MATH/MTHE 281 Notes

Introduction to Real analysis

Department of Mathematics and Statistics
Queen's University

Ole A. Nielsen R. Daniel Norman

This version: 2017/04/06

Preface for 1997 version

Most of Mathematics 281 deals with some fundamental ideas about and properties of the real numbers and continuous functions. These ideas and properties include sequences and series (which are studied in some first-year calculus), the least upper bound property, elementary topology of $\mathbb{R}$ and $\mathbb{R}^d$, convergence and uniform convergence of series of functions, and term-by-term differentiation and integration of series of functions. These notes will serve as a text for this part of the course.

For approximately twenty-five years ending with the 1996/97 academic year the last seven or so weeks of the second-year honours calculus course (Mathematics 220) dealt with the topics just described. In the 1970's there was no suitable text dealing with these topics and the first version of these notes was written by R. D. Norman; this version was really an outline and omitted some proofs and contained few examples or exercises. Over the years various other instructors added examples, comments, and other additions, with the result that the already terse notes became even harder to read. In 1992 I began revising these notes with the idea of adding exercises, an index, a table of contents, and more examples and exercises. This has proceeded since then with the assistance of Professor Norman and with students Peter Attia and Christina MacRae in 1994, Peter Chamberlain in 1995, and Greg Baker in 1996. Versions of these revised notes have been used since 1994/95.

Ole A. Nielsen
August 1997

Preface for ongoing revisions

In the Winter term of 2016, I taught MATH/MTHE 281 for the first time, and thought it a good idea to take the LaTeX for the original version of these notes, and make minor revisions. Also, the intention is to arrive at a "living" version of these course notes that can evolve, while still maintaining the suitability of the original form as a course text for MATH/MTHE 281.

Andrew D. Lewis
2017/04/06

Contents

Chapter 0

Mathematical notation and logic

In this chapter, we provide the briefest of overviews of the basic mathematical notation we shall use, and of mathematical logic, as this course is intended to introduce, not just mathematical ideas, but also techniques of proof.

0.1 Mathematical notation

Part of the power of mathematics is that it provides clear and compact notation for expression important ideas. In this section we give a rapid overview of the notation we use.

0.1.1 Sets and elementary set notation

To properly define what one means by a "set" is not an easy task. Fortunately, the work required to do this does not seem to be essential to function in mathematics. For our purposes, a ***set*** is a well-defined collection of objects. The idea is that one should be able to determine, at least in principle, whether something is a member of a set or not. For example, one might have the "set of beagles with no eyes who live in my house," which is a set whose members consist of a single beagle named "Rudy." More mathematically, one might have "the set of even integers," whose membership is pretty easily recognized, or "the set of prime numbers" whose membership is less easily recognized. A set is comprised of its ***elements*** or ***members***, i.e., all those things in the set.

Associated with constructions involving sets is some notation, which we list.

1. The ***empty set*** is the set with no elements, and is denoted by $\emptyset$.
2. Membership in a set is designated with the symbol "$\in$." Thus $x \in S$ means that x is a member of the set S. Nonmembership is denoted by "$\notin$."
3. A ***subset*** of a set A is another set, all whose elements are elements of S. If A is a subset of S, we write $A \subseteq S$. If we wish to exclude the possibility that $A = S$, we write $A \subset S$. If A is not a subset of S, then we write $A \not\subseteq S$. Note that does not mean that *some* elements of A are not in S, only that not *all* elements of A are in S.

   It is not uncommon to use $\subsetneq$ in place of our $\subset$.

