MATH2241 Introduction to Mathematical Analysis

Chun-Yin Hui

In this course, we study elementary analysis of real numbers and their real-valued functions. The key feature of this topic is the existence of a process of *taking limit*, that is fundamental to constructing objects (e.g., numbers, functions, solutions of polynomial or differential equations) and defining concepts (e.g., continuity, differentiability, Riemann integrability). All these will be made rigorous by the so called "ϵ-δ language", which is nothing more than solving inequalities. The goal of the course is to study the interplay between various concepts in mathematical analysis, e.g., differentiation of power series and the fundamental theorem of calculus.

References

[BS11] Robert G. Bartle, Donald R. Sherbert:
Introduction to Real Analysis (Wiley, 2011, Fourth Edition).

[Ro13] Kenneth A. Ross:
Elementary Analysis: The Theory of Calculus (Springer, 2013, Second Edition).

Contents

1 Real numbers **3**
- 1.1 Field . . . 3
- 1.2 Ordered Field . . . 4
- 1.3 Ordered Field with Least Upper Bound Property . . . 5

2 Sequences of numbers **8**
- 2.1 Convergent sequences and limits . . . 8
- 2.2 Subsequences, monotone sequences, limsup and liminf, and Cauchy criterion . . . 10
- 2.3 Series and absolute convergence . . . 12
- 2.4 Convergence tests for series . . . 14

3 Continuity **15**
- 3.1 Limits of functions . . . 15
- 3.2 General limits about infinity . . . 16
- 3.3 Continuity of functions . . . 17
- 3.4 Uniform continuity of functions . . . 19
- 3.5 Sequences of functions and uniform convergence . . . 20
- 3.6 Power series . . . 21

4 Differentiation **23**
- 4.1 Derivatives of functions . . . 23
- 4.2 Mean Value Theorem . . . 24
- 4.3 Derivatives vs uniform convergence . . . 25
- 4.4 L'Hospital's Rules . . . 26
- 4.5 Taylor's Theorem . . . 27

5 Riemann Integration **28**
- 5.1 Riemann integrals of functions I . . . 28
- 5.2 Riemann integrals of functions II . . . 30
- 5.3 The Fundamental Theorem of Calculus . . . 31
- 5.4 Riemann integrals vs uniform convergence . . . 33

1 Real numbers

1.1 Field

Let $\mathbb{R}$ be the set of real numbers. On the one hand, real numbers can be represented in decimal representations, e.g., $\sqrt{2} = 1.41421....$ On the other hand, if $\mathbb{R}$ is viewed as a straight line, then a real number is a point on it. Below are some distinguished subsets of $\mathbb{R}$:

- $\mathbb{N} := \{1, 2, 3, ...\}$ the set of natural numbers,
- $\mathbb{Z} := \{..., -2, -1, 0, 1, 2, ...\}$ the set of integers,
- $\mathbb{Q} := \{\frac{p}{q} : \ p, q \in \mathbb{Z}, \ q \neq 0\}$ the set of rational numbers.

Definition 1.1. A set F is a *field* if there are two binary operations $+$ and $\cdot$, called *addition* and *multiplication*, such that the following conditions are satisfied for all $a, b, c \in F$:

(**A1**) $a + b = b + a$;

(**A2**) $(a + b) + c = a + (b + c)$;

(**A3**) there exists an element $0 \in F$ (called zero) such that $0 + a = a$ and $a + 0 = a$;

(**A4**) there exists an element (denoted) $-a \in F$ (called an additive inverse) such that $a + (-a) = 0 = (-a) + a$;

(**M1**) $a \cdot b = b \cdot a$;

(**M2**) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;

(**M3**) there exists a non-zero element $1 \in F$ (called unity) such that $1 \cdot a = a = a \cdot 1$;

(**M4**) for $a \neq 0$ there exists an element (denoted) $1/a \in F$ (called a multiplicative inverse) such that $a \cdot 1/a = 1 = 1/a \cdot a$;

(**D**) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ and $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$.

The sets $\mathbb{Q}$ and $\mathbb{R}$ with usual addition and multiplication are fields. Examples of fields include also finite fields $\mathbb{F}_p$ of p (prime) elements. If F in the above definition satisfies all the conditions except (M4), then F is called a ring. For example, $\mathbb{Z}$ is a ring. By using the conditions above, one deduce the following (familiar) facts.

Proposition 1.2. *Let F be a field and $a, b \in F$.*

(i) 0 and 1 are unique.

(ii) $-a$ and $1/a$ (if $a \neq 0$) are uniquely determined by a. In particular, 0 (resp. 1) is the additive inverse (resp. multiplicative inverse) of itself.

(iii) $-(-a) = a$ and $1/(1/a) = a$ (if $a \neq 0$).

(iv) $0 \cdot a = 0$ for all $a \in F$.

(v) $-1 \cdot a = -a$.

(vi) $(-1) \cdot (-1) = 1$.

(vii) $a \cdot b = 0$ implies $a = 0$ or $b = 0$.

(viii) The map $+_a : F \to F$ sending $x \mapsto x + a$ is bijective.

(ix) If $a \neq 0$, the map $\times_a : F \to F$ sending $x \mapsto ax$ is bijective.

Proof. (i) Let 0 and $0'$ be two zeros. Then $a = a + 0$ and $b = 0' + b$ for all $a, b \in F$. By putting $a = 0'$ and $b = 0$, we obtain $0' = 0$. The uniqueness of 1 is similar.

(ii) Suppose $-a$ and b are two additive inverses of a. Then

$$-a = -a + 0 = -a + (a + b) = (-a + a) + b = 0 + b = b.$$

The uniqueness of multiplicative inverse is similar.

(iii) follows from (ii).

(iv) Since $(0 \cdot a) + (0 \cdot a) = (0 + 0) \cdot a = 0 \cdot a$, we are done by adding the additive inverse of $0 \cdot a$ on both sides.

(v) Since $(-1 \cdot a) + a = (-1 \cdot a) + (1 \cdot a) = (-1 + 1) \cdot a = 0 \cdot a = 0$ by (iv), it follows that $-1 \cdot a$ is the additive inverse of a.

(vi) By putting $a = -1$ in (v), we obtain $(-1) \cdot (-1) = -(-1) = 1$ by (iii).

(vii) Suppose $a \cdot b = 0$ and $a \neq 0$. Then

$$b = 1 \cdot b = (1/a \cdot a) \cdot b = 1/a \cdot (a \cdot b) = 1/a \cdot 0 = 0.$$

The rest is exercise. □

For $a, b, c \in F$, we adopt the convention that $a \cdot b + c$ means $(a \cdot b) + c$ to simplify notation (i.e., multiplication has priority over addition) and $a - b$ means $a + (-b)$. If $n \in \mathbb{Z}$ and $a \neq 0$, then a^n has the usual meaning (e.g., $a^3 = aaa$, $a^{-2} = (1/a)^2$, $a^0 = 1$) so that $a^m a^n = a^{m+n}$ and $(a^m)^n = a^{mn}$.

1.2 Ordered Field

Definition 1.3. A field F is an *ordered field* if for all $a, b \in F$ there is a relation $\leq$ between them, that is, either $a \leq b$ or $b \leq a$ holds and such that the following conditions are satisfied for $a, b, c \in F$:

(a) $a \leq a$;

(b) $a \leq b$ and $b \leq a$ imply $a = b$;

(c) $a \leq b$ and $b \leq c$ imply $a \leq c$;

(d) $a \leq b$ implies $a + c \leq b + c$;

(e) $a \leq b$ and $0 \leq c$ imply $ac \leq bc$.

The fields $\mathbb{Q}$ and $\mathbb{R}$ with the usual $\leq$ are ordered fields. Another example is $\mathbb{R}(x)$ the field of $\mathbb{R}$-rational functions and we say $f(x) \leq g(x)$ if $g(t) - f(t) \geq 0$ for all sufficiently large $t \in \mathbb{R}$.

In an ordered field F, we adopt the convention that $a \geq b$ is the same as $b \leq a$ and write $a < b$ if $a \leq b$ and $a \neq b$. It follows from Definition 1.3(d) that $a \leq b$ iff $a - b \leq 0$ iff $0 \leq b - a$ (resp. $a < b$ iff $a - b < 0$ iff $0 < b - a$) and one checks that Definition 1.3(c)–(e) hold if $\leq$ is replaced with $<$. We deduce some (familiar) facts.

Proposition 1.4. *Let F be an ordered field and $a, b \in F$.*

(i) $a > 0$ *iff* $-a < 0$.

(ii) $1 > 0$ *and* $-1 < 0$.

(iii) $a > 0$ *iff* $a^{-1} > 0$.

(iv) If $a > 0$, *then* $a > a/2 > 0$. *In particular, there is no smallest positive number.*

(v) If $a \geq 0$ and $a < b$ for all $b > 0$, then $a = 0$.

(vi) If $a, b > 0$, then $ab > 0$ and $a + b > 0$. Hence, $n > 0$ if $n \in \mathbb{N}$.

(vii) If $a, b < 0$, then $ab > 0$ and $a + b < 0$.

(viii) If $a > 0$ and $b < 0$, then $ab < 0$.

(ix) If $a > b > 0$, then $b^{-1} > a^{-1} > 0$.

Proof. (i) $a > 0$ implies $0 = -a + a > -a + 0 = -a$. The converse is similar.

(ii) Since $1 \neq 0$, either 1 or -1 is positive by (i). Since $1^2 = 1 = (-1)^2$ by Proposition 1.2(iv), we obtain $1 > 0$ by Definition 1.3(e) and also $-1 < 0$ by (i).

(iii) If $a > 0$ and $a^{-1} < 0$, then $-1 = a(-a^{-1}) > 0$ by Definition 1.3(e), which contradicts (ii).

The rest is exercise. □

Remark 1.5. Let F be an ordered field. There is an unique injection $\iota : \mathbb{Q} \to F$ that preserves the field structure: $\iota(a + b) = \iota(a) + \iota(b)$, $\iota(1_{\mathbb{Q}}) = 1_F$, and $\iota(ab) = \iota(a)\iota(b)$ for all $a, b \in \mathbb{Q}$ (ι is called a field embedding). Hence, we may identify $\mathbb{Q}$ as a subset (subfield) of F.

Since $F = \{a \in F : \ a \geq 0\} \cup \{a \in F : \ 0 \geq a\}$ and Definition 1.3(b) holds, elements a in F admit a *trichotomy*: $a > 0$ (positive), $a = 0$, $a < 0$ (negative).

Definition 1.6. Let F be an ordered field. Define the absolute value $|\cdot| : F \to F$ as:

$$|a| := \begin{cases} a & \text{if } a > 0 \\ 0 & \text{if } a = 0 \\ -a & \text{if } a < 0 \end{cases}$$

Proposition 1.7. *Let F be an ordered field and $a, b \in F$.*

(i) $|a| \geq 0$; $|a| = 0$ iff $a = 0$.

(ii) $|ab| = |a||b|$.

(iii) $|a|^2 = a^2$.

(iv) If $b \geq 0$, then $|a| \leq b$ iff $-b \leq a \leq b$.

(v) $|a + b| \leq |a| + |b|$ (triangle inequality).

(vi) $|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n|$ for $a_1, a_2, ..., a_n \in F$.

(vii) $||a| - |b|| \leq |a - b|$.

Proof. Since $-(|a| + |b|) \leq a + b \leq |a| + |b|$, (v) follows from (iv). The rest is exercise. □

1.3 Ordered Field with Least Upper Bound Property

Let F be an ordered field, $S \subset F$ a non-empty subset, and $b \in F$. We say that S is *bounded above* (resp. *bounded below*) by b if $a \leq b$ (resp. $a \geq b$) for all $a \in S$. In this case, we say that b is an *upper bound* (resp. *lower bound*) of S. For simplicity, we say that S is bounded above (resp. below) if it is bounded above (resp. below) by some element in F; we say that S is *bounded* if it is both bounded above and bounded below and we say S is *unbounded* if it is not bounded.

Definition 1.8. An ordered field F is said to be *complete* if it satisfies the *least upper bound property*: whenever S is a non-empty subset of F that is bounded above, then there exists $\sup(S) \in F$ satisfying the two conditions below:

(a) $\sup(S)$ is an upper bound of S and

(b) $\sup(S) \leq b$ for any upper bound b of S.

The number $\sup(S) \in F$ is called the *supremum* of S.

Remark 1.9. The supremum $\sup(S) \in F$ is unique by Definition 1.3(b).

Proposition 1.10. *Let F be a complete ordered field and $S \subset F$ be non-empty.*

(i) If S is bounded above, then for any $\epsilon > 0$ there is some $a \in S$ (depending on ϵ) such that

$$sup(S) - \epsilon < a \leq sup(S).$$

(ii) If S is bounded below, then S has the greatest lower bound, denoted $inf(S) \in F$. The number $inf(S)$ is called the infimum of S.

Proof. (i) follows directly from the definition of supremum (least upper bound). For (ii), note first that $-S := \{-a \in F : \ a \in S\}$ is bounded above and has supremum b; then argue that $-b$ is the infimum of S. $\square$

Definition 1.11. Let F be a complete ordered field and $a \leq b \in F$. Define the following subsets (called *intervals*) of F.

- $[a, b] := \{x \in F : a \leq x \leq b\}$ (closed interval, bounded).
- $(a, b) := \{x \in F : a < x < b\}$ (open interval, bounded).
- $[a, b) := \{x \in F : a \leq x < b\}$ and $(a, b] := \{x \in F : a < x \leq b\}$ (half-open interval).
- $[a, +\infty) := \{x \in F : a \leq x\}$ and $(-\infty, b] := \{x \in F : x \leq b\}$ (closed interval, unbounded)
- $(a, +\infty) := \{x \in F : a < x\}$ and $(-\infty, b) := \{x \in F : x < b\}$ (open interval, unbounded).

The following property is crucial.

Proposition 1.12. *(Archimedean property) Let F be a complete ordered field and $a \in F$. There exists $N \in \mathbb{N}$ such that $n > a$ for all $n \geq N$.*

Proof. It suffices to find $N \in \mathbb{N}$ such that $N > a$. If such $N \in \mathbb{N}$ does not exist, then $\mathbb{N}$ is bounded above by a and $\mathbb{N}$ has a supremum b. Since $n + 1 \in \mathbb{N}$ for all $n \in \mathbb{N}$, it follows that $n + 1 \leq b$ for all $n \in \mathbb{N}$. But this implies that $n \leq b - 1$ for all $n \in \mathbb{N}$, which means $b - 1$ is also an upper bound of $\mathbb{N}$ (absurd). $\square$

Corollary 1.13. *Let F be a complete ordered field and $0 < \epsilon \in F$. There exists $N \in \mathbb{N}$ such that $0 < 1/n < \epsilon$ for all $n \geq N$. In particular, the infimum of $\{1/n : \ n \in \mathbb{N}\}$ is 0.*

Proof. By Proposition 1.12, there is $N \in \mathbb{N}$ such that $n > 1/\epsilon$ for all $n \geq N$. Since $n > 0$ (Proposition 1.4(vi)) and $\epsilon > 0$, we obtain $\epsilon > 1/n > 0$ for $n \geq N$. $\square$

Corollary 1.14. *Let F be a complete ordered field and $a \in F$. There exists $n \in \mathbb{Z}$ such that $n \leq a < n + 1$.*

Proof. Apply Proposition 1.12 to a and $-a$, we find $s, t \in \mathbb{Z}$ such that $s < a < t$. Since there are only finitely many integers in $[s, t]$, the greatest integer $n \leq a$ is well-defined and we obtain $n \leq a < n + 1$. $\square$

Corollary 1.15. *($\mathbb{Q}$ is dense in F) Let F be a complete ordered field and $a < b$ be two numbers in F. Then some $p/q \in \mathbb{Q}$ satisfies $a < p/q < b$.*

Proof. By Proposition 1.12, there is $q \in \mathbb{N}$ such that $q(b - a) > 1$. Let $p \in \mathbb{Z}$ be such that $p - 1 \leq qa < p$ in Corollary 1.14. It follows that $qa < p < qb$ which is equivalent to $p/q \in (a, b)$. $\square$

Proposition 1.16. *The ordered field $\mathbb{Q}$ is not complete.*

Proof. (Sketch). Let $S := \{0 < a \in \mathbb{Q} :\ a^2 < 2\}$ (non-empty). We claim that 2 is an upper bound of S. If not, some $a \in S$ satisfies $a > 2$, which implies $a^2 > 4$ (a contradiction). Suppose $\mathbb{Q}$ is complete on the contrary and let $b > 0$ be the supremum of S. We know that $b^2 \neq 2$ (high school math). Hence, there are two cases: either $b^2 > 2$ or $b^2 < 2$. It suffices to show that both cases are absurd. We tackle only the first case as the other one is similar.

