n+1}}{c_n(z - a)^n} \right| = \alpha |z - a|$$

So for $|z - a| < 1/\alpha$, the series converges and $|z - a| > 1/\alpha$ means it diverges. But we do not know about the case $|z - a| = \alpha$. We want to show similar result in general (i.e. α may not exist).

Lemma 5.1 (Key Estimate). *Suppose p_z is the series*

$$\sum_{n=0}^{\infty} c_n(z - a)^n$$

and that p_{z_0} converges, then there is $M > 0$ such that

$$c_n \leq \frac{M}{|z_0 - a|^n}$$

Proof. We have $|c_n||z_0 - a|^n \rightarrow 0$, so there is some N such that $|c_n||z_0 - a|^n < 1$ for $n > N$, so $|c_n| \leq 1/|z_0 - a|^n$ for $n > N$, but we can take care of the rest by taking maximum since there are only finitely many terms left. \square

Theorem 5.2. *Suppose*

$$P(z_0) = \sum_{n=0}^{\infty} c_n(z - a)^n$$

converges, then $P(z)$ converges absolutely whenever $|z - a| < |z_0 - a|$.

Proof. Given z , $|z - a| < |z_0 - a|$, let $\alpha = (z - a)/(z_0 - a)$, so $|\alpha| < 1$, so choose M as in the preceding lemma gives $c_n|z - a|^n \leq M|\alpha|^n$, but $\sum_m M|\alpha|^n$ converges as a geometric series, so $P(z)$ converges absolutely. \square

Corollary 5.3. *With $P(z)$ as above, either $P(z)$ converges for all $z \in \mathbb{C}$ or there exists some $R > 0$ such that $P(z)$ converges absolutely for $|z| < R$ and diverges for $|z| > R$.*

Proof. Immediate. \square

Such an R is obviously unique, so we define

Definition 5.2. Such an R is called the radius of convergence of the power series.

Note that the series may or may not converge on the circle $|z - a| = R$.

Example 5.3. The following series have radius of convergence 1:

$$\sum_{n=0}^{\infty} nz^n, \sum_{n=1}^{\infty} \frac{z^n}{n^2}, \sum_{n=0}^{\infty} \sum_{n=1}^{\infty} \frac{z^n}{n}$$

But they exhibit different behaviours on $|z| = 1$.

Proposition 5.4. *Suppose $P(z)$ has radius of convergence R , then*

1. $c_n(z - a)^n \rightarrow 0$ for $|z| < R$
2. $c_n(z - a)^n$ is unbounded for $|z| > R$.

Proof. Trivial. \square

What we really want to do is to differentiate them. Suppose we have

$$P(z) = \sum_{n=0}^{\infty} c_n (z - a)^n$$

Then let $R > 0$ be its radius of convergence, then $P : B_R(a) \rightarrow \mathbb{C}$ is a function. If P is differentiable, we obviously want it to be

$$Q(z) = \sum_{n=1}^{\infty} n c_n (z - a)^{n-1}$$

which is also a power series. So we will want Q to have radius of convergence R , P to be differentiable and $P' = Q$.

Proposition 5.5. *Q has radius of convergence R .*

Proof. Suppose $P(z_0)$ converges with $z_0 \neq a$, then there is some $M > 0$ with $|c_n| \leq M/|z_0 - a|^n$ for all n . Then for any z with $|z - a| < |z_0 - a|$,

$$|n c_n (z - a)^{n-1}| \leq n C \alpha^{n-1}, C = M/|z_0 - a|, \alpha = |z - a|/|z_0 - a| < 1$$

The rest is trivial. □

Since we can shift stuff around, we assume $a = 0$ for the moment. The preceding proposition then says

$$P(z) = \sum_{n=0}^{\infty} c_n z^n, Q(z) = \sum_{n=1}^{\infty} n c_n z^{n-1}$$

both have radius of convergence R . What we want to show now is that P is complex differentiable on $B_R(0)$ and its derivative is Q . The optimal way to do this is to use the idea of uniform convergence, but we do not have that tool right now, so here goes the clumsy solution. For $z \in B_R(0)$, consider