4. If S and T are sets, we denote by $S \times T$ the ***product*** of S and T. Thus $S \times T$ is the set of elements we write as (x, y) for $x \in S$ and $y \in T$. This is generalized in an obvious to finite products.
5. If S and T are sets, then $S \cup T$ denotes the ***union*** of S and T, by which we mean elements of S *or* T. That is, $x \in S \cup T$ if $x \in S$ or $x \in T$.
6. If S and T are sets, then $S \cap T$ denotes the ***intersection*** of S and T, by which we mean elements of S *and* T. That is, $x \in S \cap T$ if $x \in S$ and $x \in T$.
7. The notions of union and intersection can be generalised to, not just two sets, but arbitrary numbers of sets. Thus we suppose that we have a family of sets S_λ, $\lambda \in \Lambda$, where Λ is an index set. That is, for each $\lambda \in \Lambda$, S_λ is a set. Now we denote $\cup_{\lambda \in \Lambda} S_\lambda$ to be union of the sets S_λ and $\cap_{\lambda \in \Lambda} S_\lambda$ to be the intersection of the sets S_λ. Thus $x \in \cup_{\lambda \in \Lambda} S_\lambda$ if $x \in S_\lambda$ for *some* $\lambda \in \Lambda$, and $x \in \cap_{\lambda \in \Lambda} S_\lambda$ if $x \in S_\lambda$ for *all* $\lambda \in \Lambda$.
8. For a set S and a subset $A \subseteq S$, the ***complement*** of A in S is the set $S \setminus A$ of elements of S that are not in A. We will also use the notation $A^\complement$ for complement.
9. A ***map*** f from a set S to a set T is an assignment of a single element $f(x) \in T$ to each element $x \in S$. We write $f\colon S \to T$ to denote a map from S to T. The set S is the ***domain*** of f and the set T is the ***codomain*** of f. We also denote

$$\mathrm{image}(f) = \{f(x) \mid x \in S\} \subseteq T.$$

   Note that, generally, we will not have $\mathrm{image}(f) = T$.

0.1.2 Notation for number systems

We shall sketch the construction of the real numbers in Chapter 1. Here we assume that the reader knows about real numbers, integers, rational numbers, etc., and we merely provide the notation we use for these.

By $\mathbb{Z}$ we denote the set of integers. Thus $\mathbb{Z}$ consists of the numbers $\dots, -2, -1, 0, 1, 2, \dots$. By $\mathbb{Z}_{>0}$ we denote the positive integers, i.e., the numbers $1, 2, \dots$. By $\mathbb{Z}_{\geq 0}$ we denote the nonnegative integers, i.e., $0, 1, 2, \dots$. By $\mathbb{Q}$ we denote the set of rational numbers, i.e., fractions of integers. That is, $q \in \mathbb{Q}$ if $q = \frac{n}{d}$ for some $n \in \mathbb{Z}$ and $d \in \mathbb{Z}_{>0}$. By $\mathbb{R}$ we denote the set of real numbers, i.e., the points on the number line that we learn about in our school days.

It is convenient to have special notation for certain subsets of real numbers, those called intervals. There are nine types of intervals. In the above list, a and b are real numbers with $a \leq b$.

1. $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$
2. $[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$
3. $(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$
4. $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$
5. $[a, \infty) = \{x \in \mathbb{R} \mid a \leq x\}$
6. $(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$
7. $(-\infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$
8. $(-\infty, b) = \{x \in \mathbb{R} \mid x < b\}$
9. $(-\infty, \infty) = \mathbb{R}$

0.2 Mathematical logic

In this course we will learn how to construct mathematical proofs. There are a number of elementary logical tools needed to accomplish this. In this section we overview these.

0.2.1 Propositions and symbolic logic

The basic object is a "proposition": a statement that has the truth value of either TRUE or FALSE. Here are some examples of propositions:

1. $p =$ "The year is 2017."
2. $q =$ "Pierre Elliott Trudeau is dead."
3. $r =$ "Pierre Elliott Trudeau died in 2001."
4. $s =$ "every even number larger than 2 is a sum of two primes."

We will probably agree that p is TRUE if the year is 2017 (otherwise is is FALSE), and q are TRUE, and r is FALSE (Trudeau died in 2000). Such propositions are easily proved, and do not require exotic notation to give a proof. The proposition s is one whose truth value is not known (it is the ***Goldbach Conjecture***). Proofs of more difficult propositions such as this often require complicated logical arguments, and our intention in this section is to sketch these.

First of all, given some propositions, we can create new ones by the four basic logical operations that we now explain.

1. *Negation*: For a proposition p, the proposition $\neg p$ has the opposite truth value of p: if p is TRUE then $\neg p$ is FALSE and *vice-versa*. Also $\neg(\neg p)$ has the same truth value as p.
2. *Conjunction*: For propositions p and q, the proposition $p \wedge q$ is TRUE only if *both* p and q are TRUE, otherwise $p \wedge q$ is FALSE.
3. *Disjunction*: $p \vee q$ is TRUE if either p or q (or both) is TRUE.
4. *Implication*: $p \implies q$ is TRUE only if either p is FALSE or q is TRUE, i.e., $p \implies q$ is logically equivalent to $\neg p \vee q$.