First, we find $n \in \mathbb{N}$ such that $(b - 1/n)^2 > 2$: this is possible since

$$(b - 1/n)^2 - 2 = (b^2 - 2b/n + 1/n^2) - 2 > (b^2 - 2) - 2b/n > 0 \tag{1}$$

if n is sufficiently large by Corollary 1.13. Pick $n \in \mathbb{N}$ large such that (1) holds and $b' = b - 1/n > 0$. We obtain $(b')^2 > 2 > a^2$ for all $a \in S$. Thus, $(b' - a)(b' + a) = (b')^2 - a^2 > 0$ and $b' + a > 0$ imply $b' - a > 0$. Therefore, b' is an upper of S smaller than the supremum b which is absurd. □

Theorem 1.17.

(i) The ordered field $\mathbb{R}$ is complete.

(ii) If F is a complete ordered field, then there is a unique bijective map $\iota : \mathbb{R} \to F$ that preserves the field structure and the order.

Remark 1.18. We omit the proof for this theorem since it pertains to the construction of $\mathbb{R}$ (as the set of equivalence classes of *Cauchy sequences* in $\mathbb{Q}$). By Theorem 1.17(ii), it is no loss to identify a complete ordered field F as $\mathbb{R}$. From now on, we will focus on $\mathbb{R}$ as a field with usual order that satisfies the least upper bound property (Definition 1.8) and use freely the results we have obtained so far.

The least upper bound property allows us to construct many numbers in $\mathbb{R}$ (not necessarily in $\mathbb{Q}$).

Proposition 1.19. *Let $\alpha \in \mathbb{R}$ be positive and $n \in \mathbb{N}$. There is a unique positive $b \in \mathbb{R}$ such that $b^n = \alpha$. In particular, there is $b \in \mathbb{R}\backslash\mathbb{Q}$ (irrational numbers) such that $b^2 = 2$.*

Proof. Let $S = \{0 < a \in \mathbb{Q} : a^n < \alpha\}$. First, prove that S is bounded above and $b := \sup(S)$ satisfies $b^n = \alpha$. Then prove the uniqueness of b. We leave it as an exercise. □

Remark 1.20. Another approach is to prove that the function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ that sends $x \mapsto x^n$ is bijective (one may use the *continuity* of f) and define $x^{1/n}$ to be $f^{-1}(x)$.

Lemma 1.21. *Let S, T be non-empty subsets of $\mathbb{R}$ such that $s \leq t$ for all $s \in S$ and $t \in T$. Then $sup(S) \leq inf(T)$.*

Proof. First, $\sup(S)$ and $\inf(T)$ exist. Second, $s \leq \inf(T)$ (the greatest lower bound of T) for all $s \in S$ because every $s \in S$ is a lower bound of T. But this implies that $\inf(T)$ is an upper bound of S. Finally, we obtain (the least upper bound of S) $\sup(S) \leq \inf(T)$. □

Corollary 1.22. *(Nested intervals property) Let $I_n = [a_n, b_n] \subset \mathbb{R}$ be a sequence of closed intervals such that $I_{n+1} \subset I_n$ for all $n \in \mathbb{N}$. Then the intersection $\cap_{n\in\mathbb{N}} I_n$ is non-empty.*

Proof. Let $S = \{a_n :\ n \in \mathbb{N}\}$ and $T = \{b_n :\ n \in \mathbb{N}\}$. Then S and T satisfy the condition in Lemma 1.21. By Lemma 1.21, one has $a := \sup(S) \leq \inf(T) =: b$ which implies the non-empty $[a, b] \subset I_n$ for all $n \in \mathbb{N}$. We are done. □

A set X is *countable* (e.g., $\mathbb{Q}$) if it is finite or bijective to $\mathbb{N}$. As set is *uncountable* if it is not countable. We prove below that $\mathbb{R}$ is uncountable, i.e., it has many more elements than $\mathbb{Q}$.

Corollary 1.23. *The field $\mathbb{R}$ is uncountable.*

Proof. Suppose on the contrary that $\mathbb{R}$ is countable and we enumerate the elements of $\mathbb{R}$ as a sequence: $\mathbb{R} = \{x_n \in \mathbb{R} :\ n \in \mathbb{N}\}$. Pick a closed interval $I_1 = [a_1, b_1]$ with $a_1 < b_1$ such that $x_1 \notin I_1$. Inductively for each $n \geq 2$, we pick a closed interval $I_n = [a_n, b_n]$ with $a_n < b_n$ such that $I_n \subset I_{n-1}$ and $x_n \notin I_n$. Then $\cap_{n\in\mathbb{N}} I_n$ does not contain any x_n, meaning that it is empty. This is absurd by the nested intervals property. □

2 Sequences of numbers

2.1 Convergent sequences and limits

A *sequence of real numbers* $(a_n) := (a_1, a_2, a_3, ...)$ is a map from $\mathbb{N}$ to $\mathbb{R}$ sending $n \mapsto a_n$. Given a sequence (a_n) and $N \in \mathbb{N}$, we define the N*-tail* of (a_n) to be the sequence $(a_N, a_{N+1}, a_{N+2}, ...)$. The following definition is fundamental.

Definition 2.1. (Convergent sequences) A sequence of real numbers (a_n) *converges to* $L \in \mathbb{R}$ if for every $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that $|a_n - L| < \epsilon$ for all $n \geq N_\epsilon$.

Remark 2.2. (1) The statement $|a_n - L| < \epsilon$ is equivalent to $a_n \in (L - \epsilon, L + \epsilon)$ (this open interval is called the ϵ*-neighborhood* of L). The *distance* between a_n and L is $|a_n - L|$.

(2) In other words, (a_n) converges to $L \in \mathbb{R}$ is equivalent to the statement: for any $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that the N_ϵ-tail of (a_n) is contained in $(L-\epsilon, L+\epsilon)$, the ϵ*-neighborhood* of L.

(3) Although N_ϵ depends on ϵ, we may omit the subscript ϵ of N_ϵ to simplify notation.

(4) (a_n) is said to be *convergent* if it converges to some $L \in \mathbb{R}$; otherwise (a_n) is said to be *divergent*.

(5) The statement "(a_n) converges to $L \in \mathbb{R}$" is denoted by

$$\lim_{n\to\infty} a_n = L \qquad \text{or} \qquad (a_n) \to L.$$

We call $L \in \mathbb{R}$ a *limit* of (a_n).

(6) It is obvious that the convergence of (a_n) and its N-tail are equivalent (with same limit L).

(7) It is **important** to write down the negation: (a_n) does not converge to L if there exist $\epsilon > 0$ and infinitely many $n \in \mathbb{N}$ such that $|a_n - L| \geq \epsilon$.

Proposition 2.3. *(Uniqueness of limit) A convergent sequence (a_n) has only one limit, denoted by* $\lim_{n\to\infty} a_n$.

Proof. Let $L_1 \neq L_2$ be two limits of (a_n) and take $\epsilon = \frac{|L_1 - L_2|}{2} > 0$. By definition of $(a_n) \to L_1$ and $(a_n) \to L_2$, there exists $N_1, N_2 \in \mathbb{N}$ such that $|a_n - L_1| < \epsilon$ for all $n \geq N_1$ and $|a_n - L_2| < \epsilon$ for all $n \geq N_2$. Pick some $n \geq \max\{N_1, N_2\}$. Then this is impossible (draw a picture to see) as:

$$2\epsilon = |L_1 - L_2| \leq |L_1 - a_n| + |a_n - L_2| < \epsilon + \epsilon.$$

□

Proposition 2.4. *A convergence sequence (a_n) is bounded, that is, there exists $C > 0$ such that $|a_n| \leq C$ for all $n \in \mathbb{N}$*

Proof. Obvious by Definition 2.1. □

Example 1.

(a) Let $c \in \mathbb{R}$. The constant sequence $(a_n = c)$ converges to c.

(b) The sequence $(a_n = 1/n)$ converges to 0.

(c) The sequence $(a_n = n)$ is divergent.

(d) The sequence $(a_n = (-1)^n)$ is divergent.

(e) Let $c \in \mathbb{R}$. There exists a sequence (a_n) of rational numbers converging to c.

Proof. (a) follows directly from the definition.
(b) is exactly Corollary 1.13.
(c) follows from Proposition 2.4.
(d): proof by contradiction (exercise).
(e) uses Corollary 1.15 (exercise). □

One can construct new sequences from the old ones: give two sequences (a_n), (b_n), and $c \in \mathbb{R}$ one can form $(a_n + b_n)$, (ca_n), $(a_n b_n)$ etc; if $g : \mathbb{R} \to \mathbb{R}$ is a function then one has $(g(a_n))$. We have the following arithmetic of limits.

Proposition 2.5. *Let (a_n) and (b_n) be convergent sequences with limits L_1 and L_2 respectively.*

(i) (ca_n) is convergent to cL_1 for any $c \in \mathbb{R}$.

(ii) $(a_n + b_n)$ is convergent to $L_1 + L_2$.

(iii) $(a_n b_n)$ is convergent to $L_1 L_2$.

(iv) If $b_n \neq 0$ for all n and $L_2 \neq 0$, then (a_n/b_n) is convergent to L_1/L_2.

(v) If $a_n \leq b_n$ for all n, then $L_1 \leq L_2$.

(vi) If (c_n) is a sequence such that $0 \leq c_n \leq b_n$ and $L_2 = 0$, then (c_n) converges to 0.

(vii) (Sandwich theorem) If (c_n) is a sequence such that $a_n \leq c_n \leq b_n$ for all n and $L_1 = L_2$, then (c_n) converges to L_1.

(viii) If $m \in \mathbb{N}$ and $a_n \geq 0$ for all n, then $(a_n^{1/m})$ converges to $L_1^{1/m}$.

(ix) $(|a_n|)$ converges to $|L_1|$; (a_n) converges to L_1 if and only if $(|a_n - L_1|)$ converges to 0.

Proof. For any $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that both $|a_n - L_1| < \epsilon$ and $|b_n - L_2| < \epsilon$ hold whenever $n \geq N$ (why?). We will use the triangle inequality frequently.

(i). Assume $c \neq 0$ (otherwise it is obvious). For $\epsilon > 0$, take $N := N_{\frac{\epsilon}{|c|}}$. Then for $n \geq N$ we obtain

$$|ca_n - cL_1| = |c||a_n - L_1| < |c|\frac{\epsilon}{|c|} = \epsilon.$$

(ii). For $\epsilon > 0$ we take $N := N_{\frac{\epsilon}{2}}$. Then for $n \geq N$ we obtain

$$|(a_n + b_n) - (L_1 + L_2)| \leq |a_n - L_1| + |b_n - L_2| < \epsilon/2 + \epsilon/2 = \epsilon.$$

(iii). Let $C > \max\{|L_1|, |L_2|\} \geq 0$ such that $|a_n|, |b_n| \leq C$ for all $n \in \mathbb{N}$ (Proposition 2.4). For $\epsilon > 0$, we obtain the following for all $n \geq N_{\frac{\epsilon}{2C}}$:

$$|a_n b_n - L_1 L_2| \leq |a_n b_n - L_1 b_n| + |L_1 b_n - L_1 L_2| \leq |b_n||a_n - L_1| + |L_1||b_n - L_2|$$

$$\leq C|a_n - L_1| + C|b_n - L_2| < C(\frac{\epsilon}{2C}) + C(\frac{\epsilon}{2C}) = \epsilon/2 + \epsilon/2 = \epsilon.$$

(iv). By (iii), it suffices to show $(1/b_n)$ converges to $1/L_2$. Since $L_2 \neq 0$ and $b_n \neq 0$, there exists $c > 0$ such that $|b_n| \geq c$ for all $n \in \mathbb{N}$. For $\epsilon > 0$ we take $N := N_{c|L_2|\epsilon}$. Then for $n \geq N$ we obtain

$$|1/b_n - 1/L_2| = \frac{|b_n - L_2|}{|b_n||L_2|} \leq \frac{|b_n - L_2|}{c|L_2|} < \epsilon.$$

(v). It is enough to show $b_n \geq 0$ for all n imply $L_2 \geq 0$. If $L_2 < 0$, take $0 < \epsilon < |L_2|/2$. There exists some $b_n \in (L_2 - \epsilon, L_2 + \epsilon)$, which is negative (absurd).

(vi). For $\epsilon > 0$, $\exists N_\epsilon \in \mathbb{N}$ such that $|b_n| < \epsilon$ for all $n \geq N_\epsilon$. Hence, for all $n \geq N_\epsilon$ we have

$$|c_n| \leq |b_n| < \epsilon.$$

(vii). We have $0 \leq c_n - a_n \leq b_n - a_n$ and $(b_n - a_n)$ converges to 0. Thus, $(c_n - a_n)$ converges to 0 by (vi) and $(c_n) = (c_n - a_n) + (a_n)$ converges to L_1 by (ii).

(viii). Consider two cases: $L_1 = 0$ and $L_1 > 0$ (exercise).

(ix). The first part uses Proposition 1.7(vii) and the second is definition-rewriting (exercise). $\square$

Remark 2.6. The significance of Proposition 2.5 is that for convergent sequences, taking limit commutes with arithmetic operations, e.g., (ii) says that $\lim_{n\to\infty}(a_n+b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$. Let (a_n) be convergent and $g : \mathbb{R} \to \mathbb{R}$ a *continuous* function (defined later). We will see that $(g(a_n))$ is convergent with limit $g(\lim_{n\to\infty} a_n)$.

It is **important** that students can master this ϵ-N language, which is basically solving inequalities (*backward!*). In (v), one should be careful that $a_n < b_n$ for all n does not imply $\lim a_n < \lim b_n$.

Example 2. Let c be a real number.

(a) If $|c| < 1$, then (c^n) converges to 0.

(b) The sequence $(n^{1/n})$ converges to 1.

(c) If $c > 0$, then $(c^{1/n})$ converges to 1.

(d) The sequence $\frac{c^n}{n!}$ converges to 0 (exercise).

Proof. (a). Write $1/|c| = 1 + \alpha$ where $\alpha > 0$. By Example 1(b) and Proposition 2.5(i), we obtain

$$|c|^n = \frac{1}{(1+\alpha)^n} \leq \frac{1}{n\alpha} \to 0.$$

(b). Write $n^{1/n} = 1 + a_n$ with $a_n \geq 0$. Since

$$n = (1+a_n)^n \geq 1 + \frac{n(n-1)}{2} a_n^2 \geq 1,$$

we have $\frac{2}{n} \geq a_n^2 \geq 0$. Sandwich theorem then implies that $(a_n^2) \to 0$. By Proposition 2.5(viii) with $m = 2$, we obtain $(a_n) \to 0$.

(c). It suffices to consider $c \geq 1$ (since otherwise $1/c \geq 1$). Since we have $1 \leq c^{1/n} \leq n^{1/n}$, we are done by (b) and Sandwich theorem. □

It is convenient to have the following useful definition.

Definition 2.7.

(1) A sequence of real numbers (a_n) converges to ∞ (resp. $-\infty$), denoted by $\lim_{n\to\infty} a_n = \infty$ (resp. $\lim_{n\to\infty} a_n = -\infty$), if for all $C > 0$ (resp. $C < 0$) there exists $N_C \in \mathbb{N}$ such that $a_n \geq C$ (resp. $a_n \leq C$) for all $n \geq N_C$.

(2) Let S be a non-empty subset of $\mathbb{R}$. If S is not bounded above (resp. below), we define the supremum $\sup(S)$ (resp. infimum $\inf(S)$) to be ∞ (resp. $-\infty$).

Note that if (a_n) converges to ∞ or $-\infty$, then it must be divergent.

Proposition 2.8. *Let (a_n) be a non-zero sequence. Then $(|a_n|) \to \infty$ if and only if $(1/a_n) \to 0$.*

Proof. Exercise. □

2.2 Subsequences, monotone sequences, limsup and liminf, and Cauchy criterion

In this section, we give some intrinsic criteria for a sequence to be convergent (without mentioning the limit). Let (a_n) be a sequence of real numbers. A *subsequence* of (a_n) is a sequence $(a_{n_k}) = (a_{n_1}, a_{n_2}, ...)$ such that $n_k \in \mathbb{N}$ and $n_k < n_{k+1}$ for all $k \in \mathbb{N}$.

