An Introduction to Real Analysis

LECTURES NOTES FOR REAL ANALYSIS I AT COLUMBIA UNIVERSITY

© Mitchell Faulk 2019

Preface

These notes were written during the summer of 2019 as I taught a course on Real Analysis at Columbia University.

They were originally intended to be a supplement to the textbook by Ruding, and, indeed, they follow closely the content and organization of that book. Nevertheless, throughout the writing process, my ambitions grew, and now, I believe, the notes constitute a self-contained introduction to the topic of Real Analysis, only assuming minor pre-requisites, such as elementary Calculus.

Any comments or suggestions for improvement should be directed to mitchellm-faulk@gmail.com.

Contents

1 The real numbers

1.1 Sets

We include this section only for the sake of notation. There are many interesting and subtle considerations for sets, but we won't concern ourselves with any of those here.

We will use the term **set** to mean a collection of **elements**. We will use the symbol $\varnothing$ to denote the empty set with no elements. We often use capital letters (X, Y, A, B, etc.) to refer to sets and lowercase letters (x, y, a, b, etc.) to refer to elements. We will write $x \in X$ to mean that x is an element of a set X.

By a finite set we mean a set with a finite number of elements. For example, if $x_1, \ldots, x_n$ are the (distinct) elements of set X, then we will write

$$X = \{x_1, \ldots, x_n\},$$

and we will refer to X as a finite set. An infinite set is one that is not finite.

We will use the notation $A \subset B$ to mean that A is a subset of B, that is, to mean that the following implication is true: if $x \in A$, then $x \in B$. (Some authors use the notation $A \subseteq B$ to mean the same thing.) In particular, for any set A it is true that $\varnothing \subset A$ and $A \subset A$.

Two sets A and B are equal if and only if both $A \subset B$ and $B \subset A$ are true. In such a case, we write $A = B$. Otherwise, we write $A \neq B$.

If we wish to emphasize that $A \subset B$ is true but that $A \neq B$, then we will write $A \subsetneq B$. In such a situation A is called a *proper* subset of B, although we don't often have much need to concern ourselves with such situations.

1.2 The rational numbers

Let us begin with some notation.

- By $\mathbb{N}$ we mean the set natural numbers or nonnegative integers $\mathbb{N} = \{0, 1, 2, \ldots\}$.

- By $\mathbb{Z}$ we mean the set of integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

- By $\mathbb{Q}$ we mean the set of rational numbers $\mathbb{Q} = \{p/q : p, q \in \mathbb{Z}, q \neq 0\}$.

If we wish to refer to the set of positive integers, we will write $\mathbb{Z}_{>0}$ or $\mathbb{N}^*$ or even $\mathbb{N} \setminus \{0\}$.

The set of rational numbers are, in some sense, as close together as possible: between any two distinct rational numbers x and y, there is a third, namely $(x+y)/2$.

Nevertheless, perhaps surprisingly, the set of rational numbers contains gaps. In particular, rational numbers do not always admit roots, as the following proposition demonstrates.

Proposition 1. *There is no rational number x such that $x^2 = 2$.*

Proof. Let x be a rational number, and write $x = p/q$ for some integers $p, q \in \mathbb{Z}$. We may assume that p, q are relatively prime, or else we could divide them both by their greatest common divisor to obtain another representation of x.

Suppose the relation $x^2 = 2$ is true. Then our representation of x as p/q shows that $p^2 = 2q^2$. We infer that p^2 is even and hence p is even as well (if p were odd, then p^2 would be odd). It follows that 4 divides p^2 and thus 4 also divides $2q^2$. We find that q^2 is even and thus q is even as well. But we have now found that both p and q are even, which contradicts our assumption that they are relatively prime. We conclude that the relation $x^2 = 2$ cannot be true. □

A positive number squaring to 2 would be written $\sqrt{2}$, and the above proposition would say that $\sqrt{2} \notin \mathbb{Q}$, that is, $\sqrt{2}$ would be *irrational*. However, the number $\sqrt{2}$—if such a thing exists!—would be able to be *approximated* by rational numbers in the sense that there is no rational number *closest* to $\sqrt{2}$, as the following proposition asserts.

Proposition 2. *Let A be the set of rational numbers x satisfying $x^2 < 2$. Then A contains no largest element.*

Proof. Let x be an arbitrary element of A satisfying $x > 0$. The proof will be complete if we construct a rational number y belonging to A which satisfies $x < y$.

We claim that
$$y = x - \frac{x^2 - 2}{x + 2} = \frac{2x + 2}{x + 2}$$
works. Indeed, we note that y is rational from the right-most expression of y. We also note that since x belongs to A, the numerator $x^2 - 2$ is negative, and hence $y > x$. It is also simple algebra to compute that
$$y^2 - 2 = \frac{2(x^2 - 2)}{(x + 2)^2},$$
from which we conclude that $y^2 - 2$ is negative, and hence y belongs to A as well. □

Proposition 3. *Let B be the set of positive rational numbers x satisfying $2 < x^2$. Then B contains no smallest element.*

Proof. The proof is similar to the previous result, so we omit it. □

Remark 4. The set $\mathbb{R}$ of real numbers is a way of "filling the gaps" present in the rational numbers, as we will discuss now.

1.3   Ordered sets

Definition 5. Let X be a set. A **relation** on X is a subset R of the Cartesian product $X \times X$. If the point (x, y) belongs to R, then we say x is related to y and we write xRy.

Example 6. Equality determines a relation on X given by the diagonal subset $R = \{(x, x) : x \in X\} \subset X \times X$, that is, we have xRy if and only if $x = y$. Note that this relation is symmetric (if xRy, then yRx) and transitive (if xRy and yRz, then xRz).

Example 7. Order determines a relation on $\mathbb{Q}$ by declaring $x < y$ if and only if $y - x$ is positive. Note that this relation is not symmetric (we have $1 < 2$, but $2 \not< 1$) but it is transitive (if $x < y$ and $y < z$, then $x < z$).

Definition 8. Let S be a set. An **order** on S is a relation, denoted by $<$, satisfying the following two properties.

(i) Trichotomy: If x and y belong to S, then one and only one of the following statements is true
$$x < y, \qquad x = y, \qquad y < x.$$

(ii) Transitivity: If $x, y, z \in S$ satisfy $x < y$ and $y < z$, then $x < z$.

By an **ordered set** we just mean a pair $(S, <)$ consisting of a set and an order. We often use the notation $x \leqslant y$ to mean that either $x < y$ or $x = y$ is true.

Example 9. For example, the set $\mathbb{Q}$ enjoys an order $<$ defined by writing $x < y$ if the number $y - x$ is positive. (The sets $\mathbb{N}$ and $\mathbb{Z}$ enjoy the same order.)

Definition 10. Let $(S, <)$ be an ordered set, and let E be a subset of S. An **upper bound** for E is an element β of S satisfying the following: If x belongs to E, then $x \leqslant \beta$. The notion of **lower bound** is defined similarly.

Example 11. An upper bound may or may not belong to E itself. For example, let $E \subset \mathbb{N}$ be the subset $E = \{0\}$ consisting only of the number zero itself. Then $\beta = 0$ is an upper bound for E and also belongs to E. However, any number $\beta' \in \mathbb{N}$ is also an upper bound for E.

Definition 12. Let $(S, <)$ be an ordered set, and let E be a subset of S that is bounded from above. Suppose there is a number $\alpha \in S$ satisfying the following two properties.

(i) α is an upper bound for E

(ii) α is the smallest upper bound for E in the sense that if β is any upper bound for E, then $\alpha \leqslant \beta$.

Then α is called a **least upper bound** or **supremum** of E.

Proposition 13. *The supremum of E, if it exists, is unique. That is, if α and α' are two suprema, then $\alpha = \alpha'$.*

Proof. This follows immediately from the definitions, but we feel the proof is instructive. Suppose α and α' are two suprema. Then in particular, α' is an upper bound for E, and so property (ii) applied to α asserts that $\alpha \leqslant \alpha'$. On the other hand, property (ii) applied to α' asserts that $\alpha' \leqslant \alpha$. We conclude that we must have $\alpha = \alpha'$ by trichotomy. $\qquad\square$

Notation 14. Because the supremum is unique—if it exists!—we denote it by $\sup E$. We let the reader define the anologous notion of infimum (or greatest lower bound), which is also unique (when it exists) and which we denote by $\inf E$.

Example 15. Returning to our example $E = \{0\}$ from earlier, we find that $\sup E = 0$. In particular, note that in this case $\sup E$ belongs to E itself.

Example 16. There are also examples where $\sup E$ does not belong to E.

For example, consider the subset E of $\mathbb{Q}$ defined by $E = \{x \in \mathbb{Q} : x < 0\}$. Then we claim that $\sup E = 0$.

To verify this, we must check two things: that 0 is an upper bound of E and that 0 is the smallest such upper bound. Note that 0 is upper bound for E simply by virtue of the definition of E. Suppose that β is another upper bound for E. If β were negative, then $\beta/2$ would be a number larger than β belonging to E, which would contradict the assumption that β is an upper bound. We conclude that β is non-positive, which means that $\beta \geqslant 0$. This is as desired.

On the other hand, if E' denotes the set $E' = \{x \in \mathbb{Q} : x \leqslant 0\}$, then the supremum $\sup E'$ is also equal to zero and in this case belongs to E'.

Example 17. There are even examples where the supremum of E does not exist. For example, suppose that A is the subset of $\mathbb{Q}$ considered earlier $A = \{x \in \mathbb{Q} : x^2 < 2\}$. We claim that $\sup A$ does not exist in $\mathbb{Q}$. Indeed, let α be any upper bound for A. Then α belongs to the set B defined earlier by $B = \{x \in \mathbb{Q} : x^2 > 2, x > 0\}$, and conversely any element belonging to B is an upper bound for A. Since B contains no smallest element, it is impossible for α to be the *least* upper bound for A.

This last example motivates the following definition.

Definition 18. An ordered set $(S, <)$ is said to have the **least upper bound property** if the following is true: Whenever E is a nonempty subset of S that is bounded from above, then $\sup E$ exists in S.

The preceding example shows that $S = \mathbb{Q}$ *fails* to have the least upper bound property. The real numbers $\mathbb{R}$ constitute an attempt to repair this failure, as we will see.

1.4   The real field

The rational numbers enjoy the operations of addition and multiplication in such a way that addition distributes over multiplication and several other nice properties are satisfied. For example, there is an additive (resp. multiplicative) identity, and each number (resp. nonzero number) has an additive inverse (resp. multiplicative inverse). Rudin lists all of these properties together in a collection called the **field axioms**. Any set together with two operations satisfying all of these axioms is called a **field**.

There are many elementary properties concerning algebraic manipulations involving addition and multiplication which follow from the field axioms, but we will not prove these here and instead refer the interested reader to Rudin for a more complete list.

Example 19. For a non-example, we note that although the set of integers $\mathbb{Z}$ enjoys two binary operators of addition and multiplication, it fails to be a field because the number 2 for example does not admit a multiplicative inverse.

Definition 20. By an **ordered field** we mean a field $(F, +, \cdot)$ equipped with an order $<$ in such a way that the following two properties are satisfied.

(i) If $x, y \in F$ satisfy $x < y$, then for each $z \in F$, we have $x + z < y + z$.

(ii) If $x, y \in F$ satisfy $x > 0$ and $y > 0$, then $xy > 0$.

There are also many properties concerning the relationship between the ordering and addition/multiplication which follow from these two properties, but we will not discuss these further and instead refer the reader to the textbook.

Example 21. For example, the set $\mathbb{Q}$ of rational numbers is an ordered field when equipped with its usual ordering.

The fundamental result that we will assume from this point forward is the existence of the field of real numbers.

Theorem 22. *There is an ordered field $\mathbb{R}$ which has the least upper bound property and which contains $\mathbb{Q}$ as a subfield.*

We will later discuss how to construct $\mathbb{R}$ from $\mathbb{Q}$. There are several ways of performing this construction, and the interested reader can find one such method in the appendix to the first chapter of Rudin. (We will discuss a different method.)

The following fundamental property of the real numbers is called the **Archimedean property** and the general philosophy behind this property can be phrased by saying that it is possible to travel a mile (or any distance, really) with just a ruler.

Theorem 23 (Archimedean property). *Let x be a positive real number. If y is any real number, then there is a positive integer n such that $nx > y$.*

Proof. Let E denote the subset of $\mathbb{R}$ described by $E = \{nx : n \in \mathbb{Z}_{>0}\}$. Suppose that the conclusion of the theorem is false. Then y is an upper bound for E. By the least uppper bound property, E has a *least* upper bound α. Since x is positive, the difference $\alpha - x$ is not an upper bound for E. This means that there is a positive integer m such that $mx > \alpha - x$. It follows that $(m+1)x > \alpha$, which is a contradiction to the assumption that α is an upper bound for E. We conclude that the theorem is true. $\qquad\square$

In addition, we saw earlier that the rational numbers are in some sense as close together as possible. In particular, they are what is called **dense** in $\mathbb{R}$, as made precise in the following proposition.

Proposition 24. *Between any two distinct real numbers there is a rational number.*

Proof. Let x and y be distinct real numbers. Up to relabeling, we may assume $x < y$. The difference $y - x$ is positive, so by the Archimedean property, there is a positive integer n_0 such that $n_0(y - x) > 1$, that is, $n_0 x + 1 < n_0 y$

Let E denote the set of integers given by $E = \{m \in \mathbb{Z} : m \leqslant n_0 x\}$. By definition, E is bounded above, and hence by Problem 5 of Assignment 1, we deduce that E has a largest element m_0 which satisfies $m_0 \leqslant n_0 x$, that is, $m_0 + 1 \leqslant n_0 x + 1$. Moreover, because m_0 is the greatest element of E, we have that $m_0 + 1$ satisfies $n_0 x < m_0 + 1$. We combine with the previous inequalities to find that

$$n_0 x < m_0 + 1 \leqslant n_0 x + 1 < n_0 y.$$

Upon dividing through by n_0, which is positive, we find that

$$x < \frac{m_0 + 1}{n_0} < y.$$

Thus $(m_0 + 1)/n_0$ is the rational number we seek. $\qquad\square$

Let us conclude our discussion of the real numbers by returning a construction that motivated our discussion in the first place, namely, finding a square root of 2.

Proposition 25. *There is a unique positive real number x satisfying $x^2 = 2$.*

Proof. We first show the existence of such an x. Indeed, let E denote the subset

$$E = \{t \in \mathbb{R} : t^2 < 2\}.$$

Then E is bounded from above, and hence has a supremum, which is positive. We set $x = \sup E$, and we hope to show that x has the desired property by showing that $x^2 < 2$ and $x^2 > 2$ lead to contradictions.

We require the observation that if a, b are real numbers satisfying $a < b$, then we have

$$b^2 - a^2 = (b + a)(b - a) < 2b(b - a).$$

Now assume that $x^2 < 2$. It follows that there is a positive rational number h such that
$$0 < h < \frac{2 - x^2}{2(x + 1)}.$$
We may additionally suppose that h is small enough that $h < 1$. Upon setting $a = x$ and $b = x + h$, we use the previous observation to note that

$$
\begin{aligned}
(x + h)^2 - x^2 &< 2(x + h)h && \text{previous paragraph} \\
&< 2(x + 1)h && h < 1 \\
&< 2 - x^2 && \text{choice of } h.
\end{aligned}
$$

We conclude that $x + h$ satisfies $(x + h)^2 < 2$ so that $x + h$ belongs to E. But this is a contradiction to the assumption that x is an upper bound for E.

It is similarly found that the assumption $x^2 > 2$ leads to a contradiction, so we may conclude that $x^2 = 2$.

We now deal with uniqueness. Suppose that y is another positive real number satisfying $y^2 = 2$. If it were the case that $x < y$, then we would have $x^2 = x \cdot x < x \cdot y < y \cdot y = y^2$, which is incorrect. It similarly cannot be that $y < x$. We conclude that $y = x$ by trichotomy. $\qquad \square$

One can obtain the following more general result on the existence of roots.

Theorem 26. *Let x be a positive real number and let n be a positive integer. There is a unique positive real number y such that $y^n = x$.*

1.5 Euclidean spaces

For a positive integer n, use the notation $\mathbb{R}^n$ for the n-fold Cartesian product

$$\mathbb{R}^n = \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{n \text{ copies}}$$

It is known from linear algebra that $\mathbb{R}^n$ is a vector space, and in particular, enjoys the operations of addition and scalar multiplication. For a vector $x \in \mathbb{R}^n$, we will write $x = (x_1, \ldots, x_n)$ for its components (with respect to the standard basis).

For two vectors $x, y \in \mathbb{R}^n$, we write $\langle x, y \rangle$ to denote the inner product of x and y defined by
$$\langle x, y \rangle = \sum_{k=1}^n x_k y_k,$$
and we define the norm $\|x\|$ of x to be

$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{k=1}^n x_k^2 \right)^{1/2}.$$

When x is just a real number we often write $|x| = \|x\|$.

Proposition 27. *Let x, y, z be points of $\mathbb{R}^n$ and let α be a real number. Then*

(a) $\|x\| \geqslant 0$

(b) $\|x\| = 0$ *if and only if* $x = 0$

(c) $\|\alpha x\| = |\alpha| \, \|x\|$

(d) $|\langle x, y \rangle \leqslant \|x\| \, \|y\|$

(e) $\|x + y\| \leqslant \|x\| + \|y\|$.

Proof. The properties (a), (b), and (c) are immediate, and (d) follows from the Scharz inequality. We now prove (e). We use (d) to find that

$$
\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y \rangle \\
&= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\
&\leqslant \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2,
\end{aligned}
$$

as desired. □

2 Topology

2.1 Countability

Definition 28. Let X, Y be two sets. By a **function** (or **map**) $f : X \to Y$ from X into Y we mean the data of a subset $R \subset X \times Y$ of the Cartesian product satisfying the following property: For each $x \in X$, there is one and only one $y \in Y$ such that (x, y) belongs to R. For a point $(x, y) \in R$, we will often write $y = f(x)$. The set X is called the **domain** of f and the set Y is called the **codomain**. By the **image** of f we mean the subset of Y determined by

$$
\mathrm{im}(f) = \{f(x) : x \in X\} \subset Y.
$$

Remark 29. As a remark on the terminology, some authors (including me) prefer to use the term function for a map whose codomain is $\mathbb{R}$. We will often employ this terminology.

Definition 30. Let $f : X \to Y$ be a map from X into Y.

(i) We say that f is **injective** if whenever $x_1, x_2 \in X$ satisfy $f(x_1) = f(x_2)$, then $x_1 = x_2$.

(ii) We say that f is **surjective** if $\mathrm{im}(f) = Y$.

12

(iii) We say that f is **bijective** if f is both injective and surjective.

Example 31. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $f(x) = x^2$. Then f is not injective because $f(-1) = 1 = f(1)$. Additionally, f is not surjective either because $\text{im}(f) = \mathbb{R}_{\geqslant 0}$.

However, the restriction of f to $X = \mathbb{R}_{\geqslant 0}$ is injective, meaning that the map $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}$ defined in the same way as above is injective.

On the other hand, by restricting the codomain to $Y = \mathbb{R}_{\geqslant 0}$ then we can make f surjective.

If we restrict both the domain and codomain to $X = Y = \mathbb{R}_{\geqslant 0}$, then we realize that $f : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ is now bijective.

Definition 32. A set X is called **countable** if either X is finite or there is a bijective map $f : \mathbb{N} \to X$. Otherwise, we say that X is **uncountable**.

Example 33. The set of natural numbers $\mathbb{N}$ itself is countable because the identity map $\text{id}_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}$ is bijective.

Example 34. We claim that the set $\mathbb{Z}$ of integers is countable. To see this, we will construct a bijective map $f : \mathbb{N} \to \mathbb{Z}$. We construct f piecewise in the following manner

$$f(x) = \begin{cases} x/2 & x \text{ even} \\ -(x+1)/2 & x \text{ odd} \end{cases}.$$

The first few values of f are given by

$$\begin{aligned} f(0) &= 0 \\ f(1) &= -1 \\ f(2) &= 1 \\ f(3) &= -2 \\ f(4) &= 2 \\ f(5) &= -3 \\ f(6) &= 3 \\ &\vdots \end{aligned}$$

From here, we are convinced that f will be a bijection, and it is not too difficult to verify this rigorously.

Example 35. Let A be the collection of infinite sequences consisting of 0s or 1s, that is, an element of A consists of a sequence $s = (s_0, s_1, s_2, \ldots)$ where each s_j is either 0 or 1. Then we claim that A is uncountable. Indeed, suppose that have counted the elements of A as $s^{(0)}, s^{(1)}, s^{(2)}, \ldots$, where each $s^{(j)}$ is a sequence $s^{(j)} = (s_0^{(j)}, s_1^{(j)}, s_2^{(j)}, \ldots)$. Define a new sequence $x = (x_0, x_1, x_2, \ldots)$ as follows. For each

$j = 0, 1, \ldots$, let x_j be 1 if $s_j^{(j)}$ is 0 and let x_j be 0 if $s_j^{(j)}$ is 1. Then we see that x differs from each $s^{(j)}$ in the jth slot and hence x is an element of A that we have not counted! We conclude that A is uncountable.

Example 36. The reasoning of the previous example can be used to show that the set $\mathbb{R}$ of real numbers is uncountable. Indeed, every real number admits a decimal representation, which is itself an infinite sequence of numbers. However, one must be somewhat careful with this argument because in general a decimal representation is not unqiue (for example, $1.0000\ldots = 0.9999\ldots$).

Lemma 37. *Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of finite sets, and set X to be the union*

$$X = \bigcup_{n\in\mathbb{N}} X_n.$$

Then X is countable.

Proof. Just count the elements in X_0 first, and then X_1 second, and then so on. This process can continue because each X_j is finite so the jth step will terminate in finite time. $\qquad\square$

Proposition 38. *Let X and Y be countable sets. Then the Cartesian product $X \times Y$ is countable.*

Proof. Count X as $X = \{x_0, x_1, x_2, \ldots\}$ and Y as $Y = \{y_0, y_1, y_2, \ldots\}$. Let Z be the Cartesian product $Z = X \times Y$. For each nonnegative integer $n \in \mathbb{N}$, let Z_n denote the subset of Z given by

$$Z_n = \{(x_j, y_k) : j + k = n\}.$$

Then each Z_n is a finite set. Because Z is equal to the countable union

$$Z = \bigcup_{n\in\mathbb{N}} Z_n,$$

the previous lemma asserts that Z is countable. $\qquad\square$

Example 39. It follows that the set $\mathbb{Q}$ of rational numbers is countable because we may regard $\mathbb{Q}$ as a subset of the Cartesian product $\mathbb{Z} \times \mathbb{Z}$, which is itself countable.

2.2 Metric spaces

Definition 40. Let X be a set. A **metric** on X is a map $d : X \times X \to \mathbb{R}$ satisfying the following properties.

(i) Non-degeneracy: For each point of points $p, q \in X$, we have $d(p, q) \geqslant 0$ with equality if and only if $p = q$

(ii) Symmetry: For each pair of points $p, q \in X$, we have $d(p, q) = d(q, p)$.

(iii) Triangle inequality: For each triple of points $p, q, r \in X$, we have $d(p, q) \leqslant d(p, r) + d(r, q)$.

By a **metric space** we mean a set X together with a metric d on X.

Example 41. For example, there is a standard metric on $\mathbb{R}^n$ determined by the norm:
$$d(x, y) = \|x - y\|.$$

Example 42. As another example, let X be any set, and consider the function d defined by
$$d(p, q) = \begin{cases} 0 & p = q \\ 1 & p \neq q \end{cases}.$$

Then d defines a metric on X called the discrete metric.

Example 43. If (X_1, d_2) and (X_2, d_2) are two metric spaces, then the Cartesian product $X = X_1 \times X_2$ enjoys the structure of a metric space upon defining the metric
$$d((x_1, x_2), (y_1, y_2)) = \|(d_1(x_1, y_1), d_2(x_2, y_2))\|,$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^2$. The metric d is called the product metric. One can check that if $X_1 = X_2 = \mathbb{R}$ with the standard metric, then the resulting product metric on $\mathbb{R}^2$ agrees with the one described in a previous example.

Definition 44. Let x be a point of a metric space X and let r be a positive number. Then the **open ball of radius** r **centered around** x, denoted $B_r(x)$, is the subset of X determined by
$$B_r(x) = \{y \in X : d(x, y) < r\}.$$

We sometimes also use the term **neighborhood of** x to refer to a subset N of X of the form $N = B_r(x)$ for some positive real number $r > 0$.