$$\begin{aligned} P(z+h) - P(z) - hQ(z) &= \sum_{n=0}^{\infty} c_n (z+h)^n - \sum_{n=0}^{\infty} c_n z^n - h \sum_{n=1}^{\infty} n c_n z^{n-1} \\ &= \sum_{n=0}^{\infty} c_n ((z+h)^n - z^n - n h z^{n-1}) \\ &= \sum_{n=0}^{\infty} c_n \sum_{k=2}^n \binom{n}{k} h^k z^{n-k} \\ &= \sum_{n=0}^{\infty} h^2 c_n \sum_{k=2}^n \binom{n}{k} h^{k-2} z^{n-k} \\ &= \sum_{n=0}^{\infty} h^2 c_n r_n(h; z) \end{aligned}$$

But $r_n(h; z)$ is just a polynomial. So intuitively we want to show that its growth is slow, so we claim

Lemma 5.6. For any $z, h \in \mathbb{C}$,

$$|r_n(h; z)| \leq \binom{n}{2} (|h| + |z|)^{n-2}$$

Proof. First suppose that $z, h \in \mathbb{R}_{\geq 0}$, then by using Taylor's Theorem on $f(z) = z^n$, we have

$$f(z+h) = f(z) + hf'(z) + \frac{h^2}{2}f''(y)$$

for some $y \in [z, z+h]$, so

$$r_n(h, z) = \binom{n}{2} y^{n-2} \leq \binom{n}{2} (z+h)^{n-2}$$

For the general case,

$$\begin{aligned} |r_n(h, z)| &= \left| \sum_{i=0}^n \binom{n}{2} h^{i-2} z^{n-i} \right| \\ &\leq \sum_{i=0}^n \binom{n}{2} |h|^{i-2} |z|^{n-i} \\ &= r_n(|h|, |z|) \leq \binom{n}{2} (|z| + |h|)^{n-2} \end{aligned}$$

by above. □

Theorem 5.7. P is complex differentiable on $B_R(a)$ and its derivative is Q .

Proof. Again take WLOG $a = 0$. Let $0 < \rho < R - r_1 = R - |z|$ and let $r_2 = r_1 + \rho < R$, then for $h \in B_\rho(0)$, $z+h \in B_\rho(z) \subset B_{r_2}(0) \subset B_R(0)$. Pick r_3 such that $r_2 < r_3 < R$, then $P(r_3)$ converges, so $\exists M > 0$ such that $|c_n| r_3^n \leq M$ for any n . Then

$$\begin{aligned} |c_n r_n(h; z)| &\leq \frac{M}{r_3^n} \binom{n}{2} (|h| + |z|)^{n-2} \\ &\leq \frac{M}{r_2^2} \binom{n}{2} \alpha^n, \alpha = \frac{r_2}{r_3} < 1 \end{aligned}$$

Now the series $\sum_n \binom{n}{2} \alpha^n$ converges due to the ratio test, consequently

$$\sum_{n=0}^{\infty} c_n r_n(h; z) \rightarrow w$$

For some complex number, which we assume has modulus at most $C > 0$, then

$$P(z+h) - P(z) - hQ(z) = h^2 \sum_{n=0}^{\infty} c_n r_n(h; z) = h^2 \beta(h)$$

Where $\beta(h)$ is the series sum. Let $\alpha(h) = h\beta(h)$, then $|\alpha(h)| \leq C|h|$, so $\alpha(h) \rightarrow 0$ as $h \rightarrow 0$. But also $\alpha(0) = 0$, so α is 0 at 0 and continuous at 0 and we have $P(z+h) = P(z) + hQ(z) + h\alpha(h)$, so the proof is done. □

Corollary 5.8. *A power series*

$$P(z) = \sum_{n=0}^{\infty} c_n (z - a)^n$$

with radius of convergence $R > 0$ is infinitely differentiable on $B_R(a)$ and $n!c_n = f^{(n)}(a)$.

Proof. Immediate. □

Example 5.4. Consider the sum $\sum_n n^2/2^n$. We have $\sum_n z^n = 1/(1-z)$ for $|z| < 1$, so by differentiating and building stuff we have $\sum_n n^2 z^n = z(1+z)/(1-z)^3$, so the series converges to 6 by taking $z = 1/2$.