Proposition 2.9. *If (a_n) converges to $L \in \mathbb{R} \cup \{\pm\infty\}$, then all subsequences of (a_n) converge to L.*

Proof. Obvious by Definitions 2.1 and 2.7(1). □

A sequence of real numbers (a_n) is said to be *increasing* (resp. *decreasing*) if $a_n \leq a_{n+1}$ (resp. $a_n \geq a_{n+1}$) for all $n \in \mathbb{N}$; such sequence is also called *monotone*. An increasing (resp. decreasing) sequence is bounded iff it is bounded above (resp. below).

Proposition 2.10. *(Monotone Convergence Theorem) Let* (a_n) *be an increasing (resp. decreasing) sequence that is bounded. Then* (a_n) *converges to* $\sup\{a_n : n \in \mathbb{N}\}$ *(resp.* $\inf\{a_n : n \in \mathbb{N}\}$*).*

Proof. Suppose (a_n) is increasing (the other case is similar). Since (a_n) is bounded, the supremum $L := \sup\{a_n : n \in \mathbb{N}\}$ exists (by the least upper bound property). For any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $L - \epsilon < a_N \leq L$ by the definition of supremum. As the sequence is increasing, we obtain $L - \epsilon < a_N \leq a_n \leq L$ for all $n \geq N$ which implies that $|a_n - L| < \epsilon$ for $n \geq N$. $\square$

For a bounded sequence, we give an important concept below.

Definition 2.11. Let (a_n) be a bounded sequence. Define two bounded monotone sequences (convergent by Proposition 2.10): $(x_n := \sup\{a_k : k \geq n\})$ (decreasing) and $(y_n := \inf\{a_k : k \geq n\})$ (increasing). Define the *limit-supremum* and *limit-infimum* of (a_n) as follows.

(1) $\limsup_{n\to\infty} a_n := \lim_{n\to\infty} x_n$.

(2) $\liminf_{n\to\infty} a_n := \lim_{n\to\infty} y_n$.

We write $\limsup_{n\to\infty} a_n$ as $\limsup a_n$ (also for liminf) to simplify notation. For Example 1(d), the limit-supremum (resp. limit-infimum) is 1 (resp. -1). We prove some basic properties.

Proposition 2.12. *Let* (a_n)*,* (b_n) *be bounded sequences and* $c \in \mathbb{R}$*.*

(i) $\liminf a_n \leq \limsup a_n$.

(ii) There is a subsequence of (a_n) *that converges to* $\limsup a_n$ *(resp.* $\liminf a_n$*).*

(iii) (a_n) *is convergent if and only if* $\liminf a_n = \limsup a_n$*; in this case the limit of* (a_n) *is the limit-supremum (or limit-infimum).*

(iv) If $c > 0$, $\limsup(ca_n) = c(\limsup a_n)$ *and* $\liminf(ca_n) = c(\liminf a_n)$.

(v) If $c < 0$, $\limsup(ca_n) = c(\liminf a_n)$ *and* $\liminf(ca_n) = c(\limsup a_n)$.

(vi) $\limsup(a_n + b_n) \leq \limsup a_n + \limsup b_n$.

(vii) $\liminf a_n + \liminf b_n \leq \liminf(a_n + b_n)$.

Proof. (i). By Definition 2.11, we have $y_n \leq x_n$ where (x_n) converges to limit-supremum and (y_n) converges to limit-infimum. Then Proposition 2.5(v) implies (i).

(ii). We prove that some subsequence converges to $M := \limsup_{n\to\infty} a_n$ (the other is similar). As $(x_n = \sup\{a_k : k \geq n\}) \to M$, for a given $\epsilon > 0$ there exist infinitely many $n \in \mathbb{N}$ such that $x_n \in (M - \epsilon, M + \epsilon)$, which implies the existence of infinitely many n such that $a_n \in (M - \epsilon, M + \epsilon)$. We define the subsequence (a_{n_k}) inductively as follows: pick n_1 such that $a_{n_1} \in (M - 1/2, M + 1/2)$; for $k \geq 2$ pick $n_k > n_{k-1}$ such that $a_{n_k} \in (M - 1/2^k, M + 1/2^k)$. As

$$|a_{n_k} - M| < 1/2^k$$

for all k, the subsequence converges to M by Proposition 2.5(vi),(ix) and Example 2(a).

For (iii), ($\Rightarrow$) follows from Proposition 2.9 and (ii) above and ($\Leftarrow$) follows from $y_n \leq a_n \leq x_n$ and Sandwich Theorem.

(iv) and (v). Exercise.

(vi) and (vii). We demonstrate (vi) since (vii) is similar. For each $n \in \mathbb{N}$, note that

$$\sup\{a_k + b_k : k \geq n\} \leq \sup\{a_k : k \geq n\} + \sup\{b_k : k \geq n\}.$$

The assertion holds by Proposition 2.5(ii),(v). $\square$

Corollary 2.13. *A sequence* (a_n) *is divergent if and only if one of the following assertions holds:*

(a) (a_n) *is unbounded;*

(b) There are two convergent subsequences (a_{n_k}) *and* (a_{m_k}) *with different limits.*

Proof. ($\Leftarrow$) is obvious by Propositions 2.4 and 2.9. ($\Rightarrow$) follows from Proposition 2.12(iii),(ii). $\square$

Corollary 2.14. *(The Bolzano-Weierstrass Theorem) A bounded sequence (a_n) has a convergent subsequence.*

Proof. Directly by Proposition 2.12(ii). $\square$

We define Cauchy sequences and prove that they are the same as convergent sequences.

Definition 2.15. (Cauchy sequence) A sequence of real numbers (a_n) is a *Cauchy sequence* (or just *Cauchy*) if for any $\epsilon > 0$, there is $N_\epsilon \in \mathbb{N}$ such that $|a_m - a_n| < \epsilon$ for all integers $m, n \geq N_\epsilon$.

Proposition 2.16. *Let (a_n) be a sequence of real numbers.*

(i) If (a_n) is Cauchy, then (a_n) is bounded.

(ii) (Cauchy criterion) The sequence (a_n) is convergent if and only if (a_n) is Cauchy.

Proof. (i). Obvious from the definition.

(ii). Suppose (a_n) is Cauchy and let M and M' be respectively be the limsup and liminf of (a_n). It suffices (by Proposition 2.12(iii)) to show $M = M'$. If not, take $0 < \epsilon < (M - M')/3$. Then Proposition 2.12(ii) implies that there exists infinitely many n (resp. infinitely many m) such that $a_n \in (M - \epsilon, M + \epsilon)$ (resp. $a_m \in (M' - \epsilon, M' + \epsilon)$). But this contradicts the Cauchy definition when taking $\epsilon = (M - M')/3$. Hence, we obtain ($\Leftarrow$). For ($\Rightarrow$), for $\epsilon > 0$ take $N_\epsilon \in \mathbb{N}$ such that $|a_n - L| < \epsilon/2$ for all $n \geq N_\epsilon$. Then for $m, n \geq N_\epsilon$ we have

$$|a_n - a_m| \leq |a_n - L| + |L - a_m| < \epsilon/2 + \epsilon/2 = \epsilon.$$

$\square$

2.3 Series and absolute convergence

Given a sequence (a_n), we can form a new sequence (s_n) given by partial sum

$$s_n := \sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n.$$

Such a sequence is called a *series*. If (s_n) is convergent, the limit is denoted $\sum_{i=1}^{\infty} a_i$ (infinite sum).

Example 3. Determine if the following series is convergent or divergent.

(a) $1 + r + r^2 + \cdots + r^n$, where $r \geq 0$

(b) $\frac{1}{1\times 2} + \frac{1}{2\times 3} + \cdots + \frac{1}{n(n+1)}$

(c) $1 + \frac{1}{2^k} + \frac{1}{3^k} + \cdots + \frac{1}{n^k}$, where $k \in \mathbb{N}$.

(d) $1 - \frac{1}{2} + \frac{1}{3} + \cdots + \frac{(-1)^{n+1}}{n}$.

(e) $\frac{a_1}{10} + \frac{a_2}{10^2} + \cdots + \frac{a_n}{10^n}$, where (a_n) is a sequence such that $a_n \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ for all n.

(f) $1 + \frac{2}{1!} + \frac{2^2}{2!} + \cdots + \frac{2^n}{n!}$.

(g) $\frac{1}{2\ln(2)} + \frac{1}{3\ln(3)} + \cdots + \frac{1}{n\ln(n)}$.

We give first some simple criteria for convergence of series.

Proposition 2.17. *If a series $(s_n = \sum_{i=1}^{n} a_i)$ is convergent, then (a_n) converges to 0.*

Proof. Suppose (s_n) converges to $L \in \mathbb{R}$. Since the sequence $(t_n = s_{n+1})$ is a subsequence of (s_n), it also converges to L (Proposition 2.9). Then $(a_{n+1}) = (t_n - s_n)$ converges to $L - L = 0$ (Proposition 2.5). $\square$

Lemma 2.18. *If (u_n) and (v_n) both converge to $L \in \mathbb{R}$, then the sequence $(s_n) := (v_1, u_1, v_2, u_2, v_3, u_3, ...)$ converges to L.*

Proof. Exercise. $\square$

Proposition 2.19. *(Alternating series) Suppose (a_n) satisfies $a_1 \geq a_2 \geq a_3 \geq \cdots \geq 0$ and (a_n) converges to 0. Then the alternating series*

$$s_n = a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{n+1} a_n$$

is convergent.

Proof. Consider two subsequences $(u_n) := (s_{2n})$ and $(v_n) := (s_{2n-1})$ of (s_n). Then it follows that (u_n) is increasing, (v_n) is decreasing, and $u_n \leq v_n$ for all $n \in \mathbb{N}$. By Monotone Convergence Theorem (Proposition 2.10) and Proposition 2.5(v), $(u_n) \to L_1$, $(v_n) \to L_2$, and $L_1 \leq L_2$. As $a_{2n} = v_n - u_n$ converges to 0, we obtain $L_1 = L_2$. We are done by Lemma 2.18. $\square$

Proposition 2.20. *(Cauchy criterion) A series $(s_n = a_1 + a_2 + \cdots + a_n)$ is convergent if and only if for any $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for all $m > n \geq N_\epsilon$:*

$$|a_{n+1} + a_{n+2} + \cdots + a_m| < \epsilon. \tag{2}$$

Proof. As $|s_m - s_n|$ is the L.S. of (2), this is exactly Proposition 2.16(ii) for series. $\square$

Definition 2.21. (Absolute convergence) A series $(s_n = \sum_{i=1}^n a_i)$ is said to be *absolutely convergent* if the series $(s'_n := \sum_{i=1}^n |a_i|)$ is convergent, that is,

$$\sum_{i=1}^{\infty} |a_i| = \lim_{n\to\infty} \sum_{i=1}^{n} |a_i| < \infty.$$

Proposition 2.22. *If a series $(s_n = \sum_{i=1}^n a_i)$ is absolutely convergent, then (s_n) is convergent.*

Proof. We will use Cauchy criterion (Proposition 2.20). Since $(s'_n = \sum_{i=1}^n |a_i|)$ is convergent (Cauchy), for $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that

$$|a_{n+1}| + |a_{n+2}| + \cdots + |a_m| = |s'_m - s'_n| < \epsilon \tag{3}$$

for all $m > n \geq N_\epsilon$. By (3) and triangle inequality, we obtain

$$|s_m - s_n| = |a_{n+1} + a_{n+2} + \cdots + a_m| \leq |a_{n+1}| + |a_{n+2}| + \cdots + |a_m| < \epsilon$$

for all $m > n \geq N_\epsilon$. This implies (s_n) is Cauchy (hence convergent). $\square$

Remark 2.23. Inequality (3) implies that any finite sum in the $(N_\epsilon + 1)$-tail of $(|a_n|)$ is less than ϵ.

The following fact is useful for absolute convergence series.

Proposition 2.24. *(Rearranging terms) Let $(s_n = a_1 + a_2 + \cdots + a_n)$ be an absolutely convergent series with limit L and $\phi : \mathbb{N} \to \mathbb{N}$ a bijection.*

(i) The series $(t_n = a_{\phi(1)} + a_{\phi(2)} + \cdots + a_{\phi(n)})$ is absolutely convergent.

(ii) The series (t_n) above converges to L.

Proof. By Cauchy criterion for $\epsilon > 0$, there is $N_\epsilon \in \mathbb{N}$ such that for all $m > n \geq N_\epsilon$:

$$|a_{n+1}| + |a_{n+2}| + \cdots + |a_m| < \epsilon/2. \tag{4}$$

(i). For $\epsilon > 0$, take $N' > \max\{\phi^{-1}(1), \phi^{-1}(2), ..., \phi^{-1}(N_\epsilon)\}$. Then the N'-tail of $(a_{\phi(n)})$ is a subset of the $(N_\epsilon + 1)$-tail of (a_n). By Remark 2.23 and (4), we obtain that for $m' > n' \geq N'$:

$$|a_{\phi(n'+1)}| + |a_{\phi(n'+2)}| + \cdots + |a_{\phi(m')}| < \epsilon/2.$$

(ii). It suffices to show that $(s_n - t_n) \to 0$. For $\epsilon > 0$, take $N' \in \mathbb{N}$ in (i). Then for $n \geq N'$ we have

$|s_n - t_n| \leq$ (a finite sum in $(N_\epsilon+1)$-tail of $(|a_n|)$)+(a finite sum in $(N_\epsilon+1)$-tail of $(|a_n|)$) $< \epsilon/2 + \epsilon/2 = \epsilon$

by (4) and Remark 2.23. $\square$

2.4 Convergence tests for series

A series (s_n) is said to be *positive* if $a_n \geq 0$ for all n. We give some convergence tests for series.

Proposition 2.25. *(Comparison test) Let (a_n) and (b_n) be two sequences such that $0 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$. If $(t_n = \sum_{i=1}^n b_i)$ is convergent, then $(s_n = \sum_{i=1}^n a_i)$ is convergent.*

Proof. Let L be the limit of (t_n). Since both series are increasing and $s_n \leq t_n$, it follows that $s_n \leq L$ for all $n \in \mathbb{N}$. By Monotone Convergence Theorem (Proposition 2.10), (s_n) is convergent. $\square$

Proposition 2.26. *(Root test) Let (s_n) be a series and let $\alpha = \limsup |a_n|^{1/n}$.*

(i) If $\alpha < 1$, then s_n is absolutely convergent.

(ii) If $\alpha > 1$, then s_n is divergent.

Proof. (i). Assume (s_n) is a positive series by replacing (a_n) with $(|a_n|)$. Take $\epsilon > 0$ such that $r = \alpha + \epsilon < 1$. There exists $N \in \mathbb{N}$ such that $(a_n)^{1/n} < \alpha + \epsilon$ for all $n \geq N$. Then for $n \geq N$ we obtain

$$s_n \leq (a_1 + a_2 + \cdots a_N) + r^{N+1} + \cdots + r^n$$

which is convergent.

(ii) There exists $1 < \beta < \alpha$ such that $|a_n|^{1/n} > \beta$ (equivalent to $|a_n| > \beta^n$) for infinitely many n. As $\beta^n \to \infty$, s_n cannot be convergent by Proposition 2.17. $\square$

Proposition 2.27. *(Ratio test) Let (s_n) be a series such that $a_n \neq 0$ for all n.*

(i) If $\limsup |\frac{a_{n+1}}{a_n}| < 1$, then s_n is absolutely convergent.

(ii) If $\liminf |\frac{a_{n+1}}{a_n}| > 1$, then s_n is divergent.