Example 45. For example, if $X = \mathbb{R}$ with the standard metric, then $B_r(x) = (x - r, x + r)$, where $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ denotes the open interval of real numbers between a and b.

Definition 46. Let E be a subset of a metric space X. We say that E is **bounded** if there is a point x of X and a real number $r > 0$ such that $E \subset B_r(x)$.

Example 47. For example, any open ball $B_r(x)$ is itself bounded by definition.

Example 48. As a non-example, the subset $\mathbb{N}$ of $\mathbb{R}$ is not bounded.

Definition 49. Let (X, d) be a metric space and let E be a subset of X. A point $x \in E$ is called an **interior point** of E if there is a real number $\delta > 0$ such that $B_\delta(x) \subset E$. That is, an interior point is one which admits an open ball around it that is entirely contained within E. We say that E is **open** if each point of E is an interior point of E.

Example 50. Any open ball $B_r(x)$ is itself open in X as a consequence of the triangle inequality. Indeed let y be a point of $B_r(x)$. If d_0 denotes the distance $d_0 = d(x, y)$, then set $\delta = r - d_0$. The proof will be done if we can show that $B_\delta(y) \subset B_r(x)$. To this end, let z be a point of $B_\delta(y)$. This means that $d(z, y) < \delta$. The triangle inequality then implies that

$$d(z, x) \leqslant d(z, y) + d(y, x) < \delta + d_0 = r.$$

This completes the proof of the claim.

Proposition 51. *Let A be an index set, and let $\{E_\alpha\}_{\alpha \in A}$ be a collection of open subsets of a metric space X.*

(a) The union $\cup_\alpha E_\alpha$ is open in X.

(b) If the index set A is finite, then the intersection $\cap_\alpha E_\alpha$ is open in X.

Proof. For (a), let x be a point of the union $\cup_\alpha E_\alpha$. There is an $\alpha_0 \in A$ such that $x \in E_{\alpha_0}$. Because E_{α_0} is open, the point x is an interior point of E_{α_0}, so there is a $\delta_0 > 0$ such that $B_{\delta_0}(x) \subset E_{\alpha_0}$. By definition of the union, we have $B_{\delta_0}(x) \subset (\cup_\alpha E_\alpha)$. We conclude that x is an interior point of the union $\cup_\alpha E_\alpha$.

For (b), let x be a point in the intersection $\cap_\alpha E_\alpha$. Because each E_α is open, for each $\alpha \in A$, there is a $\delta_\alpha > 0$ such that $B_{\delta_\alpha} \subset E_\alpha$. Let δ denote the minimum $\delta = \inf\{\delta_\alpha : \alpha \in A\}$, which is positive since the index set A is finite. The proof will be complete if we can show that $B_\delta(x) \subset (\cap_\alpha E_\alpha)$. To this end, let y be a point of $B_\delta(x)$. This means that $d(y, x) < \delta$. By definition of δ, we have that $d(y, x) < \delta_\alpha$ for each $\alpha \in A$. It follows that y belongs to each $B_{\delta_\alpha}(x)$, and hence y belongs to each E_α. This completes the proof. $\square$

Definition 52. Let E be a subset of a metric space X. A point $x \in X$ is called a **limit point** of E if for each $\delta > 0$, the intersection $B_\delta(x) \cap E$ is contains a point other than x. That is, a point is a limit point if every ball around it contains a point of E. We say that E is **closed** if every limit point of E belongs to E.

Example 53. For a point $x \in X$ and a real number $r > 0$, let E denote the subset

$$E = \{y \in X : d(y, x) \leqslant r\}.$$

Then we claim that E is closed in X. Indeed, let y be a limit point of E. Let $\delta > 0$ be arbitrary. The intersection $B_\delta(y) \cap E$ contains a point z other than y. Since z

belongs to E, we have $d(z, x) \leqslant r$, and since z belongs to $B_\delta(y)$, we have $d(z, y) < \delta$. The triangle inequality implies that

$$d(y, x) \leqslant d(y, z) + d(z, x) < \delta + r.$$

Since $\delta > 0$ is arbitrary, we conclude that $d(y, x) \leqslant r$. This shows that y belongs to E.

The following characterization of closed sests is useful.

Lemma 54. *A subset E of a metric space X is closed if and only if the complement $E^c = X \setminus E$ is open.*

Proof. Suppose E is closed. Let x be a point of the complement E^c. Since E is closed, x is not a limit point of E. Thus there is a $\delta > 0$ such that the ball $B_\delta(x)$ does not contain any points of E, that is, there is a $\delta > 0$ such that $B_\delta(x) \subset E^c$. This shows that x is an interior point of E^c.

Suppose E^c is open. Let x be a limit point of E. If x belongs to E^c, then since E^c is open, there is a $\delta > 0$ such that $B_\delta(x) \subset E^c$, which implies that $B_\delta(x) \cap E = \varnothing$, a contradiction to the assumption that x is a limit point of E. We conclude that x must belong to E. $\square$

The De Morgan's laws (Problem 1 of Assignment 1) together with Proposition 51 imply the following.

Proposition 55. *Let A be an index set, and let $\{E_\alpha\}_{\alpha \in A}$ be a collection of closed subsets of a metric space X.*

(a) *The intersection $\cap_\alpha E_\alpha$ is closed in X.*

(b) *If the index set A is finite, then the union $\cup_\alpha E_\alpha$ is closed in X.*

Proof. Exercise. $\square$

It is important to note that the notions of open and closed are not exclusive. In particular, some sets are both open and closed, and some sets are neither.

Example 56. The empty set $\varnothing$ set is by definition both open and closed. In addition, the whole space X is both open and closed.

Example 57. Let $E = [a, b)$ denote the half-open interval $[a, b) = \{x \in \mathbb{R} : a \leqslant x < b\}$. Then E is neither open nor closed. Indeed, the point b is a limit point of E which does not belong to E, so E is not closed. On the other hand, the point a is not an interior point of E, so E is not open.

Definition 58. Let E be a subset of a metric space X. Suppose $\bar{E}$ is a subset of X satisfying the following properties.

(i) $\bar{E}$ contains E

(ii) $\bar{E}$ is closed

(iii) If C is any closed set containing E, then $\bar{E} \subset C$.

Then $\bar{E}$ is called the **closure of** E. By construction, we see that $\bar{E}$ is the *smallest* closed subset containing E.

Proposition 59. *Let E be a subset of a metric space X. Then the closure $\bar{E}$ exists and is given by*

$$\bar{E} = E \cup E'$$

where E' denotes the set of all limit points of E in X. In addition, the set E is closed if and only if $\bar{E} = E$.

Proof. Let us write F for the union $F = E \cup E'$. Note that by definition F contains E.

We show that F is closed by showing that its complement $F^c = E^c \cap (E')^c$ is open. Let x be a point of the intersection $E^c \cap (E')^c$. Then since x is not in E', there is a neighborhood $B_\delta(x)$ of x which does not intersect E. This means that $B_\delta(x) \subset E^c$. We will have shown that x is an interior point of F if we can show that also $B_\delta(x) \subset (E')^c$. Suppose, toward a contradiction, that y is a point of $B_\delta(x)$ belonging to E'. This means that y is a limit point of E. Hence for each $\epsilon > 0$, the intersection $B_\epsilon(y) \cap E$ is nonempty, and contains a point z. In particular, choose ϵ smaller than $\delta - d(x, y)$. Then we see that

$$d(z, x) \leqslant d(z, y) + d(y, x) < \epsilon + d(x, y) < \delta.$$

This means that the ball $B_\delta(x)$ contains a point z of E, which is a contradiction.

Lastly, we show that F satisfies property (iii). Let C be any closed subset of X containing E. The inclusion $E \subset C$ implies that $E' \subset C'$. (Indeed if x is a limit point of E, then each neighborhood around x intersects E in some point other than x and hence also intersects C in some point other than x.) Also C is closed so we know the inclusion $C' \subset C$ is true. It follows that the inclusion $E' \subset C' \subset C$ is true. By assumption C contains E and so C contains F.

For the final statement in the proposition, if $\bar{E} = E$, then we have already shown that $\bar{E}$ is closed and so E is as well. On the other hand, if E is closed, then by definition we have $E' \subset E$, and hence $\bar{E} = E$. $\qquad\square$

Example 60. As an example, if $E = [a, b) \subset \mathbb{R}$, then $\bar{E} = [a, b]$.

Example 61. As another example, if $E = \mathbb{Q} \subset \mathbb{R}$, then $\bar{E} = \mathbb{R}$.

Definition 62. Let E be a subset of a metric space X. Say that E is **dense** in X if $\bar{E} = X$.

Example 63. For example, the set $\mathbb{Q}$ of rational numbers is dense in $\mathbb{R}$.

Remark 64. Let Y be a subset of a metric space X. Then Y is itself a metric space upon restricting the metric to Y. If E is a subset of Y, then it makes sense to ask whether E is open relative to Y. However, it also makes sense to ask whether E is open relative to the ambient space X. Therefore, in this situation, subsets of Y can be open in two ways. It is therefore important to specify in which way such a set is open. If $E \subset Y$ is open when regarded as a subset of Y, we will say that E is open relative to Y (and similarly for X).

Example 65. Let $Y = \{p\}$ be a single point of a metric space X, and let $E = Y = \{p\}$. Then E is open relative to Y, but closed relative to X according to Warm-up #3.

Proposition 66. *Let Y be a subset of a metric space X, and let E be a subset of Y. Then E is open relative to Y if and only if there is an open subset G of X such that $E = Y \cap G$.*

Proof. Exercise. (Or see Rudin.) $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

2.3 Compact sets

Definition 67. Let E be a subset of a metric space X. By an open cover of E we mean a collection $\{G_\alpha\}_{\alpha \in A}$ of open subsets of X whose union contains E.

Example 68. For example, the sets $(-\infty, 1)$ and $(-1, \infty)$ form an open cover of $\mathbb{R}$.

As an example involving an infinite number of subsets, for each $n \in \mathbb{N}$, let G_n be the open subset of $\mathbb{R}$ defined by $(-\infty, n)$. Then the collection $\{G_n\}_{n \in \mathbb{N}}$ forms an open cover of $\mathbb{R}$.

Definition 69. Let E be a subset of a metric space X. We say that E is **compact** if every open cover of E admits a finite subcover. More precisely, the requirement is that if $\{G_\alpha\}_{\alpha \in A}$ is an open cover of E, then there are finitely many indices $\alpha_1, \ldots, \alpha_n$ such that
$$E \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}.$$

Example 70. For example, the space $X = \mathbb{R}$ is not compact, for the following reason. Let G_n be the open cover defined earlier by $G_n = (-\infty, n)$. If $\{G_{n_1}, \ldots, G_{n_r}\}$ is any finite subcover, then the union
$$\bigcup_{k=1}^{r} G_{n_k}$$
is equal to $(-\infty, N)$, where N is the maximum of $n_1, \ldots, n_r$. It follows that no finite subcover of $\{G_n\}_{n \in \mathbb{N}}$ contains $\mathbb{R}$.

Example 71. Any *finite* subset of a metric space is compact by Warm-Up #4.

It turns out that the notion of compactness is crucial in Analysis, and we will soon prove the existence of a large class of infinite compact subsets of $\mathbb{R}^n$.

Lemma 72. *Let E be a subset of a metric space X. If E is compact, then E is closed.*

Proof. Suppose E is compact. Fix a point y in the complement of E. For each point x of E, let $\delta(x) = d(x, y)/2$. Then the collection $\{B_{\delta(x)}(x)\}_{x \in E}$ forms an open cover of E. Since E is compact, there are a finite number of points $x_1, \ldots, x_n \in E$ such that $\{B_{\delta(x_i)}(x_i)\}_{i=1}^n$ forms an open cover of E. If $\delta = \inf\{\delta_i : i = 1, \ldots, n\}$, then the ball $B_\delta(y)$ does not intersect E. It follows that y is an interior point of E^c. We conclude that E^c is open. $\square$

Lemma 73. *Let E be a subset of a metric space X. If X is compact and E is closed, then E is compact.*

Proof. Let $\{G_\alpha\}_{\alpha \in A}$ be an open cover of E. Since E is closed, the complement E^c is open. It follows that $\{G_\alpha\}_{\alpha \in A} \cup \{E^c\}$ is an open cover of X. Since X is compact, there are a finite number of indices $\alpha_1, \ldots, \alpha_k$ such that

$$X \subset (E^c \cup G_{\alpha_1} \cup \cdots \cup G_{\alpha_k}).$$

It follows that

$$E \subset (G_{\alpha_1} \cup \cdots \cup G_{\alpha_k}).$$

We are done. $\square$

Proposition 74. *Let $\{E_n\}_{n \geqslant 1}$ be a sequence of nonempty subsets of a metric space X, and let E be the intersection $E = \cap_n E_n$. Suppose each E_n is compact and the collection is nested in the sense that $E_n \supset E_{n+1}$. Then the intersection E is nonempty.*

Proof. Let $G_n = E_n^c$. Then each G_n is open by the previous proposition, and they are nested in the reverse way that $G_n \subset G_{n+1}$.

Assume that no point $x \in E_1$ belongs to each E_n. Then the collection $\{G_n\}_{n \geqslant 2}$ forms an open cover of E_1. Indeed, if it didn't, then there would be a point $x \in E_1$ in the complement

$$\left(\bigcup_{n \geqslant 2} G_n \right)^c = \bigcap_{n \geqslant 2} E_n$$

which is contradictory to our assumption.

Since E_1 is compact, it follows that there are a finite number of indices $i_1 < \cdots < i_k$ such that

$$E_1 \subset G_{i_1} \cup \cdots \cup G_{i_k}.$$

This says that

$$E_1 \subset (E_{i_1} \cap \cdots \cap E_{i_k})^c,$$

that is, that the intersection

$$E_1 \cap E_{i_1} \cap \cdots \cap E_{i_k} = E_{i_k}$$

is empty. This is a contradiction to the assumption that E_{i_k} is nonempty. $\square$

Our next goal is to show that any closed interval $[a, b]$ of $\mathbb{R}$ is compact. From the previous proposition, we see that the following lemma is necessary (but not sufficient).

Lemma 75. *Let $\{I_n\}_{n \geqslant 1}$ be a sequence of closed intervals of $\mathbb{R}$, and let I be the intersection $I = \cap_n I_n$. If the intervals are nested in the sense that $I_n \supset I_{n+1}$, then I is nonempty.*

Proof. Write $I_n = [a_n, b_n]$, and let E be the set of all a_n. Then E is nonempty and bounded from above by b_1. It follows that $\sup E$ exists, and let us write x for $\sup E$. If m and n are nonnegative integers, then the nesting property guarantees that

$$a_n \leqslant a_{n+m} \leqslant b_{n+m} \leqslant b_m.$$

This means that b_m bounds E from above, and so $x \leqslant b_m$ for each m. On the other hand, x is an upper bound for E and hence satisfies $a_m \leqslant x$ for each m. We conclude that x satisfies

$$a_m \leqslant x \leqslant b_m \qquad \text{for each } m$$

and hence x belongs to I_m for each m. This concludes the proof. $\qquad \square$

Definition 76. Let k be a positive integer. By a k-cell we mean a subset of $\mathbb{R}^k$ of the form

$$I_1 \times \cdots \times I_k$$

where each I_j is a closed interval of $\mathbb{R}$.

Corollary 77. *Let k be a positive integer. Let $\{I_n\}$ be a sequence of k-cells, and let I be the intersection $I = \cap_n I_n$. If the k-cells are nested in the sense that $I_n \supset I_{n+1}$, then I is nonempty.*

Proof. Write the k-cell I_n as

$$I_n = I_{n,1} \times \cdots \times I_{n,k}.$$

Apply the previous lemma to the sequence $\{I_{n,j}\}_{n \in \mathbb{N}}$ to obtain a real number x_j in the intersection $\cap_n I_{n,j}$. Conclude that $x = (x_1, \ldots, x_k)$ belongs to I. $\qquad \square$

Theorem 78. *Each k-cell is compact.*

Proof. Let I be a k-cell and write

$$I = [a_1, b_1] \times \cdots \times [a_k, b_k].$$

Let δ be the length of the diagonal of I given by

$$\delta = \left(\sum_{j=1}^{k} (b_j - a_j)^2 \right)^{1/2}.$$

Then if $x, y \in I$, we have $\|x - y\| \leqslant \delta$.

Suppose, toward a contradiction, that there is an open cover $\{G_\alpha\}_{\alpha \in A}$ of I which contains no finite subcover. If c_j denotes the midpoint of $[a_j, b_j]$ given by $c_j = (a_j + b_j)/2$, then the intervals $[a_j, c_j]$ and $[c_j, b_j]$ determine 2^k k-cells whose union is I. At least one of these k-cells, call it I_1, cannot be covered by any finite subcollection of $\{G_\alpha\}$. We then divide I_1 in a similar way, and repeat the process. As a result, we obtain a sequence $\{I_n\}_{n \in \mathbb{N}}$ of k-cells with the following properties.

(i) $I := I_0 \supset I_1 \supset I_2 \supset \cdots$

(ii) I_n is not covered by any finite subcollection of $\{G_\alpha\}$

(iii) if $x, y \in I_n$, then we have $\|x - y\| \leqslant 2^{-n}\delta$.

By (i) and the previous result, there is a point x which lies in each I_n. Because $\{G_\alpha\}$ is a cover, there is an α such that $x \in G_\alpha$. Since G_α is open, there is a $r > 0$ such that $B_r(x) \subset G_\alpha$. By the Archimedean property, we can choose a positive integer n so large that $2^{-n}\delta < r$. Then with this choice of n, (iii) implies that $I_n \subset G_\alpha$. But this contradicts (ii). $\qquad \square$

Theorem 79 (Heine-Borel Theorem). *Let E be a subset of the metric space $\mathbb{R}^k$. Then the following are equivalent.*

(i) E is closed and bounded.

(ii) E is compact.

Proof. Suppose E is closed and bounded. Since E is bounded, there is a k-cell I such that $E \subset I$. Then E is compact by Lemma 73.

Suppose that E is compact. Then Lemma 72 shows that E is closed, so it remains to show that E is bounded. The collection $\{B_1(x)\}_{x \in E}$ of open balls of radius 1 centered at points of E forms an open cover of E. Since E is compact, there are a finite number of points $x_1, \ldots, x_n \in E$ such that

$$E \subset (B_1(x_1) \cup \cdots \cup B_1(x_n)).$$

Let M be the maximum of $d(x_i, x_j)$ for $i, j \in \{1, \ldots, n\}$. If x, y are points of E, then each is contained in one of the balls, say $x \in B_1(x_i)$ and $y \in B_1(x_j)$, and then the triangle inequality says that

$$d(x, y) \leqslant d(x, x_i) + d(x_i, x_j) + d(x_j, y) \leqslant 1 + M + 1 = M + 2.$$

We conclude that E is bounded. $\qquad \square$

3 Sequences and series

3.1 Convergent sequences

Definition 80. A **sequence** in a metric space is a map $f : \mathbb{N} \setminus \{0\} \to X$. We often write $x_n = f(n)$ and say that $\{x_n\}$ is a sequence.

Definition 81. A sequence $\{x_n\}$ is said to **converge** if there is a point $x \in X$ satisfying the following: for each $\epsilon > 0$ there is an integer $N > 0$ such that if $n \geqslant N$, then $d(x_n, x) < \epsilon$. In this case, we say that $\{x_n\}$ converges to x or that x is the **limit** of the sequence $\{x_n\}$ and we write $x_n \to x$ or

$$\lim_{n \to \infty} x_n = x.$$

If $\{x_n\}$ does not converge, then it is said to **diverge**.

Example 82. Let $\{x_n\}$ be the sequence of $\mathbb{R}$ given by $x_n = 1/n$. Then $\{x_n\}$ is bounded, and we claim that $\{x_n\}$ converges to 0. Indeed, let $\epsilon > 0$ be given. The Archimedian property says that there is a positive integer N such that $N > 1/\epsilon$. For this choice of N, if $n \geqslant N$, then we find that

$$d(x_n, 0) = \frac{1}{n} < \frac{1}{N} < \epsilon.$$

Example 83. Let $\{x_n\}$ be the sequence of $\mathbb{R}$ given by $x_n = n$. Then $\{x_n\}$ is not bounded and it is routine to show that x_n does not converge, which we leave as an exercise.

Example 84. Let $\{x_n\}$ be the sequence of $\mathbb{R}$ given by $x_n = (-1)^n$. Then $\{x_n\}$ is bounded, but still diverges. Indeed, let x be any point of $\mathbb{R}$. There are two cases for x.

- Suppose $x \geqslant 0$. Given any $N > 0$, let n be an *odd* number larger than N, and note that $x_n = -1$ so that $d(x_n, x) \geqslant 1 > \epsilon$.

- Suppose $x \leqslant 0$. Given any $N > 0$, let n be an *even* number larger than N, and note that $x_n = 1$ so that $d(x_n, x) \geqslant 1 > \epsilon$.

Proposition 85. *Let $\{x_n\}$ be a sequence of a metric space X.*

(i) *If x_n converges to x and converges to x', then $x = x'$.*

(ii) *If x_n converges, then $\{x_n\}$ is bounded.*

Proof. For (i), let $\epsilon > 0$ be arbitrary. There is an $N > 0$ such that if $n \geqslant N$, then $d(x_n, x) < \epsilon/2$. There is an $N' > 0$ such that if $n \geqslant N'$, then $d(x_n, x') < \epsilon/2$. The triangle inequality implies that if n is bigger than both N and N', then

$$d(x, x') \leqslant d(x, x_n) + d(x_n, x') < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we conclude that $d(x, x') = 0$, and hence $x = x'$.

For (ii), suppose that $\{x_n\}$ converges to x. There is an $N > 0$ such that if $n \geqslant N$, then $d(x_n, x) < 1$. Set

$$r = \max\{1, d(x_0, x), d(x_1, x), \ldots, d(x_N, x)\}.$$

Then we have that
$$d(x_n, x) \leqslant r \qquad \text{for each } n > N.$$

This shows that $\{x_n\}$ is bounded. $\qquad\square$

Theorem 86. *Let x_n and y_n be sequences in $\mathbb{R}$, and suppose that x_n converges to x and y_n converges to y.*

(i) $\lim_{n \to \infty}(x_n + y_n) = x + y$

(ii) *If c is any real number, then* $\lim_{n \to \infty} c x_n = cx$.

(iii) $\lim_{n \to \infty} x_n y_n = xy$

(iv) *If $x \neq 0$, then the sequence x_n is nonzero for n sufficiently large, and also* $\lim_{n \to \infty} \frac{1}{x_n} = \frac{1}{x}$.

Proof. For (i), let $\epsilon > 0$ be arbitrary. Let N_x be such that if $n \geqslant N_x$ then $d(x_n, x) < \epsilon/2$. Let N_y be such that if $n \geqslant N_y$ then $d(y_n, y) < \epsilon/2$. Set $N = \max\{N_x, N_y\}$. Then if $n \geqslant N$, we have

$$d(x_n + y_n, x + y) = |(x_n - x) + (y_n - y)| \leqslant |x_n - x| + |y_n - y| < \epsilon/2 + \epsilon/2 = \epsilon.$$

For (ii), let $\epsilon > 0$ be arbitrary. If $c = 0$, then $c x_n \to cx$ trivially. Otherwise, let N be large enough such that if $n \geqslant N$, then $|x_n - x| < \epsilon/|c|$. Then if $n \geqslant N$, we have

$$|c x_n - cx| = |c||x_n - x| < \epsilon.$$

For (iii), let $\epsilon > 0$ be given. There are integers N_x and N_y such that

$$n > N_x \text{ implies } |x_n - x| < \sqrt{\epsilon}$$
$$n > N_y \text{ implies } |y_n - y| < \sqrt{\epsilon}.$$

If we set $N = \max\{N_1, N_2\}$, then whenever $n \geqslant N$, we have

$$|(x_n - x)(y_n - y)| < \epsilon,$$

24

which implies that
$$\lim_{n \to \infty} (x_n - x)(y_n - y) = 0.$$

Now use the identity
$$x_n y_n - xy = (x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x).$$

and parts (i) and (ii) to conclude that each term on the right tends to 0 and hence
$$\lim_{n \to \infty} (x_n y_n - xy) = 0.$$

For (iv), since $x \neq 0$, there is an $N_0 \in \mathbb{N}$ such that if $n \geqslant N_0$, we have $|x_n - x| < \frac{1}{2}|x|$ so that in particular

$$|x_n| > \frac{1}{2}|x| \qquad \text{for } n > N_0 \qquad\qquad (1)$$

and hence $x_n \neq 0$ for $n \geqslant N_0$. Given $\epsilon > 0$, there is an $N \geqslant N_0$ such that if $n \geqslant N$, then
$$|x_n - x| < \frac{1}{2}|x|^2 \epsilon.$$

We then find that for $n \geqslant N$, we have
$$
\begin{aligned}
\left| \frac{1}{x_n} - \frac{1}{x} \right| &= \frac{|x - x_n|}{|x x_n|} \\
&< \frac{2|x - x_n|}{|x|^2} \qquad\qquad \text{by (1)} \\
&< \epsilon.
\end{aligned}
$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.2 Subsequences

Definition 87. Let x_n be a sequence in a metric space X. By a **subsequence of** x_n we mean we are given the data of a sequence of positive integers $n_1 < n_2 < \cdots$ and the subsequence so determined is the sequence x_{n_k} now indexed by k.