The kind of implications we will dealing with are often called "material implication" by logicians because it is not assumed that there is a causal connection between p and q. In our examples above, we would agree that $(p \wedge r) \implies q$ is TRUE although we admit that r is FALSE.

Using the four basic operations, some additional constructions immediately follow and will be useful.

1. *Logical Equivalence*: We have just seen that $p \implies q$ means $\neg p \vee q$, i.e., these assertions are logically equivalent which we write $(p \implies q) \iff (\neg p \vee q)$. This means that $p \implies q$ and $\neg p \vee q$ have identical truth values. There are also relations between $\vee$ and $\wedge$, namely $(p \wedge q) \iff \neg(\neg p \vee \neg q)$, and $(p \vee q) \iff \neg(\neg p \wedge \neg q)$. If we were being careful, we would have defined $\implies$ and $\vee$ in terms of $\neg$ and $\wedge$.

However, we leave these details for a proper course in mathematical logic. Note that using these rules $\neg(p \implies q) \iff (p \wedge \neg q)$.

2. *Negation of an implication*: An important application of the rules above occurs when we wish to negate an implication. $\neg(p \implies q) \iff \neg(\neg p \vee q) \iff p \wedge \neg q$.

0.2.2 Predicates

Let X be a set. A ***predicate*** p in X assigns to each $x \in X$ the value TRUE or FALSE. Note that, for each $x \in X$, $p(x)$ is a proposition since it is either TRUE or FALSE. Thus a predicate can be called a "propositional function." The variable x is the ***subject*** of the predicate. Often $p(x)$ is thought of as being a "property" of x, as will be clear from the examples below.

Predicates are often used to form conditional propositions, i.e., "if/then" statements. Thus, if p and q are predicates in X, one has a proposition $p(x) \implies q(x)$. This means that, if $x \in X$ is such that $p(x)$ is TRUE, then so too is $q(x)$ TRUE. As a proposition itself, $p(x) \implies q(x)$ may be TRUE or FALSE.

Here are some examples.

0.1 Examples: 1. Take $X = \mathbb{Z}_{>0}$ and consider the predicates p and q defined by $p(n)$="n is odd" and $q(n)$="n is prime". Thus $p(1)$=TRUE, $p(3)$=TRUE, $q(1)$=FALSE, and $q(3)$=TRUE. Note that both of the propositions $p(n) \implies q(n)$ and $q(n) \implies p(n)$ are FALSE.

2. Let X be the set of maps $f\colon [a,b] \to \mathbb{R}$. Consider the two predicates $p(f)$="f is continuous" and $q(f)$="there exists $c \in (a,b)$ such that $f(c) = 0$". The intermediate value theorem says that $p(f) \implies q(f)$ is TRUE.

3. Let $p(x)$ be the predicate in $\mathbb{R}$ given by "x is such that $1 - x^2 = 0$." This is TRUE if $x = \pm 1$ and FALSE otherwise. •

0.2.3 Quantifiers

Associated with predicates are some useful logical constructions called "quantifiers."

1. *Existential quantifier:* $\exists\, x\ p(x)$ means that there is x such that $p(x)$ is TRUE. For example, $\exists\, x\ x^2 - 1 = 0$ means that there is an x such that $x^2 - 1 = 0$. We might find it necessary to be more specific about where x is as in $\exists\, x\ (x \in \mathbb{R}) \wedge (x^2 - 1 = 0)$, which we shall write as $\exists\, x \in \mathbb{R}, x^2 - 1 = 0$. We shall also write $\nexists x \in \mathbb{R}, x^2 + 1 = 0$ for $\neg(\exists\, x \in \mathbb{R}, x^2 + 1 = 0)$. Finally, we shall write $\exists!\, x$ to mean there exists a unique x, as in $\exists!\, x \in \mathbb{R}, x^3 - 1 = 0$.
2. *Universal quantifier:* $\forall\, x\ p(x)$ means that, for all x, $p(x)$ is TRUE. For example, $\forall\, x \in \mathbb{R}\ x^2 \geq 0$ means that the square of every real number is nonnegative.

In constructions such as $p(x)$, x is a free variable; whereas in $\forall\, x\ p(x)$, x is a bound variable. A bound variable is like a dummy variable in calculus: the proposition $\forall\, x\ p(x)$ is logically equivalent to $\forall\, y\ p(y)$.