Proof. It follows directly from Proposition 2.26 and the lemma below. $\square$

Lemma 2.28. *Let (a_n) be a sequence with $a_n > 0$ for all n. Then*

$$\liminf \frac{a_{n+1}}{a_n} \leq \liminf (a_n)^{1/n} \leq \limsup (a_n)^{1/n} \leq \limsup \frac{a_{n+1}}{a_n}.$$

Proof. The middle inequality is Proposition 2.12(i). We prove the last one as the first is similar. Assume $\limsup \frac{a_{n+1}}{a_n} = L \geq 0$ is finite (otherwise the inequality is obvious). It suffices to show $\limsup (a_n)^{1/n} \leq L'$ for all $L' > L$. By definition, there exists $N \in \mathbb{N}$ (depending on L') such that

$$a_{n+1} \leq L' a_n$$

for all $n \geq N$. Hence for all $n \geq N$ we obtain

$$a_n \leq (L')^{n-N} a_N,$$

which implies

$$(a_n)^{1/n} \leq (L')^{1-N/n} (a_N)^{1/n} = L' c^{1/n}$$

for some $c > 0$. Since $\limsup (a_n)^{1/n} \leq \limsup (L' c^{1/n}) = \lim (L' c^{1/n}) = L'$ (Example 2(c)), we are done. $\square$

Proposition 2.29. *(Integral test) Let $f : [1, \infty) \to \mathbb{R}_{\geq 0}$ be a decreasing function. Then the positive series $(s_n = f(1) + f(2) + \cdots + f(n))$ is convergent if and only if*

$$\lim_{n \to \infty} \int_1^n f(x) dx < \infty.$$

Proof. This is due to $\int_1^{n+1} f(x) dx \leq s_n \leq f(1) + \int_1^n f(x) dx$ and Monotone Convergence Theorem. $\square$

3 Continuity

3.1 Limits of functions

We first use ϵ-δ language to define limits of functions. Let I be an open interval (Definition 1.11, i.e., (c, d) with $c < d$ or (c, ∞) or $(-\infty, d)$), $\alpha \in \mathbb{R}$ that either belongs to I or is an end-point of I (i.e., $\alpha = c$ or d), and $f : I\backslash\{\alpha\} \to \mathbb{R}$ a function. For $\delta > 0$, the subset $(\alpha - \delta, \alpha + \delta)\backslash\{\alpha\}$ is called a *δ-deleted neighborhood* of α. Such formulation is necessary for defining derivative of function later.

Definition 3.1. Let $f : I\backslash\{\alpha\} \to \mathbb{R}$ be a function. We say that f is *convergent* to a number $L \in \mathbb{R}$ as x tends to α if for any $\epsilon > 0$ there exists $\delta := \delta_\epsilon > 0$ such that $|f(x) - L| < \epsilon$ for all $x \in I$ satisfying $0 < |x - \alpha| < \delta$. This statement is also represented as follows in three cases:

(1) When $\alpha \in I$, write "$\lim_{x\to\alpha} f(x) = L$" or "$f(x) \to L$ as $x \to \alpha$", and call L a *limit.*

(2) When α is the left end-point of I, write "$\lim_{x\to\alpha^+} f(x) = L$" or "$f(x) \to L$ as $x \to \alpha^+$", and call L a *right limit.*

(3) When α is the right end-point of I, write "$\lim_{x\to\alpha^-} f(x) = L$" or "$f(x) \to L$ as $x \to \alpha^-$", and call L a *left limit.*

Remark 3.2.

(1) For the symbol $* \in \{\alpha, \alpha^+, \alpha^-\}$, $\lim_{x\to*} f(x) = L \in \mathbb{R}$ iff for any $\epsilon > 0$, there is $\delta_\epsilon > 0$ such that

$$f((\delta_\epsilon\text{-deleted neighborhood of } \alpha) \cap I) \subset (\epsilon\text{-neighborhood of } L).$$

Draw a graph of f to see the meaning. Thus, $\lim_{x\to*} f(x) = L$ implies f is bounded near α.

(2) The existence and value of $\lim_{x\to*} f$ depend only on the values of f on arbitrary small deleted neighborhood of α (even if $f(\alpha)$ is well-defined and not equal to L). If f does not converge at α, we say that f is *divergent* at α.

(3) When $\alpha \in I$, we can define f_1 (resp. f_2) as the restriction of f to $I_1 := \{x \in I : \ x < \alpha\}$ (resp. $I_1 := \{x \in I : \ x > \alpha\}$). Then $\lim_{x\to\alpha} f = L$ iff $\lim_{x\to\alpha^-} f_1 = L = \lim_{x\to\alpha^+} f_2$ (exercise).

The following facts are basic.

Proposition 3.3. *Let I be an open interval, $\alpha \in \mathbb{R}$ either belongs to I or is an end-point of I, and $f, g : I\backslash\{\alpha\} \to \mathbb{R}$ be two functions. For the symbol $* \in \{\alpha, \alpha^+, \alpha^-\}$, if $\lim_{x\to*} f(x) = L_1$ and $\lim_{x\to*} g(x) = L_2$ exist in $\mathbb{R}$, then the following assertions hold.*

(i) $\lim_{x\to*} f$ *is unique.*

(ii) $\lim_{x\to*} cf = cL_1$ *for any* $c \in \mathbb{R}$.

(iii) $\lim_{x\to*}(f + g) = L_1 + L_2$.

(iv) $\lim_{x\to*}(fg) = L_1L_2$.

(v) If $L_2 \neq 0$, *then* $\lim_{x\to*}(f/g) = L_1/L_2$.

(vi) If $f(x) \leq g(x)$ *for all* x, *then* $L_1 \leq L_2$.

(vii) (Sandwich theorem) If h is another function on $I\backslash\{\alpha\}$ such that $f(x) \leq h(x) \leq g(x)$ for all x and $L_1 = L_2$, then $\lim_{x\to*} h = L_1$.

Proof. We demonstrate (i) and (iv) using ϵ-δ language. The other is exercise.

(i). Suppose $L_1 \neq L_1'$ are two limits. For an ϵ such that $|L_1 - L_1'|/2 > \epsilon > 0$, we obtain $\delta, \delta' > 0$ such that $|f(x) - L_1| < \epsilon$ for all $0 < |x - \alpha| < \delta$ and $|f(x) - L_1'| < \epsilon$ for all $0 < |x - \alpha| < \delta'$. Take $\delta_1 = \min\{\delta, \delta'\} > 0$ and $x \in I$ such that $0 < |x - \alpha| < \delta_1$, we get a contradiction (draw a picture to see):

$$2\epsilon < |L_1 - L_1'| \leq |f(x) - L_1| + |f(x) - L_1'| < \epsilon + \epsilon.$$

(iv). Since the limits of f and g at α exist, there exist a constant $C > \max\{|L_1|, |L_2|\} \geq 0$ and $\delta' > 0$ such that $|f(x)|, |g(x)| \leq C$ for all $0 < |x - \alpha| < \delta'$. For $\epsilon > 0$, pick $0 < \delta_{\epsilon/2C}$ with $\delta_{\epsilon/2C} < \delta'$ such that $|f(x) - L_1| < \epsilon/2C$ and $|g(x) - L_2| < \epsilon/2C$ hold for all $0 < |x - \alpha| < \delta_{\epsilon/2C}$. Then for all $x \in I$ such that $0 < |x - \alpha| < \delta_{\epsilon/2C}$, we obtain

$$\begin{aligned}|f(x)g(x) - L_1L_2| &\leq |f(x)g(x) - g(x)L_1| + |L_1g(x) - L_1L_2| = |g(x)||f(x) - L_1| + |L_1||g(x) - L_2| \\ &\leq C|f(x) - L_1| + C|g(x) - L_2| < C(\epsilon/2C) + C(\epsilon/2C) = \epsilon.\end{aligned}$$

□

Remark 3.4. Proposition 3.3 asserts that if limits exist, then taking limit commutes with arithmetic operations. It is **important** that students can master this ϵ-δ language (similar to ϵ-N), which is basically solving inequalities (backward!). Hence, students should prove the rest of the proposition themselves using this language.

We give an alternate (short) proof of Proposition 3.3 by proving an important result that relates limit of function (Definition 3.1) to limit of sequence (Definition 2.1).

Theorem 3.5. *(Sequential criterion) Let $f : I\backslash\{\alpha\} \to \mathbb{R}$ be a function and $L \in \mathbb{R}$. For $* \in \{\alpha, \alpha^+, \alpha^-\}$, $\lim_{x\to *} f(x) = L$ if and only if for any sequence $(a_n) \subset I\backslash\{\alpha\}$ converging to α, $(f(a_n))$ converges to L.*

Proof. ($\Rightarrow$). Given $\epsilon > 0$, there is $\delta > 0$ such that $|f(x) - L| < \epsilon$ for all $0 < |x - \alpha| < \delta$. As $(a_n) \to \alpha$, there exists N_δ such that $|a_n - \alpha| < \delta$ for all $n \geq N_\delta$. Hence, we obtain $|f(a_n) - L| < \epsilon$ for all $n \geq N_\delta$.

($\Leftarrow$). $\lim_{x\to *} f(x) \neq L$ (negation) if there exists $\epsilon > 0$ such that for any $\delta > 0$, there is $0 < |x_\delta - \alpha| < \delta$ such that $|f(x_\delta) - L| \geq \epsilon$. By taking e.g. $\delta = 1/n$, we obtain a sequence $(x_{1/n}) \to \alpha$ such that $(f(x_{1/n}))$ does not converge to L. □

Draw a graph of f to see the meaning of Theorem 3.5.

Corollary 3.6. *Proposition 3.3 follows from Theorem 3.5 and Proposition 2.5.*

Proof. We demonstrate Proposition 3.3(iii) since the others follow the same idea. Let $* \in \{\alpha, \alpha^+, \alpha^-\}$ and $(a_n) \subset I\backslash\{\alpha\}$ be converge to α. Since $f(x) \to L_1$ and $g(x) \to L_2$ as $x \to *$, Theorem 3.5 asserts that $\lim_{n\to\infty} f(a_n) = L_1$ and $\lim_{n\to\infty} g(a_n) = L_2$. Thus, Proposition 2.5(ii) implies that

$$\lim_{n\to\infty} (f + g)(a_n) = \lim_{n\to\infty} [f(a_n) + g(a_n)] = L_1 + L_2$$

for all such sequences (a_n) in $I\backslash\{\alpha\}$. Therefore, Theorem 3.5 implies $\lim_{x\to *}(f + g) = L_1 + L_2$. □

Example 4. If $f : \mathbb{R} \to \mathbb{R}$ is a function below, then $\lim_{x\to\alpha} f(x) = f(\alpha)$.

(a) $f(x) = c$ constant function.

(b) $f(x) = x$.

(c) $f(x) = |x|$ (use $||x| - |y|| \leq |x - y|$).

(d) $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ a real polynomial (use Proposition 3.3).

3.2 General limits about infinity

Let I be an open interval and $\alpha \in \mathbb{R}$ either belongs to I or is an end-point of I.

Definition 3.7.

(1) Let $f : I\backslash\{\alpha\} \to \mathbb{R}$ be a function. For $* \in \{\alpha, \alpha^+, \alpha^-\}$, we say that $\lim_{x\to *} f(x) = \infty$ (resp. $-\infty$) if for all $C > 0$ (resp. $C < 0$) there exists $\delta_C > 0$ such that $f(x) > C$ (resp. $f(x) < C$) for all $x \in I$ such that $0 < |x - \alpha| < \delta_C$.

(2) Let $f : (c, \infty) \to \mathbb{R}$ (resp. $f : (-\infty, d) \to \mathbb{R}$) be a function and $L \in \mathbb{R}$.

(a) We say that $\lim_{x\to\infty} f(x) = L$ (resp. $\lim_{x\to-\infty} f(x) = L$) if for all $\epsilon > 0$ there exists $M > 0$ (resp. $M < 0$) such that $|f(x) - L| < \epsilon$ for all $x > M$ (resp. $x < M$).

(b) We say that $\lim_{x\to\infty} f(x) = \infty$ (resp. $\lim_{x\to-\infty} f(x) = \infty$) if for all $C > 0$ there exists $M > 0$ (resp. $M < 0$) such that $f(x) > C$ for all $x > M$ (resp. $x < M$).

(c) We say that $\lim_{x\to\infty} f(x) = -\infty$ (resp. $\lim_{x\to-\infty} f(x) = -\infty$) if for all $C < 0$ there exists $M > 0$ (resp. $M < 0$) such that $f(x) < C$ for all $x > M$ (resp. $x < M$).

Example 5. Let $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{j=0}^{m} b_j x^j$ be two real polynomials and $g(x)$ is not identically zero. For the symbol $* \in \{a^+, a^- : \ a \in \mathbb{R}\} \cup \{\pm\infty\}$, $\lim_{x\to *} \frac{f(x)}{g(x)}$ exists in $\mathbb{R} \cup \{\pm\infty\}$.

The sequential criterion (Theorem 3.5) can be generalized to the general limit setting.

Theorem 3.8. *(Sequential criterion II) Let $f : J \to \mathbb{R}$ be some function in Definition 3.7(1),(2),(3). Consider the symbols $* \in \{a, a^+, a^-, b^-, \infty, -\infty\}$ and $L \in \mathbb{R} \cup \{\pm\infty\}$. Then $\lim_{x\to *} f(x) = L$ if and only if for all sequence (a_n) in J that converges to $*$, we have $(f(a_n)) \to L$.*

Proof. With Definition 2.7(1) available, the proof is essentially the same as Theorem 3.5. We demonstrate the case for $\lim_{x\to\infty} f = \infty$:

($\Rightarrow$). Given $M > 0$, there is $C > 0$ such that $f(x) > M$ for all $x > C$. As $(a_n) \to \infty$, there exists N_C such that $a_n > C$ for all $n \geq N_C$. Hence, we obtain $f(a_n) > M$ for all $n \geq N_C$.

($\Leftarrow$). $\lim_{x\to\infty} f(x) \neq \infty$ (negation) if there exists $M > 0$ such that for any $C > 0$, there is $x_C > C$ such that $f(x_C) \leq M$. By taking e.g. $C = n$, we obtain a sequence $(x_n) \to \infty$ such that $(f(x_n)) \not\to \infty$. □

Remark 3.9. Intuitively, it seems reasonable to extend some (not all!) arithmetic of $\mathbb{R}$ to $\mathbb{R} \cup \{\pm\infty\}$, (e.g., define $\infty + \infty = \infty$ and $\infty \cdot \infty = \infty$) so that a general version of Proposition 2.5 can be proven. Then this and Theorem 3.8 imply a general version of Proposition 3.3, e.g., if $\lim_{x\to *} f = \infty$ and $\lim_{x\to *} g = -2$ then $\lim_{x\to *} f/g = -\infty$. Note that the following arithmetic for limits has no definite answer:

$$\infty - \infty, \quad \frac{\infty}{\infty}, \quad \infty \cdot 0, \quad \frac{0}{0}, \quad \infty^0, \quad 0^0, \quad 1^\infty.$$

For example, if $f(x) = 1/x^2$ and $g(x) = 1/x^4$ both $\to \infty$ as $x \to 0$ but $\lim_{x\to 0} f/f = 1$ while $\lim_{x\to 0} f/g = 0$. Later, we will see that these indefinite forms could be handled by L'Hospital's rules.

3.3 Continuity of functions

In this section, $I, J \subset \mathbb{R}$ denote some non-empty intervals in Definition 1.11.

Definition 3.10. Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ a function. We say that f is *continuous* at $\alpha \in I$ if for any $\epsilon > 0$ there is $\delta_\epsilon > 0$ such that $|f(x) - f(\alpha)| < \epsilon$ for all $x \in I$ satisfying $|x - \alpha| < \delta_\epsilon$. We call f a *continuous function* if it is continuous at every $\alpha \in I$. Denote by $C(I)$ the set of continuous functions on I.

The continuity of f at α depends only on the function f on a neighborhood of α. Definitions 3.10 and 3.1 look very similar (the only differences are the addition of the point $x = \alpha$ and requiring $L = f(\alpha)$)! The relationship is given below.

Proposition 3.11. *Let $I \subset \mathbb{R}$ be an interval, $\alpha \in I$, and $f : I \to \mathbb{R}$ a function.*

(i) If I is singleton (i.e., $I = [\alpha, \alpha]$), then f is continuous at α.

(ii) Suppose I is not singleton.

(a) If α is the left end-point of I, then f is continuous at α iff $\lim_{x\to\alpha^+} f(x) = f(\alpha)$.

(b) If α is the right end-point of I, then f is continuous at α iff $\lim_{x\to\alpha^-} f(x) = f(\alpha)$.

(c) If α is not an end-point of I, then f is continuous at α iff $\lim_{x\to\alpha} f(x) = f(\alpha)$.

Proof. (i) and (ii) are obvious from Definitions 3.10 and 3.1. □

We have a sequential criterion for continuity, which is useful for computing limit of sequence.

Theorem 3.12. *(Sequential criterion III) Let $I \subset \mathbb{R}$ be an interval, $\alpha \in I$, and $f : I \to \mathbb{R}$ a function. Then f is continuous at α iff for any sequence (a_n) in I converging to α:*

$$\lim_{n\to\infty} f(a_n) = f(\alpha),$$

i.e., $\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n)$ *(exchange of limit and f).*

Proof. It follows from (i),(ii) and Theorem 3.5 (sequential criterion II). □

We prove some basic properties for continuity.