Lemma 88. *Let E be a subset of a metric space X. If x is a limit point of E and $\delta > 0$, then the intersection $B_\delta(x) \cap E$ is infinite.*

Proof. Suppose the intersection is not infinite but equal to $\{x, x_1, \ldots, x_n\}$. Let r be the number
$$r = \min_{1 \leqslant i \leqslant n} d(x, x_i),$$

which is positive because the indexing set is finite. Then the ball $B_r(x)$ contains no points of E other than x itself, which is a contradiction to the assumption that x is a limit point of E. $\qquad\qquad\square$

Proposition 89. *Let p_n be a sequence in a metric space X. If X is compact, then p_n has a convergent subsequence.*

Proof. Let $E = \{p_n : n \geqslant 1\}$ be the subset of X determined by the sequence. Either E is finite or infinite.

Suppose E is finite. Then there is a point $p \in E$ and an infinite number of positive integers $n_1 < n_2 < \cdots$ such that $x_{n_k} = p$. It follows that the subsequence so determined converges.

Suppose E is infinite. We claim that E has a limit point in X. Indeed suppose not. Then for each $x \in X$, there is a $\delta_x > 0$ such that $B_{\delta_x}(x)$ contains no points of E other than possibly x. Since E is infinite, the cover $\{B_{\delta_x}(x)\}_{x \in E}$ of E admits no finite subcover. Since E is a subset of X, the cover $\{B_{\delta_x}(x)\}_{x \in X}$ of X similarly admits no finite subcover. This is a contradiction to the assumption that X is compact.

Let x be a limit point of E. There is a point x_1 in $B_1(x) \cap E$ which is not equal to X. Because x_1 belongs to E, we can write $x_1 = p_{n_1}$ for a positive integer n_1. By the previous lemma, the intersection $B_{1/2}(x) \cap E$ is infinite, so there is an $n_2 > n_1$ such that $d(p_{n_2}, x) < 1/2$. Inductively we obtain a sequence $n_1 < n_2 < \cdots$ such that $d(p_{n_k}, x) < 1/k$. It follows that the subsequence we determined converges to x. $\qquad\square$

Corollary 90. *If p_n is a bounded sequence in $\mathbb{R}^k$, then p_n has a convergent subsequence.*

Proof. The subset $E = \{p_n\}$ is bounded in $\mathbb{R}^k$ and hence is a subset of some k-cell, which is compact. The result then follows from the previous result. $\qquad\square$

3.3 Cauchy sequences

Definition 91. Let x_n be a sequence in a metric space X. Say that x_n is a **Cauchy** sequence if the following is true. For each $\epsilon > 0$, there is a positive integer N such that if $m \geqslant N$ and $n \geqslant N$, then $d(x_n, x_m) < \epsilon$.

Heuristically, a Cauchy sequence is one in which the terms in the sequence become close to *one another*. This differs slightly from the notion of a convergent sequence where the terms become close to a specific point in the metric space.

Theorem 92. *Let X be a metric space.*

(i) *Every convergent sequence is a Cauchy sequence.*

(ii) *If X is compact and if p_n is a Cauchy sequence of X, then p_n converges to a point of X.*

Proof. For (i), suppose p_n converges to p. Let $\epsilon > 0$ be given. There is an $N > 0$ such that if $n \geqslant N$, then $d(p_n, p) < \epsilon/2$. Then for $m \geqslant N$ and $n \geqslant N$, the triangle inequality implies that

$$d(p_n, p_m) \leqslant d(p_n, p) + d(p, p_m) < \epsilon/2 + \epsilon/2 = \epsilon.$$

For (ii), let p_n be a Cauchy sequence in the compact space X. For each positive integer N, let E_N be the subset of X consisting of the points $p_N, p_{N+1}, p_{N+2}, \ldots$. The closure $\bar{E}_N$ is a closed subset of a comapact space, and hence $\bar{E}_N$ is compact itself. Also we have the nesting property that $E_N \supset E_{N+1}$, which implies also that $\bar{E}_N \supset \bar{E}_{N+1}$. Proposition 74 implies that the intersection $\cap_n \bar{E}_n$ is nonempty. Let p be a point in this intersection.

Let $\epsilon > 0$ be given. Because p_n is Cauchy, there is an $N > 0$ such that if $m, n \geqslant N$, then $d(p_n, p_m) < \epsilon/2$. Because p belongs to $\bar{E}_N$, there is an $n_0 \geqslant N$ such that $d(p, p_{n_0}) < \epsilon/2$. Then if $m \geqslant N$, we have

$$d(p, p_m) \leqslant d(p, p_{n_0}) + d(p_{n_0}, p_m) < \epsilon/2 + \epsilon/2 = \epsilon.$$

We conclude that p_m converges to p. $\qquad\square$

Corollary 93. *Every Cauchy sequence in $\mathbb{R}^k$ converges.*

Proof. Let x_n be a Cauchy sequence of $\mathbb{R}^k$. Let $E_N \subset \mathbb{R}^k$ be the set of points $x_N, x_{N+1}, x_{N+2}, \ldots$. Because x_n is Cauchy, there is an $N_0 > 0$ such that if $p, q \in E_{N_0}$, then $d(p, q) < 1$, which implies E_{N_0} is bounded. The subset E of $\mathbb{R}^k$ determined by x_n is then also bounded because it consists of E_{N_0} together with the finite set of points $\{x_1, \ldots, x_{N_0-1}\}$. The closure $\bar{E}$ is then closed and bounded, and hence compact by Heine-Borel. Then the result follows from the previous result part (ii). $\qquad\square$

Definition 94. A metric space (X, d) in which each Cauchy sequence converges is called **complete**.

Example 95. The previous corollary shows that $\mathbb{R}^n$ is complete. The preceding result also shows than any compact metric space is complete.

Example 96. The set of rational numbers $\mathbb{Q}$ together with the usual metric is not complete, because there exist Cauchy sequences in $\mathbb{Q}$ which do not converge to any point of $\mathbb{Q}$. (For example, consider a sequence of rational numbers converging to an irrational number.)

3.4 Sequences of real numbers

Definition 97. A sequence s_n of real numbers is said to be

(a) **monotonically increasing** if $s_n \leqslant s_{n+1}$ for each $n = 1, 2, \ldots$.

(b) **monotonically decreasing** if $s_n \geqslant s_{n+1}$ for each $n = 1, 2, \ldots$.

Proposition 98. *Suppose s_n is monotonic (either decreasing or increasing). Then s_n converges if and only if it is bounded.*

Proof. Without loss of generality, we may assume that s_n is increasing so that $s_n \leqslant s_{n+1}$. Let E be the subset of $\mathbb{R}$ determined by the sequence s_n.

Suppose E is bounded. Let $s = \sup E$. Since s is an upper bound, we have $s_n \leqslant s$ for each $n = 1, 2, \ldots$. Given $\epsilon > 0$, Problem 4 of Assignment 1 says there is an integer N such that

$$s - \epsilon < s_N \leqslant s$$

and hence for each $n \geqslant N$, the increasing nature of the sequence implies that

$$s - \epsilon < s_n \leqslant s.$$

Since $\epsilon > 0$ is arbtirary, we find that s_n converges to s.

Conversely, suppose that s_n converges. Then E is bounded by Proposition 85 (ii). $\qquad\square$

Definition 99. Let s_n be a sequence of real numbers that is bounded. For a positive integer N, let E_N be the set $E_N = \{s_N, s_{N+1}, s_{N+2}, \ldots\}$. Then for each $N > 0$, the set E_N is bounded from above and hence admits a supremum. In addition, because $E_N \supset E_{N+1}$, the sequence of these suprema is monotonically decreasing and bounded from below, so we may define

$$\limsup_{n \to \infty} s_n = \lim_{N \to \infty} \sup E_N.$$

The previous proposition in fact shows that

$$\limsup_{n \to \infty} s_n = \inf_{N > 0} \sup_{n \geqslant N} s_n.$$

In a similar way, we can define

$$\liminf_{n \to \infty} s_n = \lim_{N \to \infty} \inf E_N = \sup_{N > 0} \inf_{n \geqslant N} s_n.$$

More generally, if s_n is any sequence of real numbers, not necessarily bounded, then define

$$\limsup_{n \to \infty} s_n = \lim_{N \to \infty} \sup E_N$$

as above, provided the limit exists and is finite. The existence of the limit is equivalent to the statement that the sequence s_n is bounded from above. In the case that the limit does not exist, that is, when the sequence is unbounded from above, we set

$$\limsup_{n \to \infty} s_n = \infty.$$

We may similarly define

$$\liminf_{n \to \infty} s_n = \lim_{N \to \infty} \inf E_N$$

provided this limit exists, and otherwise set

$$\liminf_{n \to \infty} s_n = -\infty$$

in which case the sequence is unbounded from below.

Corollary 100. *It is apparent from the definitions that if s_n is any sequence of real numbers then*

$$\liminf_{n\to\infty} s_n \leqslant \limsup_{n\to\infty} s_n$$

(provided we understand $\pm\infty$ as satisfying the inequalities in the obvious ways).

Example 101. Let s_n be the sequence $s_n = (-1)^n$. Then for each $N > 0$, we have $\sup E_N = 1$, and so

$$\limsup_{n\to\infty} s_n = \lim_{N\to\infty} \sup E_N = 1.$$

We also find that

$$\liminf_{n\to\infty} s_n = -1.$$

Example 102. Let s_n be the sequence

$$s_n = \frac{(-1)^n}{1 + \frac{1}{n}}.$$

Then we have that

$$s_{2n} = \frac{1}{1 + \frac{1}{2n}} = \frac{2n}{2n + 1}$$

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

The sequence $t_n := s_{2n}$ is monotonically decreasing with limit 1, and the sequence $r_n := s_{2n+1}$ is monotonically increasing with limit -1. One can check that

$$\limsup_{n\to\infty} s_n = 1$$

and

$$\liminf_{n\to\infty} s_n = -1.$$

Proposition 103. *Let s_n be a sequence of real numbers. Then s_n converges to s if and only if*

$$\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = s.$$

Proof. Exercise. □

Proposition 104. *Let s_n and t_n be two sequences of real numbers. If there is a positive integer N such that $s_n \leqslant t_n$ for each $n \geqslant N$, then*

$$\limsup_{n\to\infty} s_n \leqslant \limsup_{n\to\infty} t_n$$

and

$$\liminf_{n\to\infty} s_n \leqslant \liminf_{n\to\infty} t_n.$$

Proof. Exercise. □

29

3.5 Examples of sequences

Example 105. For $p > 0$, let $s_n = 1/n^p$. Then we claim that $s_n \to 0$. Indeed, let $\epsilon > 0$ be given. By the Archimedean property, there is a positive integer N such that $N > (1/\epsilon)^{1/p}$. Then if $n > N > (1/\epsilon)^{1/p}$, we have that $n^p > 1/\epsilon$, and so

$$s_n = \frac{1}{n^p} < \epsilon.$$

Example 106. For $p > 0$, let $s_n = \sqrt[n]{p}$. Then we claim that $s_n \to 1$. There are three cases for p.

- If $p = 1$, then $s_n = 1$, so the result is trivial.

- Suppose $p > 1$. We will prove that $x_n = \sqrt[n]{p} - 1$ converges to 0. Note that $x_n > 0$. The binomial theorem says that

$$p = (1 + x_n)^n = \sum_{j=0}^{n} \binom{n}{j} x_n^j > \overbrace{1}^{j=0} + \overbrace{n x_n}^{j=1},$$

  because $x_n > 0$. We conclude that

$$x_n < \frac{p-1}{n},$$

  and hence $x_n \to 0$.

- Suppose $p < 1$. Then upon setting

$$y_n = \frac{1}{s_n} - 1,$$

  we see that $y_n > 0$. The binomial theorem says in this case that

$$\frac{1}{p} = (1 + y_n)^n > 1 + n y_n$$

  so that

$$y_n < \frac{\frac{1}{p} - 1}{n}$$

  and we conclude $y_n \to 0$.

Example 107. Let $s_n = \sqrt[n]{n}$. Then we claim that $s_n \to 1$. Indeed let $x_n = s_n - 1$. Note that each x_n satisfies $x_n \geqslant 0$ and the binomial theorem says that

$$n = (1 + x_n)^n = \sum_{j=0}^{n} \binom{n}{j} x_n^j \geqslant \overbrace{\frac{n(n-1)}{2}}^{j=2} x_n^2.$$

30

Algebraic manipulations show that

$$0 \leqslant x_n \leqslant \sqrt{\frac{2}{n-1}}.$$

We conclude that $x_n \to 0$.

Example 108. For $p > 0$ and $\alpha \in \mathbb{R}$, let

$$s_n = \frac{n^\alpha}{(1+p)^n}.$$

Then we claim that $s_n \to 0$. Indeed let k be a positive integer such that $k > \alpha$. Then for n satisfying $n > 2k$, we have

$$
\begin{aligned}
(1+p)^n &\geqslant \binom{n}{k} p^k & &\text{binomial theorem} \\
&= \frac{n(n-1)\cdots(n-k+1)}{k!} p^k & &\text{definition of } \binom{n}{k} \\
&\geqslant \left(\frac{n}{2}\right)^k \frac{p^k}{k!} & &n - j > \frac{n}{2} \text{ for } j < k.
\end{aligned}
$$

We then find that

$$
\begin{aligned}
0 &\leqslant \frac{n^\alpha}{(1+p)^n} \\
&\leqslant n^\alpha \frac{2^k k!}{n^k p^k} & &\text{previous inequality} \\
&= \frac{2^k k!}{p^k} n^{\alpha - k}.
\end{aligned}
$$

Because $\alpha - k < 0$, we conclude that the right-hand side converges to 0 by a previous example.

Example 109. As a special case of the previous example, suppose x is a real number satisfying $0 < x < 1$, and set $s_n = x^n$. Then we find that $s_n \to 0$ because we may take $\alpha = 0$ in the previous example (and $p > 0$ satisfying $1 + p = x^{-1}$ so that in addition $(1+p)^{-n} = x^n$).

3.6   Series

In this section, we will consider only sequences and series that are real-valued, unless otherwise explicitly stated. Rudin considers complex-valued series, and indeed all of our proofs will carry over to that case, so we content ourselves only with the real case.

Definition 110. For a sequence a_k of real numbers, the n**th partial sum** is the number

$$s_n = \sum_{k=1}^{n} a_k = a_1 + \cdots + a_n.$$

The partial sums form a sequence s_n in their own right. If the sequence s_n converges to a number s, then we write

$$\sum_{k=1}^{\infty} a_k = s,$$

and we call the symbol on the left-hand side a **series**.

As a consequence of the Cauchy criterion for convergence of sequences of real numbers, we have the following result.

Corollary 111. *The series $\sum_k a_k$ converges if and only if for each $\epsilon > 0$ there is an integer N such that*

$$\left| \sum_{k=n}^{m} a_k \right| < \epsilon \qquad \text{whenever } m \geqslant n > N.$$

In particular, taking $m = n$ in the previous result, we find that

$$|a_n| < \epsilon \qquad \text{whenever } n > N.$$

Corollary 112. *If the series $\sum_k a_k$ converges, then the sequence a_k converges to 0.*

The following result asserts that in certain situations if the terms of two series are "comparable," then the convergence (resp. divergence) of one implies the convergence (resp. divergence) of the other.

Theorem 113 (Comparison test)**.** *Let a_k, b_k be sequences.*

(i) *If there is a positive integer N such that $|a_k| \leqslant b_k$ for each $k \geqslant N$ and if $\sum_k b_k$ converges, then $\sum_k a_k$ converges.*

(ii) *If there is a positive integer N such that $a_k \geqslant b_k \geqslant 0$ for each $k \geqslant N$ and if $\sum_k b_k$ diverges, then $\sum_k a_k$ diverges.*

Proof. For (i), let $\epsilon > 0$ be given. By the Cauchy criterion, there is an $M > 0$ such that whenever $m \geqslant n > M$, we have

$$\sum_{k=n}^{m} b_k < \epsilon$$

and then the triangle inequality implies that

$$\left| \sum_{k=n}^{m} a_k \right| \leqslant \sum_{k=n}^{m} |a_k| \leqslant \sum_{k=n}^{m} b_k < \epsilon.$$

The result follows by the Cauchy criterion.

Note that (ii) follows from (i) because if $\sum_k a_k$ converges, then so must $\sum_k b_k$. $\qquad \square$

3.7 Examples of series

Example 114 (Geometric series). Let x be a real number. If $0 \leqslant x < 1$, then we have

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

(It is important to note here that the series includes the $n = 0$ term.) If $x \geqslant 1$, then the series diverges.

Indeed, let us prove the claims of the previous paragraph. Suppose $x \neq 1$. Then it is straightforward to see that in the multiplication

$$(1-x)(1+x+x^2+\cdots+x^n)$$

all of the cross terms cancel, so we are left with

$$(1-x)\sum_{k=0}^{n} x^k = 1 - x^{n+1},$$

that is,

$$\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x} - \frac{x^{n+1}}{1-x}.$$

Taking $n \to \infty$, we deduce the claims of the previous paragraph, except for the case $x = 1$. In this case, the series is

$$1 + 1 + \cdots$$

which evidently diverges.

The following result is useful and it says that a series of nonnegative decreasing terms converges if and only if a corresponding series constructed from a rather "thin" subsequence of the terms converges. We first require a small lemma.

Lemma 115. *Let a_k be a sequence of nonnegative terms. Then $\sum_k a_k$ converges if and only if the sequence of partial sums is bounded.*

Proof. The sequence of partial sums is monotonically increasing, so the result follows from Proposition 98. $\qquad\square$

Theorem 116. *Suppose a_k is a sequence satisfying $a_1 \geqslant a_2 \geqslant a_3 \geqslant \cdots \geqslant 0$. Then the series $\sum_k a_k$ converges if and only if the series*

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$$

converges.

Proof. By the lemma, it suffices to consider the boundedness of the partial sums, which we will denote by

$$s_n = a_1 + a_2 + \cdots + a_n$$
$$t_m = a_1 + 2a_2 + \cdots + 2^m a_{2^m}.$$

If $n < 2^m$, then we find that

$$
\begin{aligned}
s_n &= a_1 + \cdots + a_n && \\
&\leqslant a_1 + \cdots + a_{2^m} && n < 2^m, 0 \leqslant a_k \\
&\leqslant a_1 + (a_2 + a_3) + \cdots + (a_{2^m} + \cdots + a_{2^{m+1}-1}) && 0 \leqslant a_k \\
&\leqslant a_1 + 2a_2 + \cdots 2^m a_{2^m} && a_{k+1} \leqslant a_k \\
&= t_m
\end{aligned}
$$

so that

$$s_n \leqslant t_m \qquad \text{for } n \leqslant 2^m.$$

On the other hand, if $n > 2^m$, we have

$$
\begin{aligned}
s_n &= a_1 + \cdots + a_n && \\
&\geqslant a_1 + \cdots + a_{2^m} && n > 2^m, a_k \geqslant 0 \\
&= a_1 + \overbrace{a_2}^{j=1} + \overbrace{(a_3 + a_4)}^{j=2} + \cdots + \overbrace{(a_{2^{m-1}+1} + \cdots + a_{2^m})}^{j=m} && \\
&\geqslant \frac{1}{2}a_1 + a_2 + 2a_4 + \cdots + 2^{m-1}a_{2^m} && a_k \geqslant a_{k+1} \\
&= \frac{1}{2}t_m
\end{aligned}
$$

so that

$$2s_n \geqslant t_m \qquad \text{for } n > 2^m.$$

Suppose the sequence s_n is bounded. This means that there is a real number $A \geqslant 0$ such that $0 \leqslant s_n \leqslant A$ for each $n = 1, 2, \ldots$. If m is any positive integer, then choose n large enough so that $n > 2^m$, and then the previous paragraph implies that

$$t_m \leqslant 2s_n \leqslant 2A.$$

Whence the sequence t_m is bounded.

On the other hand, suppose that the sequence t_m is bounded. This means there is a real number $B \geqslant 0$ such that $0 \leqslant t_m \leqslant B$ for each $m = 1, 2, \ldots$. If n is any positive intger, then let m be a positive integer such that $2^m > n$, and then the previous paragraph implies that

$$s_n \leqslant t_m \leqslant B.$$

Whence the sequence s_n is bounded. $\qquad\square$

34

Example 117 (*p*-series)**.** The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$ and diverges if $p \leqslant 1$. One case for p is that $p \leqslant 0$, in which case the terms $(1/n)^p$ do not converge to zero, and the series therefore diverges. For the case $p > 0$, we can use the previous result. Indeed, consider the corresponding series

$$\sum_{k=0}^{\infty} 2^k \frac{1}{2^{kp}} = \sum_{k=0}^{\infty} 2^{(1-p)k}.$$

This is a geometric series which converges if and only if 2^{1-p} satisfies

$$0 \leqslant 2^{1-p} < 1,$$

and this condition is equivalent to

$$1 - p < 0.$$

The comparison test then implies the desired result.

Example 118. The series

$$\sum_{k=0}^{\infty} \frac{1}{k!}$$

converges. Indeed, note that the sequence of partial sums satisfy

$$s_n = 1 + 1 + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \cdots + \frac{1}{1 \cdot 2 \cdots n}$$
$$< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} < 3,$$

and are hence bounded (and monotonic).

Definition 119. We let e denote the real number given by the limit of the previous series

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}.$$

Nore that if x is any real number, then the number e^x is positive. If y is any positive real number, let $\log(y)$ denote that unique real number such that $e^{\log(y)} = y$. In particular, the number $\log(y)$ can be defined by

$$\log(y) = \sup\{z \in \mathbb{R} : e^z < y\}.$$

(See Exercise 7 of Chapter 1 of Rudin or Problem 1 of Assignment 3.)