Remark. In general, it is false to say $f_n(z) \rightarrow f(z)$ implying $f'_n(z) \rightarrow f'(z)$. For example one can take $f_n(x) = \sin(nx)/n$ converges to 0 (even uniformly), but $f'_n(x) = \cos(nx)$ certainly does not converge to 0.

5.2 The Exponentials, etc.

Definition 5.3. For $z \in \mathbb{C}$, the exponential function $\exp(z)$ is defined by

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

which converges for any $z \in \mathbb{C}$ by the ratio test.

We have $\exp' = \exp$ by Theorem 5.7. Also $\exp(0) = 1$.

Proposition 5.9. $\forall z, w \in \mathbb{C}, \exp(z+w) = \exp(z)\exp(w)$.

Proof. Fix $c \in \mathbb{C}$ and consider $f(x) = \exp(x)\exp(c-x)$, then by differentiation $f' = 0$, so f is constant. So $\exp(x)\exp(c-x) = \exp(0)\exp(c-0) = \exp(c)$, so taking $x = z, c = z+w$ gives the result. □

In particular, $\exp(z)\exp(-z) = 1$, so $\exp(z)$ is never zero. And $\exp(-z) = 1/\exp(z)$. By induction, we can easily see that $\exp(nz) = \exp(z)^n$ for all $n \in \mathbb{Z}$. For $x \in \mathbb{R}$, $\exp x$ is real since every term of it is real.

Lemma 5.10. $\exp(\mathbb{R}) = (0, \infty)$ and it is a strictly increasing bijection $\mathbb{R} \rightarrow (0, \infty)$.

Proof. Obvious. □

The lemma guarantees a function $\log : (0, \infty) \rightarrow \mathbb{R}$ as the inverse of $\exp : \mathbb{R} \rightarrow (0, \infty)$. So we have $\log(1) = 0, \log(ab) = \log(a) + \log(b)$ for $a, b > 0$. As for the derivative, by the inverse function theorem, $\log'(x) = 1/x$.

Proposition 5.11.

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

for $|y| < 1$.

Proof. Let $p(y)$ be the series in the right hand side, then p does converge for $|y| < 1$ by the ratio test. Also $p'(y) = 1/(1-y)$ by differenating term-by-term. So $f(y) = p(y) + \log(1-y)$ has zero derivative, so f is constant. Also when $y = 0$, we have $f(0) = 0$, so $f \equiv 0$, hence $-\log(1-y) = p(y)$. \square

Definition 5.4. For $a \in (0, \infty)$ and $x \in \mathbb{R}$, we define the power a^x is defined by $\exp(x \log a)$.

It follows instantly that the power is well-defined, differentiable, and has the series

$$a^x = \sum_{n=0}^{\infty} \frac{(\log a)^n}{n!} x^n$$

Also we have $a^{x+y} = a^x a^y$ and $(a^x)^y = a^{xy}$. Define the constant $e = \exp(1)$, so $e^x = \exp(x)$.

Definition 5.5. For $z \in \mathbb{C}$, we define

$$\sin z = \frac{1}{2i}(e^z - e^{-z}), \cos z = \frac{1}{2}(e^z + e^{-z})$$

Then it follows that they are both differentiable and $\sin' = \cos, \cos' = -\sin$. Also they have power series

$$\sin z = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} z^{2n+1}, \cos z = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} z^{2n}$$

The familiar trigonometric identities all follows from definition. In particular, $\sin^2 z + \cos^2 z = 1$, so \cos, \sin has range in $[0, 1]$ whenever $z \in \mathbb{R}$. We also have Euler's formula: $e^{iz} = \cos z + i \sin z$.

Now we recall the alternating series test: If $a_{n+1} \geq a_n \geq 0$ and $a_n \rightarrow 0$, then $\sum_n (-1)^n a_n$ converges. Furthermore, if we let s_n be the n^{th} partial sum of the alternating series and s be the limit, then

Proposition 5.12. For any $n \in \mathbb{N}$, $s_{2n} \geq s \geq s_{2n+1}$.

Proof. s_{2n} decreases to s and s_{2n+1} increases to s , so the proposition follows immediately. \square

Lemma 5.13. $\sin x > 0$ for $x \in (0, 2]$.