Proposition 3.13. *Let $f, g : I \to \mathbb{R}$ be continuous at $\alpha \in I$.*

(i) cf is continuous at α for all $c \in \mathbb{R}$.

(ii) $f + g$ is continuous at α.

(iii) fg is continuous at α.

(iv) If $g(x) \neq 0$ for all $x \in I$, then f/g is continuous at α.

(v) (Sandwich theorem) If $h : I \to \mathbb{R}$ satisfies $f(x) \leq h(x) \leq g(x)$ for all x and $f(\alpha) = g(\alpha)$, then h is continuous at α.

(vi) If $f(I) \subset J$ and $h : J \to \mathbb{R}$ is continuous at $f(\alpha) \in J$, then the composition $h \circ f : I \to \mathbb{R}$ is continuous at α.

Proof. (i)–(v) follows directly from Proposition 3.11(i),(ii) and Proposition 3.3. For (vi), if (a_n) be a sequence in I converging to α then $f(a_n) \to f(\alpha)$ by f continuous at α and Theorem 3.12. By h continuous at $f(\alpha)$ and Theorem 3.12, we obtain $h(f(a_n)) \to h(f(\alpha))$. Hence, $h \circ f$ is continuous at α by Theorem 3.12 again. □

Remark 3.14. The proposition implies that the sum, difference, product, quotient (if possible), and composition of continuous functions are continuous.

Example 6. (a),(b) are continuous on $\mathbb{R}$; (c) is continuous only at $x = 0$; (d) is nowhere continuous.

(a) Real polynomials, $|x|$, $\frac{1}{x^2+1}$.

(b) $\max\{f, g\} = \frac{f+g}{2} + \frac{|f-g|}{2}$ and $\min\{f, g\} = \frac{f+g}{2} - \frac{|f-g|}{2}$, where f and g are continuous on $\mathbb{R}$.

(c) $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = x$ if $x \in \mathbb{Q}$ and $f(x) = -x$ otherwise.

(d) $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = 1$ if $x \in \mathbb{Q}$ and $f(x) = 0$ otherwise.

We prove some fundamental results for continuous functions on intervals.

Theorem 3.15. *Let $f : I = [a, b] \to \mathbb{R}$ be a continuous function on a closed bounded interval. Then $f(I)$ is also a closed bounded interval.*

Proof. First, we show that $f(I)$ is bounded. If not, WLOG we find a sequence (x_n) in $[a, b]$ such that $f(x_n) \to \infty$. By Bolzano-Weierstrass Theorem (Corollary 2.14), a subsequence (x_{n_k}) of (x_n) converges to some $\alpha \in [a, b]$ (as $a \leq x_{n_k} \leq b$ for all n). Theorem 3.12 implies that $f(x_{n_k}) \to f(\alpha) \in \mathbb{R}$, which contradicts that $f(x_n) \to \infty$.

Let M (resp. M') be the supremum (resp. infimum) of $f(I)$. We prove that $M, M' \in f(I)$. We treat the first case. By definition, there exists a sequence (x_n) in $[a, b]$ such that $f(x_n) \to M$ as $n \to \infty$. By using Bolzano Weierstrass Theorem as above, we may assume $(x_n) \to \alpha \in [a, b]$. Then Theorem 3.12 implies $f(x_{n_k}) \to f(\alpha)$ which is equal to M.

It remains to show that if $M' < c < M$, then $c \in f(I)$. By above, we find $x_1, y_1 \in I$ such that $f(x_1) = M$ and $f(y_1) = M'$. We construct two sequences $x_n, y_n \in I$ inductively for $n \geq 2$ as follows. Given x_n and y_n, there are two cases to consider:

(1) if $f(\frac{x_n+y_n}{2}) \geq c$, define $x_{n+1} = \frac{x_n+y_n}{2}$ and $y_{n+1} = y_n$;

(2) if $f(\frac{x_n+y_n}{2}) < c$, define $x_{n+1} = x_n$ and $y_{n+1} = \frac{x_n+y_n}{2}$.

Since two sequences (x_n) and (y_n) in I are monotone and $|x_n - y_n| \to 0$ as $n \to \infty$ (as $|x_{n+1} - y_{n+1}| = \frac{|x_n-y_n|}{2}$ for all n), they both converge to the same $\alpha \in I$. Since $f(x_n) \geq c < f(y_n)$ for all n, we obtain $f(\alpha) \leq c \leq f(\alpha)$ after taking limit. This implies that $f(\alpha) = c$. □

Corollary 3.16. *(Maximum and mimimum achieved) Let $f : I = [a, b] \to \mathbb{R}$ be a continuous function on a closed bounded interval. Then there exists $\alpha, \beta \in I$ such that $f(\beta) \leq f(x) \leq f(\alpha)$ for all $x \in I$.*

Proof. Obvious. □

Corollary 3.17. *(Intermediate Value Theorem) Let $f : I = [a, b] \to \mathbb{R}$ be a continuous function on a closed bounded interval. Then any number in between $f(a)$ and $f(b)$ belongs to $f(I)$.*

Proof. Obvious. □

Remark 3.18. If $p(x)$ is a real polynomial of odd degree, then $p(x)$ has a (real) root by Intermediate Value Theorem.

Corollary 3.19. *Let $f : I \to \mathbb{R}$ be a continuous function, where I is any interval. Then $f(I)$ is an interval.*

Proof. Note that the open interval $(\inf(f(I)), \sup(f(I))) \subset f(I)$ (exercise). □

Corollary 3.20. *Let $f : I \to \mathbb{R}$ be an injective continuous function.*

(i) The function f is either strictly increasing or decreasing.

(ii) The inverse $f^{-1} : J \to I \subset \mathbb{R}$ is also continuous, where $J := f(I)$ is the image interval.

Proof. (i). If the statement is not true, WLOG we may assume there exist $a < b < c$ in I such that $f(a) < f(b)$ and $f(b) > f(c)$. Pick $\epsilon > 0$ such that $f(b) - \epsilon > \max\{f(a), f(c)\}$. Then Intermediate Value Theorem implies that there exist $a' \in (a, b)$ and $c' \in (b, c)$ such that $f(a') = f(b) - \epsilon = f(c')$, contradicting injectivity.

(ii). By Corollary 3.19, $J = f(I)$ is an interval. WLOG, assume f is strictly increasing and $\beta := f(\alpha)$. We will prove that f^{-1} is continuous at $\beta \in J$. We treat the case α is not an end-point of I (the other is similar). Given small $\epsilon > 0$, define

$$\delta := \min\{f(\alpha) - f(\alpha - \epsilon), f(\alpha + \epsilon) - f(\alpha)\} > 0.$$

Then $|f^{-1}(y) - f^{-1}(\beta)| < \epsilon$ whenever $|y - \beta| < \delta$ (draw a picture to see). □

Question 1. Suppose $f : I \to J \subset \mathbb{R}$ is bijective onto J and continuous at $\alpha \in I$. Is f^{-1} continuous at $f(\alpha) \in J$?

Example 7. The following functions are continuous.

(a) $f(x) = x^{1/n}$ defining on $\mathbb{R}$, where $n \in \mathbb{N}$ is odd.

(b) $f(x) = x^{1/n}$ defining on $\mathbb{R}_{\geq 0}$, where $n \in \mathbb{N}$ is even.

3.4 Uniform continuity of functions

Definition 3.21. Let I be an interval and $f : I \to \mathbb{R}$ be a function.

(1) f is said to be *uniform continuous* if for any $\epsilon > 0$, there exists $\delta_\epsilon > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $x, y \in I$ satisfy $|x - y| < \delta_\epsilon$.

(2) f is said to be *Lipschitz continuous* if there exists $C > 0$ such that $|f(x) - f(y)| \leq C|x - y|$ for all $x, y \in I$.

Remark 3.22. It follows that Lipschitz continuous functions are uniform continuous and uniform continuous functions are continuous.

Proposition 3.23. *Let $f : I = [a, b] \to \mathbb{R}$ be a continuous function on a closed bounded interval. Then f is uniformly continuous.*

Proof. If not true, then there exists $\epsilon > 0$ such that for any $\delta > 0$, there exist $x_\delta, y_\delta \in I$ such that $|x_\delta - y_\delta| < \delta$ and $|f(x_\delta) - f(y_\delta)| \geq \epsilon$. Take $\delta = 1/n$ for $n \in \mathbb{N}$. We obtain two sequences $(x_{1/n})$ and $(y_{1/n})$ such that

$$\lim_{n\to\infty} |x_{1/n} - y_{1/n}| = 0. \tag{5}$$

Since I is bounded, we may assume both sequences are convergent by replacing them with subsequences and Bolzano Weierstrass Theorem. Then (5) implies that they converge to the same limit $\alpha \in I$. As f is continuous, we obtain a contradiction

$$0 = |f(\alpha) - f(\alpha)| = \lim_{n\to\infty} |f(x_{1/n}) - f(y_{1/n})| \geq \epsilon.$$

□

Example 8. (a) $f(x) = 1/x$ defining on $[\epsilon, \infty)$ is uniform continuous, where $\epsilon > 0$.

(b) $f(x) = 1/x$ defining on $(0, \infty)$ is not uniform continuous.

3.5 Sequences of functions and uniform convergence

Let I be an interval and $(f_n : I \to \mathbb{R})$ a sequence of functions for all $n \in \mathbb{N}$. We would like to study the convergence of (f_n).

Definition 3.24. Let $(f_n : I \to \mathbb{R})$ be a sequence of functions and $f : I \to \mathbb{R}$ another function.

(1) We say that (f_n) converges to f *pointwise* if $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in I$. Denote this by $(f_n) \to f$.

(2) We say that (f_n) converges to f *uniformly* if for any $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq N_\epsilon$ and for all $x \in I$. Denote this by $(f_n) \rightrightarrows f$.

(3) We say that (f_n) is *uniformly Cauchy* if for any $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that $|f_m(x) - f_n(x)| < \epsilon$ for all $m > n \geq N_\epsilon$ and for all $x \in I$.

Example 9. Let $(f_n : [0, 1] \to \mathbb{R})$ be a sequence of (continuous) functions with $f_n(x) = x^n$. Then it converges pointwise to (non-continuous) f with $f(x) = 0$ for $0 \leq x < 1$ and $f(1) = 1$. By Theorem 3.26 below, the convergence is not uniform.

Example 10. Let $f_n(x) = \frac{x}{1+nx^2}$ on $\mathbb{R}$. Then $(f_n(x)) \to 0$ for all $x \in \mathbb{R}$. To show that $(f_n) \rightrightarrows 0$, for $\epsilon > 0$ we need to find $N_\epsilon \in \mathbb{N}$ (independent of $x \in \mathbb{R}$) such that

$$\frac{|x|}{1 + nx^2} < \epsilon \tag{6}$$

for all $n \geq N_\epsilon$ and $x \in \mathbb{R}$. Since (6) holds for $x = 0$, we may assume $x \neq 0$ and solve $\frac{1}{|x|^2} - \frac{1}{\epsilon|x|} + n > 0$ for n. By considering (discriminant) $\Delta = \frac{1}{\epsilon^2} - 4n < 0$, we obtain $n > \frac{1}{4\epsilon^2}$. Hence, we can take N_ϵ to be any integer $> \frac{1}{4\epsilon^2}$.

It follows from the definition that uniform convergence implies pointwise convergence. We also have the equivalence between uniform convergence and uniform Cauchy.

Proposition 3.25. *(Cauchy criterion) The sequence (f_n) converges uniformly to f if and only if (f_n) is uniformly Cauchy.*

Proof. ($\Rightarrow$). For any $\epsilon > 0$, there exists $N_\epsilon > 0$ such that $|f_n(x) - f(x)| < \epsilon/2$ for all $x \in I$ and $n \geq N_\epsilon$. Then for $m > n \geq N_\epsilon$ and $x \in I$, we obtain

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f(x)| + |f_n(x) - f(x)| \leq \epsilon/2 + \epsilon/2 = \epsilon.$$

($\Leftarrow$). Since uniform Cauchy implies pointwise Cauchy, $(f_n) \to f$ pointwise by Proposition 2.16(ii) (Cauchy criterion for sequence). For any $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that

$$|f_m(x) - f_n(x)| < \epsilon \tag{7}$$

for all $m > n \geq N_\epsilon$ and for all $x \in I$. Take limit $m \to \infty$, we obtain

$$|f(x) - f_n(x)| < \epsilon$$

for all $n \geq N_\epsilon$ and for all $x \in I$. We are done. $\square$

Theorem 3.26. *Let (f_n) be a sequence of continuous functions. If (f_n) converges uniformly to f, then f is continuous.*

Proof. Let (a_n) be a sequence in I converging to $\alpha \in I$. It suffices to show that $(f(a_n)) \to f(\alpha)$ by sequential criterion (Theorem 3.12). For any $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| < \epsilon/3 \tag{8}$$

for all $n \geq N_\epsilon$ and $x \in I$. Since f_{N_ϵ} is continuous at α, sequential criterion implies that there is $N'_\epsilon \in \mathbb{N}$ such that

$$|f_{N_\epsilon}(a_m) - f_{N_\epsilon}(\alpha)| < \epsilon/3 \tag{9}$$

for all $m \geq N'_\epsilon$. Hence, (8) and (9) imply that

$$|f(a_m) - f(\alpha)| \leq |f(a_m) - f_{N_\epsilon}(a_m)| + |f_{N_\epsilon}(a_m) - f_{N_\epsilon}(\alpha)| + |f_{N_\epsilon}(\alpha) - f(\alpha)| < 3(\epsilon/3) = \epsilon$$

for all $m \geq N'_\epsilon$. $\square$

Question 2. Let (f_n) and (g_n) be two sequences in $C(I)$ that are uniformly convergent to f and g respectively. Are (cf_n), $(f_n + g_n)$, $(f_n g_n)$, (f_n/g_n) (if $g(x) \neq 0$ for all x) uniformly convergent? What if I is a closed bounded interval (hint: same proof as Proposition 2.5 and note that $f, g \in C(I)$)?

An important result about continuous functions on $[a, b]$ is below (analogous to $\mathbb{Q}$ is dense in $\mathbb{R}$!).

Theorem 3.27. *(Weierstrass's Approximation Theorem, see [BS11, §5.4] or [Ro13, §27]) Let $f : [a, b] \to \mathbb{R}$ be a continuous function defined on a closed bounded interval. Then there exists a sequence of polynomials $(p_n(x))$ that converges uniformly to f on $[a, b]$.*

3.6 Power series

This section connects §3.5 uniform convergence of functions with §2.4 convergence tests.

Given a sequence $(a_0, a_2, a_2, ...)$ and a point $x_0 \in \mathbb{R}$, we can form the power series

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + \cdots, \tag{10}$$

which can be viewed as the limit of the series $s_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n$ of functions if it converges. Denote the set $\mathbb{R}_{\geq 0} \cup \{\infty\}$ by $[0, \infty]$.

Theorem 3.28. *Let $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ be a power series. Let $\beta = \limsup |a_n|^{1/n} \in [0, \infty]$ and $R = 1/\beta \in [0, \infty]$. The the following assertions hold.*

(i) The power series $f(x)$ converges absolutely (pointwise) for all x such that $|x - x_0| < R$.

(ii) The power series $f(x)$ converges uniformly on $I = [x_0 - R', x_0 + R']$ if $0 \leq R' < R$. In particular, $f(x)$ is a continuous function on I.

(iii) The power series $f(x)$ diverges for all x such that $|x - x_0| > R$.

(iv) If the power series $f(x)$ converges at some point $x_1 \neq x_0$, then it converges uniformly on $[x_0 - R', x_0 + R']$ if $0 \le R' < |x_1 - x_0|$.

Proof. For simplicity, assume $x_0 = 0$. By the root test (Proposition 2.26(i)), $f(x)$ converges if

$$\limsup_{n\to\infty} |a_n x^n|^{1/n} < 1,$$

which is equivalent to $|x| < R$. Thus we obtain (i), and (iii) is similar by the root test (Proposition 2.26(ii)). For (ii), it suffices to show that f is uniformly Cauchy (Proposition 3.25). Pick $t \in (R', R)$ so that $R'/t < 1$. As $\limsup |a_n|^{1/n} = 1/R < 1/t$, there exists $N \in \mathbb{N}$ such that $|a_n| < 1/t^n$ for all $n \ge N$. For $m > n \ge N$ and $|x| \le R'$, we have

$$|s_m(x) - s_n(x)| = |a_{n+1}x^{n+1} + \cdots + a_m x^m| \le (\frac{R'}{t})^{n+1} + \cdots + (\frac{R'}{t})^m. \tag{11}$$

Since $R'/t < 1$, there exists $N' \in \mathbb{N}$ such that the R.S. of (11) is $< \epsilon$ for all $m > n \ge N'$ and $|x| \le R'$. By Theorem 3.26, f is continuous on I since s_n (polynomial) is continuous. (iv) follows from (i),(ii),(iii). □

Remark 3.29. The *number* $R \in [0, \infty]$ is called the *radius of convergence* of the power series $f(x)$. The convergence of f at $x_0 \pm R$ is not conclusive.