Corollary 120. *The function $x \mapsto \log(x)$ is increasing in the sense that if x, y are positive numbers satisfying $x < y$, then $\log(x) < \log(y)$.*

Theorem 121. *We have*

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Proof. Let s_n denote the nth partial sum

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}$$

and let t_n denote the nth term in the sequence

$$t_n = \left(1 + \frac{1}{n}\right)^n.$$

The proof will be complete if we can show that

$$\limsup_{n \to \infty} t_n \leqslant e \quad \text{and} \quad \liminf_{n \to \infty} t_n \geqslant e.$$

The binomial theorem implies that

$$
\begin{aligned}
t_n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{k!} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n} \frac{1}{k!} \\
&\leqslant \sum_{k=0}^{n} \frac{1}{k!} \\
&= s_n.
\end{aligned}
$$

It then follows that

$$\limsup_{n \to \infty} t_n \leqslant \limsup_{n \to \infty} s_n = e.$$

Now let m and n be positive integers such that $n \geqslant m$. Then again use the binomial theorem to find that

$$
\begin{aligned}
t_n &= \sum_{k=0}^{n} \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-k+1}{n} \frac{1}{k!} \\
&= \sum_{k=0}^{n} 1 \cdot \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \frac{1}{k!} \\
&\geqslant \sum_{k=0}^{m} 1 \cdot \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \frac{1}{k!},
\end{aligned}
$$

36

so that

$$t_n \geqslant 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \cdots + \frac{1}{m!}\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{m-1}{n}\right).$$

Fixing m, and taking $n \to \infty$ shows that

$$\liminf_{n\to\infty} t_n \geqslant 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!},$$

that is,

$$\liminf_{n\to\infty} t_n \geqslant s_m.$$

Now taking the limit as $m \to \infty$, we find that

$$\liminf_{n\to\infty} \geqslant e,$$

as desired. $\qquad\square$

Theorem 122. *The number e is irrational.*

Proof. If s_n denotes the nth partial sum, then note that

$$\begin{aligned}
e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots \\
&< \frac{1}{(n+1)!}\left(1 + \frac{1}{(n+1)} + \frac{1}{(n+1)^2} + \cdots\right) \\
&= \frac{1}{(n+1)!}\frac{1}{1 - \frac{1}{n+1}} \\
&= \frac{1}{(n+1)!}\frac{n+1}{n} \\
&= \frac{1}{n!n}.
\end{aligned}$$

Suppose that $e = m/n$, with m and n positive integers. Then $n!e$ is a positive integer, and also

$$n!s_n = n!\left(1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!}\right)$$

is a positive integer. However,

$$0 < n!(e - s_n) < \frac{1}{n},$$

by the above. $\qquad\square$

Example 123. If $p > 1$, then the series

$$\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$$

converges. If $p \leqslant 1$, then the series diverges. Indeed the sequence

$$a_k = \frac{1}{k(\log k)^p}$$

is nonnegative and monotonically decreasing, so we may apply Theorem 116. We therefore consider the series

$$\sum_{n=2}^{\infty} 2^n \frac{1}{2^n(\log 2^n)^p} = \sum_{n=2}^{\infty} \frac{1}{(n\log 2)^p} = \frac{1}{(\log 2)^p} \sum_{n=2}^{\infty} \frac{1}{n^p},$$

which converges if and only if $p > 1$.

3.8 Tests for series

Theorem 124 (Root test)**.** *Let a_n be a sequence of real numbers, and set $\alpha = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.*

   (a) If $\alpha < 1$, then the series $\sum_n a_n$ converges.

   (b) If $\alpha > 1$, then the series $\sum_n a_n$ diverges.

   (c) If $\alpha = 1$, then no conclusion can be made.

Proof. For (a), suppose $\alpha < 1$. Let ϵ satisfy $0 < \epsilon < 1 - \alpha$. Then since

$$\alpha = \inf \left\{ \sup_{n \geqslant N} \sqrt[n]{|a_n|} : N > 0 \right\},$$

there is an $N > 0$ such that

$$\sup_{n \geqslant N} \sqrt[n]{|a_n|} < \alpha + \epsilon$$

by the infimum analogue of Problem 4 of Assignment 1. In other words, if $n \geqslant N$, then

$$|a_n| < (\alpha + \epsilon)^n.$$

Since $\alpha + \epsilon < 1$, the series

$$\sum_{n \geqslant N} (\alpha + \epsilon)^n$$

converges. The convergence of $\sum_n a_n$ now follows by comparison.

For (b), suppose $\alpha > 1$. Then by definition of α, there is a subsequence a_{n_k} such that

$$\lim_{k \to \infty} \sqrt[n_k]{|a_{n_k}|} = \alpha > 1.$$

It follows that $|a_n| > 1$ for infinitely many values of k, and so the sequence a_n cannot converge to zero. Thus the series $\sum_n a_n$ cannot converge.

For (c), we exhibit two series with $\alpha = 1$ where one converges and the other diverges. Indeed the series

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges (as a p-series), but the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges. $\qquad\square$

Theorem 125 (Ratio test). *Let a_n be a sequence of nonzero terms and let b_n be the ratio*

$$b_n = \frac{|a_{n+1}|}{|a_n|}$$

of subsequent terms.

(a) *If $\limsup_{n \to \infty} b_n < 1$, then the series $\sum_n a_n$ converges.*

(b) *If there is a positive integer N such that $b_n \geqslant 1$ for $n \geqslant N$, then the series diverges.*

Proof. For (a), because $\limsup_{n \to \infty} b_n < 1$, there is a $\beta < 1$ and an $N > 0$ such that $b_n < \beta$ for each $n \geqslant N$. This means that in particular we have

$$|a_{N+1}| < \beta|a_N|$$
$$|a_{N+2}| < \beta|a_{N+1}| < \beta^2|a_N|$$
$$\vdots$$
$$|a_{N+p}| < \beta^p|a_N|.$$

In other words, we have

$$|a_n| < \beta^{n-N}|a_N| = \beta^{-N}|a_N|\beta^n$$

for each $n \geqslant N$. It then follows that $\sum_n a_n$ converges by comparison with the series $\sum_n \beta^n$, which converges because $\beta < 1$.

For (b), suppose that $b_n \geqslant 1$ for each $n \geqslant N$. This means that $|a_{n+1}| \geqslant |a_n|$ for each $n \geqslant N$. It follows that the sequence a_n does not tend to zero, and hence the series $\sum_n a_n$ cannot converge. $\qquad\square$

Example 126. Consider the series

$$\sum_{n=0}^{\infty} a_n$$

given by

$$1 + 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{2^3} + \frac{1}{3^3} + \cdots .$$

Then for a nonnegative integer n, we have that

$$\frac{|a_{2n+1}|}{|a_{2n}|} = \left(\frac{2}{3}\right)^n$$

$$\frac{|a_{2n+2}|}{|a_{2n+1}|} = \left(\frac{1}{2}\right)^{n+1} \cdot 3^n = \frac{1}{2}\left(\frac{3}{2}\right)^n .$$

It follows from this that

$$\limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = \lim_{n\to\infty} \frac{1}{2}\left(\frac{3}{2}\right)^n = +\infty$$

is not finite, and so the ratio test fails.

For the root test, we note that

$$|a_{2n}|^{1/2n} = \left(\frac{1}{2^n}\right)^{1/2n} = \frac{1}{\sqrt{2}}$$

$$|a_{2n+1}|^{1/2n+1} = \left(\frac{1}{3^n}\right)^{1/(2n+1)} = \frac{1}{3^{n/(2n+1)}}$$

and so

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}$$

$$\liminf_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \frac{1}{3^{n/(2n+1)}} = \frac{1}{\sqrt{3}} .$$

Since $\sqrt{2} > 1$, the root test indicates convergence.

In general, the root test is more useful than the ratio test, as the following theorem illustrates.

Theorem 127. *Let c_n be a be a sequence of positive numbers. Then*

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \limsup_{n\to\infty} \frac{c_{n+1}}{c_n} .$$

Hence if the ratio test indicates convergence, then so does the root test. (And if the root test is inconclusive, then so is the ratio test.)

Proof. Set

$$\alpha = \limsup_{n\to\infty} \frac{c_{n+1}}{c_n}.$$

If $\alpha = \infty$, then we are done. Otherwise, let β be any number satisfying $\beta > \alpha$. There is a positive integer N such that

$$\frac{c_{n+1}}{c_n} \leqslant \beta \qquad \text{for each } n \geqslant N.$$

In particular, for any $p > 0$, we have

$$c_{N+p} \leqslant \beta^p c_N,$$

that is,

$$c_n \leqslant \beta^{n-N} c_N = \beta^{-N} c_N \cdot \beta^n \qquad \text{for each } n \geqslant N.$$

Taking the nth root of both sides gives that

$$\sqrt[n]{c_n} \leqslant \sqrt[n]{c_N \beta^{-N}} \cdot \beta \qquad \text{for each } n \geqslant N.$$

The number $c_N \beta^{-N}$ is positive, so we have

$$\lim_{n\to\infty} \sqrt[n]{c_N \beta^{-N}} = 1.$$

It follows that

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \beta.$$

Since β was any number satisfying $\beta > \alpha$, we conlude that

$$\limsup_{n\to\infty} \sqrt[n]{c_n} \leqslant \alpha,$$

as desired. $\qquad\square$

We would next like to prove a convergence result for series whose subsequent terms have different signs, that is, for alternating series. We first require a useful "summation by parts" formula.

Theorem 128 (Summation by parts). *Let a_n, b_n be two sequences. For $n \geqslant 0$, set*

$$A_n = \sum_{k=0}^{n} a_k$$

and set $A_{-1} = 0$. Then if $0 \leqslant p \leqslant q$, we have

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p.$$

41

Proof. Note that

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q} (A_n - A_{n-1}) b_n$$

$$= \sum_{n=p}^{q} A_n b_n - \sum_{n=p}^{q} A_{n-1} b_n$$

$$= \sum_{n=p}^{q} A_n b_n - \sum_{m=p-1}^{q-1} A_m b_{m+1}$$

$$= \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p,$$

as desired. □

Theorem 129. *In the notation of the previous theorem, suppose*

(a) the partial sums A_n form a bounded sequence

(b) $b_0 \geqslant b_1 \geqslant b_2 \geqslant \cdots$

(c) $\lim_{n \to \infty} b_n = 0$.

Then $\sum_n a_n b_n$ converges.

Proof. Let $\epsilon > 0$ be arbitrary. Because A_n is bounded, there is an M such that $|A_n| \leqslant M$ for all n. Because $b_n \to 0$, there is an integer N such that

$$b_n \leqslant \frac{\epsilon}{2M} \qquad \text{for each } n \geqslant N.$$

Then for $N \leqslant p \leqslant q$, we have

$$\left| \sum_{n=p}^{q} a_n b_n \right| = \left| \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p \right|$$

$$\leqslant M \left| \sum_{n=p}^{q-1} (b_n - b_{n+1}) + b_q + b_p \right| \qquad -M \leqslant A_n \leqslant M \text{ and } b_n - b_{n+1} \geqslant 0$$

$$= M \left| (b_p - b_q) + b_q + b_p \right| \qquad \text{telescoping sum}$$

$$= 2M b_p$$

$$\leqslant \epsilon.$$

The convergence of the series now follows from the Cauchy criterion. □

Corollary 130 (Alternating series test). *Let c_n be a sequence of nonnegative terms and suppose c_n satisfies*

(a) $c_0 \geqslant c_1 \geqslant c_2 \geqslant \cdots$ *and*

(b) $\lim_{n \to \infty} c_n = 0$.

Then the series

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

converges.

Proof. Apply the previous result with $a_n = (-1)^n$ and $b_n = c_n$. $\qquad\square$

Lemma 131. *Let a_n be a sequence. If $\sum_n |a_n|$ converges, then so does $\sum_n a_n$.*

Proof. We note that

$$\left| \sum_{k=n}^{m} a_k \right| \leqslant \sum_{k=n}^{m} |a_k|$$

so the Cauchy criterion implies the result. $\qquad\square$

Definition 132. A series $\sum_n a_n$ is said to **converge absolutely** if the series $\sum_n |a_n|$ converges.

Example 133. As a p-series, the **harmonic series**

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

diverges. However the alternating series test implies that the series

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

converges. This convergence is therefore non-absolute.

3.9   Addition and multiplication of series

Theorem 134. *If $\sum_n a_n = A$ and $\sum_n b_n = B$, then $\sum_n (a_n + b_n) = A + B$ and $\sum_n c a_n = cA$ for any fixed real number c.*

Proof. If A_n denotes the sequence of partial sums

$$A_n = \sum_{k=1}^{n} a_k$$

and similarly for B_n, then we find that

$$\sum_{k=1}^{n}(a_k + b_k) = A_n + B_n.$$

The converges now follows from

$$\lim_{n\to\infty}(A_n + B_n) = A + B.$$

For the second claim, we have

$$\sum_{k=1}^{n}(ca_k) = cA_n,$$

and the convergence follows from

$$\lim_{n\to\infty} cA_n = cA.$$

$\square$

Definition 135. Given two series $\sum_n a_n$ and $\sum_n b_n$, let c_n denote the sequence

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Then the series $\sum_n c_n$ is called the **product** of the given series.

Remark 136. This can be motivated as follows. If $p(x), q(x)$ are two polynomials

$$p(x) = \sum_{k=0}^{r} a_k x^k$$

$$q(x) = \sum_{k=0}^{s} a_k x^k,$$

then the coefficient of x^n in the product $p(x)q(x)$ is given by

$$\sum_{k+j=n} a_j b_k.$$

Remark 137. If $\sum_n a_n = A$ and $\sum_n b_n = B$, it is not clear that $\sum_n c_n = AB$, and in fact, this is in general not true.

Example 138. The series

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{n+1}}$$

converges by the alternating series test. Consider the product of this series with itself. The terms of that series are given by

$$c_n = \sum_{k=0}^{n} \frac{(-1)^k}{\sqrt{k+1}} \frac{(-1)^{n-k}}{\sqrt{n-k+1}}$$

$$= (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{(n-k+1)(k+1)}}.$$

Note that

$$(n-k+1)(k+1) = \left[\left(\frac{n}{2}+1\right) + \left(\frac{n}{2}-k\right)\right]\left[\left(\frac{n}{2}+1\right) - \left(\frac{n}{2}-k\right)\right]$$

$$= \left(\frac{n}{2}+1\right)^2 - \left(\frac{n}{2}-k\right)^2$$

$$\leqslant \left(\frac{n}{2}+1\right)^2.$$

It follows that each c_n satisfies

$$|c_n| \geqslant \sum_{k=0}^{n} \frac{1}{\frac{n}{2}+1} = \sum_{k=0}^{n} \frac{2}{n+2} = \frac{2(n+1)}{n+2}.$$

It follows that the sequence c_n does not converge to 0, and so the series $\sum_n c_n$ cannot converge.

Theorem 139. *Suppose $\sum_n a_n = A$ and $\sum_n b_n = B$. If $\sum_n a_n$ converges absolutely, then the product $\sum_n c_n$ converges to AB.*

Proof. See Rudin. $\square$

Remark 140. Note that the condition of absolute convergence is needed. The example before the theorem is a case of the product of two non-absolutely convergent series, which we noted is divergent.

3.10 A word on rearrangements of series

Definition 141. Let $f : \mathbb{N} \setminus 0 \to \mathbb{N} \setminus 0$ be a bijective function. If we are given a sequence a_n, and we set

$$b_n = a_{f(n)},$$

then the series $\sum_n b_n$ is called a **rearrangement** of the series $\sum_n a_n$.

Remark 142. If s_n denotes the partial sums of $\sum_n a_n$ and t_n denotes the partial sums of the rearrangement $\sum_n b_n$, then in general, the sets $\{s_n\}$ and $\{t_n\}$ consist of entirely different numbers. Thus, if the series $\sum_n a_n$ converges to A, then the series $\sum_n b_n$ may converge to a different number, and in fact, may not even converge at all.

Example 143. If

$$a_n = \frac{(-1)^n}{n+1} \qquad n = 0, 1, 2, \ldots,$$

then the alternating series test shows that $\sum_n a_n$ converges. In fact, if A denotes the sum of the series, then note that A satisfies

$$A = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \cdots$$

$$= 1 - \frac{1}{2} + \frac{1}{3} - \left(\frac{1}{4} - \frac{1}{5}\right) - \left(\frac{1}{6} - \frac{1}{7}\right) - \left(\frac{1}{8} - \frac{1}{9}\right) - \cdots$$

$$< 1 - \frac{1}{2} + \frac{1}{3} = \frac{5}{6}.$$

Consider the rearrangement $\sum_n b_n$ given by

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$$

in which two positive terms are always followed by one negative term. Consider

$$\overbrace{\left(1 + \frac{1}{3} - \frac{1}{2}\right)}^{k=1} + \overbrace{\left(\frac{1}{5} + \frac{1}{7} - \frac{1}{4}\right)}^{k=2} + \overbrace{\left(\frac{1}{9} + \frac{1}{11} - \frac{1}{6}\right)}^{k=3} + \cdots$$

$$= \sum_{k=1}^{\infty} \left(\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k}\right).$$

In other words, if t_n denotes the sequence of partial sums, then the subsequence t_{3n} is given by

$$t_{3n} = \sum_{k=1}^{n} \left(\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k}\right).$$

Because

$$\frac{1}{4k-3} + \frac{1}{4k-1} - \frac{1}{2k} = \frac{2k(4k-1) + 2k(4k-3) - (4k-3)(4k-1)}{2k(4k-3)(4k-1)}$$

$$= \frac{8k^2 - 2k + 8k^2 - 6k - (16k^2 - 16k + 3)}{2k(4k-3)(4k-1)}$$

$$= \frac{8k-3}{2k(4k-3)(4k-1)},$$

$$> 0$$

46

we see that the subsequence t_{3n} satisfies $t_3 < t_6 < t_9 < \cdots$. Hence we have that

$$\limsup_{n \to \infty} t_n \geqslant t_3 = \frac{5}{6},$$

so that the sequence t_n does not converge to A. In other words, rearranging has changed the sum. That this new sum $B = \sum_n b_n$ is still finite can be checked by noting that

$$B = 1 + \frac{1}{3} - \left(\frac{1}{2} - \frac{1}{5} - \frac{1}{7} \right) - \left(\frac{1}{4} - \frac{1}{9} - \frac{1}{11} \right) - \cdots$$
$$= 1 + \frac{1}{3} - \sum_{j=1}^{\infty} \left(\frac{1}{2j} - \frac{1}{4j+1} - \frac{1}{4j+3} \right),$$

and then checking that

$$\begin{aligned}
\frac{1}{2j} - \frac{1}{4j+1} - \frac{1}{4j+3} &= \frac{(4j+1)(4j+3) - 2j(4j+3) - 2j(4j+1)}{2j(4j+1)(4j+3)} \\
&= \frac{16j^2 + 16j + 3 - 8j^2 - 6j - 8j^2 - 2j}{2j(4j+1)(4j+3)} \\
&= \frac{8j+3}{2j(4j+1)(4j+3)} \\
&> 0,
\end{aligned}$$

so that $B < 4/3$.

The previous example shows that by rearranging a series we can change its sum. In fact, the following more general result is true, which is due to Riemann.

Theorem 144. *Let $\sum_n a_n$ be a series of real numbers that converges non-absolutely. Let α and β be real numbers satisfying $\alpha \leqslant \beta$. Then there is a rearrangement $\sum_n b_n$ with partial sums t_n which satisfy*

$$\liminf_{n \to \infty} t_n = \alpha \qquad and \qquad \limsup_{n \to \infty} t_n = \beta.$$

Proof. For the proof, see Rudin. □

On the other hand, if the series converges absolutely, then every rearrangement has the same sum.

Theorem 145. *If $\sum_n a_n$ converges absolutely, then every rearrangement converges too, and each converges to the same sum.*

Proof. Let $\sum_n b_n$ be a rearrangement with partial sums t_n. Say that b_n is determined by $f : \mathbb{N} \setminus 0 \to \mathbb{N} \setminus 0$ in the sense that $b_n = a_{f(n)}$ Let $\epsilon > 0$ be arbitrary. There is an $N > 0$ such that if $m \geqslant n \geqslant N$, then

$$\sum_{k=n}^{m} |a_k| < \frac{\epsilon}{2}.$$

Choose p large enough that

$$\{1, 2, \ldots, N\} \subset \{f(1), \ldots, f(p)\}.$$

Then we note that if $m > p$, the numbers $a_1, \ldots, a_N$ cancel in the difference

$$s_m - t_m = \sum_{k=1}^{m} (a_k - a_{f(k)}).$$

It follows that the difference only consists of numbers from the set $\{a_{N+1}, \ldots, a_m\}$, and so we find that

$$|s_m - t_m| \leqslant 2 \sum_{k=N+1}^{m} |a_k| < \epsilon.$$

It follows that the rearrangement has the same sum as $\sum_n a_n$. $\qquad\square$

4 Continuity

4.1 Limits of funtions

Let X and Y be metric spaces, and let E be a subset of X. Let $f : E \to Y$ be a function defined on the subspace E. Let p be a limit point of E, and let q be a point in Y.

Definition 146. We write

$$\lim_{x \to p} f(x) = q$$

to mean the following: for each $\epsilon > 0$, there is a $\delta > 0$ such that whenever $x \in E$ satisfies $0 < d_X(x, p) < \delta$, then $d_Y(f(x), q) < \epsilon$. In such a situation, we say that f approaches q as x approaches p. We also say that q is the **limit of f as x approaches** p.

Note that f need not be defined at the particular point $p \in X$ in order for the limit of f near p to be defined; we only need f to be defined near p in the sense that p needs to be a limit point of the domain E of f.

Theorem 147. *We have*

$$\lim_{x \to p} f(x) = q$$

if and only if

$$\lim_{n \to \infty} f(p_n) = q$$

for each sequence p_n in E such that

$$p_n \neq p \qquad and \quad p_n \to p.$$

Proof. Suppose that $f(x) \to q$ as $x \to p$. Let p_n be a sequence in E satisfying $p_n \neq p$ and $p_n \to p$. Let $\epsilon > 0$ be arbitrary. Then by assumption, there is a $\delta > 0$ such that if $x \in E$ satisfies $d_X(x, p) < \delta$, then $d_Y(f(x), q) < \epsilon$. Since $p_n \to p$, there is an $N > 0$ such that if $n \geqslant N$, then $d_X(p_n, p) < \delta$. It follows that if $n \geqslant N$, then $d(f(x_n), q) < \epsilon$. We conclude that $f(p_n) \to q$.

Conversely, suppose that

$$\lim_{x \to p} f(x) = q$$

is false. This means that there is an $\epsilon > 0$ such that for each $\delta > 0$, there is an $x \in E$ satisfying $0 < d_X(x, p) < \delta$ and $d_Y(f(x), q) \geqslant \epsilon$. In particular, letting $\delta = 1/n$, we obtain a sequence $p_n \in E$ satisfying

$$0 < d_X(p_n, p) < \frac{1}{n}.$$

From this, it follows that $p_n \neq p$ and also that $p_n \to p$. However, we have that

$$d_Y(f(p_n), q) \geqslant \epsilon,$$

so that $f(p_n)$ does not converge to q. $\square$

Corollary 148. *If f has a limit at p, then this limit is unique.*

Proof. This follows from the previous result together with the fact that the limit $f(p_n)$ is unique. $\square$

Theorem 149. *Suppose f and g are real-valued and*

$$\lim_{x \to p} f(x) = A \qquad and \qquad \lim_{x \to p} g(x) = B.$$

Then

(a) $\lim_{x \to p}(f + g)(x) = A + B$

(b) $\lim_{x \to p} c \cdot f(x) = c \cdot A$ *for each real number $c \in \mathbb{R}$*

(c) $\lim_{x \to p}(fg)(x) = AB$

(d) $\lim_{x \to p} \left(\frac{f}{g}\right)(x) = \frac{A}{B}$, *if $B \neq 0$.*

Proof. All of these properties follow from the corresponding properties for sequences of real numbers. $\square$

4.2 Continuous functions

Let X and Y be metric spaces, and let E be a subset of X. Let $f : E \to Y$ be a function defined on the subset E. Let p be a point of E.

Definition 150. We say that f is **continuous at** p if for each $\epsilon > 0$ there is a $\delta > 0$ such that whenever $x \in E$ satisfies $d_X(x, p) < \delta$, then $d_Y(f(x), f(p)) < \epsilon$. We say that f is **continuous** if f is continuous at each point of its domain.

Note that in order for f to be continuous at p, we need f to be defined at p in the first place.