Proof. For $x \in (0, 2]$, $\sin x$ is an alternating series, so $\sin x \geq x - x^3/6 = x(1 - x^2/6) > 0$. \square

Lemma 5.14. $\cos x > 0$ for $x \in [0, 1]$ and there is a unique $\alpha \in (1, 2]$ with $\cos \alpha = 0$.

Proof. For $x \in [0, 1]$, \cos is an alternating series, so $\cos x \geq 1 - x^2/2 \geq 1/2 > 0$ for $x \in [0, 1]$.

Now the series $\cos 2$ is alternating from the third term, so $\cos 2 \leq 1 - 2^2/2 + 2^4/24 = -1/3 < 0$. So there is some $\alpha \in (1, 2)$ such that $\cos \alpha = 0$ by the intermediate value theorem. To see uniqueness, observe that $\cos' = -\sin < 0$ in $[1, 2]$ by the preceding lemma, so \cos must be strictly monotone hence injective in $[1, 2]$, therefore α is unique. \square

Definition 5.6. Let α be as in the preceding lemma, then we define $\pi = 2\alpha$.

Consequently $\sin \alpha = 1$ by $\sin^2 z + \cos^2 z = 1$, and $\cos(x + \alpha) = -\sin x$, $\sin(x + \alpha) = \cos x$. So \sin, \cos are periodic with period 2π , and we know the trends in between.

6 Integration

6.1 The Riemann Integral

Our goal is the following: for a “sufficiently nice” function $f : [a, b] \rightarrow \mathbb{R}$ we want to define its integral

$$\int_a^b f(x) dx$$

We would like to think of this as the area under the graph of $y = f(x)$, which motivates the Riemann integral.

Definition 6.1. We say $f : [a, b] \in \mathbb{R}$ is bounded if $\exists M \in \mathbb{R}, |f(x)| \leq M$ for all $x \in [a, b]$.

Definition 6.2. A dissection of the interval $[a, b]$ is a finite subset $D \subset [a, b]$ such that $a, b \in D$. If D is a dissection, we can write $D = \{a_0, \dots, a_n\}$ where $a = a_0 < a_1 < \dots < a_n = b$.

Note that if D, D' are both dissections, so is $D \cup D'$.

Definition 6.3. If $f : [a, b] \rightarrow \mathbb{R}$ is bounded and a dissection D consists of $a = a_0 < a_1 < \dots < a_n = b$, then we define the upper and lower Riemann sums wrt D by

$$U(f, D) = \sum_{i=1}^n (a_i - a_{i-1}) \sup_{a_{i-1} < x < a_i} f(x), L(f, D) = \sum_{i=1}^n (a_i - a_{i-1}) \inf_{a_{i-1} < x < a_i} f(x)$$

respectively. They exist as f is bounded.

Since f is bounded, U, L are bounded over all dissections D .

Lemma 6.1. If $D \subset D'$ are dissections of $[a, b]$, then $U(f, D) \geq U(f, D')$ and $L(f, D) \leq L(f, D')$.

Proof. Obvious. □

Lemma 6.2. If D_1, D_2 are dissections of $[a, b]$, then $U(f, D_1) \geq L(f, D_2)$.

Proof. Trivial. □

Consider the nonempty sets

$$U = \{U(f, D) : D \text{ dissects } [a, b]\}, L = \{L(f, D) : D \text{ dissects } [a, b]\}$$

It then follows by the preceding lemma that U is bounded below and L is bounded above, so

Definition 6.4. We define

$$U(f) = \inf U = \inf_{D \text{ dissects } [a,b]} U(f, D), L(f) = \sup L = \sup_{D \text{ dissects } [a,b]} L(f, D)$$

as the upper and lower Riemann sums of f over $[a, b]$.

Lemma 6.3. $U(f) \geq L(f)$.