Example 11. What is the radius of convergence of the following series?

(a) $1 + x + x^2 + x^3 + \cdots$

(b) $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \cdots$

(c) $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \cdots$

(d) $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$

(e) $(x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \frac{(x-1)^5}{5} - \cdots$

Definition 3.30. Define $\exp(x) := 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$ with radius of convergence $R = \infty$.

Proposition 3.31. *The following assertions hold about exp.*

(i) $exp(0) = 1$ and $exp(1) = e$ Euler's number.

(ii) exp is continuous on $\mathbb{R}$.

(iii) $exp(x + y) = exp(x)exp(y)$ for all $x, y \in \mathbb{R}$.

(iv) $exp(-x) = 1/exp(x)$.

(v) $exp(x) > 0$ for all $x \in \mathbb{R}$ and $exp(x)$ is strictly increasing (hence injective).

(vi) $exp(x)$ is bijective from $\mathbb{R}$ onto $\mathbb{R}_{>0}$; the inverse function denoted $\ln(x) : \mathbb{R}_{>0} \to \mathbb{R}$ is continuous.

Proof. (i) is obvious; (ii) follows from Theorem 3.28(ii); (iii) is **good** exercise.

Note that (iv) follows from (i) and (iii) by putting $y = -x$. For (v), first note that $\exp(x) > 1$ if $x > 0$ (as $(a_n) \ge 0$). As (iv) asserts that $\exp(-x) = 1/\exp(x) > 0$ for all $x > 0$, the first assertion holds. For strictly increasing, note that if $x - y > 0$ then $\exp(x)/\exp(y) = \exp(x - y) > 1$.

(vi). Note that $e > 2$ and (iii) imply that $\exp(n) = e^n \to \infty$ as $n \to \infty$. Then (v) and Corollary 3.19 imply that $\exp([0, \infty)) = [1, \infty)$. By (iv), we obtain $\exp((-\infty, 0]) = (0, 1]$. Together with (v), the first assertion holds. The second one follows from Corollary 3.20. □

Remark 3.32. Actually, $\exp(x)$ is just e^x we used to *know*. Example 11(a),(b),(c),(e) on their domains of convergence are respectively equal to $\frac{1}{1-x}$, $\sin(x)$, $\cos(x)$, and $\ln(x)$. The definitions of $\sin(x)$ and $\cos(x)$ are not easy as they involve arc length of unit circle. You may take (b),(c) as definitions of $\sin(x)$ and $\cos(x)$ and assume the usual properties about them that you learnt in high school.

Definition 3.33. For $x, y \in \mathbb{R}_{>0}$, define $x^y := \exp(y \ln(x))$. The definition coincides with usual x^y when $x \in \mathbb{R}_{>0}$ and $y \in \mathbb{Q}_{>0}$.

4 Differentiation

4.1 Derivatives of functions

Definition 4.1. Let I be an open interval, $\alpha \in I$, and $f : I \to \mathbb{R}$. The limit

$$f'(\alpha) := \lim_{x\to\alpha} \frac{f(x) - f(\alpha)}{x - \alpha}, \tag{12}$$

if exists, is called the *derivative* of f at α. In this case, we say that f is *differentiable* at α. We say that f is *differentiable* on I if it is differentiable at every $\alpha \in I$. In this case, $f' : I \to \mathbb{R}$ is a function.

Remark 4.2. The existence of $f'(\alpha)$ depends only on the function f on a neighborhood of α. Geometrically, $f'(\alpha)$ is the slope of the tangent line of the graph of f at the point $(\alpha, f(\alpha))$. Similarly, one defines left- or right-derivative $\lim_{x\to\alpha^\pm} \frac{f(x)-f(\alpha)}{x-\alpha}$ of f at α (if exists). It follows that f is differentiable at α with derivative $f'(\alpha)$ iff both the left and right derivatives at α exist and equal to $f'(\alpha)$ (Remark 3.2(3)). If α is an end-point of a closed bounded interval, one defines one-side derivative if it exists.

Proposition 4.3. *If $f : I \to \mathbb{R}$ is differentiable at $\alpha \in I$, then f is continuous at α.*

Proof. By Remark 3.2(1), $\frac{f(x)-f(\alpha)}{x-\alpha}$ is bounded near α, which means the existence of $C > 0$ and $\delta > 0$ such that $|\frac{f(x)-f(\alpha)}{x-\alpha}| \leq C$ for all $0 < |x - \alpha| < \delta$. Hence, we obtain

$$0 \leq |f(x) - f(\alpha)| \leq C|x - \alpha|$$

for all $|x-\alpha| < \delta$. By taking limit $x \to \alpha$ and Sandwich Theorem, we obtain $\lim_{x\to\alpha} f(x) = f(\alpha)$. $\square$

We establish a basic lemma and then some arithmetic of derivatives.

Lemma 4.4. *Suppose $f : I \to \mathbb{R}$ is differentiable at $\alpha \in I$.*

(i) If $f'(\alpha) > 0$, $\exists \delta > 0$ such that $f(\alpha) < f(x)$ for $x \in (\alpha, \alpha+\delta)$ and $f(x) < f(\alpha)$ for $x \in (\alpha-\delta, \alpha)$.

(ii) If $f'(\alpha) < 0$, $\exists \delta > 0$ such that $f(\alpha) > f(x)$ for $x \in (\alpha, \alpha+\delta)$ and $f(x) > f(\alpha)$ for $x \in (\alpha-\delta, \alpha)$.

Hence, if α is a maximum (or minimum) on $(\alpha - \delta, \alpha + \delta)$ for some $\delta > 0$, then $f'(\alpha) = 0$.

Proof. We demonstrate the first part of (i) since others are similar. If not true, there exists a sequence $(a_n) \to \alpha^+$ such that $f(\alpha) \geq f(a_n)$ for all n. By sequential criterion, we obtain $f'(\alpha) = \lim_{n\to\infty} \frac{f(a_n)-f(\alpha)}{a_n-\alpha} \leq 0$ which is absurd. $\square$

Proposition 4.5. *Let $f, g : I \to \mathbb{R}$ be two functions that are differentiable at $\alpha \in I$.*

(i) cf is differentiable at α with derivative $cf'(\alpha)$.

(ii) $f + g$ is differentiable at α with derivative $f'(\alpha) + g'(\alpha)$.

(iii) fg is differentiable at α with derivative $f'(\alpha)g(\alpha) + f(\alpha)g'(\alpha)$.

(iv) If $g(x) \neq 0$ for $x \in I$, then f/g is differentiable at α with derivative $\frac{f'(\alpha)g(\alpha)-f(\alpha)g'(\alpha)}{g^2(\alpha)}$.

(v) (Chain rule) If $f(I) \subset J$ an open interval and $h : J \to \mathbb{R}$ is differentiable at $f(\alpha)$, then the composition $h \circ f$ is differentiable at α with derivative $h'(f(\alpha))f'(\alpha)$.

(vi) If $f : I \to \mathbb{R}$ is differentiable and bijective onto J (open interval) and $f'(x) \neq 0$ for all $x \in I$, then $f^{-1} : J \to \mathbb{R}$ is differentiable with derivative function $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$.

Proof. (i) and (ii) are obvious by the formula (12) and Proposition 3.3. We demonstrate (iii):

$$\frac{f(x)g(x) - f(\alpha)g(\alpha)}{x - \alpha} = g(x)(\frac{f(x) - f(\alpha)}{x - \alpha}) + f(\alpha)(\frac{g(x) - g(\alpha)}{x - \alpha}) \to f'(\alpha)g(\alpha) + f(\alpha)g'(\alpha)$$

as $x \to \alpha$. (iv) is similar (using also Proposition 4.3).

For (v), we first consider $f'(\alpha) \neq 0$. Lemma 4.4 implies that $f(x) \neq f(\alpha)$ for $0 < |x - \alpha| < \delta$ for some δ. By the continuity of h at $f(\alpha)$ (Proposition 4.3), we obtain as $x \to \alpha$ that

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

Next consider $f'(\alpha) = 0$, we need to show that $(h \circ f)'(\alpha) = 0$. Since h is differentiable at $f(\alpha)$, there exist $C > 0$ and $\delta_1 > 0$ such that $|\frac{h(y)-h(f(\alpha))}{y-f(\alpha)}| \leq C$ for $0 < |y - f(\alpha)| < \delta_1$. Since f is continuous at α, there exists $\delta_2 > 0$ such that $|f(x) - f(\alpha)| < \delta_1$ for all $|x - \alpha| < \delta_2$. Hence, for all $|x - \alpha| < \delta_2$ we have

$$|\frac{h(f(x)) - h(f(\alpha))}{x - \alpha}| \leq C|\frac{f(x) - f(\alpha)}{x - \alpha}|. \tag{13}$$

Note that if $f(x) = f(\alpha)$ then both sides of (13) are zero. Since $f'(\alpha) = 0$, for a given $\epsilon > 0$ there exists $0 < \delta < \delta_2$ such that R.S. of (13) $< \epsilon$ for all $|x - \alpha| < \delta$.

(vi). Since f is continuous (Proposition 4.3) and strictly monotone (Corollary 3.20(i)), the interval $J = f(I)$ (Corollary 3.19) is an open interval. Suppose $\beta = f(\alpha)$ and (y_n) is any sequence in $J\backslash\{\beta\}$ converging to β. Then the continuity of f^{-1} (Corollary 3.20(ii)) and sequential criterion (Theorem 3.12) imply that $(x_n := f^{-1}(y_n))$ is a sequence in $I\backslash\{\alpha\}$ converging to α. Since

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(\beta)}{y_n - \beta} = \lim_{n\to\infty} \frac{x_n - \alpha}{f(x_n) - f(\alpha)} = \frac{1}{f'(\alpha)}$$

holds, we are done by sequential criterion (Theorem 3.5). □

Remark 4.6. The proposition implies that the sum, difference, product, quotient (if possible), composition, and inverse (if derivatives non-zero) of differentiable functions are differentiable.

Example 12. Is f below differentiable and what is the derivative?

(a) f is constant function, $f(x) = mx + c$ where $m, c \in \mathbb{R}$, $f(x) = x^n$ where $n \in \mathbb{Z}$.

(b) $f(x) = |x|$.

(c) $f(x) = x^{1/3}$ (what about $f(x) = x^r := e^{r\ln(x)}$, where $r \in \mathbb{R}$?).

(d) $0 \leq f(x) \leq x^2$ for all x (draw a picture).

4.2 Mean Value Theorem

We prove a basic and very useful theorem for differentiable functions.

Theorem 4.7. *Let $f, g : [a, b] \to \mathbb{R}$ be continuous functions that are differentiable on (a, b).*

(i) (Rolle's Theorem) If $f(a) = f(b)$, then there exists $\zeta \in (a, b)$ such that $f'(\zeta) = 0$.

(ii) (Mean Value Theorem) There exists $\zeta \in (a, b)$ such that

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

(iii) (Cauchy Mean Value Theorem) There exists $\zeta \in (a, b)$ such that

$$f'(\zeta)(g(b) - g(a)) = g'(\zeta)(f(b) - f(a)).$$

Proof. (i). If f is constant, then $f' \equiv 0$ and we are done. Otherwise, either the maximum or minimum (exist by Corollary 3.16) of f occurs at some $\zeta \in (a, b)$ and we obtain $f'(\zeta) = 0$ by Lemma 4.4.

(ii). Let $L(x) = mx+c$ be the function whose graph is a straight line joining $(a, f(a))$ and $(b, f(b))$, and note that $L'(x) \equiv m = \frac{f(b)-f(a)}{b-a}$. Then $F(x) = f(x) - L(x)$ is continuous on $[a, b]$, differentiable on (a, b), and satisfies $F(a) = 0 = F(b)$. Rolle's Theorem implies that $0 = F'(\zeta) = f'(\zeta) - \frac{f(b)-f(a)}{b-a}$ for some $\zeta \in (a, b)$.

(iii). Apply Rolle's Theorem to $h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)]$. □

Note that (iii)=(ii) if $g(x) = x$. The following corollaries are immediate.

Corollary 4.8. *Let $f : I \to \mathbb{R}$ be differentiable on open interval I with $f' \equiv 0$. Then f is constant.*

Corollary 4.9. *Let $f : I \to \mathbb{R}$ be differentiable on open interval I.*

(i) If $f'(x) > 0$ ($f'(x) < 0$) for all $x \in I$, then f is strictly increasing (resp. strictly decreasing).

(ii) If $f'(x) \geq 0$ ($f'(x) \leq 0$) for all $x \in I$, then f is increasing (resp. decreasing).

Corollary 4.10. *If $f : [a, b] \to \mathbb{R}$ is continuous that is differentiable on (a, b) such that f' is bounded, then f is Lipschitz continuous (Definition 3.21(2)).*

Theorem 4.11. *(Intermediate Value Theorem for Derivatives)[Ro13, Theorem 29.8] Let $f : I \to \mathbb{R}$ be differentiable on open interval I and $x_1 < x_2$ belong to I. For any $c \in \mathbb{R}$ that lies between $f'(x_1)$ and $f'(x_2)$, there exists $\zeta \in (x_1, x_2)$ such that $f'(\zeta) = c$.*

4.3 Derivatives vs uniform convergence

We prove a general theorem on uniform convergence of derivatives.

Theorem 4.12. *Let I be an open interval and (f_n) a sequence of function on I that converges uniformly to f. Suppose f_n is differentiable on I for all n and (f'_n) converges uniformly to g. Then f is differentiable on I and $f' = g$.*

Proof. Let $\alpha \in I$. We would like to prove that f is differentiable at α with derivative $g(\alpha)$. By uniform convergence of (f'_n), for $\epsilon > 0$ we pick $N_1 \in \mathbb{N}$ such that

$$|f'_m(\zeta) - f'_n(\zeta)| < \epsilon/3 \tag{14}$$

for all $m > n \geq N_1$ and $\zeta \in I$. For any $x \in I$, the Mean Value Theorem and (14) imply that

$$|\frac{f_m(x) - f_m(\alpha)}{x - \alpha} - \frac{f_n(x) - f_n(\alpha)}{x - \alpha}| = |\frac{(f_m(x) - f_n(x)) - (f_m(\alpha) - f_n(\alpha))}{x - \alpha}| = |f'_m(\zeta) - f'_n(\zeta)| < \epsilon/3, \tag{15}$$

where ζ is in between x and α. Letting $m \to \infty$ in (16), for $n \geq N_1$ and $x \in I$ we obtain

$$|\frac{f(x) - f(\alpha)}{x - \alpha} - \frac{f_n(x) - f_n(\alpha)}{x - \alpha}| \leq \epsilon/3. \tag{16}$$

Consider

$$|\frac{f(x) - f(\alpha)}{x - \alpha} - g(\alpha)| \leq |\frac{f(x) - f(\alpha)}{x - \alpha} - \frac{f_n(x) - f_n(\alpha)}{x - \alpha}| + |\frac{f_n(x) - f_n(\alpha)}{x - \alpha} - f'_n(\alpha)| + |f'_n(\alpha) - g(\alpha)|. \tag{17}$$

For a large enough $n \geq N_1$, the last term on the R.S of (17) is $< \epsilon/3$. Then pick $\delta > 0$ such that for all $0 < |x - \alpha| < \delta$ the middle term on the R.S of (17) is $< \epsilon/3$. Together with (16), we conclude that L.S. of (17) is $< \epsilon$ for $0 < |x - \alpha| < \delta$. We are done. ☐

Definition 4.13. Let I be an open interval and $f : I \to \mathbb{R}$. For $n \in \mathbb{N}$, denote the nth derivative of f by $f^{(n)} : I \to \mathbb{R}$ or $\frac{d^n f}{dx^n}$ if it exists. Define $C^n(I) := \{f : I \to \mathbb{R} : \ f^{(n)} \text{ exists and is continuous}\}$ and $C^\infty(I) := \cap_{n \in \mathbb{N}} C^n(I)$. Functions in $C^\infty(I)$ are called *smooth* or *infintely differentiable.*

Theorem 4.14. *(Derivative of power series) Let $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ be a power series with radius of convergence $R \in (0, \infty]$.*

(i) The series $g(x) = \sum_{n=0}^{\infty} a_{n+1}(n + 1)(x - x_0)^n$ has same radius of convergence R.