The quantifiers $\forall$ and $\exists$ are related by the rule $\neg(\exists\, x\; p(x)) \iff \forall\, x\; \neg p(x)$ and $\neg(\forall\, x\; p(x)) \iff \exists\, x\; \neg p(x)$. Indeed, one can *define* $\exists$ in terms of $\forall$ and $\neg$.

Let us give an example that illustrates how a well-known notion, that of continuity of a function (see Section 4.4), can be expressed using our above symbolic logic terminology.

0.2 Example: Let f be a function and $p(x)$ be the proposition

$$\forall\, \epsilon > 0\; \exists\, \delta > 0\; \forall\, y\; |x - y| < \delta \implies |f(x) - f(y)| < \epsilon$$

This is the way of saying that f is continuous at x.

Let us use our rules to find $\neg p(x)$:

$$\begin{aligned}
\neg p(x) &\iff \neg(\forall\, \epsilon > 0\; \exists\, \delta > 0\; \forall\, y\; |x - y| < \delta \implies |f(x) - f(y)| < \epsilon) \\
&\iff \exists\, \epsilon > 0\; \neg(\exists\, \delta > 0\; \forall\, y\; |x - y| < \delta \implies |f(x) - f(y)| < \epsilon) \\
&\iff \exists\, \epsilon > 0\; \forall\, \delta > 0\; \neg(\forall\, y\; |x - y| < \delta \implies |f(x) - f(y)| < \epsilon) \\
&\iff \exists\, \epsilon > 0\; \forall\, \delta > 0\; \exists\, y \neg(|x - y| < \delta \implies |f(x) - f(y)| < \epsilon) \\
&\iff \exists\, \epsilon > 0\; \forall\, \delta > 0\; \exists\, y\; |x - y| < \delta \wedge |f(x) - f(y)| \geq \epsilon
\end{aligned}$$

Thus $\neg p(x)$ is how we write the proposition that f is *not* continuous at x. •

0.2.4 Assertions and proofs

In this section we consider propositions as assertions we wish to prove. While each proof of a proposition tends to be its own entity, with its own particular personality, there are some broad strategies for proving propositions that are common enough that they are worth describing. Here we mention some of these, and give examples of how they are used.

Types of logical assertions one might prove

There are, broadly, four type of assertions we will prove.

1. Prove that a proposition p is TRUE.
2. Prove that a proposition p is FALSE.
3. Given propositions p and q, prove that $p \implies q$. This is, prove that, if p is TRUE, then q is TRUE.
4. Given propositions p and q, prove that $p \iff q$. That is, prove that, if p is TRUE, then q is TRUE *and* that, if q is TRUE, then p is TRUE.
5. Suppose that we have the proposition $p \implies q$. The ***converse*** proposition is $q \implies p$. It is very frequent in informal speech to confuse $q \implies p$ with $p \implies q$; in mathematics this is a fatal error.

To be clear, 2 3, and 4 are special cases of 1 applied to the propositions $p' = \neg p$, $p' = (p \implies q)$, and $p'' = (p \iff q)$, respectively. Sometimes, however, it is convenient to make use of the special structure of the proposition in these cases.

Next we describe some of the sorts of proofs one may employ.

Direct proof

Here we wish to prove that, given propositions p and q, $p \implies q$. The method of direct proof is that one uses p the fact that p is TRUE, with no other assumptions, to prove that q is TRUE.

We illustrate this with a few simple examples.

0.3 Examples: 1. Consider the proposition p given by "4 is an even number." To prove that p is TRUE, we note that $4 = 2 \cdot 2$, and so 4 is even, being an integer multiple of 2.

2. Consider the proposition p given by "4 is a prime number." To prove that p is FALSE, we note that $4 = 2 \cdot 2$, and so 4 has a factor other than 1 and 4.

3. We take $p(n)$ to be the predicate with integer subject "n is a square of odd numbers" and $q(n)$ to be the predicate with integer subject "n is an odd number." We will show that $p(n) \implies q(n)$. Suppose that n is such that $p(n)$ is TRUE. Then $n = m^2$ for an odd number m. Thus $m = 2k + 1$ for an integer k and

$$n = m^2 = (2k + 1)^2 = 2(2k^2 + 2k) + 1$$

and so $q(n)$ is TRUE. •

Proof by contradiction

For any proposition p, the proposition $p \wedge \neg p$ is always false. If, in a proof, one shows that a certain hypothesis implies the proposition $p \wedge \neg p$ for some p, then we say we have reached a ***contradiction*** and our hypothesis must be false. Sometimes, when we wish to prove p is TRUE, we assume that p is FALSE and then reach a contradiction, say $q \wedge \neg q$. That is, we assume $\neg p$ and then reach a contradiction; thus we have proved $\neg(\neg p))$, i.e., p is TRUE.