It is possible to reformulate the condition of continuity as follows. The function f is continuous at p if and only if for each $\epsilon > 0$ there is a $\delta > 0$ such that $f(B_\delta(p)) \subset B_\epsilon(f(p))$, where
$$f(B_\delta(p)) = \{f(x) : x \in B_\delta(p)\}.$$

Example 151. Let $f_i : \mathbb{R}^k \to \mathbb{R}$ be the component function $f(x) = x_i$. Then we claim that f_i is continuous. Indeed, let p be a point of $\mathbb{R}^k$, and let $\epsilon > 0$ be arbitrary. Choose $\delta = \epsilon$. Then if $x \in \mathbb{R}^k$ satisfies
$$\|x - p\| < \delta = \epsilon,$$
then this means that
$$\left(\sum_{i=1}^{k} |x_i - p_i|^2 \right)^{1/2} < \epsilon,$$
and so we see in particular that
$$|x_i - p_i| < \epsilon,$$
that is
$$|f_i(x) - f_i(p)| < \epsilon.$$

Example 152. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^2 + 1$. Then we claim that f is continuous. Let us check at the point $p = 2$ in particular, and then the reader can verify for other points p. Let $\epsilon > 0$ be arbitrary. Then choose $\delta = \min\{1, \epsilon/5\}$. Suppose that x satisfies $|x - 2| < \delta$. Then on the one hand, $|x - 2| < 1$, and so the triangle inequality implies that
$$|x + 2| = |x - (-2)| \leqslant |x - 2| + |2 - (-2)| < 1 + 4 = 5.$$

On the other hand, $|x - 2|$ satisfies
$$|x - 2| < \epsilon/5.$$

Then using these inequalities, we find that

$$\begin{aligned} |f(x) - f(2)| &= |(x^2 + 1) - (2^2 + 1)| \\ &= |x^2 - 2^2| \\ &= |x + 2||x - 2| \\ &< 5 \cdot \frac{\epsilon}{5} = \epsilon. \end{aligned}$$

From the presentation, it might not be clear how δ was chosen. Note that the choice of δ is allowed to depend on ϵ and p. We would like the inequality

$$|f(x) - f(2)| < \epsilon$$

to hold. Note that this inequality is equivalent to

$$|x^2 + 1 - (2^2 + 1)| < \epsilon,$$

that is,

$$|x^2 - 2^2| < \epsilon.$$

We can rewrite this as

$$|x - 2||x + 2| < \epsilon.$$

By choosing δ, we are allowed to control the size of the term $|x - 2|$ and make it as small as we want. Thus to ensure that the left-hand side is small, we only need to ensure additionally that the other term $|x + 2|$ is not too large. In particular, if it were bounded by a constant, this would suffice. Note that we can in fact ensure that this term $|x + 2|$ is bounded near $x = 2$ because the triangle inequality guarantees that

$$|x + 2| = |x - (-2)| \leqslant |x - 2| + |2 - (-2)| \leqslant |x - 2| + 4,$$

so that our countrol of $|x - 2|$ gives us control of $|x + 2|$. In particular, if we are assuming that $|x - 2| < \delta$, then we see that $|x + 2| < \delta + 4$. Thus if we make δ smaller than some fixed constant, say smaller than 1, then we ensure that $|x + 2|$ is bounded by a fixed constant, say 5. Thus whenever $|x - 2| < \delta \leqslant 1$, we have ensured that

$$|x - 2||x + 2| \leqslant \delta \cdot 5.$$

The choice of $\delta = \min\{\epsilon/5, 1\}$ is now justified.

Theorem 153. *Suppose f is continuous at p and g is continuous at $f(p)$. Then the composition $g \circ f$ is continuous at p.*

Proof. Let $\epsilon > 0$ be arbitrary. Because g is continuous at $f(p)$, there is an $\eta > 0$ such that whenever y satisfies $d(y, f(p)) < \eta$, then $d(g(y), g(f(p))) < \epsilon$. Because f is continuous at p, there is a $\delta > 0$ such that whenever x satisfies $d(x, p) < \delta$, then $d(f(x), f(p)) < \eta$. Thus, if x satisfies $d(x, p) < \delta$, then $d(f(x), f(p)) < \eta$, which implies $d(g(f(x)), g(f(p))) < \epsilon$. $\qquad \square$

Theorem 154. *Let $f : X \to Y$ be a function. Then f is continuous if and only if for each open subset $V \subset Y$, the preimage $f^{-1}(V)$ is open in X.*

Proof. Suppose that f is continuous. Let V be an open subset of Y. Let x be a point of $f^{-1}(V)$. This means that $f(x) \in V$. Because V is open, there is an $\epsilon > 0$ such that $B_\epsilon(f(x)) \subset V$. Because f is continuous at x, there is a $\delta > 0$ such that $f(B_\delta(x)) \subset B_\epsilon(f(x))$. We can rewrite this latter condition as $B_\delta(x) \subset f^{-1}(B_\epsilon(f(x)))$. But because $B_\epsilon(f(x)) \subset V$, we infer that $f^{-1}(B_\epsilon(f(x))) \subset f^{-1}(V)$ and therefore

$$B_\delta(x) \subset f^{-1}(B_\epsilon(f(x))) \subset f^{-1}(V).$$

This shows that x is an interior point of $f^{-1}(V)$.

Suppose that the latter condition in the statement of the theorem is satisfied. Let x be a point of X. Let $\epsilon > 0$ be arbitrary. The set $B_\epsilon(f(x))$ is open in Y. By assumption, therefore, the subset $f^{-1}(B_\epsilon(f(x)))$ is open in X. The point x belongs to the set $f^{-1}(B_\epsilon(f(x)))$. It follows that there is a $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(B_\epsilon(f(x)))$. This implies that $f(B_\delta(x)) \subset B_\epsilon(f(x))$. We infer that f is continuous at x. $\quad\square$

Theorem 155. *Let f, g be continuous real-valued functions on X. Then $f + g, fg$, and f/g are continuous wherever they are defined.*

Corollary 156. *Any polynomial in the k variables $x_1, \ldots, x_k$ is a continuous real-valued function on $\mathbb{R}^k$.*

Theorem 157. *Let $f_1, \ldots, f_k$ be real-valued functions on X, and let $f : X \to \mathbb{R}^k$ denote the function whose components are given by*

$$f(x) = (f_1(x), \ldots, f_k(x)).$$

Then f is continuous if and only if each f_j is.

Proof. Note that

$$\|f(x) - f(p)\| = \left(\sum_{j=1}^{k} |f_j(x) - f_j(p)|^2 \right)^{1/2}.$$

Hence if we can make each $|f_j(x) - f_j(p)|$ smaller than ϵ, then we can make $\|f(x) - f(p)\|$ smaller than ϵ as well. On the other hand, if we can make $\|f(x) - f(p)\|$ smaller than ϵ, then we can certainly make each $|f_j(x) - f_j(p)|$ smaller than ϵ. $\quad\square$

4.3 Continuity and compactness

Theorem 158. *Let $f : X \to Y$ be a continuous mapping. If X is compact, then so is $f(X)$.*

Proof. Let $\{V_\alpha\}$ be an open cover of $f(X)$. Since f is continuous, each $f^{-1}(V_\alpha)$ is open, and so the collection $\{f^{-1}(V_\alpha)\}$ forms an open cover of X. Since X is compact, there are finitely many indices $\alpha_1, \ldots, \alpha_n$ such that

$$X \subset (f^{-1}(V_{\alpha_1}) \cup \cdots \cup f^{-1}(V_{\alpha_n})).$$

Applying f to both sides and using Warm-Up 11 (which says that $f(f^{-1}(V_\alpha)) \subset V_\alpha$), we find that

$$f(X) \subset (V_{\alpha_1} \cup \cdots \cup V_{\alpha_n}).$$

This completes the proof. $\square$

Corollary 159 (Extreme Value Theorem). *Let $f : X \to \mathbb{R}$ be a continuous mapping. Suppose X is compact. If*

$$M = \sup_{x \in X} f(x) \qquad \text{and} \qquad m = \inf_{x \in X} f(x),$$

then there are points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

Proof. The previous theorem says that $f(X)$ is a compact subset of $\mathbb{R}$. Hence, $f(X)$ is closed and bounded. Because $f(X)$ is bounded, both $M = \sup f(X)$ and $m = \inf f(X)$ exist. Moreover, because $f(X)$ is closed, we know that both $M = \sup f(X)$ and $m = \inf(X)$ belong to $f(X)$. $\square$

Theorem 160. *Let $f : X \to Y$ be a continuous function. If X is compact and f is bijective, then the inverse function $f^{-1} : Y \to X$ defined by*

$$f^{-1}(f(x)) = x$$

is continuous as well.

Proof. For notation's sake, let $g : Y \to X$ denote the inverse. Let U be an open subset of X. Then the complement U^c is closed in X, which is compact, and hence U^c is compact too. Since f is continuous, the image $f(U^c)$ is compact in Y. The fact that f is bijective implies that

$$f(U)^c = f(U^c),$$

and so $f(U)$ is open in Y. But the relation $g \circ f = \text{id}_X$ implies that

$$f(U) = g^{-1}(U)$$

and so g is continuous by Theorem 154.

$\square$

Definition 161. Let $f : X \to Y$ be a function. We say that f is **uniformly continuous** if for each $\epsilon > 0$, there is a $\delta > 0$ such that whenever $x, y \in X$ satisfy $d_X(x, y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$.

Remark 162. Note that if f is uniformly continuous, then f is continuous. However, in general, the converse is not true because the choice of δ for the notion of continuity depends on the point p. On the other hand, the choice of δ for the notion of uniform continuity is uniform and independent of the particular point in the domain.

Theorem 163. *Let $f : X \to Y$ be a continuous function. If X is compact, then f is uniformly continuous.*

Proof. Let $\epsilon > 0$ be given. Because f is continuous, for each point $p \in X$, there is a real number $\delta(p) > 0$ such that $f(B_{\delta(p)}(p)) \subset B_{\epsilon/2}(f(p))$. The collection $\{B_{\delta(p)/2}(p)\}_{p \in X}$ of balls with even smaller radii forms an open cover of X. Because X is compact, there are finitely many points $p_1, \ldots, p_n$ of X together with radii $\delta_i = \delta(p_i)$ such that

$$X \subset (B_{\delta_1/2}(p_1) \cup \cdots \cup B_{\delta_n/2}(p_n)). \tag{2}$$

Set

$$\delta = \frac{1}{4} \min_{1 \leqslant i \leqslant n} \delta_i.$$

Let x, y be points of X satisfying $d(x, y) < \delta$. By (2), there is an $i \in \{1, \ldots, n\}$ such that $x \in B_{\delta_i/2}(x_i)$. Then by the triangle inequality, we have that

$$d(x_i, y) \leqslant d(x_i, x) + d(x, y) < \delta_i/2 + \delta < \delta_i.$$

We then again use the triangle inequality to find that

$$d(f(x), f(y)) \leqslant d(f(x), f(x_i)) + d(f(x_i), f(y)) < \epsilon/2 + \epsilon/2 = \epsilon.$$

$\square$

Example 164. Let $X = (0, 1) \subset \mathbb{R}$, and let $f : X \to \mathbb{R}$ be the function $f(x) = 1/x$. Then we let the reader check that f is continuous at each point of its domain. However, note that X is not compact, so the previous result does not apply, and indeed we verify now that f is not uniformly continuous. Choose $\epsilon = 1$, and let $\delta > 0$ be arbitrary. By the Archimedean property, there is a positive integer n satisfying $n > 2/\delta$. Let $x = 1/n$ and let $y = 1/(n+2)$. Then note that

$$|x - y| = \left| \frac{1}{n} - \frac{1}{n+2} \right| = \frac{2}{n(n+2)} < \frac{2}{n} < \delta.$$

However, note also that

$$|f(x) - f(y)| = |n - (n+2)| = 2 \geqslant \epsilon.$$

Example 165. If $X = \mathbb{R}$, then X is not compact, and we let the reader verify that $f : X \to \mathbb{R}$ defined by $f(x) = x^2$ is continuous but not uniformly continuous.

4.4   Continuity and connectedness

Definition 166. Let X be a metric space. We say that X is **connected** if the following implication is true: If A is a nonempty subset of X that is both open and closed, then $A = X$. Otherwise, we say that X is **disconnected**.

Theorem 167. *Let X be a nonempty subspace of $\mathbb{R}$. Then X is connected if and only if the following property is satisfied: If $x, y \in X$ satisfy $x < y$, then $(x, y) \subset X$.*

Proof. Suppose the latter property is not satisfied. This means that there are $x, y \in X$ satisfying $x < y$, and there is a $z \in \mathbb{R}$ satisfying $x < z < y$ such that $z \notin X$. Let A denote the subset of X determined by

$$A = X \cap (-\infty, z).$$

Note that x belongs to A, so A is nonempty. Also note that $y \notin A$, so $A \neq X$. Note also that A is open relative to X by construction. Since $z \notin X$, we compute that

$$X \setminus A = X \cap (z, \infty),$$

and so $X \setminus A$ is open as well. This implies that A is closed. We conclude that X is not connected.

Suppose the latter property is satisfied. Let A be a nonempty subset of X and suppose that A is both open and closed relative to X. Let B be the complement $B = X \setminus A$, which is both open and closed relative to X as well. Let y be a point of A. If B is empty, then we are done. If not, let x be a point of B, and we hope to obtain a contradiction. Up to relabeling x and y, we may assume that $x < y$. Let B_y be the subset of B given by

$$B_y = \{w \in B : w < y\}.$$

Then B_y is nonempty because $x \in B_y$. Also B_y is bounded from above, so B_y admits a supremum. Set $z = \sup B_y$. Note that z satisfies $x \leqslant z \leqslant y$ because y is an upper bound for B_y. There are three cases for z.

- Suppose that $z \notin X$. Then we have contradicted our assumption that X satisfies the latter property in the statement of the theorem.

- Suppose that $z \in A$. Since A is open relative to X, there is a real number $\epsilon > 0$ such that $(z - \epsilon, z + \epsilon) \cap X \subset A$. It follows that $(z - \epsilon, z] \cap B = \varnothing$, which is a contradiction to the assumption that z is the supremum of B, according to Problem 4 of Assignment 1.

- Suppose that $z \in B = X \setminus A$. Since B is open relative to X, there is an $\epsilon > 0$ such that $(z - \epsilon, z + \epsilon) \cap X \subset B$. Since y does not belong to B, we find that $z < y$. By assumption on X, every element between z and y belongs to X as well. In particular, this means that $z + \epsilon/2$ belongs to X, and hence belongs to B, but this is a contradiction to the assumption that z is an upper bound for B.

We are done. □

Corollary 168. *The space $\mathbb{R}$ is connected.*

Proposition 169. *Let $f : X \to Y$ be a continuous function. If X is connected, then $f(X)$ is connected.*

Proof. Let $g : X \to f(X)$ denote the same mapping as f but with the codomain restricted to $f(X)$. By doing so, we have ensured that g is surjective. Also note that g is continuous since f is continuous.

Let A be a nonempty subset of $f(X)$ that is both open and closed relative to $f(X)$. Since A is nonempty, the preimage $g^{-1}(A)$ is a nonempty subset of X. Since g is continuous, the preimage $g^{-1}(A)$ is open relative to X. Also since g is continuous, the preimage $g^{-1}(A)$ is closed relative to X. Since X is connected, we find that $g^{-1}(A) = X$. Since g is surjective, we conclude that

$$A = g(g^{-1}(A)) = g(X) = f(X).$$

We are done. □

Theorem 170 (Intermediate Value Theorem). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. If c is any number between $f(a)$ and $f(b)$, then there is a point $x \in [a, b]$ such that $f(x) = c$.*

Proof. The image $f([a, b])$ is connected by the previous result. Without loss of generality, we may assume that $f(a) < f(b)$. Then Theorem 167 shows that every point c satisfying $f(a) < c < f(b)$ also satisfies $c \in f([a, b])$. It follows that there is a point $x \in [a, b]$ such that $f(x) = c$. We are done. □

4.5   Discontinuities

For this section, let f be a real-valued function defined on a subset E of $\mathbb{R}$.

Definition 171. Fix a real number p.

(a) Suppose there is a $b > p$ such that the domain of f contains (p, b). Then we write

$$\lim_{x \to p^+} f(x) = q$$

to mean whenever p_n is a sequence in (p, b) satisfying $p_n \to p$, then

$$\lim_{n \to \infty} f(p_n) = q.$$

(b) Suppose there is an $a < p$ such that the domain of f contains (a, p). Then we write

$$\lim_{x \to p^-} f(x) = q$$

to mean whenever p_n is a sequence in (a, p) satisfying $p_n \to p$, then

$$\lim_{n \to \infty} f(p_n) = q.$$

Remark 172. We let the reader check that if f is defined on (a, b) and $p \in (a, b)$, then

$$\lim_{x \to p} f(x) = q$$

if and only if

$$\lim_{x \to p^+} f(x) = \lim_{x \to p^-} f(x) = q.$$

(The forward direction is immediate.)

Example 173. Let $f : \mathbb{R} \to \mathbb{R}$ be the piecewise function

$$f(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & x \text{ irrational} \end{cases}.$$

Let p be any point of $\mathbb{R}$. Then we claim that

$$\lim_{x \to p^+} f(x)$$

does not exist.

Indeed, let q be any number. There are two cases for q: either $q \neq 1$ or $q = 1$.

- Suppose $q \neq 1$. Let p_n be a sequence of rational numbers satisfying $p_n > p$. Then we see that each $f(p_n) = 1$, and so $f(p_n) \to 1$. In particular, we do not have $f(p_n) \to q$.

- Suppose $q = 1$. Then let p_n be a sequence of irrational numbers satisfying $p_n > p$. Then we see that each $f(p_n) = 0$, and so we do not have $f(p_n) \to q$ in this case either.

It is similarly true that $\lim_{x \to p^-} f(x)$ does not exist either.

Example 174. Let $f : \mathbb{R} \to \mathbb{R}$ be the piecewise function

$$f(x) = \begin{cases} x & x \text{ rational} \\ 0 & x \text{ irrational} \end{cases}.$$

We first claim that f is continuous at $p = 0$. Indeed, it suffices to prove that

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

Let $\epsilon > 0$ be arbitrary. Let p_n be any sequence of real numbers satisfying $p_n \to 0$ and $p_n \neq 0$. There is an $N > 0$ such that if $n \geqslant N$ then $|p_n| < \epsilon$. From the definition of f, either $f(p_n) = p_n$ or $f(p_n) = 0$. In either case, if $n \geqslant N$, then we also have $|f(p_n)| < \epsilon$. The claim now follows from Theorem 147.

We also claim that if $p \neq 0$, then neither $\lim_{x \to p^+} f(x)$ nor $\lim_{x \to p^-} f(x)$ exist. Indeed, the proof is similar to the previous example, so we omit it, and leave it as an exercise for the reader.

Example 175. Define $f : \mathbb{R}_{>0} \to \mathbb{R}$ by the rule

$$f(x) = \sin\left(\frac{1}{x}\right).$$

We will check that $\lim_{x \to 0^+} f(x)$ does not exist.

Let x_n be the sequence $x_n = 1/n$. Note that $f(x_n) = \sin(n)$. We claim that $f(x_n)$ does not converge to any point of $[-1, 1]$.

In Assignment 4, we show that the subset

$$E = \{n + 2m\pi : n, m \in \mathbb{Z}\}$$

is dense in $\mathbb{R}$. We will show later that $x \mapsto \sin(x)$ is a continuous function. It follows then also from Assignment 4 that $\sin(E)$ is dense in $[-1, 1]$. The 2π-periodicity of $\sin(x)$ implies that

$$\sin(E) = \{\sin(n) : n \in \mathbb{Z}\}.$$

Let q be any point of $[-1, 1]$. Let $\epsilon > 0$ be arbitrary. Let B denote the subset of $[-1, 1]$ given by

$$B = (B_\epsilon(q)) \cup B_\epsilon(-q)) \cap [-1, 1].$$

Note that B satisfies the property that if $x \in B$, then $-x \in B$. Because $\sin(E)$ is dense in $[-1, 1]$, the intersection $B \cap \sin(E)$ is infinite. We will construct a subsequence x_{n_k} such that $f(x_{n_k}) \in B$ for each k. Indeed there is a nonzero integer n_1 such that $\sin(n_1) \in B$. We may assume that n_1 is positive because if not, then $-n_1$ is positive and $\sin(-n_1) = -\sin(n_1)$ is still in B. Once $n_1 < n_2 < \cdots < n_k$ are chosen, then the intersection

$$B \cap \{\sin(n) : |n| > n_k\}$$

is still infinite, so there is an $n_{k+1} > n_k$ such that $\sin(n_{k+1}) \in B$. In this way, we obtain a subsequence x_{n_k} such that $f(x_{n_k}) \in B$.

Suppose $f(x_n)$ did converge. Then the previous paragraph implies that $f(x_n)$ converges either to q or $-q$. But q was also arbitrary, and since limits are unique, we obtain a contradiction.

5   Differentiation

5.1   Derivatives

Definition 176. Let f be a real-valued function on $[a, b]$, and let x_0 be a point of (a, b). We say that f is **differentiable at** x_0 to mean that the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists, and in such a case, we denote the value of this limit by $f'(x_0)$.

Example 177. Let $f(x) = |x|$. Let $x_0 = 0$. Then we claim that f is not differentiable at $x_0 = 0$. Indeed, we consider that for $x \geqslant 0$, we have $f(x) = x$ and so

$$\lim_{x \to 0^+} \frac{f(x) - f(x_0))}{x - x_0} = \lim_{x \to 0^+} \frac{x}{x} = 1.$$

On the other hand, for $x \leqslant 0$, we have $f(x) = -x$ and so

$$\lim_{x \to 0^-} \frac{f(x) - f(x)}{x - x_0} = \lim_{x \to 0^-} \frac{-x}{x} = -1.$$

On the other hand, the function f is differentiable at points $x_0 \neq 0$, which we let the reader verify. Indeed, more precisely, we have

$$f'(x_0) = \begin{cases} 1 & x_0 > 0 \\ -1 & x_0 < 0 \end{cases}.$$

The following lemma asserts that we may "change variables" in limits, provided certain hypotheses are met.

Lemma 178. *Let* $u : T \to X$, *let* E *be a subset of* X, *and let* $g : E \to Y$. *Let* t_0 *be a limit point of* T, *and set* $x_0 = u(t_0)$. *Suppose*

(a) x_0 *is a limit point of* E

(b) u *is continuous*

(c) u *is injective*

(d) $\lim_{x \to x_0} g(x)$ *exists.*

Then

$$\lim_{t \to t_0} g(u(t)) = \lim_{x \to x_0} g(x).$$

Proof. Exercise. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Corollary 179. *Suppose f is differentiable at x_0. Then*

$$f'(x_0) = \lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t}.$$

Proof. Let $u(t) = x_0 + t$, which is both continuous and injective. Let $g(x)$ be the difference quotient

$$g(x) = \frac{g(x) - g(x_0)}{x - x_0}$$

defined near x_0. Then

$$g(u(t)) = \frac{g(x_0 + t) - g(x_0)}{(x_0 + t) - x_0} = \frac{g(x_0 + t) - g(x_0)}{t}.$$

The previous lemma then implies the result. $\qquad\square$

Corollary 180. *If f is differentiable at x_0, then f is continuous at x_0.*

Proof. We investigate

$$\begin{aligned}
\lim_{x \to x_0} (f(x) - f(x_0)) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \\
&= f'(x_0) \lim_{x \to x_0} (x - x_0) \\
&= 0.
\end{aligned}$$

We conclude that

$$\lim_{x \to x_0} f(x) = f(x_0),$$

so we are done. $\qquad\square$

Corollary 181. *Let f be a function on $[a, b]$ and x_0 a point of (a, b). The following are equivalent.*

(a) *The function f is differentiable at x_0.*

(b) *There is a number $f'(x_0)$ such that the error function $E(h)$ defined for h near zero by*

$$f(x_0 + h) - f(x_0) = h[f'(x_0) + E(h)]$$

satisfies $E(h) \to 0$ as $h \to 0$.

Proof. Suppose f is differentiable at x_0. If we define E as required, then we just need to check that the statement about the limit is true. For $h \neq 0$ we have that

$$\frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0) + E(h), \tag{3}$$

and then taking the limit as $h \to 0$ shows that

$$f'(x_0) = f'(x_0) + \lim_{h \to 0} E(h).$$

We are done with this direction of the proof.

For the other direction, suppose that (b) is satisfied. Taking the limit of equation (3) as $h \to 0$ shows that

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0)$$

because of the hypothesis on $E(h)$. This means that f is differentiable at x_0. We are done. $\qquad\square$

This corollary says that $f(x_0 + h) - f(x_0)$ is approximated well by a linear function of h. Indeed the mapping $h \mapsto f'(x_0)h$ is a linear map and it only differs from $h \mapsto f(x_0 + h) - f(x_0)$ by an error term $h \cdot E(h)$ that goes to zero relatively fast as $h \to 0$ (indeed faster than h as h goes to zero).

If f is differentiable at x_0, then the error term $E(h)$ is defined for h near zero. We may extend the domain of E to $h = 0$ by just declaring $E(0) = 0$. In so doing, we note that we have ensured that E is continuous at 0 (because $E(h) \to 0$ as $h \to 0$).