(ii) f is differentiable on $I = (x_0 - R, x_0 + R)$ and $f' = g$.

(iii) f is a smooth function on I and $f^{(n)}(x_0) = a_n n!$.

Proof. (i) follows since $\limsup |a_n|^{1/n} = \limsup |a_{n+1}(n+1)|^{1/n}$ (exercise). For (ii), we need to show that $f'(\alpha) = g(\alpha)$ for $\alpha \in I$. Pick a closed bounded interval $[a, b] \subset I$ such that $a < \alpha < b$. Then we are done by Theorem 3.28(ii), Theorem 4.14(i), and Theorem 4.12. (iii) follows from Theorem 4.14 (i),(ii) and Definition 4.13. $\square$

Example 13. Find the derivatives of the power series in Example 11 and the functions below.

(a) $\ln(x)$ as inverse function of $\exp(x) = e^x$.

(b) $f(x) = x^r := e^{r \ln(x)}$, where $r \in \mathbb{R}$.

4.4 L'Hospital's Rules

We develop powerful tools that compute limits of indefinite forms in Remark 3.9. Let $f : (a, b) \to \mathbb{R}$ be a continuous function. If $\lim_{x\to a^+} f(x) = c$ exists, then f can be extended to be a continuous function on $[a, b)$ by defining $f(a) = c$.

Theorem 4.15. *(L'Hospital's Rule I) Let f and g be differentiable functions defined on an open interval I. Suppose $* \in \{a^+, a^-, \infty, -\infty\}$ is an end-point of I and the following conditions hold:*

(a) $g(x) \neq 0$ and $g'(x) \neq 0$ for all $x \in I$ near $*$.

(b) $\lim_{x\to *} f(x) = 0 = \lim_{x\to *} g(x)$.

(c) $\lim_{x\to *} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{\pm\infty\}$.

Then $\lim_{x\to *} \frac{f(x)}{g(x)} = L$.

Proof. It suffices to handle the case $* = a^+$ and $* = \infty$. For the first case, we may assume f and g are continuous at a with values 0 by (b). Then Cauchy Mean Value Theorem and (c) imply that

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\zeta_x)}{g'(\zeta_x)} \to L$$

as $x \to a^+$ (note that $\zeta_x \in (a, x)$). For the second case, we do a change of variable $u = 1/x$. Then

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{u\to 0^+} \frac{f(1/u)}{g(1/u)} = \lim_{u\to 0^+} \frac{f'(1/u)(-1/u^2)}{g'(1/u)(-1/u^2)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = L \tag{18}$$

by the first case. $\square$

Theorem 4.16. *(L'Hospital's Rule II) Let f and g be differentiable functions defined on an open interval I. Suppose $* \in \{a^+, a^-, \infty, -\infty\}$ is an end-point of I and the following conditions hold:*

(a) $g'(x) \neq 0$ for all $x \in I$ near $*$.

(b) $\lim_{x\to *} f(x) \in \{\pm\infty\}$ and $\lim_{x\to *} g(x) \in \{\pm\infty\}$.

(c) $\lim_{x\to *} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{\pm\infty\}$.

Then $\lim_{x\to *} \frac{f(x)}{g(x)} = L$.

Proof. It suffices to handle the case $* = a^+$ and $* = \infty$. For the first case, take $b \in I$ and consider $x \in (a, b)$. Cauchy Mean Value Theorem implies that

$$g'(\zeta_x)[f(x) - f(b)] = f'(\zeta_x)[g(x) - g(b)]$$

for some $\zeta_x \in (x, b)$. By rearranging terms, we obtain

$$\frac{f(x)}{g(x)} = \frac{f'(\zeta_x)}{g'(\zeta_x)}(1 - \frac{g(b)}{g(x)}) + \frac{f(b)}{g(x)}. \tag{19}$$

We demonstrate the case $L \in \mathbb{R}$ (the others are similar). By (c), for small $\epsilon > 0$ pick $b \in I$ such that $|\frac{f'(x)}{g'(x)} - L| < \epsilon$ for all $x \in (a, b)$. By (b), pick $(b-a) > \delta > 0$ such that $|\frac{g(b)}{g(x)}| < \epsilon$ and $|\frac{f(b)}{g(x)}| < \epsilon$ for all $x \in (a, a+\delta)$. By using (19) and these estimates, for all $x \in (a, a+\delta)$ we have

$$|\frac{f(x)}{g(x)} - L| \leq |\frac{f'(\zeta_x)}{g'(\zeta_x)} - L| + \epsilon|\frac{f'(\zeta_x)}{g'(\zeta_x)}| + \epsilon \leq \epsilon(|L| + \epsilon) + 2\epsilon$$

which can be made to be arbitrary small. The first case is done. For the second, it is the same as (18) by using a change of variable and the first case. □

Example 14. Find the following limits. You may assume the usual properties of $\sin(x), \cos(x)$.

(a) $\lim_{x\to 0} \frac{\sin(x)}{x}$ and $\lim_{x\to 0}(\frac{1}{x} - \frac{1}{\sin(x)})$.

(b) $\lim_{x\to\infty} \frac{\ln(x)}{x}$ and $\lim_{x\to\infty} \frac{x^2}{e^x}$.

(c) $\lim_{x\to\infty} x^{1/x}$, $\lim_{x\to 0^+} x^x$, and $\lim_{x\to\infty}(1 + \frac{1}{x})^x$.

4.5 Taylor's Theorem

Definition 4.17. Let $f : I \to \mathbb{R}$ be a function defined on an open interval I and $x_0 \in I$.

(1) If f is smooth (Definition 4.13), the *Taylor series* of f at x_0 is defined as

$$Tay(f; x_0)(x) := f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \cdots \quad (20)$$

(2) If the $(n-1)$th derivative $f^{(n-1)}$ exists on I (for some $n \in \mathbb{N}$), the *remainder term* $R_n(x)$ is defined as

$$R_n(x) := f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k. \quad (21)$$

If a power series $\sum_n a_n(x - x_0)^n$ converges uniformly to f on some $[x_0 - r, x_0 + r]$ $(r > 0)$, then it must be the Taylor series (20) by Theorem 4.14(iii). Let R be the radius of convergence of (20) and suppose $R \in (0, \infty]$. It is not necessarily true that (20) converges uniformly to f on $[x_0 - R', x_0 + R']$ for $0 < R' < R$. For example, define f so that $f(x) = 0$ for $x \leq 0$ and $f(x) = e^{-1/x^2}$ for $x > 0$; then $f(0) = 0$ and $f^{(n)}(0) = 0$ for all n. We give a condition in terms of the remainder $R_n(x)$.

Proposition 4.18. *If there exists $r > 0$ such that the Taylor series (20) of a smooth f converges pointwise to $f(x)$ (i.e., $\lim R_n(x) = 0$) for all $x \in (x_0 - r, x_0 + r)$, then it converges uniformly to f on $[x_0 - R', x_0 + R']$ for all $0 < R' < r$.*

Proof. It follows directly from Theorem 3.28(iv). □

The remainder $R_n(x)$ has a formula below whose proof is a repeated use of Rolle's Theorem (note that the $n = 1$ case is just Mean Value Theorem).

Theorem 4.19. *(Taylor's Theorem, see [Ro13, Theorem 31.3]) Let $f : I \to \mathbb{R}$ be a function on an open interval I such that $f^{(n)}$ exists on I for some $n \in \mathbb{N}$ and let $x_0 \in I$. For any $x \neq x_0$ in I, there exists ζ in between x_0 and x such that*

$$R_n(x) = \frac{f^{(n)}(\zeta)}{n!}(x - x_0)^n.$$

Corollary 4.20. *Let $f : I \to \mathbb{R}$ be smooth function and let $C > 0$ such that $|f^{(n)}(x)| \leq C$ for all $x \in I$ and $n \in \mathbb{N}$. Then $\lim_{n\to\infty} R_n(x) = 0$ for all $x \in I$.*

Proof. Use Taylor's Theorem and Example 2(d). □

Remark 4.21. The functions $\sin(x), \cos(x), e^x$ satisfy the conditions of Corollary 4.20 for bounded I.

5 Riemann Integration

5.1 Riemann integrals of functions I

Given a closed bounded interval $[a, b]$ (with $a < b$), a *partition* $\mathcal{P}$ of $[a, b]$ is a finite strictly increasing sequence of points in $[a, b]$ that starts from a and ends at b:

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b, \tag{22}$$

where $n \in \mathbb{N}$. A partition $\mathcal{Q}$ ($a = y_0 < y_1 < \cdots < y_m = b$) is called a *refinement* of $\mathcal{P}$ (denoted $\mathcal{P} \subset \mathcal{Q}$) if the points $\{x_i\} \subset \{y_j\}$. A *tagged partition* $\dot{\mathcal{P}}$ is a partition $\mathcal{P}$ together with a set of points $t_i \in [x_{i-1}, x_i]$ for all $1 \leq i \leq n$. The *norm* of a partition $\mathcal{P}$ is defined as

$$||\mathcal{P}|| := \max\{x_i - x_{i-1} : 1 \leq i \leq n\}.$$

The norm $||\dot{\mathcal{P}}||$ of a tagged partition $\dot{\mathcal{P}}$ is defined as $||\mathcal{P}||$. Let $f : [a, b] \to \mathbb{R}$ be a function on a closed bounded interval. The *Riemman sum* of f corresponding to $\dot{\mathcal{P}}$ is defined as

$$S(f; \dot{\mathcal{P}}) := \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}). \tag{23}$$

Definition 5.1. A function $f : [a, b] \to \mathbb{R}$ is said to be *Riemann integrable* if there exists $L \in \mathbb{R}$ such that for any $\epsilon > 0$, there exists $\delta_\epsilon > 0$ such that

$$|S(f; \dot{\mathcal{P}}) - L| < \epsilon \tag{24}$$

for all tagged partitions $\dot{\mathcal{P}}$ with norm $||\dot{\mathcal{P}}|| < \delta_\epsilon$. In this case, L is called a *Riemann integral* of f.

Proposition 5.2. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable.*

(i) The Riemman integral L of f is unique and denoted $\int_a^b f(x)dx$ or $\int_a^b f$.

(ii) f is bounded on $[a, b]$.

Proof. (i). If $L_1 \neq L_2$ are two Riemann integrals of f, we take $\epsilon = \frac{|L_1 - L_2|}{2} > 0$ and some $\dot{\mathcal{P}}$ with $||\dot{\mathcal{P}}|| < \delta_\epsilon$ to get a contradiction.

(ii). Suppose f is unbounded. Since $[a, b]$ is bounded, Bolzano-Weierstrass Theorem implies that there is a (WLOG) strictly increasing sequence (a_n) in $[a, b]$ converging to $\alpha \in [a, b]$ such that (WLOG) $\lim f(a_n) = \infty$. Fix $\epsilon > 0$, then (24) holds for all tagged partition $\dot{\mathcal{P}}$ with $||\dot{\mathcal{P}}|| < \delta_\epsilon$. Fix a $\dot{\mathcal{P}}_0$ with $||\dot{\mathcal{P}}_0|| < \delta_\epsilon$. Suppose $\alpha \in [x_{m-1}, x_m]$ in $\mathcal{P}_0$ such that $x_{m-1} < \alpha$. For all k sufficiently large, we have $a_k \in [x_{m-1}, x_m]$ and we let $\dot{\mathcal{P}}_k$ be the tagged partition given by replacing only the tag t_m in $\dot{\mathcal{P}}_0$ with a_k. Since $||\dot{\mathcal{P}}_k|| < \delta_\epsilon$ for all k and $S(f; \dot{\mathcal{P}}_k) \to \infty$ as $k \to \infty$, we get a contradiction. $\square$

We will develop some intrinsic criteria for a bounded function to be Riemann integrable without referencing the Riemann integral L (analogous to Cauchy sequence).

Definition 5.3. Let $f : [a, b] \to \mathbb{R}$ be a bounded function and $\mathcal{P}$ (22) a partition of $[a, b]$. The *upper Darboux sum* of f with respect to $\mathcal{P}$ is defined as

$$U(f; \mathcal{P}) := \sum_{i=1}^{n} \sup(f([x_{i-1}, x_i]))(x_i - x_{i-1}) \tag{25}$$

and the *lower Darboux sum* of f with respect to $\mathcal{P}$ is defined as

$$L(f; \mathcal{P}) := \sum_{i=1}^{n} \inf(f([x_{i-1}, x_i]))(x_i - x_{i-1}). \tag{26}$$

The *upper Darboux integral* $U(f)$ and *lower Darboux integral* $L(f)$ of f are defined as

$$U(f) := \inf\{U(f; \mathcal{P}) : \mathcal{P} \text{ partition}\} \quad \text{and} \quad L(f) := \sup\{L(f; \mathcal{P}) : \mathcal{P} \text{ partition}\}. \tag{27}$$

Since f is bounded, all four terms are finite. It is obvious that $L(f;\mathcal{P}) \leq U(f;\mathcal{P})$. We prove some basic estimates.

Proposition 5.4. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $\mathcal{P} \subset \mathcal{Q}$ two partitions of $[a,b]$.*

(i) The following holds:

$$-\infty < L(f;\mathcal{P}) \leq L(f;\mathcal{Q}) \leq L(f) \leq U(f) \leq U(f;\mathcal{Q}) \leq U(f;\mathcal{P}) < \infty.$$

(ii) $U(f;\mathcal{P}) = sup\{S(f;\dot{\mathcal{P}}) :\ \dot{\mathcal{P}}$ is a tagged partition on $\mathcal{P}\}$.

(iii) $L(f;\mathcal{P}) = inf\{S(f;\dot{\mathcal{P}}) :\ \dot{\mathcal{P}}$ is a tagged partition on $\mathcal{P}\}$.

Proof. (i). Since f is bounded, the terms defined in Definition 5.3 are all finite. By the definition,

$$-\infty < L(f;\mathcal{P}) \leq L(f) \leq U(f) \leq U(f;\mathcal{P}) < \infty$$

holds. It remains to show $L(f;\mathcal{P}) \leq L(f;\mathcal{Q})$ and $U(f;\mathcal{Q}) \leq U(f;\mathcal{P})$ but they are obvious by the definition of $\mathcal{P} \subset \mathcal{Q}$.

(ii) and (iii) follow from definition. □

Theorem 5.5. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. The following assertions are equivalent.*

(a) f is Riemann integrable.

(b) For any $\epsilon > 0$, there is $\delta_\epsilon > 0$ such that $U(f;\mathcal{P}) - L(f;\mathcal{P}) < \epsilon$ for all partitions $\mathcal{P}$ with $||\mathcal{P}|| < \delta_\epsilon$.

(c) For any $\epsilon > 0$, there is a partition $\mathcal{P}$ such that $U(f;\mathcal{P}) - L(f;\mathcal{P}) < \epsilon$.

(d) The upper and lower Darboux sums coincide: $U(f) = L(f)$.

In this case, $U(f) = L(f)$ is the Riemann integral $\int_a^b f$.

Proof. $(c) \iff (d)$. Follows directly from (27) and Proposition 5.4.

$(a) \Rightarrow (b)$. Given $\epsilon > 0$, there exists $\delta_\epsilon > 0$ such that for all $\dot{\mathcal{P}}$ with $||\dot{\mathcal{P}}|| < \delta_\epsilon$:

$$L - \frac{\epsilon}{3} < S(f;\dot{\mathcal{P}}) < L + \frac{\epsilon}{3}. \tag{28}$$

Thus by (28) and Proposition 5.4(ii),(iii), we obtain for all $\mathcal{P}$ with $||\mathcal{P}|| < \delta_\epsilon$ that

$$L - \frac{\epsilon}{3} \leq L(f;\mathcal{P}) \leq U(f;\mathcal{P}) \leq L + \frac{\epsilon}{3}.$$

$(b) \Rightarrow (c)$. Obvious.