We give an example of proof by contradiction.

0.4 Example: We prove that $\sqrt{2}$ is irrational. Suppose that $\sqrt{2} = p/q$ for some integers p and q with p and q relatively prime, i.e., having no common factor. Then $p^2 = 2q^2$. So p^2 is even and hence p is even and thus p is divisible by 2. Hence $p^2 = (p')^2 \cdot 2^2$ for $p' = p/2$. We can now cancel a 2 to obtain $2(p')^2 = q^2$ and conclude that q is also even. This contradicts our assumption that p and q are relatively prime. Hence we have reached a contradiction. Thus our assumption that there were integers p and q, relatively prime, such that $\sqrt{2} = p/q$ was FALSE, i.e., $\sqrt{2}$ is irrational. •

Proof of contrapositive

We saw that $p \implies q$ was logically equivalent to $\neg p \vee q$, which is the same as $\neg(\neg q) \vee \neg p$, which finally is the same as $\neg q \implies \neg p$. This last expression is called the ***contrapositive*** of $p \implies q$ and we have just seen that $(p \implies q) \iff (\neg q \implies \neg p)$.

Let us give an example of proof by contrapositive.

0.5 Example: Let $X = \mathbb{Z}_{>0} \times \mathbb{Z}_{>0} \times \mathbb{Z}_{>0}$, and consider the predicates $p(n_1, n_2, n_3)$ and $q(n_1, n_2, n_3)$ in X given by "the product of n_1, n_2, and n_3 is greater than 1000" and "at least one of n_1, n_2, and n_3 is greater than 10." We will prove that $p(n_1, n_2, n_3)$ implies $q(n_1, n_2, n_3)$ by proving the contrapositive. Then we assume that $\neg q(n_1, n_2, n_3)$ is TRUE, meaning that we assume that $n_1, n_2, n_3 \leq 10$. Then

$$n_1 \cdot n_2 \cdot n_3 \leq 10 \cdot 10 \cdot 10 = 1000.$$

Thus $\neg p(n_1, n_2, n_3)$ is TRUE. •

Here is a more complicated example where the contrapositive notion is easier to prove.

0.6 Example: Suppose we wish to prove that for two real numbers x and y we have $x = y$. Quite often it will happen that we will have obtained one of the numbers, x say, as the reult of some kind of limiting process. This makes it difficult to show equality; so instead we show that $x \leq y$ and $y \leq x$. However, sometimes we find even showing $x \leq y$ too difficult, so we cut ourselves some further slack and prove that, for every $\epsilon > 0$, we have $x \leq y + \epsilon$. Let us now show that this implies that $x \leq y$. Let p be the statement $\forall\, \epsilon > 0\ x \leq y + \epsilon$ and q the statement $x \leq y$. We have to show $p \implies q$. This is logically equivalent to the contrapositive $\neg q \implies \neg p$. Let us expand this out:

$$\begin{aligned}(\neg q \implies \neg p) &\iff (\neg(x \leq y) \implies \neg(\forall\, \epsilon > 0\ x \leq y + \epsilon)) \\ &\iff (x > y) \implies (\exists\, \epsilon > 0\ x > y + \epsilon)\end{aligned}$$

Now, for the last statement, we can give a direct proof: given $x > y$, let $\epsilon = (x - y)/2$, then $\epsilon > 0$ and $y + \epsilon = (x + y)/2 < x$ as required. •

Proof by induction

It is common to have situations where one has an assertion that is said to hold for every $n \in \mathbb{Z}_{>0}$. That is, one has a predicate p in $\mathbb{Z}_{>0}$, and one wishes to prove the proposition $\forall\, n \in \mathbb{Z}_{>0}\ p(n)$. Frequently such propositions are proved by ***induction***. The way this works is as follows. The assertion $p(1)$ is proved, somehow, someway. Then the proposition $p(n)$ is assumed; this is called the ***induction hypothesis***. Using the fact that $p(n)$ is TRUE, $p(n+1)$ is proved to be TRUE, somehow, someway. In this way we prove the proposition $\forall\, n \in \mathbb{Z}_{>0}\ p(n)$. Another strategy that is rather similar is ***strong induction***, where the induction hypothesis is replaced with the ***strong induction hypothesis***, which is that $p(k)$ is TRUE for $k \in \{1, \dots, n\}$, and then it is proved that $p(n + 1)$ is TRUE.