Theorem 182 (Leibniz rule). *Suppose that f and g are differentiable at x_0. Let h be the product $h = fg$. Then h is differentiable at x_0 and we have*

$$h'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Proof. Note that

$$\begin{aligned}
h(x) - h(x_0) &= f(x)g(x) - f(x_0)g(x_0) \\
&= f(x)[g(x) - g(x_0)] + g(x_0)[f(x) - f(x_0)].
\end{aligned}$$

It follows that for $x \neq x_0$, we have

$$\frac{h(x) - h(x_0)}{x - x_0} = f(x)\frac{g(x) - g(x_0)}{x - x_0} + g(x_0)\frac{f(x) - f(x_0)}{x - x_0}.$$

Taking the limit as $x \to x_0$ shows that

$$h'(x_0) = \left(\lim_{x \to x_0} f(x) \right) g'(x_0) + g(x_0)f'(x_0).$$

Since f is continuous at x_0, the claim follows. $\qquad\square$

Theorem 183 (Quotient rule). *Suppose that f and g are differentiable at x_0. Let h be the quotient $h = f/g$. Then, provided $g(x_0) \neq 0$, the quotient h is differentiable at x_0 and*

$$h'(x_0) = \frac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Proof. The proof is of similar nature to the proof of the previous result, so we omit it and direct the reader to Rudin for more details. □

Theorem 184 (Chain rule). *Suppose $h = g \circ f$ is defined. Suppose f is differentiable at x_0 and g is differentiable at $f(x_0)$. Then h is differentiable at x_0 at*

$$h'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. If we define error functions F and G by

$$f(x_0 + t) - f(x_0) = t[f'(x_0) + F(t)]$$
$$g(f(x_0) + s) - g(f(x_0)) = s[g'(f(x_0)) + G(s)],$$

then $F(t) \to 0$ as $t \to 0$ and $G(s) \to 0$ as $s \to 0$. (We also define $G(0) = 0$, which ensures that G is continuous at zero.) We then compute

$$\begin{aligned}
h(x_0 + t) - h(x_0) &= g(f(x_0 + t)) - g(f(x_0)) \\
&= g[f(x_0) + \underbrace{t(f'(x_0) + F(t))}_{s}] - g(f(x_0)) \\
&= \underbrace{t[f'(x_0) + F(t)]}_{s}[g'(f(x_0)) + G(\underbrace{t[f'(x_0) + F(t)]}_{s})] \\
&= t[f'(x_0) + F(t)][g'(f(x_0)) + G(f(x_0 + t) - f(x_0))].
\end{aligned}$$

It follows that for $t \neq 0$, we have

$$\frac{h(x_0 + t) - h(x_0)}{t} = [f'(x_0) + F(t)][g'(f(x_0)) + G(f(x_0 + t) - f(x_0))].$$

Taking the limit of this as $t \to 0$, we find that

$$\lim_{t \to 0} \frac{h(x_0 + t) - h(x_0)}{t} = f'(x_0)g'(f(x_0))$$

because $F(t) \to 0$ and because the function

$$t \mapsto f(x_0 + t) - f(x_0)$$

is continuous at $t = 0$ and $G(s)$ is continuous at $s = 0$. □

5.2   Mean value theorem

Definition 185. Let f be a real function defined on a metric space X. Say that f has a **local maximum at** $p \in X$ if there is a $\delta > 0$ such that whenever $q \in X$ satisfies $d(p, q) < \delta$, then $f(q) \leqslant p$. The notion of local minimum is defined similarly.

Theorem 186. *Let f be defined on $[a, b]$. Suppose f is differentiable at $x \in (a, b)$ and f has a local maximum at x. Then $f'(x) = 0$.*

Proof. Exercise. Hint: compute the left-hand and right-hand limits separately. $\square$

Let $f : [a, b] \to \mathbb{R}$, and let $G \subset \mathbb{R}^2$ be the graph of f given by

$$G = \{(x, f(x)) : x \in [a, b]\}.$$

Note that the secant line connecting the points $(a, f(a))$ and $(b, f(b))$ has slope given by

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

The mean value theorem asserts that if f is differentiable on $[a, b]$, then this slope is equal to the slope of some tangent line.

Theorem 187 (Mean Value Theorem). *Let f be continuous on $[a, b]$ and differentiable on (a, b). Then there is a point $c \in (a, b)$ such that*

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

Proof. Let $h(t)$ be the function defined for $t \in [a, b]$ by

$$h(t) = t(f(b) - f(a)) - (b - a)f(t).$$

Then h is continuous on $[a, b]$ and differentiable on (a, b). We also have that

$$h(b) = b(f(b) - f(a)) - (b - a)f(b) = a \cdot f(b) - b \cdot f(a)$$

and

$$h(a) = a(f(b) - f(a)) - (b - a)f(a) = a \cdot f(b) - b \cdot f(a),$$

which means that $h(a) = h(b)$. If h is constant, then we are done because we can let c be any point. Otherwise, there is a point $t \in (a, b)$ such that $h(t) \neq h(a)$. Without loss of generality we may assume that $h(t) > h(a)$. Because h is continuous, there is a point $x \in (a, b)$ where h achieves its maximum. The previous result then shows that $h'(x) = 0$, which is what we require. $\square$

There is a generalized version of the previous result too.

Theorem 188 (Generalized Mean Value Theorem). *Let f, g be continuous on $[a, b]$ and differentiable on (a, b). Then there is a point $x \in (a, b)$ such that*

$$g'(x)[f(b) - f(a)] = f'(x)[g(b) - g(a)].$$

Proof. Replicate the proof above but with

$$h(x)[f(b) - f(a)]g(t) - [g(b) - g(a)]f(t)$$

instead. □

Definition 189. Let f be defined on $[a, b]$. We say that

(a) f is **monotonically increasing** if whenever $x \leqslant y$, then $f(x) \leqslant f(y)$

(b) f is **monotonically decreasing** if whenever $x \leqslant y$, then $f(x) \geqslant f(y)$.

Corollary 190. *Suppose f is differentiable on (a, b).*

(a) If $f'(x) \geqslant 0$ for each $x \in (a, b)$, then f is increasing.

(b) If $f'(x) \leqslant 0$ for each $x \in (a, b)$, then f is decreasing.

(c) If $f'(x) = 0$ for each $x \in (a, b)$, then f is constant.

Proof. The proofs follow from the equation

$$f'(c)(x - y) = f(x) - f(y)$$

which holds for each $x < y$ and for some c satisfying $x < c < y$. □

5.3 L'Hospital's rule

Theorem 191 (L'Hospital's Rule: Version 1)**.** *Let f and g be differentiable on (a, b). Suppose that*

(a) $g'(x) \neq 0$ for each $x \in (a, b)$

(b) $\lim_{x \to a} \frac{f'(x)}{g'(x)} = A$

(c) $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$.

Then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = A.$$

Proof. We proceed in steps.
Assertion 1. If q satisfies $A < q$, then there is a point $c \in (a, b)$ such that whenever $a < x < c$ we have

$$\frac{f(x)}{g(x)} < q.$$

64

Proof of Assertion 1. Let r be a number satisfying $A < r < q$. Because $f'(x)/g'(x) \to A$, there is a point $c \in (a, b)$ such that whenever $a < x < c$ we have

$$\frac{f'(x)}{g'(x)} < r.$$

If x, y satisfy $a < x < y < c$, then the generalized mean value theorem shows that there is a point $t \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r.$$

We then take the limit of this inequality as $x \to a$ and use hypothesis (c) to find that

$$\frac{f(y)}{g(y)} \leqslant r < q$$

for each $a < y < c$. This completes the proof of Assertion 1.

Assertion 2. If p satisfies $p < A$, then there is a point $d \in (a, b)$ such that whenever $a < x < d$ we have

$$p < \frac{f(x)}{g(x)}.$$

Proof of Assertion 2. The proof is similar to that of Assertion 1.

The proof now follows from Assertions 1 and 2 together. $\qquad\square$

The following version is also useful. Note that hypothesis (c) is replaced by a slightly different hypothesis (c') involving only the limit of $g(x)$ as $x \to a$.

Theorem 192 (L'Hospital's Rule: Version 2)**.** *Let f and g be differentiable on (a, b). Suppose that*

(a) $g'(x) \neq 0$ *for each* $x \in (a, b)$

(b) $\lim_{x \to a} \frac{f'(x)}{g'(x)} = A$

(c') $\lim_{x \to a} g(x) = \infty.$

Then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = A.$$

Proof. See Rudin. $\qquad\square$

5.4   Taylor's Theorem

Definition 193. If f has a derivative f' on an interval and if f' is itself differentiable, then we will denote the derivative of f' by f''. If we can continue this process, then we obtain functions

$$f, f', f'', f^{(3)}, \ldots, f^{(n)}$$

where $f^{(n)}$ is the nth derivative of f.

Theorem 194. *Let f be defined on $[a, b]$. Suppose that $f^{(n-1)}$ is continuous on $[a, b]$ and $f^{(n)}(t)$ exists for each $t \in (a, b)$. Let α and β be distinct points of $[a, b]$ and define the polynomial*

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k.$$

Then there is a point x between α and β such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n.$$

Remark 195. For the case $n = 1$, the polynomial $P(t)$ is just constant $P(t) = f(\alpha)$, and the statement is the mean value theorem.

Proof. Let M be the number defined by

$$f(\beta) = P(\beta) + M(\beta - \alpha)^n.$$

Note that our goal is to find an x between α and β such that $f^{(n)}(x) = Mn!$. Define a function g by the rule

$$g(t) = f(t) - P(t) - M(t - \alpha)^n.$$

Note that the nth derivative of g satisfies

$$g^{(n)}(t) = f^{(n)}(t) - 0 - Mn!.$$

It follows that the proof will be complete if we can find a point x between α and β such that $g^{(n)}(x) = 0$.

Specializing to the point $t = \alpha$, we note that because $f^{(k)}(\alpha) = P^{(k)}(\alpha)$ for each $k = 1, \ldots, n - 1$, we have that

$$g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0.$$

At the point $t = \beta$, the choice of M implies directly that $g(\beta) = 0$. The mean value theorem asserts that there is a point x_1 between α and β such that $g'(x_1) = 0$. For the same reason, there is a point x_2 between α and x_1 such that $g''(x_2) = 0$. Iteratively, we obtain a point x_n between α and x_{n-1} such that $g^{(n)}(x_n) = 0$. $\qquad \square$

6 Integration

6.1 Riemann integrals

Definition 196. By a **partition** P of $[a, b]$ be mean a finite set of points $\{x_0, \ldots, x_n\}$ such that $a = x_0 < x_1 < \cdots < x_n = b$. Let us use the notation Δx_i to mean

$$\Delta x_i = x_i - x_{i-1}, \qquad \text{for } i = 1, \ldots, n.$$

Definition 197. Let f be a bounded function on $[a, b]$ and let P be a partition of $[a, b]$. Set

$$M_i = \sup\{f(x) : x_{i-1} \leqslant x \leqslant x_i\}$$
$$m_i = \inf\{f(x) : x_{i-1} \leqslant x \leqslant x_i\}$$

and then define

$$U(f, P) = \sum_{i=1}^{n} M_i \Delta x_i$$
$$L(f, P) = \sum_{i=1}^{n} m_i \Delta x_i.$$

Also define

$$U(f) = \inf\{U(f, P) : P \text{ is a partition of } [a, b]\}$$
$$L(f) = \sup\{I(f, P) : P \text{ is a partition of } [a, b]\}.$$

These numbers are called the **upper and lower Riemann integrals of** f respectively. If these two numbers are equal, then we denote their common value by

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

and we say that f is **Riemann integrable on** $[a, b]$.

Definition 198. Given two partitions P, P' of $[a, b]$, we say that P' is a **refinement** of P if $P \subset P'$. Any two partitions P_1, P_2 have a common refinement $P' = P_1 \cup P_2$.

Lemma 199. *If P' refines P, then*

$$L(f, P) \leqslant L(f, P')$$

and

$$U(f, P') \leqslant U(f, P).$$

Proof. Since P and P' differ only by a finite number of points (they are both finite sets), it suffices to prove the claim for the case when P' has only one more point than P. Say that this extra point is x' and it occurs between x_{i-1} and x_i. Set

$$w_1 = \inf\{f(x) : x_{i-1} \leqslant x \leqslant x'\}$$
$$w_2 = \inf\{f(x) : x' \leqslant x \leqslant x_i\}.$$

Then we have both $m_i \leqslant w_1$ and $m_i \leqslant w_2$. We then compute that

$$L(f, P') - L(f, P) = w_1(x' - x_{i-1}) + w_2(x_i - x') - m_i(x_i - x_{i-1})$$
$$= (w_1 - m_i)(x' - x_{i-1}) + (w_2 - m_i)(x_i - x').$$

This last quantity is a sum of nonnegative terms. We deduce the result concerning the lower Riemann sums. The proof of the result for the upper Riemann sums is analogous. $\square$

Theorem 200. *We have*
$$L(f) \leqslant U(f).$$

Proof. Since $L(f)$ is defined to be a supremum, it suffices to check that $U(f)$ is an upper bound for the set

$$E = \{L(f, P) : P \text{ is a partition of } [a, b]\}.$$

Let P_1 be an arbitrary partition. Let $\epsilon > 0$ be arbitrary. By definition of the infimum, there is a partition P_2 such that

$$U(f, P_2) < U(f) + \epsilon.$$

Let $P' = P_1 \cup P_2$ be the common refinement of P_1 and P_2. Then the previous lemma shows that
$$L(f, P_1) \leqslant L(f, P') \leqslant U(f, P') < U(f) + \epsilon.$$

Since ϵ is arbtrary, we conclude that

$$L(f, P_1) \leqslant U(f).$$

This is as desired. $\square$

Theorem 201. *A function f is Riemann integrable on $[a, b]$ if and only if for each $\epsilon > 0$ there is a partition P such that*

$$U(f, P) - L(f, P) < \epsilon.$$

Proof. Suppose the latter condition is true. Let $\epsilon > 0$ be arbitrary. There is a partition P such that $U(f, P) - L(f, P) < \epsilon$. From the previous result, we have

$$L(f, P) \leqslant L(f) \leqslant U(f) \leqslant U(f, P).$$

It follows that $U(f) - L(f) < \epsilon$. Since ϵ was arbitrary, we conclude that $U(f) = L(f)$, and therefore f is Riemann integrable.

Suppose that f is Riemann integrable. Let $\epsilon > 0$ be arbitrary. By definition of the supremum and infimum, there are partitions P_1 and P_2 such that

$$\int_a^b f(x)\, dx < L(f, P_1) + \frac{\epsilon}{2}$$

$$U(f, P_2) < \int_a^b f(x)\, dx + \frac{\epsilon}{2}.$$

Let $P' = P_1 \cup P_2$ denote the common refinement of P_1 and P_2. Using the previous result, we have

$$
\begin{aligned}
U(f, P') &\leqslant U(f, P_2) \\
&< \int_a^b f(x)\, dx + \frac{\epsilon}{2} \\
&< \left(L(f, P_1) + \frac{\epsilon}{2} \right) + \frac{\epsilon}{2} \\
&\leqslant L(f, P') + \epsilon.
\end{aligned}
$$

We are done. $\qquad\square$

Theorem 202. *If f is continuous on $[a, b]$, then f is Riemann integrable on $[a, b]$.*

Proof. Let $\epsilon > 0$ be arbitrary. Choose η so small that

$$\eta(b - a) < \epsilon.$$

Since f is uniformly continuous on $[a, b]$, there is a $\delta > 0$ such that whenever $|x - y| < \delta$ we have $|f(x) - f(y)| < \eta$. Let P be a partition so fine that $\Delta x_i < \delta$ for each $i = 1, 2, \ldots, n$. Then the choice of δ shows that

$$M_i - m_i \leqslant \eta \qquad \text{for each } i = 1, 2, \ldots, n.$$

We then compute that

$$U(f, P) - L(f, P) = \sum_{i=1}^n (M_i - m_i)\Delta x_i \leqslant \eta \sum_{i=1}^n \Delta x_i = \eta(b - a) < \epsilon.$$

It follows that f is Riemann integrable by the previous result. $\qquad\square$

Theorem 203. *IF f is monotonic on $[a, b]$, then f is Riemann integrable.*

Proof. Without loss of generality, we assume that f is increasing. Let $\epsilon > 0$ be given. Let n be a positive integer so large that

$$\frac{(f(b) - f(a))(b - a)}{n} < \epsilon.$$

Choose a partition P of $[a, b]$ such that each Δx_i satisfies $\Delta x_i = (b - a)/n$. Since f is increasing, we know that the extremal values of f are achieved at the endpoints of the subintervals determined by the partition. In particular, we have

$$M_i = f(x_i), \qquad m_i = f(x_{i-1}).$$

We then have that

$$U(f, P) - L(f, P) = \frac{b - a}{n} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) = \frac{(b - a)(f(b) - f(a))}{n} < \epsilon.$$

We are done. $\square$

Theorem 204. *Suppose that the composition $h = \varphi \circ f$ makes sense. Suppose that f is Riemann integrable, f is bounded, and φ is continuous. Then h is Riemann integrable.*

Proof. Let $\epsilon > 0$ be given. Since φ is uniformly continuous, there is a δ satisfying $0 < \delta < \epsilon$ such that whenever x, y satisfy $|x - y| < \delta$ we have $|\varphi(x) - \varphi(y)| < \epsilon$. Since f is Riemann integrable, there is a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \delta^2.$$

Let M_i, m_i denote the max and min of f over $[x_i, x_{i-1}]$, and let M_i^*, m_i^* denote the max and min of h. Let A denote the subset of $\{1, \dots, n\}$ defined by

$$A = \{i : M_i - m_i < \delta\}.$$

Let B be the complement of A, so that for $i \in B$, we have $\delta \leqslant M_i - m_i$.

- If $i \in A$, then the choice of δ shows that $M_i^* - m_i^* < \epsilon$.

- Suppose $i \in B$. Because f is bounded, we may ensure that the image of f is contained in some compact set, and since φ is continuous, the restriction of φ to this compact set is bounded, say by the number $K > 0$. It follows immediately from the definition of h that $M_i^* - m_i^* \leqslant 2K$. By the definition of B and the choice of partition P, we have

$$\delta \sum_{i \in B} \Delta x_i \leqslant \sum_{i \in B} (M_i - m_i) \Delta x_i < \delta^2.$$

70

We infer that
$$\sum_{i \in B} \Delta x_i < \delta.$$

It follows that

Putting together the above, we have that

$$U(h, P) - L(h, P) = \sum_{i \in A}(M_i^* - m_i^*)\Delta x_i + \sum_{i \in B}(M_i^* - m_i^*)\Delta x_i$$
$$\leqslant \epsilon \sum_{i \in A} \Delta x_i + 2K \sum_{i \in B} \Delta x_i$$
$$< \epsilon(b - a) + 2K\delta$$
$$< \epsilon[(b - a) + 2K].$$

Since $\epsilon > 0$ was arbitrary, we obtain the result. $\qquad \square$

Lemma 205. *Let A, B be nonempty bounded subsets of $\mathbb{R}$. Let $A + B$ denote the set defined by*
$$A + B = \{a + b : a \in A, b \in B\}.$$

Then

$$\sup(A + B) = \sup(A) + \sup(B)$$
$$\inf(A + B) = \inf(A) + \inf(B).$$

Proof. Since $\sup(A)$ is an upper bound for A and $\sup(B)$ is an upper bound for B, it follows immediately that $\sup(A) + \sup(B)$ is an upper bound for $A + B$, and so we obtain the inequality $\sup(A + B) \leqslant \sup(A) + \sup(B)$.

For the other inequality, let $\epsilon > 0$ be given. By Problem 4 of Assignment 1, there are $a \in A$ and $b \in B$ such that

$$\sup(A) < a + \epsilon/2$$
$$\sup(B) < b + \epsilon/2.$$

Since $\sup(A + B)$ is an upper bound for $A + B$, it follows that

$$\sup(A) + \sup(B) < (a + b) + \epsilon \leqslant \sup(A + B) + \epsilon.$$

Since ϵ is arbitrary, we obtain $\sup(A) + \sup(B) \leqslant \sup(A + B)$.

The proof for the infimum is similar. $\qquad \square$

Theorem 206. *Suppose f and g are integrable on $[a, b]$. Then $f + g$ is integrable on $[a, b]$ and moreover*

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

Proof. Set $h = f + g$. Let $\epsilon > 0$ be given. Since f and g are integrable, there are partitions P_1, P_2 such that

$$U(f, P_1) - L(f, P_1) < \epsilon$$
$$U(g, P_2) - L(g, P_2) < \epsilon.$$

Let $P = P_1 \cup P_2$ be the common refinement. Then by a previous result, we also have

$$U(f, P) - L(f, P) < \epsilon$$
$$U(g, P) - L(g, P) < \epsilon.$$

Note that
$$M_i^h = \sup\{f(x) + g(x) : x_{i-1} \leqslant x \leqslant x_i\}.$$

The preceding lemma therefore implies that

$$M_i^h = M_i^f + M_i^g.$$

Similarly we have $m_i^h = m_i^f + m_i^g$. It follows that

$$U(h, P) = U(f, P) + U(g, P)$$
$$L(h, P) = L(f, P) + L(g, P).$$

Thus we find that

$$U(h, P) - L(h, P) = [U(f, P) - L(f, P)] + [U(g, P) - L(g, P)] < 2\epsilon.$$

Since ϵ is arbitrary, we conclude that h is integrable.

To show the statement about the integrals, it suffices to show that

$$U(h) \leqslant U(f) + U(g)$$
$$L(f) + L(g) \leqslant L(h).$$

We prove only the first, claiming that the proof of the second is similar. There are partitions P_1 and P_2 such that

$$U(f, P_1) < U(f) + \epsilon$$
$$U(g, P_2) < U(g) + \epsilon.$$

Let $P = P_1 \cup P_2$ be the common refinement. Then since $U(h)$ is an infimum, we have

$$U(h) \leqslant U(h, P) = U(f, P) + U(g, P)$$
$$\leqslant U(f, P_1) + U(g, P_2)$$
$$< U(f) + U(g) + 2\epsilon.$$

Since ϵ is arbitrary, we are done. $\qquad\qquad\square$

There are many other properties of the integral, some of which we state below, without proof.

Theorem 207. *Let f and g be integrable on $[a, b]$. The following are true*

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

(b) $\int_a^b k \cdot f(x)\, dx = k \int_a^b f(x)\, dx$ *for each constant* k

(c) *If* $f \leqslant g$ *on* $[a, b]$, *then* $\int_a^b f(x)\, dx \leqslant \int_a^b g(x)\, dx$

(d) *If* c *is a point satisfying* $a < c < b$, *then* $\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx$.

(e) $\int_a^b f(x + x_0)\, dx = \int_{a+x_0}^{b+x_0} f(x)\, dx$.

Corollary 208. *If f and g are integrable on $[a, b]$, then so is fg.*

Proof. If $\varphi(t) = t^2$, which is continuous, then Theorem 204 shows that f^2 is integrable. The result then follows from the identity

$$4fg = (f + g)^2 - (f - g)^2.$$

$\square$

Corollary 209. *If f is integrable, then so is $|f|$ and moreover*

$$\left| \int_a^b f(x) dx \right| \leqslant \int_a^b |f(x)|\, dx.$$

Proof. Theorem 204 shows that $|f|$ is integrable because $\varphi(t) = |t|$ is continuous. The statement about the inequality follows from the identity

$$-|f| \leqslant f \leqslant |f|$$

and part (c) of Theorem 207.

$\square$

6.2   Integration and differentiation

Theorem 210 (Fundamental Theorem of Calculus I). *Let f be a bounded integrable function on $[a, b]$. For x satisfying $a \leqslant x \leqslant b$, define*

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

Then F is continuous on $[a, b]$. Moreover, if f is continuous at a point $x_0 \in [a, b]$, then F is differentiable at x_0 and

$$F'(x_0) = f(x_0).$$

Proof. Let $\epsilon > 0$ be given. Since f is bounded, there is an $M > 0$ such that $|f(x)| \leqslant M$ for each $x \in [a, b]$. Let δ be smaller than ϵ/M. Then we compute that whenever $a \leqslant x < y \leqslant b$ and $y - x < \delta$, we have

$$|F(x) - F(y)| = \left| \int_x^y f(t)\, dt \right| \leqslant \int_x^y |f(t)|\, dt = M(y - x) < \epsilon.$$

This shows that F is continuous.