Proof. Follows from definition. □

So we will want to define

Definition 6.5. If $L(f) = U(f)$, then we say f is (Riemann) integrable on $[a, b]$ and write

$$\int_a^b f(x) dx = L(f) = U(f)$$

Example 6.1. 1. Consider $f(x) = c$ for a constant c for all $x \in [a, b]$, then consider $D = \{a, b\}$, then $U(f, D) = c(b-a) = L(f, D)$, hence $U(f) = c(b-a) = L(f)$, therefore f is Riemann integrable and

$$\int_a^b f(x) dx = c(b-a)$$

2. Take $f : [a, b] \rightarrow \mathbb{R}$ by $f(x) = 0$ if $x \in \mathbb{Q}$ and $f(x) = 1$ if $x \in \mathbb{R} \setminus \mathbb{Q}$. Note that if $a_{i-1} < a_i$, then $[a_{i-1}, a_i]$ contains both rational and irrational numbers, so it follows that $U(f, D) = b-a$ and $L(f, D) = 0$ for any dissection D of $[a, b]$, therefore $U(f) = b-a \neq 0 = L(f)$, hence f is not Riemann integrable.

6.2 Properties of the Integral

Proposition 6.4. $f : [a, b] \rightarrow \mathbb{R}$ is integrable on $[a, b]$ and

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

iff for every $\epsilon > 0$, there is a dissection D of $[a, b]$ with $U(f, D) < I + \epsilon$ and $L(f, D) > I - \epsilon$

Proof. Immediate. □

Proposition 6.5. $f : [a, b] \rightarrow \mathbb{R}$ is integrable on $[a, b]$ iff for any $\epsilon > 0$, there is a dissection D of $[a, b]$ such that $U(f, D) - L(f, D) < \epsilon$

Proof. Consider $\inf_D U(f, D) - L(f, D)$. □

Definition 6.6. If $D = \{a = a_0 < a_1 < \dots < a_n = b\}$ is a dissection, then the mesh of D is the difference $m(D) = \max\{a_i - a_{i-1}\}$.

Observe that if $\epsilon > 0$, then one can find dissection with mesh less than ϵ .

Theorem 6.6. If $f : [a, b] \rightarrow \mathbb{R}$ is increasing, then f is integrable.

Proof. Since f is increasing, $f(a_i) \geq f(x)$ whenever $x \leq a_i$. So

$$\sup_{a_{i-1} < x < a_i} f(x) \leq f(a_i), \quad \inf_{a_{i-1} < x < a_i} f(x) \geq f(a_{i-1})$$

Hence

$$U(f, D) = \sum_{i=1}^n (a_i - a_{i-1}) \sup_{a_{i-1} < x < a_i} f(x) \leq \sum_{i=1}^n (a_i - a_{i-1}) f(a_i)$$

and

$$L(f, D) = \sum_{i=1}^n (a_i - a_{i-1}) \inf_{a_{i-1} < x < a_i} f(x) \geq \sum_{i=1}^n (a_i - a_{i-1}) f(a_{i-1})$$

So by subtracting

$$U(f, D) - L(f, D) \leq \sum_{i=1}^n (a_i - a_{i-1}) (f(a_i) - f(a_{i-1})) \leq m(D)(f(b) - f(a)) < \epsilon$$

If $f(b) = f(a)$ then f is constant and we know that it is integrable. Otherwise we take D with $m(D) < \epsilon / (f(b) - f(a))$, then by the preceding lemma, f is integrable on $[a, b]$. \square

Proposition 6.7. *Suppose $f, g : [a, b] \in \mathbb{R}$ are integrable, then*

1. *If $\lambda \in \mathbb{R}$, then $f + \lambda g$ is integrable and*

$$\int_a^b f(x) + \lambda g(x) dx = \int_a^b f(x) dx + \lambda \int_a^b g(x) dx$$

2. *For any $c \in [a, b]$, $f|_{[a,c]}$, $f|_{[c,b]}$ are integrable and*

$$\int_a^c f(x) dx + \int_c^b f(x) dx = \int_a^b f(x) dx$$

3. *If $f \geq g$ on $[a, b]$, then*

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx$$

4. *$|f|$ is integrable and*

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

Proof. Trivial but need patience. \square

6.3 Fundamental Theorem of Calculus

Recall that continuous function on closed interval is bounded.

Theorem 6.8. *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, then f is integrable.*

Lemma 6.9. Suppose $U(f, D) - L(f, D) > c > 0$, then there is $a_i \in D$ such that $\exists x, y \in (a_i, a_{i+1})$ with $f(x) - f(y) \geq c/(b - a)$.