Let's look at a couple of examples.

0.7 Examples: 1. Consider the predicate p in $\mathbb{Z}_{>0}$ given by

$$p(n) = \text{“}1 + \cdots + n = \tfrac{1}{2}n(n + 1)\text{”}.$$

We prove this by induction. For $n = 1$ we are asked to prove that $p(1) =$ "$1 = \frac{1}{2} \cdot 1 \cdot 2 = 1$", which is pretty clearly TRUE. The induction hypothesis is now that, for $n \in \mathbb{Z}_{>0}$, we assume that

$$p(n) = \text{“}\sum_{k=1}^{n} k = \tfrac{1}{2}n(n+1)\text{”}$$

is TRUE, and we are to show that

$$p(n+1) = \text{“}\sum_{k=1}^{n+1} k = \tfrac{1}{2}(n+1)(n+2)\text{”}$$

is TRUE. We prove this as follows:

$$\sum_{k=1}^{n+1} k = \Big(\sum_{k=1}^{n} k\Big) + (n+1) = \tfrac{1}{2}n(n+1) + (n+1) = \tfrac{1}{2}(n+1)(n+2).$$

2. As an example of a proof by strong induction, we offer the following. We $X = \{n \in \mathbb{Z}_{>0} \mid n \geq 2\}$ let p be the predicate

$$p(n) = \text{“}n \text{ is a product of prime numbers”}.$$

In this case, although the index for $p(n)$ is not $\mathbb{Z}_{>0}$, we can clearly still use an inductive proof. For $n = 2$ we have $p(2) =$ "$2 = 2$", giving 2 as a product of primes. Now the strong induction hypothesis is that, for $k \in \{1, \dots, n\}$,

$$p(k) = \text{“}k \text{ is a product of primes”}$$

is TRUE, and we are to show that

$$p(n+1) = \text{“}n+1 \text{ is a product of primes”}$$

is TRUE. To show that this is the case, there are two possibilities.

(a) $n + 1$ is prime: In this case, $p(n + 1)$ is TRUE immediately.

(b) $n + 1$ is not prime: In this case, the definition of "not prime" means that $n + 1 = a \cdot b$ for $a, b \in X$. We claim that $a, b \leq n$. Suppose otherwise and that one of a or b, say a, exceeds n. Then

$$ab \geq 2a > 2n = n + n > n + 1,$$

which contradicts the fact that $ab = n + 1$. Thus we can suppose that $a, b \leq n$. In this case, the strong induction hypothesis gives

$$a = p_1 \cdots p_k, \quad b = q_1 \cdots q_l$$

for primes $p_1, \dots, p_k, q_1, \dots, q_l$. Thus

$$n + 1 = p_1 \cdots p_k q_1 \cdots q_l,$$

showing that $p(n + 1)$ is TRUE. •

Nonconstructive proof

In a nonconstructive proof, one proves a statement about the existence of something, but without giving a specific instance of it, or an algorithm for finding it.

Here are two examples.

0.8 Examples: 1. If there are 367 people in a room, then at least two of the people share the same birthday. A nonconstructive proof is immediate: if the mapping from "person" to "birth date" is one-to-one, this would mean that there are at least 367 days in a year, which cannot be. However, in this case a constructive proof is also possible by interviewing all people in the room and establishing which have the same birthday.

2. The intermediate value theorem asserts that, if $f\colon [a,b] \to \mathbb{R}$ is a continuous function for which $f(a) < 0$ and $f(b) > 0$, then there exists $c \in (a,b)$ such that $f(c) = 0$. While this theorem is true, a proof does not give any indication of what c is or how to find it. •

Proof by counterexample

A counterexample is typically used to show that a proposition is not TRUE. Normally a counterexample is most useful in a situation where the proposition is of the form $p(x)$ for a predicate in a set X. To show that $\forall\, x \in X\ p(x) = \text{TRUE}$ is TRUE, one must do exactly that: show that it is TRUE for all x. To show that $\forall\, x\ p(x)$ is false, one need only show that $\exists\, x \in X\ p(x) = \text{FALSE}$, i.e., a single $x \in X$ is all that is required, and this is said to be a ***counterexample***.

Here are some examples.