Suppose that f is continuous at x_0. Let $\epsilon > 0$ be arbitrary. There is a $\delta > 0$ such that whenever h satisfies $|h| < \delta$, then $|f(x_0 + h) - f(x_0)| < \epsilon/2$. We have that

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h} \int_{x_0}^{x_0+h} f(t)\, dt = \frac{1}{h} \int_0^h f(t + x_0)\, dt$$

and also that

$$f(x_0) = \frac{1}{h} \int_0^h f(x_0)\, dt.$$

We then compute that if h satisfies $|h| < \delta$, then we have

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| = \frac{1}{h} \left| \int_0^h (f(t + x_0) - f(x_0))\, dt \right|$$

$$\leqslant \frac{1}{h} \int_0^h |f(t + x_0) - f(x_0)|\, dt$$

$$\leqslant \frac{1}{h} \cdot h \cdot \left(\frac{\epsilon}{2} \right) < \epsilon.$$

It follows that $F'(x_0) = f(x_0)$. $\qquad\square$

Theorem 211 (Fundamental Theorem of Calculus II). *Let f be integrable on $[a, b]$. If F is any function such that $F' = f$, then*

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

Proof. Let $\epsilon > 0$ be given. Because f is integrable, there is a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \epsilon.$$

Because $F' = f$, the mean value theorem asserts that there are points $t_i \in [x_{i-1}, x_i]$ such that

$$F(x_i) - F(x_{i-1}) = f(t_i)\Delta x_i \tag{4}$$

Because $m_i \leqslant f(t_i) \leqslant M_i$, we have that the sum $\sum_i f(t_i)\Delta x_i$ satisfies

$$L(f, P) \leqslant \sum_i f(t_i)\Delta x_i \leqslant U(f, P).$$

74

It follows from our choice of P that

$$\left| \sum_i f(t_i)\Delta x_i - \int_a^b f(x)\,dx \right| < \epsilon.$$

But the sum $\sum_i f(t_i)\Delta x_i$ is telescoping from relation (4) and is equal to

$$F(b) - F(a) = \sum_i f(t_i)\Delta x_i$$

so that in fact we have

$$\left| (F(b) - F(a)) - \int_a^b f(x)\,dx \right| < \epsilon.$$

Because ϵ is arbitrary, we are done. $\qquad\square$

Theorem 212 (Change of variables). *Let $\varphi : [a,b] \to [c,d]$ and $f : [c,d] \to \mathbb{R}$. Assume*

(a) *f is a continuous*

(b) *the derivative $\varphi'(x)$ exists, is positive everywhere, and is integrable on $[a,b]$*

(c) *φ is surjective onto $[c,d]$.*

Then the composition $g(x) = f(\varphi(x))$ is integrable on $[a,b]$ and

$$\int_a^b g(x)\,\varphi'(x)dx = \int_c^d f(y)\,dy.$$

Proof. Let $F(y) = \int_a^y f(t)\,dt$. Then the first fundamental theorem tells us that $F'(y) = f(y)$ because f is assumed to be continuous by hypothesis (a). Let $G(x)$ be the composition $G(x) = F(\varphi(x))$. Then the chain rule implies that

$$G'(x) = F'(\varphi(x))\varphi'(x) = f(\varphi(x))\varphi'(x) = g(x)\varphi'(x)$$

by the definition of g. We then just apply the fundamental theorem twice to obtain

$$\begin{aligned}
\int_a^b g(x)\,\varphi'(x)\,dx &= G(b) - G(a) \\
&= F(\varphi(b)) - F(\varphi(a)) \\
&= \int_{\varphi(a)}^{\varphi(b)} f(y)\,dy \\
&= \int_c^d f(y)\,dy,
\end{aligned}$$

where the equalities $c = \varphi(a)$ and $d = \varphi(b)$ follow from the fact that φ is bijective and increasing. $\qquad\square$

Theorem 213 (Integration by parts). *Suppose F and G are differentiable functions on $[a, b]$ and their derivatives f and g are integrable. Then*

$$\int_a^b F(x)g(x)\,dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)\,dx.$$

Proof. Let $H(x)$ be the product $H(x) = F(x)G(x)$. Note that the derivative of H is given by

$$H'(x) = f(x)G(x) + F(x)g(x).$$

Then the fundamental theorem asserts that

$$\int_a^b H'(x) = H(b) - H(a).$$

Rewriting this, we find that

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

This is the result. $\qquad\square$

6.3 Arc length of curves

Definition 214. By a **curve** in $\mathbb{R}^k$ we mean a continuous mapping $\gamma : [a, b] \to \mathbb{R}^k$. If γ is injective, we say that γ is an **arc**. If $\gamma(a) = \gamma(b)$ we say that γ is a **closed curve**.

Example 215. Let $\gamma : [0, \pi]$ be the curve in $\mathbb{R}^2$ given by $\gamma(t) = (\cos(t), \sin(t))$. Then γ is an arc.

If the domain of γ is extended to $[0, 2\pi]$, then γ is a closed curve.

Definition 216. To each partition P of $[a, b]$ associate the number

$$\Lambda(\gamma, P) = \sum_{i=1}^n |\gamma(x_i) - \gamma(x_{i-1})|.$$

The number $\Lambda(\gamma, P)$ represents the length of the polygonal path approximating the image of γ with vertices at the points $\gamma(x_i)$. We define the **length of** γ to be

$$\Lambda(\gamma) = \sup\{\Lambda(\gamma, P) : P \text{ a partition of } [a, b]\},$$

provided this supremum is finite. In the case that the supremum is finite, we say that γ is **rectifiable**.

Theorem 217. *Suppose γ admits a continuous derivative γ' on $[a, b]$. Then γ is rectifiable and*

$$\Lambda(\gamma) = \int_a^b |\gamma'(t)| \, dt.$$

Proof. From the fundamental theorem of calculus applied to each component of γ, we have

$$|\gamma(x_i) - \gamma(x_{i-1})| = \left| \int_{x_{i-1}}^{x_i} \gamma'(t) \, dt \right| \leqslant \int_{x_{i-1}}^{x_i} |\gamma'(t)| \, dt.$$

It follows therefore that

$$\Lambda(\Lambda) \leqslant \int_a^b |\gamma'(t)| \, dt.$$

Let's prove the other inequality.

Let $\epsilon > 0$ be given. Since γ' is uniformly continuous, there is a $\delta > 0$ such that $|\gamma'(s) - \gamma'(t)| < \epsilon$ whenever $|s - t| < \delta$. Let P be a partition so fine that $\Delta x_i < \delta$ for each i. Then if t satisfies $x_{i-1} \leqslant t \leqslant x_i$, we have that

$$|\gamma'(t)| \leqslant |\gamma'(x_i)| + \epsilon. \tag{5}$$

Upon integrating both sides of this inequality along $[x_{i-1}, x_i]$ we find that

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)| \, dt \leqslant |\gamma'(x_i)| \Delta x_i + \epsilon \Delta x_i = \left| \int_{x_{i-1}}^{x_i} [\gamma'(t) + \gamma'(x_i) - \gamma'(t)] \, dt \right| + \epsilon \Delta x_i.$$

An application of the triangle inequality shows that

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)| \, dt \leqslant \left| \int_{x_{i-1}}^{x_i} \gamma'(t) \, dt \right| + \left| \int_{x_{i-1}}^{x_i} [\gamma'(x_i) - \gamma'(t)] \, dt \right| + \epsilon \Delta x_i.$$

Now we use the fundamental theorem of calculus and (5) again to find that

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)| \, dt \leqslant |\gamma(x_i) - \gamma(x_{i-1})| + \epsilon \Delta x_i + \epsilon \Delta x_i.$$

If we add up all such inequalities, we find that

$$\int_a^b |\gamma'(t)| \, dt \leqslant \Lambda(\gamma, P) + 2\epsilon(b - a).$$

Since $\epsilon > 0$ was arbitrary, we obtain

$$\int_a^b |\gamma'(t)| \, dt \leqslant \Lambda(\gamma, P)$$

as desired. $\square$

7 Sequences and series of functions

7.1 Pointwise convergence

Definition 218. Let E be a subset of a metric space X, and let f_n be a sequence of real-valued functions defined on E. (This means that each f_n is a function $f_n : E \to \mathbb{R}$.) We say that f_n **converges (pointwise)** if there is a function $f : E \to \mathbb{R}$ with the property that for each $x \in E$, the sequence of real numbers $f_n(x)$ converges to $f(x)$. In this case, the function f is called the **limit** of the sequence f_n.

Example 219. Let $f_n : [0, 1] \to \mathbb{R}$ be the sequence

$$f_n(x) = x^n.$$

Note that whenever $x \neq 1$, we have $x^n \to 0$ as $n \to \infty$. On the other hand, for $x = 1$, we have $1^n \to 1$ as $n \to \infty$. It follows that the limit of the sequence f_n is given by the function

$$f(x) = \begin{cases} 1 & x = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Remark 220. It is also important to remark that a more precise version of the definition of convergence of functions can be written as follows. For each point $x \in E$ and each $\epsilon > 0$, there is an $N > 0$ such that $|f_n(x) - f(x)| \leqslant \epsilon$ whenever $n \geqslant N$. In particular, in this version of the definition, it is apparent that the choice of $N = N(x, \epsilon)$ can depend both on ϵ *and* the point x. Indeed, for the previous example, we see that as x approaches 1, the choice of N must become larger to ensure that x^n is small.

Remark 221. The previous example also illustrates that in general fairly disastrous things can happen to the limit function. In particular, each of the functions f_n in the example are continuous, but the limiting function f fails to be so. The next example presents an even worse situation.

Example 222. For a positive integer m, let $f_m : \mathbb{R} \to \mathbb{R}$ be defined by

$$f_m(x) = \lim_{n \to \infty} (\cos(m!\pi x))^{2n}.$$

We note that if $m!x$ is an integer, then $f_m(x) = 1$. For all other values of x, we have $f_m(x) = 0$. We summarize as follows

$$f_m(x) = \begin{cases} 1 & m!x \text{ is an integer} \\ 0 & \text{otherwise.} \end{cases}$$

Now, when x is rational, there is a positive integer M such that $m!x$ is an integer whenever $m \geqslant M$. It follows that for x rational, we have $\lim_{m \to \infty} f_m(x) = 1$. On the

other hand, for x irrational, we see that $f_m(x) = 0$, and so $\lim_{m \to \infty} f_m(x) = 0$ in this case. We realize that the limit function f is defined by

$$f(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & \text{otherwise.} \end{cases}$$

In particular, f is not Riemann integrable on any closed subinterval, even though each f_m is.

Example 223. Let $f_n : [0, 1] \to \mathbb{R}$ be the sequence of integrable functions

$$f_n(x) = \begin{cases} n & 0 < x < 1/n \\ 0 & \text{otherwise.} \end{cases}$$

Note that the integral of each f_n satisfies

$$\int_0^1 f_n(x)\, dx = 1.$$

Also whenever $x \in (0, 1]$, the sequence of values $f_n(x)$ is eventually zero (provided $n > 1/x$.) It follows that the sequence converges to the zero function $f(x) = 0$. However, note that in this case, the sequence of integrals do not converge because

$$\int_0^1 f(x)\, dx = 0.$$

This example shows that even though the limiting function may be integrable, its integral may not be the expected number.

7.2 Uniform convergence

Definition 224. Let f_n be a sequence of functions on E, and let f be a function on E. We say that f_n **converges uniformly** to f if for each $\epsilon > 0$ there is an $N > 0$ such that whenever $n \geqslant N$, we have

$$|f_n(x) - f(x)| < \epsilon \qquad \text{for each } x \in E.$$

In other words, the choice of N is only allowed to depend on ϵ.

Corollary 225. *If f_n converges uniformly to f, then f_n converges (pointwise) to f.*

Proof. Immediate. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Example 226. For each n, let $f_n : [0, 1] \to \mathbb{R}$ be

$$f_n(x) = \frac{x^2}{x^2 + (1 - nx)^2}.$$

Then for each fixed $x \in [0, 1]$, the denominator of $f_n(x)$ grows large as $n \to \infty$, and so we find that the (pointwise) limit function is

$$f(x) = 0.$$

However, we claim that the convergence is not uniform. Indeed let $\epsilon = 1/2$, and let $N > 0$ be arbitrary. Then at the point $x = 1/N$, we have that

$$f_N \left(\frac{1}{N} \right) = 1,$$

and so

$$\left| f \left(\frac{1}{N} \right) - f_N \left(\frac{1}{N} \right) \right| \geq 1 > \epsilon.$$

Lemma 227. *Let f_n be a sequence of functions on E. There there is a function f on E which is the uniform limit of the sequence f_n if and only if for each $\epsilon > 0$ there is an $N > 0$ such that whenever $m, n \geq N$ we have*

$$|f_n(x) - f_m(x)| < \epsilon \qquad \text{for each } x \in E.$$

Proof. Suppose that $f_n \to f$ uniformly. Let $\epsilon > 0$ be given. There is an $N > 0$ such that whenever $n \geq N$ we have

$$|f_n(x) - f(x)| < \epsilon/2 \qquad \text{for each } x \in E.$$

It then follows from the triangle inequality that for $m, n \geq N$ and $x \in E$ have

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \epsilon.$$

Suppose on the other hand that the latter condition in the statement of the theorem is satisfied. Let $\epsilon > 0$ be arbitrary. There is an $N > 0$ such that whenever $m, n \geq N$, we have

$$|f_n(x) - f_m(x)| < \epsilon/2 \qquad \text{for each } x \in E. \tag{6}$$

It follows that for each fixed x, the sequence of numbers $f_n(x)$ is Cauchy in $\mathbb{R}$, and hence converges to some number $f(x)$. This defines a limiting function $f(x)$. We prove now that the convergence is uniform. Keeping n fixed and letting $m \to \infty$ in (6) we have that

$$|f_n(x) - f(x)| \leq \epsilon/2 < \epsilon \qquad \text{for each } x \in E.$$

This completes the proof. $\qquad \square$

Theorem 228 (Weierstrass M-test)**.** *Let f_n be a sequence of functions defined on E. Suppose that there is a sequence of real numbers M_n such that*

$$|f_n(x)| \leqslant M_n \qquad \text{for each } x \in E.$$

If the series $\sum_n M_n$ converges, then the series $\sum_n f_n$ converges uniformly on E.

Proof. Suppose $\sum_n M_n$ converges. Let $\epsilon > 0$ be given. If m and n are large enough, we have

$$\left| \sum_{k=n}^{m} f_k(x) \right| \leqslant \sum_{k=n}^{m} M_k < \epsilon.$$

We conclude that the series $\sum_n f_n$ converges uniformly from Lemma 227. $\qquad \square$

7.3   Uniform convergence and continuity

Theorem 229. *Suppose f_n converges uniformly to f and each f_n is continuous. Then f is continuous.*

Proof. Let x_0 be a fixed point of E. We will show that f is continuous at x_0. Let $\epsilon > 0$ be given. Because $f_n(x_0) \to f(x_0)$, there is an $N_1 > 0$ such that

$$|f_n(x_0) - f(x_0)| < \epsilon/3 \qquad \text{whenever } n \geqslant N_1.$$

Because $f_n \to f$ uniformly, there is an N_2 such that

$$|f_n(x) - f(x)| < \epsilon/3 \qquad \text{for each } x \in E \text{ whenever } n \geqslant N_2.$$

Let n_0 be larger than N_1 and N_2. Then because f_{n_0} is continuous at x_0, there is a $\delta > 0$ such that

$$|f_{n_0}(x) - f_{n_0}(x_0)| < \epsilon/3 \qquad \text{whenever } |x - x_0| < \delta.$$

We then have for $|x - x_0| < \delta$ that

$$|f(x) - f(x_0)| \leqslant |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)|$$
$$< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

We conclude that f is continuous at x_0. $\qquad \square$

Example 230. Let $f_n : [0, 1] \to \mathbb{R}$ be $f_n(x) = x^n$. We saw that $f_n \to f$ where

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

Since f is discontinuous, the previous result implies that the convergence is not uniform.

Definition 231. For a metric space X, let $C(X)$ denote the set of all real-valued continuous and bounded functions on X. The set $C(X)$ is eqipped with a metric defined as follows. For $f \in X$, we define the **supremum norm** to be the number

$$\|f\| = \sup_{x \in X} |f(x)|.$$

We then set

$$d(f, g) = \|f - g\|.$$

Corollary 232. *A sequence f_n in $C(X)$ converges to f if and only if f_n converges uniformly to f on X.*

Proof. This is just a restatement of the definitions. □

Theorem 233. *The space $C(X)$ is complete.*

Proof. Let f_n be a Cauchy sequence in X. It follows from Lemma 227 that there is a function f defined on X such that $f_n \to f$ uniformly on X. This means also that $f_n \to f$ in $C(X)$. By Theorem 229, the function f is continuous.

It remains to show that f is bounded, which we do now. There is an $n > 0$ such that

$$|f_n(x) - f(x)| < 1 \qquad \text{for each } x \in E.$$

It therefore follows that

$$\|f\| \leqslant \|f_n\| + 1.$$

We are done. □

7.4   Uniform convergence and integration

Theorem 234. *Let f_n be a sequence of functions on $[a, b]$. Suppose each f_n is integrable and $f_n \to f$ uniformly on $[a, b]$. Then f is integrable and moreover*

$$\int_a^b f(x)\,dx = \lim_{n \to \infty} \int_a^b f_n(x)\,dx.$$

Proof. Let $\epsilon > 0$. There is an $N > 0$ such that for $n \geqslant N$ we have

$$f_n(x) - \epsilon f(x) \leqslant f_n(x) + \epsilon \qquad \text{for } x \in [a, b].$$

From this, it follows that for $n \geqslant N$, we have

$$U(f) \leqslant \int_a^b (f_n + \epsilon)\,dx = \epsilon(b - a) + \int_a^b f_n(x)\,dx.$$

$$L(f) \geqslant \int_a^b (f_n - \epsilon)\,dx = -\epsilon(b - a) + \int_a^b f_n(x)\,dx.$$

Taking $\epsilon \to 0$, we conclude that $U(f) = L(f)$ and f is integrable.

In addition, we find that for $n \geqslant N$, we have

$$\left| \int_a^b f(x) \, dx - \int_a^b f_n(x) \, dx \right| < \epsilon(b - a).$$

The statement about the integrals now follows. □

Corollary 235. *If each f_n is integrable on $[a, b]$ and the series*

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

converges uniformly on $[a, b]$, then f is integrable and

$$\int_a^b f(x) \, dx = \sum_{n=1}^{\infty} \int_a^b f_n(x) \, dx.$$

7.5 Uniform convergence and differentiation

Lemma 236. *Let E be a subset of X, and let x_0 be a limit point of E. Suppose $f_n \to f$ uniformly on E. Then*

$$\lim_{x \to x_0} f(x) = \lim_{n \to \infty} \lim_{x \to x_0} f(x).$$

Proof. Exercise or see Rudin. (Hint: Follow the proof of 229.) □

Theorem 237. *Let f_n be a sequence of differentiable functions on $[a, b]$, and let x_0 be a point of $[a, b]$. Suppose*

(a) the derivatives f_n' converge uniformly on $[a, b]$

(b) the sequence of values $f_n(x_0)$ converges.

Then f_n converges uniformly on $[a, b]$ to a function f such that $f_n' \to f'$.

Proof. Let $\epsilon > 0$ be given. Hypotheses (a) and (b) guarantee the existence of an N such that

$$|f_n(x_0) - f_m(x_0)| < \epsilon/2 \qquad \text{whenever } n \geqslant N$$
$$|f_n'(t) - f_m'(t)| < \frac{\epsilon}{2(b - a)} \qquad \text{whenever } n \geqslant N \text{ and } t \in [a, b].$$

The second inequality implies by the mean value theorem that

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leqslant \frac{|x - t|\epsilon}{2(b - a)} \leqslant \frac{\epsilon}{2} \qquad \text{for each } x, t \in [a, b].$$

83

But then we also have from the triangle inequality that

$$|f_n(x) - f_m(x)| \leqslant |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|.$$

Putting together all of the inequalities, we find that

$$|f_n(x) - f_m(x)| \leqslant \epsilon \qquad \text{for } m, n \geqslant N \text{ and } x \in [a, b].$$

This shows that f_n converges uniformly on $[a, b]$.

Now let's prove the statement about the derivatives. Let f be the limit of the f_n's. Let x be a fixed point in $[a, b]$. Define a sequence of difference quotients

$$\varphi_n(t) = \frac{f_n(t) - f_n(x)}{t - x}$$

and the limiting difference quotient

$$\varphi(t) = \frac{f(t) - f(x)}{t - x}.$$

The functions φ_n and φ are defined near x with x being a limit point in their domains. If we could show that $\varphi_n \to \varphi$ away from x, then we would be done by the lemma. But a previous inequality implies that for each $\epsilon > 0$ there is an N such that

$$|\varphi_n(t) - \varphi_m(t)| \leqslant \frac{\epsilon}{2(b - a)} \qquad \text{for } n \geqslant N \text{ and } t \in [a, b].$$

This means that the sequence φ_n is uniformly Cauchy, and hence converges to a limiting function. At least when $t \neq x$, we conclude from the definition of φ and the fact that $f_n \to f$ that we in fact have $\varphi_n \to \varphi$. We are done. $\qquad \square$

Theorem 238. *There is a continuous function on $\mathbb{R}$ which is nowhere differentiable.*

Proof. Define $\varphi : [-1, 1] \to \mathbb{R}$ by $\varphi(x) = |x|$. Extend the definition of φ to all real x by requiring that

$$\varphi(x + 2) = \varphi(x) \qquad \text{for each } x \in \mathbb{R}.$$

Then φ satisfies

$$|\varphi(x) - \varphi(y)| \leqslant |x - y| \tag{7}$$

and so in particular φ is continuous on $\mathbb{R}$. Because $|\varphi(x)| \leqslant 1$, the Weierstrass M-test shows that the series

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x)$$

converges uniformly on $\mathbb{R}$. By Theorem 229 the function f is continuous on $\mathbb{R}$.

We show now that f is nowhere differentiable. Let x be a point of $\mathbb{R}$. For each positive integer m, let δ_m be the number

$$\delta_m = \frac{1}{2 \cdot 4^m}.$$

Because $4^m \delta_m = 1/2$, it is impossible that both of the intervals $[4^m x, 4^m(x + \delta_m)]$ and $[4^m(x - \delta_m), 4^m x]$ contain an integer. Let σ_m be the sign ($+$ or $-$) such that no integer lies between $4^m x$ and $4^m(x + \sigma_m \delta_m)$. Set

$$\gamma_{n,m} = \frac{\varphi(4^n(x + \sigma_m \delta_m)) - \varphi(4^n x)}{\sigma_m \delta_m}.$$

When $n > m$, the number $4^n \delta_m$ is an even integer, so that $\gamma_{n,m} = 0$ in this case. When n satisfies $0 \leqslant n \leqslant m$, then equation (7) implies that

$$|\gamma_{n,m}| \leqslant \frac{|4^n(x + \sigma_m \delta_m) - 4^n x|}{\delta_m} = 4^n.$$

When $n = m$, we have that

$$|\gamma_{m,m}| = \frac{|4^m(x + \sigma_m \delta m) - 4^m x|}{\delta_m} = 4^m.$$

Then let us compute

$$\left| \frac{f(x + \sigma_m \delta_m) - f(x)}{\sigma_m \delta_m} \right| = \left| \sum_{n=0}^{\infty} \left(\frac{3}{4} \right)^n \gamma_{n,m} \right|$$

$$= \left| \sum_{n=0}^{m} \left(\frac{3}{4} \right)^n \gamma_{n,m} \right|$$

$$= \left| \left(\frac{3}{4} \right)^m \gamma_{m,m} - \left(-\sum_{n=0}^{m} \left(\frac{3}{4} \right)^n \gamma_{n,m} \right) \right|$$

$$\geqslant 3^m - \sum_{n=0}^{m-1} 3^n$$

$$= 3^m - \frac{1 - 3^m}{1 - 3}$$

$$= 3^m - \frac{3^m - 1}{2}$$

$$= \frac{3^m + 1}{2}.$$

This shows that the difference quotient is bounded below as we vary m. But as $m \to \infty$, we have $\delta_m \to 0$. This implies that f cannot be differentiable at x. Since x is arbitrary, we conclude that f is nowhere differentiable. $\qquad \square$

7.6 Arzela-Ascoli

Recall that $C([a, b])$ denotes the metric space of continuous functions on $[a, b]$ together with the supremum norm. The question we turn to now is which subsets $A \subset C([a, b])$ are compact.

The celebrated Arzela-Ascoli theorem gives sufficient conditions for a subset A to be compact.