Proof. Basically just pigeonhole principle. \square

Proof of Theorem 6.8. If f is not integrable, then there is some $\epsilon > 0$ such that $U(f, D) - L(f, D) > \epsilon$ for all dissections D of $[a, b]$. Let $0 < \alpha < \epsilon/(b - a)$. By the preceding lemma, if D is a dissection of $[a, b]$, then there is some $a_i \in D$ such that $\exists x, y \in (a_i, a_{i+1}), f(x) - f(y) \geq \alpha$. Consider the dissections D_n with mesh $m(D_n) < 1/n$, then we can find $a_i^{(n)} \in D_n$ and $x_n, y_n \in (a_i^{(n)}, a_{i+1}^{(n)})$, $f(x_n) - f(y_n) \geq \alpha$. Since $m(D_n) < 1/n$, we know that $|x_n - y_n| < 1/n$, so $x_n - y_n \rightarrow 0$. Now by Bolzano-Weierstrass, we can find a converging subsequence $x_{n_k} \rightarrow x$ of x_n . Also $x \in [a, b]$ since $[a, b]$ is closed. Note also that since $x_n - y_n \rightarrow 0$, we have $y_{n_k} \rightarrow x$. But f is continuous, so $0 < \alpha \leq f(x_{n_k}) - f(y_{n_k}) \rightarrow 0$, contradiction. \square

Theorem 6.10 (Fundamental Theorem of Calculus, Version 1). Suppose $f : [a, b] \rightarrow \mathbb{R}$ is continuous, then define

$$G(x) = \int_a^x f(t) dt$$

for $x \in [a, b]$. Then G is differentiable and $G' = f$.

Note that given $x \in [a, b]$, since f is continuous in $[a, b]$, then it is continuous and hence integrable in $[a, x]$, so $G(x)$ is always well-defined.

Proof. We need to show that, as $h \rightarrow 0$,

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

for fixed $x \in [a, b]$. Assume for the moment that $h > 0$, then

$$G(x+h) - G(x) = \int_x^{x+h} f(t) dt$$

Given $\epsilon > 0$, we can find $\delta > 0$ such that $|f(t) - f(x)| < \epsilon$ whenever $|t - x| < \delta$. So whenever $0 < h < \delta$, we have $f(x) - \epsilon < f(t) < f(x) + \epsilon$ for $t \in [x, x+h]$, so

$$f(x)h - \epsilon h = \int_x^{x+h} f(x) - \epsilon dt \leq \int_x^{x+h} f(t) dt \leq \int_x^{x+h} f(x) + \epsilon dt = f(x)h + \epsilon h$$

hence

$$\left| \frac{G(x+h) - G(x)}{h} - f(x) \right| \leq \epsilon$$

for $0 < h < \delta$, hence

$$\lim_{h \rightarrow 0^+} \frac{G(x+h) - G(x)}{h} = f(x)$$

Using exactly the same way, we also have

$$\lim_{h \rightarrow 0^-} \frac{G(x+h) - G(x)}{h} = f(x)$$

Combining them gives the result. \square

Definition 6.7. If $F, f : [a, b] \rightarrow \mathbb{R}$ with $F' = f$, we say F is an antiderivative of f .

So the theorem says every continuous function on closed interval has an antiderivative.

Theorem 6.11 (Fundamental Theorem of Calculus, Version 2). *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous and $F : [a, b] \rightarrow \mathbb{R}$ is differentiable with $F' = f$, then*

$$\int_a^b f(x) dx = F(b) - F(a)$$

Proof. Let

$$G(x) = \int_a^x f(t) dt$$

Then G is differentiable and $G' = f = F'$, so $(G - F)' = 0$, hence $G - F$ is constant hence and equals $G(a) - F(a) = -F(a)$. Hence

$$\int_a^b f(x) dx = G(b) = F(b) + G(b) - F(b) = F(b) - F(a)$$

Which is what we want. □

Note that we have used here many theorems about the real numbers, like the mean value theorem.