0.9 Examples: We show that the converses to some propositions proved above are false, by finding counterexamples.

1. We consider the predicates in $\mathbb{Z}_{>0}$ given by

$$p(n) = \text{“}n \text{ is a square of odd numbers”},$$
$$q(n) = \text{“}n \text{ is an odd number”},$$

and recall that in Example 0.3–3 we showed that $p(n) \implies q(n)$ is TRUE. We now show by counterexample that $q(n) \implies p(n)$ is FALSE. Indeed, $q(7)$ is TRUE, while $p(7)$ is FALSE, and so $q(n) \implies p(n)$ is FALSE.

2. We take X to be the set of mappings $f\colon [a,b] \to \mathbb{R}$ and consider the two predicates $p(f)$=“f is continuous” and $q(f)$=“there exists $c \in (a,b)$ such that $f(c) = 0$”. As remarked in Example 0.1–2, the proposition $p(f) \implies q(f)$ is TRUE. We will show by counterexample that the proposition $q(f) \implies p(f)$ is FALSE. Indeed, the mapping $f \in X$ defined by

$$f(x) = \begin{cases} 1, & x = \frac{1}{2}(a+b), \\ 0, & \text{otherwise} \end{cases}$$

satisfies $q(f) = \text{TRUE}$ and $p(f) = \text{FALSE}$, and so the proposition $q(f) \implies p(f)$ is FALSE. (Note: We should assume that $a < b$ for this counterexample to be valid. Indeed, if $a > b$, then $X = \emptyset$, and so $q(f) \implies p(f)$ is TRUE, since $q(f)$ is never TRUE. Also, the proposition is actually true if $a = b$.) •

Exercises

E0.1 Let A be a set of real numbers. What does the following proposition say?

$$\forall\, x \in A\ \exists\, y \in A\ (x < y) \wedge (\forall\, z \in A\ (x < z) \implies (y \leq z))$$

Is there a set A for which this proposition is TRUE? Is there a set A for which it is FALSE?

E0.2 Let $A = [0, 1]$. Let $p(x)$ be the proposition

$$(\forall\, a \in A\ a \leq x) \wedge (\forall\, b \in \mathbb{R}\ (\forall\, a \in A\ a \leq b) \implies x \leq b)$$

For which x is $p(x)$ TRUE?

E0.3 (a) For each of the following, give (if possible) an example of a function $f\colon \mathbb{R} \to \mathbb{R}$ that makes the proposition TRUE.

1. $\forall\, x \in \mathbb{R}\ \exists\, y \in \mathbb{R}$ s.t. $f(x) = y$
2. $\forall\, y \in \mathbb{R}\ \exists\, x \in \mathbb{R}$ s.t. $f(x) = y$
3. $\exists\, y \in \mathbb{R}$ s.t.$\forall\, x \in \mathbb{R}\ f(x) = y$

(b) Give the negations of the above propositions.

E0.4 For some $A \subseteq \mathbb{R}$, consider the following proposition:

$$\forall\, x \in A\ \exists\, y \in A \text{ s.t.} (x < y) \wedge (\forall z \in A\, (x < z \implies y \leq z))$$

(a) Describe simply what the proposition is saying.

(b) Find, if possible, a set $A \subseteq \mathbb{R}$ for which the above proposition is

1. TRUE
2. FALSE

Chapter 1

The real numbers

The basic ingredient in "real analysis" is the set of real numbers. While we may well have some understanding of what these are—we are taught about the "number line" in school—a full and complete understanding of the crucial attributes of the real numbers may not be entirely familiar. While we understand that the algebraic properties of addition, subtraction, multiplication, and division are important, these fall into the category of "algebra." What concerns us here is "analysis," and this requires the property of the real numbers known of as "completeness." This property can be arrived at in a variety of ways, and we shall sketch two of these, the "least upper bound property" and "convergence of Cauchy sequences." Ultimately we arrive at the former of these as the starting point for the properties of the real numbers that allows one to do analysis. We do this by giving an overview, without all details, of the manner in which one builds the real numbers "from nothing." To do this properly takes more space and effort than we can give here, but to not mention this at all is rather unsatisfying. Thus we attempt to tread the line between making the construction clear, but without making it excessively burdensome.