Definition 239. Let A be a family of continuous functions on $[a, b]$.

- We say that A is **uniformly bounded** if there is a constant $M > 0$ such that

$$|f(x)| \leqslant M \qquad \text{whenever } f \in A \text{ and } x \in [a, b].$$

- We say that A is **equicontinuous** if for each $\epsilon > 0$ there is a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon \qquad \text{whenever } |x - y| < \delta \text{ and } f \in A.$$

Corollary 240. *If A is equicontinuous, then each $f \in A$ is uniformly continuous.*

Proof. The proof is immediate from the definitions. $\qquad\qquad\square$

Example 241. Let us turn to our favorite sequence of $f_n : [0, 1] \to \mathbb{R}$ defined by $f_n(x) = x^n$. Let A denote the family $A = \{f_n : n = 1, 2, \ldots\}$. Then this family is uniformly bounded (by the constant 1). However, we claim that A is not equicontinuous. Indeed, let $\epsilon = 1/2$. Let $\delta > 0$ be arbitrary. Let $x \in [0, 1]$ be a point satisfying $0 < |x - 1| < \delta$. Because $x^n \to 0$, there is an n such that $|x^n - 1| \geqslant 1/2 = \epsilon$.

Theorem 242. *If a sequence f_n in $C([a, b])$ is uniformly bounded and equicontinuous, then it has a uniformly convergent subsequence.*

Proof. One proof consists of a diagonalization argument along the countable dense subset given by the rational points in $[a, b]$. Let x_n be an enumeration of the set $\mathbb{Q} \cap [a, b]$.

Consider the point x_1. Because the sequence f_n is uniformly bounded, the sequence of values $f_n(x_1)$ is bounded, and hence admits a convergent subsequence. It follows that by passing to a subsequence $f_{1,n}$ of f_n, we may ensure that $f_{1,n}$ converges at x_1.

Suppose for a fixed $k \geqslant 1$ that we have a subsequence $f_{k,n}$ of f_n such that $f_{k,n}$ converges at the points $x_1, \ldots, x_k$. Because the sequence of values $f_{k,n}(x_{k+1})$ is bounded, it has a convergent subsequence. Hence, by passing to another subsequence $f_{k+1,n}$, we may ensure that $f_{k+1,n}$ converges at the points $x_1, \ldots, x_{k+1}$.

Thus, we have obtained the following. For each positive integer k, there is a subsequence $f_{k,n}$ of f_n which converges at the points $x_1, \ldots, x_k$. Moreover, we have ensured that the subsequences are nested in the sense that

$$\{f_n\} \supset \{f_{1,n}\} \supset \{f_{2,n}\} \supset \cdots$$

Construct a the new subsequence given by the diagonal subsequence $\{f_{n,n}\}$. By construction, the subsequence $f_{n,n}$ converges at each x_ℓ, and hence at each rational point in $[a, b]$.

Because the sequence $f_{n,n}$ is equicontinuous, there is a $\delta > 0$ such that

$$|f_{n,n}(x) - f_{n,n}(y)| < \epsilon/3 \qquad \text{whenever } |x - y| < \delta \text{ for each } n = 1, 2, \ldots.$$

Choose a finite number of rational points $q_1, \ldots, q_\ell$ such that the collection $B_\delta(q_j)$ covers $[a, b]$ (which we can do because $[a, b]$ is compact). Because $f_{n,n}$ converges at each q_j, there is an N_j such that

$$|f_{n,n}(q_j) - f_{m,m}(q_j)| < \epsilon/3 \qquad \text{whenever } n \geqslant N_j.$$

Let $N = \max_{1 \leqslant j \leqslant \ell} N_j$. Let n be any positive integer satisfying $n \geqslant N$ and let x be any point of $[a, b]$. Because the balls cover $[a, b]$, there is a q_j such that x belongs to $B_\delta(q_j)$. Then we have that

$$
\begin{aligned}
&|f_{n,n}(x) - f_{m,m}(x)| \\
&\leqslant |f_{n,n}(x) - f_{n,n}(q_j)| + |f_{n,n}(q_j) - f_{m,m}(q_j)| + |f_{m,m}(q_j) - f_{m,m}(x)| \\
&< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.
\end{aligned}
$$

This shows that the sequence $f_{n,n}$ is uniformly Cauchy, as required. $\qquad \square$

Using some basic topology, namely, that a space is compact if and only if it is sequentially compact, we can obtain the following restatement of the Arzela-Ascoli theorem.

Corollary 243. *Let A be a subset of $C([a, b])$. If A is uniformly bounded and equicontinuous, then the closure of A is compact. In particular, if A is is uniformly bounded and equicontinuous and closed with respect to the supremum norm, then A is compact.*

Proof. We won't prove this result here but the proof just follows from the basic topological fact that any sequentially compact space is also compact. $\qquad \square$

Definition 244. Let us say that a function $f : [a, b] \to \mathbb{R}$ belongs to $C^1([a, b])$ if f is differentiable and the derivative f' is continuous on $[a, b]$. For this space $C^1([a, b])$, let us define the norm

$$\|f\|_{C^1} = \sup_{x \in [a,b]} |f(x)| + \sup_{x \in [a,b]} |f'(x)|.$$

This norm induces a metric on $C^1([a, b])$ in the usual way.

Corollary 245. *Let f_n be a sequence in $C^1([a, b])$. If f_n is uniformly bounded with respect to the C^1-norm, then there is a subsequence f_{n_k} that converges in $C^0([a, b])$.*

Proof. Suppose $M > 0$ is such that

$$\|f_n\|_{C^1} \leqslant M \qquad \text{for each } n = 1, 2, \ldots.$$

From the definition of the C^1-norm, we have that

$$|f_n(x)| \leqslant M \qquad \text{for each } x \in [a, b] \text{ and each } n = 1, 2, \ldots$$
$$|f_n'(x)| \leqslant M \qquad \text{for each } x \in [a, b] \text{ and each } n = 1, 2, \ldots.$$

It follows that the sequence f_n is uniformly bounded with respect to the C^0-norm. The proof would follow from the Arzela-Ascoli theorem if we could show that the sequence f_n is equicontinuous, which we do now.

Let $\epsilon > 0$ be given. The mean value theorem together with the previous paragraph implies that

$$|f_n(x) - f_n(y)| \leqslant M|x - y| \qquad \text{for each } x, y \in [a, b] \text{ and each } n = 1, 2, \ldots.$$

If we choose $\delta = \epsilon/2M$, then it follows that

$$|f_n(x) - f_n(y)| \leqslant \epsilon/2 < \epsilon \qquad \text{whenever } |x - y| < \delta \text{ for each } n = 1, 2, \ldots,$$

as desired. $\qquad \square$

Definition 246. More generally, say that $f : [a, b] \to \mathbb{R}$ belongs to $C^k([a, b])$ if f admits k derivatives $f', f^{(2)}, \ldots, f^{(k)}$ each of which is continuous. Define the C^k-norm by

$$\|f\|_{C^k} = \sum_{\ell=0}^{k} \sup_{x \in [a,b]} |f^{(\ell)}(x)|.$$

Corollary 247. *Let f_n be a sequence in $C^{k+1}([a, b])$. If the sequence f_n is uniformly bounded with respect to the C^{k+1}-norm, then there is a subsequence f_{n_m} that converges in $C^k([a, b])$.*

Proof. Exercise. $\qquad \square$

The Arzela-Ascoli theorem has a generalization to any compact metric space X.

Theorem 248. *Let X be a compact metric space, and let A be a subset of $C(X)$. If A is uniformly bounded and equicontinuous, then the closure of A is compact.*

Proof. The proof will be analogous to the case of $[a, b]$ if we can show that there is a countable dense subset E of X. So this is what we will show.

For each n, the balls $\{B_{1/n}(x)\}_{x \in X}$ cover X, and since X is compact, there is a finite subset $E_n \subset X$ of centers such that

$$X \subset \bigcup_{x \in E_n} B_{1/n}(x).$$

Define

$$E = \bigcup_{n=1}^{\infty} E_n.$$

Then E is countable as it is the countable union of finite sets. Moreover, we claim that E is dense in X. Indeed, suppose that $E \neq X$ and let x_0 be a point in $X \setminus E$. Let $\epsilon > 0$ be given. Choose n_0 so large that $n > 1/\epsilon$. Then since

$$X \subset \bigcup_{x \in E_n} B_{1/n}(x)$$

there is a $x \in E_n$ such that $d(x, x_0) < 1/n < \epsilon$. Since $x_0 \notin E_n$, we see that x_0 is a limit point of E_n and hence of E. It follows that $\overline{E} = X$. We are done. $\qquad\square$

7.7 Stone-Weierstrass Theorem

Lemma 249. *For any positive integer n, we have*

$$(1 - x^2)^n \geqslant 1 - nx^2$$

for $-1 < x < 1$.

Proof. Let $f(x)$ be the function

$$f(x) = (1 - x^2)^n - 1 + nx^2.$$

To prove the claim, it suffices to prove that f is nonnegative on $(-1, 1)$. Since f is even, it actually suffices to prove that f is nonnegative on $[0, 1)$. We note that the derivative satisfies

$$f'(x) = n(1 - x^2)^{n-1} + 2nx.$$

It follows that for $x \in (0, 1)$, the number $f'(x)$ is positive. From the mean value theorem, we infer that f is monotonically increasing on $[0, 1)$. Since $f(0) = 0$, we conclude that f is nonnegative on $[0, 1)$, as desired. $\qquad\square$

Lemma 250. *Let $f : [0, 1] \to \mathbb{R}$ be a continuous function and suppose that $f(0) = f(1) = 0$. Then there is a sequence of polynomials p_n on $[0, 1]$ converging uniformly to f on $[0, 1]$.*

Proof. We may extend f in a continuous manner to the whole real line by declaring f to be zero outside of $[0, 1]$. In this way, f becomes uniformly continuous on $\mathbb{R}$.

For each $n = 1, 2, \ldots$, let

$$c_n = \int_{-1}^{1} (1 - x^2)^n \, dx,$$

89

and let q_n be the polynomial

$$q_n(x) = c_n^{-1}(1 - x^2)^n.$$

In this way, we have ensured that $\int_{-1}^{1} q_n(x)\, dx = 1$.

We claim that c_n satisfies $c_n > 1/\sqrt{n}$. Indeed, we compute using the lemma that

$$
\begin{aligned}
c_n &= \int_{-1}^{1} (1 - x^2)^n\, dx \\
&= 2 \int_{0}^{1} (1 - x^2)^n\, dx \\
&\geqslant 2 \int_{0}^{1/\sqrt{n}} (1 - x^2)^n\, dx \\
&\geqslant 2 \int_{0}^{1/\sqrt{n}} (1 - nx^2)\, dx \\
&= \frac{4}{3\sqrt{n}} \\
&> \frac{1}{\sqrt{n}}.
\end{aligned}
$$

We claim it readily follows from the previous paragraph that

$$q_n(x) \leqslant \sqrt{n}(1 - \delta^2)^n \qquad \text{for } \delta \leqslant x \leqslant 1.$$

In particular, we can make $|q_n(x)|$ as small as we want independently of x whenever $\delta \leqslant x \leqslant 1$. This implies that the sequence q_n converges uniformly to 0 on the closed subinterval $[\delta, 1]$.

For x satisfying $0 \leqslant x \leqslant 1$, define

$$p_n(x) = \int_{-1}^{1} f(x + t) q_n(t)\, dt.$$

Then the fact that $f(t)$ is zero for t outside of $[0, 1]$ implies that

$$p_n(x) = \int_{-x}^{1-x} f(x + t) q_n(t)\, dt.$$

Also by changing variables $(u = x + t)$, we find that

$$p_n(x) = \int_{0}^{1} f(u) q_n(u - x)\, du.$$

In particular, we recognize that $p_n(x)$ is a polynomial in x (because q_n is).

We will show that p_n converges uniformly to f. Let $\epsilon > 0$ be given. Since f is uniformly continuous, there is a $\delta > 0$ such that

$$|f(x) - f(y)| < \epsilon/2 \qquad \text{whenever } |x - y| < \delta.$$

Let $M = \sup |f(x)|$. Now we try to estimate $|p_n(x) - f(x)|$. Because the integral of q_n is 1, we have that

$$f(x) = \int_{-1}^{1} q_n(t) f(x) \, dt.$$

It follows that

$$|p_n(x) - f(x)| = \left| \int_{-1}^{1} [f(x+t) - f(x)] q_n(t) \right| \leqslant \int_{-1}^{1} |f(x+t) - f(x)| |q_n(t)| \, dt.$$

We now split the interval $[-1, 1]$ into the three pieces $[-1, -\delta] \cup [-\delta, \delta] \cup [\delta, 1]$, since we have control of q_n on two of these pieces. We find that

$$|p_n(x) - f(x)| \leqslant 2M \int_{-1}^{-\delta} |q_n(t)| \, dt + \frac{\epsilon}{2} \int_{-\delta}^{\delta} |q_n(t)| \, dt + 2M \int_{\delta}^{1} |q_n(t)| \, dt$$

and we use the inequality $q_n(t) \leqslant \sqrt{n}(1 - \delta^2)^n$ to find that

$$|p_n(x) - f(x)| \leqslant 4M\sqrt{n}(1 - \delta^2)^n + \epsilon/2.$$

If n is large enough, we can make $(1 - \delta^2)^n$ as small as we want. It follows that

$$|p_n(x) - f(x)| < \epsilon \qquad \text{for all } x \in [0, 1]$$

for n sufficiently large. $\qquad \square$

Corollary 251. *Let f be a continuous function on $[0, 1]$. Then there is a sequence of polynomials p_n on $[0, 1]$ converging uniformly to f on $[0, 1]$.*

Proof. Define g on $[0, 1]$ by

$$g(x) = f(x) - f(0) - x[f(1) - f(0)].$$

Then it is routine to check that g satisfies $g(0) = g(1) = 0$. Hence by the previous result, there is a sequence q_n of polynomials converging uniformly to g on $[0, 1]$. If we define

$$p_n(x) = q_n(x) + f(0) + x[f(1) - f(0)],$$

then each $p_n(x)$ is a polynomial on $[0, 1]$. Moreover, we check that

$$|f(x) - p_n(x)| = |(g(x) + f(0) + x[f(1) - f(0)]) - (q_n(x) + f(0) + x[f(1) - f(0)])|$$
$$= |g(x) - q_n(x)|.$$

It follows that the sequence p_n converges uniformly to f on $[0, 1]$. $\qquad \square$

Lemma 252. *Let φ be a continuous bijection from $[c, d]$ to $[a, b]$ with a continuous inverse. Let f be a function on $[a, b]$, and let g denote the composition $g = f \circ \varphi$. Suppose that g_n is a sequence of functions on $[c, d]$ converging uniformly to g. Then $g_n \circ \varphi^{-1}$ is a sequence of functions converging uniformly to f on $[a, b]$.*

Proof. Exercise. ☐

Corollary 253. *Let f be a continuous function on $[a, b]$. Then there is a sequence of polynomials p_n converging uniformly to f on $[a, b]$.*

Proof. Define $\varphi : [0, 1] \to [a, b]$ by

$$\varphi(x) = a + x(b - a).$$

Check that φ is a continuous from $[0, 1]$ to $[a, b]$ with a continuous inverse, and then apply the lemma to the function $g = f \circ \varphi : [0, 1] \to \mathbb{R}$, which is the uniform limit of a sequence of polynomials by the previous result. ☐

Corollary 254. *In particular, let $f : [-a, a] \to \mathbb{R}$ be the function $f(x) = |x|$. Then there is a sequence of polynomials p_n on $[-a, a]$ satisfying*

(a) p_n *converges uniformly to f on $[-a, a]$*

(b) *each p_n satisfies $p_n(0) = 0$.*

Proof. By the previous result, there is a sequence of polynomials p_n^* converging uniformly to f on $[-a, a]$. Define the polynomials $p_n(x) = p_n^*(x) - p_n^*(0)$. The sequence of constant functions $p_n^*(0)$ converges uniformly to the zero function on $[-a, a]$. It follows that p_n converges uniformly to $f - 0 = f$ by a homework problem. ☐

Rudin then isolates those properties of the polynomials that make the approximation theorem possible.

Definition 255. Let X be a metric space. A family $\mathcal{A}$ of real-valued functions on X is called an **algebra** if

(a) $f + g \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$

(b) $fg \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$

(c) $cf \in \mathcal{A}$ whenever $c \in \mathbb{R}$ and $f \in \mathcal{A}$.

In other words, the requirements are that $\mathcal{A}$ be closed under addition, multiplication, and scalar multiplication.

Example 256. Let $X = [a, b]$. The family $\mathcal{P}$ of polynomials on $[a, b]$ is an algebra.

Example 257. Let $X = \mathbb{R}$. The family $\mathcal{T}$ of trigonometric polynomials on $[0, 2\pi]$ is an algebra. Here a function f is a trigonometric polynomial if

$$f = A_0 + \sum_{k=1}^{n} A_k \sin(kx) + \sum_{\ell=1}^{m} B_\ell \cos(\ell x)$$

for some constants A_k, B_ℓ.

Definition 258. We say that $\mathcal{A}$ is **uniformly closed** if whenever f_n is a sequence in $\mathcal{A}$ that converges uniformly to some function f on X then f also belongs to $\mathcal{A}$. The **uniform closure of** $\mathcal{A}$ consists of all uniform limits of sequences from $\mathcal{A}$.

Example 259. The algebra $\mathcal{P}$ of polynomials $[a, b]$ is not uniformly closed because any continuous function belongs to the uniform closure of $\mathcal{P}$. In fact, the uniform closure of $\mathcal{P}$ consists of all continuous functions on $[a, b]$.

Theorem 260. *Assume $\mathcal{A}$ is an algebra consisting of bounded functions. Then the uniform closure of $\mathcal{A}$ is itself an algebra.*

Proof. Conditions (a) and (c) are satisfied for the uniform closure regardless of whether $\mathcal{A}$ consists of bounded functions or not. The hypothesis that $\mathcal{A}$ consists of bounded functions ensures that condition (b) is satisfied. $\qquad\square$

Definition 261. Let $\mathcal{A}$ be an algebra of functions on X.

(a) We say that $\mathcal{A}$ **separates points of** X if for each pair of distinct points $x_1, x_2 \in X$, there is a function f in $\mathcal{A}$ such that $f(x_1) \neq f(x_2)$.

(b) We say that $\mathcal{A}$ **vanishes at no points of** X if for each point $x \in X$, there is a function f in $\mathcal{A}$ such that $f(x) \neq 0$.

Example 262. The algebra of even polynomials on $[-1, 1]$ does not separate points (it does not separate the points $-x$ and x for example). The algebra of odd polynomials on $[-1, 1]$ vanishes at the point $x = 0$.

Example 263. The algebra of trigonometric polynomials on $\mathbb{R}$ does not separate points (it does not separate the points 0 and 2π for example).

Lemma 264. *Let $\mathcal{A}$ be an algebra of functions on X, let x_1 and x_2 be distinct points of X, and let c_1, c_2 be any two real numbers. Suppose that $\mathcal{A}$ separates points of X and vanishes at no points of X. Then there is a function f in $\mathcal{A}$ such that*

$$f(x_i) = c_i \qquad i = 1, 2.$$

Proof. Because $\mathcal{A}$ vanishes at no points of X, there are functions h_1 and h_2 in $\mathcal{A}$ such that $h_i(x_i) \neq 0$. Because $\mathcal{A}$ separates points of X, there is a function g in $\mathcal{A}$ such that $g(x_1) \neq g(x_2)$. Let u be the function defined by

$$u = h_1 \cdot (g - g(x_2)).$$

Then we see from construction that $u(x_1) \neq 0$ and $u(x_2) = 0$. Let v be the function

$$v = h_2 \cdot (g - g(x_1))$$

so that $v(x_2) \neq 0$ and $v(x_1) = 0$. Let f be the function defined by

$$f = \frac{c_1}{u(x_1)} u + \frac{c_2}{v(x_2)} v.$$

Then f satisfies the requirements. $\qquad\square$

Theorem 265. *Let $\mathcal{A}$ be an algebra of continuous functions on a compact set K. If $\mathcal{A}$ separates points on K and if $\mathcal{A}$ vanishes at no points of K, then the uniform closure of $\mathcal{A}$ consists of all continuous functions on K.*

Proof. We omit the full proof because of time constraints. Instead we give examples. $\qquad\square$

Example 266. Let $K = [a, b]$. Then the algebra $\mathcal{P}$ of polynomials on $[a, b]$ separates points and vanishes at no points of $[a, b]$. By the theorem, the uniform closure of $\mathcal{P}$ consists of all continuous functions on $[a, b]$.

Example 267. Let $K = S^1$ be the unit circle in $\mathbb{R}^2$. There is a natural bijective correspondence between $C(K)$ and the set of continuous 2π-periodic functions on $\mathbb{R}$ in such a way that the metrics on the two spaces correspond. In this way, the algebra $\mathcal{T}$ of trigonometric polynomials on $\mathbb{R}$ corresponds to an algebra of continuous functions on S^1. This algebra separates points and vanishes at no points of S^1. By the theorem, the uniform closure of $\mathcal{T}$ consists of all continuous functions on S^1, that is, all continuous 2π-periodic functions on $\mathbb{R}$.

8   Appendix

Example 268. Let $f_k(x)$ be defined on $\mathbb{R} \setminus 0$ by

$$f_k(x) = \begin{cases} x^k/k! & x > 0 \\ -x^k/k! & x < 0 \end{cases}.$$

We compute that the ℓth derivative of f_k satisfies

$$f_k^{(\ell)}(x) = \begin{cases} x^{k-\ell}/(k-\ell)! & x > 0 \\ -x^{k-\ell}/(k-\ell)! & x < 0 \end{cases}.$$

In particular, the kth derivative is given by

$$f_k^{(k)}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}.$$

Such a function, because we have excluded $x = 0$ from the domain, is continuous.

Lemma 269. *Let k be a positive integer, and let f be k-times differentiable in a neighborhood of $x = 0$ except possibly at $x = 0$. If f satisfies $|f(x)| \leqslant |x|^{k+1}$, then $f^{(k)}(0) = 0$.*

Proof. Assume the claim is true for some positive integer k. Assume that f satisfies $|f(x)| \leqslant |x|^{k+2}$. For x between -1 and 1, we have that $|x|^{k+2} \leqslant |x|^{k+1}$. It follows that $|f(x)| \leqslant |x|^{k+1}$. By the inductive hypothesis, we infer that $f^k(0) = 0$. Now we have also that for $x \neq 0$ that

$$\left| \frac{f(x)}{x^{k+1}} \right| \leqslant |x| \qquad x \neq 0.$$

The squeeze theorem implies that

$$\lim_{x \to 0} \frac{f(x)}{x^{k+1}} = 0.$$

L'Hospital's rule applied k times in a row implies that

$$\lim_{x \to 0} \frac{f^{(k)}(x)}{(k+1)!x} = 0.$$

We deduce that

$$\lim_{x \to 0} \frac{f^{(k)}(x) - f^{(k)}(0)}{x} = 0.$$

This means that $f^{(k+1)}(0) = 0$. The proof is complete by induction. □

Example 270. For a positive integer k, let $f_k(x)$ be defined by

$$f_k(x) = \begin{cases} x^{k+1} \sin(1/x) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

By the lemma, we have that $f^{(k)}(0) = 0$. However, we claim that $f^{(k)}$ is not continuous at zero.

Let us investigate first the case that $k = 1$. In this case, we have that

$$f_1(x) = x^2 \sin(1/x) \qquad x \neq 0.$$

Taking the first derivative, we find that

$$f_1'(x) = 2x \sin(1/x) - \cos(1/x).$$

As $x \to 0$, the first term tends to zero, but the second term has no limit.

Perhaps it is also instructive to look at the $k = 2$ case. In this case, we have that

$$f_2'(x) = 3x^2 \sin(1/x) - x \cos(1/x) \qquad x \neq 0.$$

For the second derivative, we have that

$$f_2''(x) = 6x \sin(1/x) - 4 \cos(1/x) - x^{-1} \sin(1/x) \qquad x \neq 0.$$

As $x \to 0$ in this expression, we see that only the first term has a limit. In fact, for any real number $M > 0$ and each $\delta > 0$, there is an x satisfying $0 < x < \delta$ such that $|f_2''(x)| > M$.

For other positive integers k, it is similarly possible to show that $f_k^{(k)}(x)$ has no finite limit as $x \to 0$.