Corollary 6.12. *Suppose $f : [a, b] \rightarrow \mathbb{R}$ is continuous and $g : [c, d] \rightarrow \mathbb{R}$ is C^1 , and $g(c) = a, g(d) = b$, then*

$$\int_a^b f(t) dt = \int_c^d f(g(s))g'(s) ds$$

Proof. Chain rule and FTC. □

Corollary 6.13 (Integration by Part). *Suppose $f, g : [a, b] \rightarrow \mathbb{R}$ are both C^1 , then*

$$\int_a^b g'(t)h(t) dt = g(b)h(b) - g(a)h(a) - \int_a^b g(t)h'(t) dt$$

Proof. Product rule and FTC. □

Recall that if $f : [a, b] \rightarrow \mathbb{R}$ is C^k , the k^{th} Taylor polynomial centered at a is

$$P_k(x) = \sum_{i=0}^k \frac{f^{(i)}(a)}{i!} (x - a)^i$$

Theorem 6.14 (Integral Form of Taylor's Theorem). *If $f : [a, b] \rightarrow \mathbb{R}$ is C^k and $x \in [a, b]$, then $R_k(x) = f(x) - P_k(x)$ has*

$$R_k(x) = \int_a^x \frac{(x - t)^{k-1}}{(k - 1)!} f^{(k)}(t) dt$$

Proof. Induction by using integration by part. □

Corollary 6.15. *If $f : [a, b] \rightarrow \mathbb{R}$ is C_k , then there is a constant $M \in \mathbb{R}$ such that $|f(x) - P_{k-1}(x)| \leq M|x - a|^k$.*

Proof. Since f is C^k , $|f^{(k)}|$ is bounded. The rest follows instantly. \square

If $x > 0, s \in \mathbb{R}$, we defined $x^s = \exp(s \log x)$, so $dx^s/dx = sx^{s-1}$. Now consider $f(x) = (1+x)^s$. It has Taylor polynomials (centered at 0)

$$P_k(x) = \sum_{i=0}^k \binom{s}{i} x^i, \binom{s}{i} = \frac{1}{i!} \prod_{j=0}^{i-1} (s-j)$$

We write $P(x)$ to be the formal sum when $k \rightarrow \infty$. By the ratio test, P converges whenever $|x| < 1$.

Theorem 6.16 (Newton's Binomial Theorem). *For $x \in (-1, 1)$, $P(x)$ converges to $(1+x)^s$.*

Lemma 6.17.

$$k \binom{s}{k} = s \binom{s-1}{k-1}, \binom{s}{k-1} + \binom{s}{k} = \binom{s+1}{k}$$

Proof. Trivial. \square

Proof of Theorem 6.16. By termwise differentiation,

$$(1+x)P'(x) = (1+x) \sum_{i=0}^{\infty} i \binom{s}{i} x^{i-1} = sP(x)$$

by the preceding lemma. So consider $h(x) = (1+x)^{-s}P(x)$, so $h'(x) = 0$ by the differential equation we obtained above, hence h is constant in $(-1, 1)$, hence $h \equiv h(0) = 1$. Therefore $P(x) = (1+x)^s$ for each $x \in (-1, 1)$. \square

6.4 Improper Integral

Our definition of integral

$$\int_a^b f(x) dx$$

requires $a, b < \infty$ and f is defined and bounded in $[a, b]$. But we'd also like to think about things like

$$\int_1^{\infty} \frac{dx}{x^2+1}, \int_0^1 \frac{dx}{\sqrt{x}}$$

which are not defined in our previous definition of integral. Suppose $f : [a, b) \rightarrow \mathbb{R}$ is continuous and $b \in \mathbb{R} \cup \{\infty\}$. Then if $x \in [a, b)$, then

$$\int_a^x f(t) dt$$

is well-defined since f is continuous on $[a, x]$.

Definition 6.8. If $f : [a, b) \rightarrow \mathbb{R}$ is continuous and

$$\lim_{x \rightarrow b^-} \int_a^x f(t) dt$$

exists and equals to some $c \in \mathbb{R}$, we define

$$\int_a^b f(t) dt = c$$

and say that this integral converges. If this limit does not exist, we say the integral diverges.

Similarly, if $f : (a, b] \rightarrow \mathbb{R}$ is continuous and

$$\lim_{x \rightarrow a^+} \int_x^b f(t) dt$$

exists and equals to c , then we define

$$\int_a^b f(t) dt = c$$

And say that this integral converges. Otherwise we say this integral diverges.

Note that by FTC this definition (whenever the integral converges) does coincide with our original definition of integrals.