$(c) \Rightarrow (a)$. By (c) and Proposition 5.4(i), we have $L(f) = U(f)$ and denote it by L. Given $\epsilon > 0$, (c) implies the existence of a partition $\mathcal{P}$ such that

$$U(f;\mathcal{P}) - L(f;\mathcal{P}) < \epsilon/2. \tag{29}$$

Suppose $\mathcal{P}$ has $n+2$ points ($n \in \mathbb{N}$) and $|f|$ is bounded by some $C > 0$. Pick $0 < \delta < \frac{\epsilon}{4Cn}$ and let $\dot{\mathcal{Q}}$ to be a tagged partition with norm $< \delta$. Choose a tagged partition $\mathcal{Q} \dot{\cup} \mathcal{P}$ such that

(I) the partition is given by the union $\mathcal{Q} \cup \mathcal{P}$ and

(II) the tags of $\dot{\mathcal{Q}}$ is a subset of the tags of $\mathcal{Q} \dot{\cup} \mathcal{P}$.

Then we obtain

$$|S(f;\dot{\mathcal{Q}}) - S(f;\mathcal{Q} \dot{\cup} \mathcal{P})| \leq n(2C\delta) < \epsilon/2. \tag{30}$$

Since $\mathcal{Q} \cup \mathcal{P}$ is a refinement of $\mathcal{P}$, we obtain

$$|S(f;\mathcal{Q} \dot{\cup} \mathcal{P}) - L| < \epsilon/2 \tag{31}$$

by (29) and Proposition 5.4(i),(ii),(iii). Therefore, we obtain $|S(f;\dot{\mathcal{Q}}) - L| < \epsilon$ for all $\dot{\mathcal{Q}}$ with norm $< \delta$ by triangle inequality, (30) and (31). □

Corollary 5.6. *Let $f : [a, b] \to \mathbb{R}$ be a Riemann integrable function. If $(\dot{\mathcal{P}}_n)$ is a sequence of tagged partitions such that $||\dot{\mathcal{P}}_n|| \to 0$, then*

$$\lim_{n\to\infty} L(f; \mathcal{P}_n) = \lim_{n\to\infty} S(f; \dot{\mathcal{P}}_n) = \lim_{n\to\infty} U(f; \mathcal{P}_n) = \int_a^b f.$$

Proof. By Proposition 5.4(ii),(iii), we obtain

$$L(f; \mathcal{P}_n) \le S(f; \dot{\mathcal{P}}_n) \le U(f; \mathcal{P}_n)$$

for all n. Since Theorem 5.5 ((a)⇒(b)) implies that the first and third terms tend to $L(f) = U(f) = \int_a^b f$ as $n \to \infty$, we are done by Sandwich theorem. □

5.2 Riemann integrals of functions II

Theorem 5.7. *If $f : [a, b] \to \mathbb{R}$ is monotone, then f is Riemann integrable.*

Proof. WLOG, suppose f is increasing. Then $f(a) \le f(x) \le f(b)$ for all x. Partition $[a, b]$ into n closed intervals of equal widths and denote it $\mathcal{P}_n$. Then $U(f; \mathcal{P}_n) - L(f; \mathcal{P}_n) = \frac{(f(b)-f(a))(b-a)}{n} \to 0$. We are done by Theorem 5.5. □

Theorem 5.8. *If $f : [a, b] \to \mathbb{R}$ is continuous, then f is Riemann integrable.*

Proof. By Corollary 3.16, f is bounded. Since f is uniformly continuous (Proposition 3.23), there exists $\delta > 0$ such that $|f(x) - f(y)| < \frac{\epsilon}{b-a}$ for all $x, y \in [a, b]$ such that $|x - y| < \delta$. Take a partition $\mathcal{P}$ with norm $< \delta$. Then $U(f; \mathcal{P}) - L(f; \mathcal{P}) < \frac{\epsilon}{b-a}(b - a) = \epsilon$. □

We prove some properties for Riemann integrable functions.

Proposition 5.9. *Let $f, g : [a, b] \to \mathbb{R}$ be Riemann integrable.*

(i) cf is Riemann integrable and $\int_a^b (cf) = c \int_a^b f$.

(ii) $f + g$ is Riemann integrable and $\int_a^b (f + g) = \int_a^b f + \int_a^b g$.

(iii) f^2 and fg are Riemann integrable.

(iv) If $f(x) \le g(x)$ for all x, then $\int_a^b f \le \int_a^b g$.

(v) $|f|$ is Riemann integrable and $|\int_a^b f| \le \int_a^b |f|$.

(vi) If $h : [a, b] \to \mathbb{R}$ is a function and $c \in (a, b)$, then h is Riemann integrable on $[a, b]$ iff h is Riemann integrable on $[a, c]$ and $[c, b]$. In this case, $\int_a^b h = \int_a^c h + \int_c^b h$.

Proof. All the functions we consider are bounded so that the criteria in Theorem 5.5 can be applied.
(i). Treat only the case $c > 0$ ($c \le 0$ is exercise). Note that $U(cf) = cU(f) = cL(f) = L(cf)$.
(ii). Note that $U(f) = L(f)$, $U(g) = L(g)$, and $L(f)+L(g) \le L(f+g) \le U(f+g) \le U(f)+U(g)$. Indeed,

$$U(f+g) = \inf\{U(f+g; \mathcal{P}) : \mathcal{P}\} \le \inf\{U(f; \mathcal{P})+U(g; \mathcal{P}) : \mathcal{P}\} \le \lim_{n\to\infty}[U(f; \mathcal{P}_n)+U(g; \mathcal{P}_n)] = U(f)+U(g)$$

by Corollary 5.20 where $||\mathcal{P}_n|| \to 0$.
(iii). By $fg = (f + g)^2/4 - (f - g)^2/4$ and (i),(ii), it suffices to consider f^2. Hint: if $|f| \le C$, show

$$U(f^2; \mathcal{P}) - L(f^2; \mathcal{P}) \le 2C[U(f; \mathcal{P}) - L(f; \mathcal{P})].$$

(iv). Note that $U(f) \le U(g)$.
(v). Since $U(|f|; \mathcal{P})-L(|f|; \mathcal{P}) \le U(f; \mathcal{P})-L(f; \mathcal{P})$ for any partition $\mathcal{P}$, $|f|$ is Riemann integrable. Then $|\int_a^b f| \le \int_a^b |f|$ follows from $-|f| \le f \le |f|$ and (i),(iv).
(vi). If h is bounded, show that $U(h) = U(h|_{[a,c]}) + U(h|_{[c,b]})$ and $L(h) = L(h|_{[a,c]}) + L(h|_{[c,b]})$ (exercise). □

Example 15. Find the Riemann integral (if exists) of the following $f : [0, 1] \to \mathbb{R}$.

(a) $f(x) \equiv c$ constant.

(b) $f(x) = x$.

(c) $f(x) = 1$ if $x \in \mathbb{Q}$ and $f(x) = 0$ otherwise.

(d) $f(x) = 1$ if $x = \frac{1}{2^n}$ for all $n \in \mathbb{N}$ and $f(x) = 0$ otherwise.

Theorem 5.10. *(Mean Value Theorem for integration) Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there exists $\zeta \in [a, b]$ such that*

$$f(\zeta) = \frac{1}{b-a} \int_a^b f.$$

Proof. Let M and M' be respectively the maximum and minimum of f on $[a, b]$ (Corollary 3.16). Since $m \le \frac{1}{b-a} \int_a^b f \le M$ (Proposition 5.9(iv)), we are done by the Intermediate Value Theorem (Corollary 3.17). □

Definition 5.11. A subset $S \subset \mathbb{R}$ is said to be of *measure zero* if for any $\epsilon > 0$, there exists a union of countably many open intervals $I = \cup_{n\in\mathbb{N}}(a_n, b_n)$ such that $S \subset I$ and $\sum_{n=1}^{\infty}(b_n - a_n) < \epsilon$.

For example, a countable subset of $\mathbb{R}$ is of measure zero. It is **good** to know the following criterion.

Theorem 5.12. *(Lebesgue's Integrability Criterion, see [BS11, 7.3.12])*
A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if the set

$$\{\alpha \in [a, b] : \ f \text{ is not continuous at } \alpha\}$$

is of measure zero.

5.3 The Fundamental Theorem of Calculus

We study the relationship between differentiation and Riemann integration, which enables us to compute Riemann integrals of functions. We say that a function $f : (a, b) \to \mathbb{R}$ is Riemann integrable if some extension $\widetilde{f}$ of f to $[a, b]$ is Riemann integrable. Then we write $\int_a^b f := \int_a^b \widetilde{f}$, which is independent of extension (why?).

Theorem 5.13. *(Fundamental Theorem of Calculus I) Let $F : [a, b] \to \mathbb{R}$ be a continuous function that is differentiable on (a, b). If $f := F'$ is Riemann integrable on (a, b), then*

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

Proof. Given $\epsilon > 0$, there exists a partition $\mathcal{P} = (a = x_0 < x_1 < \cdots < x_n = b)$ such that $U(f; \mathcal{P}) - L(f; \mathcal{P}) < \epsilon$. For $1 \le i \le n$, Mean Value Theorem implies

$$f(t_i) = \frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}}$$

for some $t_i \in (x_{i-1}, x_i)$. Then the Riemann sum

$$S(f; \dot{\mathcal{P}}) = \sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) = F(b) - F(a),$$

where $\dot{\mathcal{P}}$ is the tagged partition given by using $\{t_i\}$ as tags. Hence, both numbers $F(b) - F(a)$ and $\int_a^b f$ belong to $[L(f; \mathcal{P}), U(f; \mathcal{P})]$ of length $< \epsilon$. As $\epsilon > 0$ is arbitrary, the two numbers are equal. □

Corollary 5.14. *(Integration by parts) Let u and v be continuous functions on $[a,b]$ that are differentiable on (a,b). If u' and v' are integrable on (a,b), then*

$$\int_a^b uv' + \int_a^b u'v = u(b)v(b) - u(a)v(a).$$

Proof. Use $(uv)' = u'v + uv'$, Proposition 5.9(iii), and Theorem 5.13. □

Example 16. Find $\int_a^b f$ for the following f.

(a) $f(x) = x^n$ where $n \in \mathbb{Z}$.

(b) $f(x) = xe^x$.

(c) $f(x) = \ln|x|$.

Definition 5.15. Let $f : [a,b] \to \mathbb{R}$ be a Riemann integrable function and $x, y \in [a,b]$. Define $\int_x^y f := -\int_y^x f$ if $x > y$ and $\int_x^x f := 0$.

Check that $\int_c^x f - \int_c^y f = \int_y^x f$ for any $c, x, y \in [a,b]$ (by Proposition 5.9(vi)).

Theorem 5.16. *(Fundamental Theorem of Calculus II) Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable and $c \in [a,b]$.*

(i) The function $F(x) := \int_c^x f$ is continuous on $[a,b]$.

(ii) If f is continuous at $x_0 \in (a,b)$, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$. If f is continuous at a (resp. b), then the right derivative of F at a (resp. left derivative of F at b) is $f(a)$ (resp. $f(b)$).

Proof. (i). Pick $C > 0$ such that $|f| \le C$. For $x, y \in [a,b]$, Proposition 5.9(v) implies that

$$|F(y) - F(x)| = |\int_x^y f| \le \pm \int_x^y |f| \le C|y - x|,$$

which is Lipschitz continuity.

(ii) and (iii). Suppose f is continuous at $x_0 \in [a,b]$. For $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ for all $|x - x_0| < \delta$. Then

$$|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)| = \frac{|\int_{x_0}^x (f(t) - f(x_0))dt|}{|x - x_0|} < \epsilon$$

for all $0 < |x - x_0| < \delta$. Hence, we have $F'(x_0) = f(x_0)$. □

Let $C[a,b]$ denote the $\mathbb{R}$-vector space of continuous functions on $[a,b]$ and let $C^1[a,b]$ denote the $\mathbb{R}$-vector space of differentiable functions F on $[a,b]$ (taking one-sided derivatives at the end-points) such that the derivative F' is continuous on $[a,b]$. Assuming basic notion of linear algebra.

Corollary 5.17. *The map $\frac{d}{dx} : C^1[a,b]/\{constants\} \to C[a,b]$ that sends $F \mapsto F'$ is an $\mathbb{R}$-vector space isomorphism. The inverse map is $f \mapsto \int_a^x f(t)dt + C$.*

Proof. By Corollary 4.8 and Theorem 5.16. □

Remark 5.18. It is difficult to describe the space $\{\int_a^x f(t)dt : f$ Riemman integrable on $[a,b]\}$. There is a satisfactory answer if we replace "Riemann integrable" with "Lebesgue integrable".

Corollary 5.19. *(Change of variable) Let I, J be open interval, $f \in C(I)$, and $u \in C^1(J)$ such that $u(J) \subset I$. Then for any $a, b \in J$:*

$$\int_a^b f(u(x))u'(x)dx = \int_{u(a)}^{u(b)} f(u)du.$$

Proof. Let $F(u) = \int_c^u f(t)dt$ for some $c \in I$. If $g(x) = F(u(x))$ then Theorem 5.16(ii) implies that $g'(x) = F'(u(x))u'(x) = f(u(x))u'(x)$, which is continuous on J. Then Theorem 5.13 (and Definition 5.15) imply that

$$\int_a^b f(u(x))u'(x)dx = g(b) - g(a) = F(u(b)) - F(u(a)) = \int_c^{u(b)} f(t)dt - \int_c^{u(a)} f(t)dt = \int_{u(a)}^{u(b)} f(t)dt.$$

□

5.4 Riemann integrals vs uniform convergence

We prove first a general theorem on Riemann integrals of a sequence of functions.

Theorem 5.20. *Let (f_n) be a sequence of Riemann integrable functions on $[a,b]$ that converges uniformly to f. Then f is Riemann integrable and $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$.*

Proof. Given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f(x) - f_n(x)| < \epsilon$ for all $n \geq N$ and all $x \in [a,b]$. Thus, f is bounded and for $n \geq N$:

$$L(f_n) - \epsilon(b-a) \leq L(f) \leq U(f) \leq U(f_n) + \epsilon(b-a). \tag{32}$$

Since $L(f_n) = U(f_n)$ by Theorem 5.5, we obtain $0 \leq U(f) - L(f) \leq 2\epsilon(b-a)$. As this is true for all $\epsilon > 0$, we get $U(f) = L(f)$ and f is Riemann integrable (Theorem 5.5). Finally, $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$ follows from (32) for all $n \geq N$. □

Theorem 5.21. *(Integration of power series) Let $f(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k = \lim_{n\to\infty} f_n(x)$ be a power series with radius of convergence $R \in (0,\infty]$, where $f_n(x) := \sum_{k=0}^{n} a_k(x-x_0)^k$.*

(i) The series $h(x) = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-x_0)^{k+1}$ has same radius of convergence R.

(ii) The series f is Riemann integrable on any closed bounded interval in $(x_0 - R, x_0 + R)$.

(iii) The series $h(x)$ converges uniformly to $\int_{x_0}^{x} f(t)dt$ on any closed bounded interval in $(x_0 - R, x_0 + R)$.

Proof. (i) is similar to the proof of Theorem 4.14(i); (ii) follows directly from Theorem 3.28(ii), Riemann integrability of $f_n(x)$ (by continuity), and Theorem 5.20; (iii) follows from Theorem 5.20 and Theorem 3.28(iv). □

If $f \in C^n(I)$ for some $n \in \mathbb{N}$, there is a Cauchy's formula for remainder $R_n(x)$ (21), whose proof is a repeated use of integration by parts (note that the $n = 1$ case is just the fundamental theorem of calculus).

Theorem 5.22. *([Ro13, Theorem 31.5]) Let I be an open interval, $f : I \to \mathbb{R}$ be a $C^n(I)$-function for some $n \in \mathbb{N}$, and $x_0 \in I$. For any $x \neq x_0$ in I, we have*

$$R_n(x) = \int_{x_0}^{x} \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)dt.$$

Corollary 5.23. *Let f be the function in Theorem 5.22. Then for $x \neq x_0$ in I there is ζ between x_0 and x such that*

$$R_n(x) = \frac{f^{(n)}(\zeta)}{(n-1)!}(x-\zeta)^{n-1}(x-x_0).$$

Proof. Since $\frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)$ is continuous in t, we are done by Theorem 5.22 and Theorem 5.10. □

This better formula for remainder $R_n(x)$ (than Theorem 4.19) enables us to prove the following.

Theorem 5.24. *(Binomial Series Theorem, see [Ro13, Theorem 31.7]) If $r \in \mathbb{R}$ and $|x| < 1$, then*

$$(1+x)^r = 1 + \sum_{k=1}^{\infty} \frac{r(r-1)\cdots(r-k+1)}{k!} x^k.$$