Example 6.2. Consider the integral

$$\int_1^\infty t^{-s} dt$$

Then for $s > 1$, then integral converges and equals $1/(s-1)$, and it diverges for $s \leq 1$. Similarly

$$\int_0^1 t^{-s} dt$$

converges iff $s < 1$.

More generally,

Definition 6.9. If $f : (a, b) \rightarrow \mathbb{R}$ is continuous, then choose $c \in (a, b)$, then if

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

both converge, then we say the integral of $f(t)$ over (a, b) converges and

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

Note that this definition (both the convergence and the value of the integral) does not depend on the choice of c .

Example 6.3. We have

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \int_{-\infty}^0 \frac{dx}{1+x^2} + \int_0^{\infty} \frac{dx}{1+x^2} = \pi$$

But

$$\int_{-\infty}^{\infty} t \, dt$$

does not converge since

$$\int_0^{\infty} f(t) \, dt$$

does not converge.

$$\int_0^{\infty} t^{-s} \, dt = \int_0^1 t^{-s} \, dt + \int_1^{\infty} t^{-s} \, dt$$

never converges for any value of s .

Proposition 6.18 (The Integral Test). *Suppose $f : [1, \infty) \rightarrow [0, \infty)$ is continuous and decreasing, then*

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

converges iff

$$\int_1^{\infty} f(t) \, dt$$

converges.

Example 6.4. $\sum_n n^{-s}$ converges iff $s > 1$ due to our discussion on the integral of t^{-s} over $[1, \infty)$.

Proof. Since $f(x) \geq 0$,

$$F(x) = \int_1^x f(t) \, dt$$

is increasing. Let $A = F([1, \infty))$, then if A is bounded above iff the integral converges. Similarly since $F(n) \geq 0$, the sequence s_n of partial sums is increasing, so the series converges iff $B = \{s_n\}$ is bounded above. Now let $D_n = \{1, 2, \dots, n\}$ be a dissection of $[1, n]$, then

$$s_n - f(1) = L(f, D_n) \leq F(n) \leq U(f, D_n) = s_n - f(n) \leq s_n - c$$

where $c = \inf_x f(x)$ (exists as f is bounded below by 0) since f is decreasing. This implies the claim. \square

Remark. As we see in the proposition, improper integrals behave like series in a certain sense. We also have a comparison test for improper integrals. If $f, g : [a, b) \rightarrow \mathbb{R}$ satisfies $|g| \leq f$ and

$$\int_a^b f(t) \, dt$$

converges, then

$$\int_a^b g(t) \, dt$$

converges.

Using improper integrals, we can define certain interesting functions.

Definition 6.10. The gamma function $\Gamma : (0, \infty) \rightarrow \mathbb{R}$ is defined by

$$\Gamma(s) = \int_0^{\infty} x^{s-1} e^{-x} dx$$

Note that the integral always converges by whatever triviality, so this function is always well-defined. Note also that by integration by part we know that for $s > 1$, we have $\Gamma(s+1) = s\Gamma(s)$, hence $\Gamma(n+1) = n!$ for $n \in \mathbb{Z}_{\geq 1}$. One of the motivation of defining this gamma function is because of the following (here many informal arguments are used): Now consider an n -dimensional sphere $S^{n-1} = \{v \in \mathbb{R}^n : |v| = 1\}$. Let $V(S^{n-1})$ be the volume (or area, etc.) of S^{n-1} , so for example $V(S_1) = 2\pi$. Observe that

$$\int_{\mathbb{R}^n} f(|v|) dv = \int_0^{\infty} V(S^{n-1}) r^{n-1} f(r) dr$$

Now consider $f(r) = e^{-r^2}$ and let c_n be the integral as above, then

$$c_n = \prod_{i=1}^n \int_{\mathbb{R}} e^{-x^2} dx = (c_1)^n$$

But on the other hand c_n equals to

$$\int_0^{\infty} V(S^{n-1}) r^{n-1} e^{-r^2} dr = \frac{V(S^{n-1})}{2} \Gamma\left(\frac{n}{2}\right)$$

So

$$V(S^{n-1}) = \frac{2c_1^n}{\Gamma(n/2)}$$

We know that $c_1 = \sqrt{\pi}$ since we know $V(S_1) = 2\pi$, and this gives the formula.