1.1 Construction of the nonnegative integers

The first construction we make is of the nonnegative integers. We do this recursively, defining first 0, then 1, then 2, etc., as sets. To be precise, we define

$$\begin{aligned}
0 &= \emptyset \\
1 &= 0 \cup \{0\} = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\} \\
2 &= 1 \cup \{1\} = \{0\} \cup \{1\} = \{0, 1\} \\
3 &= 2 \cup \{2\} = \{0, 1\} \cup \{2\} = \{0, 1, 2\} \\
&\vdots
\end{aligned}$$

Note that 0 is a set with zero elements, 1 is a set with one elements, and so on. By $\mathbb{Z}_{\geq 0}$ we denote the set whose elements are all the results of the preceding recursive construction. We call this the ***set of nonnegative integers***. Note that the language here is confusing, since it seems to presuppose that we know what the integers are. However, keep in mind that it is the "nonnegative integers" (all one notion) we are defining. The integers will come shortly.

We now define the familiar operations on $\mathbb{Z}_{\geq 0}$. First we define "less than" and "less than or equal to." We write $j < k$ if $j \subset k$ and $j \leq k$ if $j \subseteq k$. One can check that these operations have the expected interpretation and properties. We also denote $\mathbb{Z}_{>0} = \mathbb{Z}_{\geq 0} \setminus \{0\}$, i.e., the set of numbers $1, 2, 3, \ldots$.

Now we need to define addition and multiplication in $\mathbb{Z}_{\geq 0}$. Note that our definition of $\mathbb{Z}_{\geq 0}$ essentially defines the operation of "adding 1" by $k + 1 = k \cup \{k\}$. Given $k \in \mathbb{Z}_{\geq 0}$ we define two maps $a_k : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ and $m_k : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ to be "addition of k" and "multiplication by k." The definition we give for these operations is recursive. We define $a_k(0) = k$ and $m_k(0) = 0$. Assume now that we have defined $a_k(j)$ and $m_k(j)$ for $j \leq m$. We then define

$$a_k(m+1) = a_k(m) + 1, \quad m_k(m+1) = m_k(m) + k.$$

One can tediously verify that these definitions of addition and multiplication have the expected commutativity, associativity, and distributivity properties.

1.2 Construction of the integers

Now we describe how one builds the integers from the set $\mathbb{Z}_{\geq 0}$ of nonnegative integers. Here is the idea. Given $m \in \mathbb{Z}_{\geq 0}$, one would like "$-m$" to have the property that, if $j \geq m$, then $k = j + (-m)$ is that nonnegative integer which, when added to m, gives j. There are many possible j's and k's for which this holds, and we *define* $-m$ to be the set of all of these. Here's how we do this precisely.

We say that two pairs (j_1, k_1) and (j_2, k_2) of nonnegative integers are ***equivalent*** if $j_1 + k_2 = j_2 + k_1$. Given a pair (j, k) of nonnegative integers, we denote by $[(j, k)]$ the set of pairs of nonnegative integers that are equivalent to (j, k). We *define* the set of ***integers*** to be the set $\{[(j, k)] \mid j, k \in \mathbb{Z}_{\geq 0}\}$ of equivalent pairs. One can then easily prove the following.

1.1 Lemma: *If (j, k) is a pair of nonnegative integers, then exactly one of the following statements holds:*

(i) (j, k) *is equivalent to* $(m, 0)$ *for a unique* $m \in \mathbb{Z}_{>0}$*;*

(ii) (j, k) *is equivalent to* $(0, m)$ *for a unique* $m \in \mathbb{Z}_{>0}$*;*

(iii) (j, k) *is equivalent to* $(0, 0)$.

With this in mind, we denote by m the set of pairs of nonnegative integers equivalent to $(m, 0)$, by $-m$ the set of pairs of nonnegative integers equivalent to $(0, m)$, and by 0 the set of pairs of nonnegative integers equivalent to $(0, 0)$. This now makes it clear how we assign to our abstract notion of the set of integers to the one that we think about in practice. We should also define the operations of addition and multiplication in $\mathbb{Z}$. This we do as follows:

$$\begin{aligned}
[(j_1, k_1)] + [(j_2, k_2)] &= [(j_1 + j_2, k_1 + k_2)], \\
[(j_1, k_1)] \cdot [(j_2, k_2)] &= [(j_1 \cdot j_2 + k_1 \cdot k_2, j_1 \cdot k_2 + j_2 \cdot k_1)].
\end{aligned}$$

Thus we define addition and multiplication only knowing addition and multiplication in $\mathbb{Z}_{\geq 0}$. Note that we can now define subtraction by

$$[(j_1, k_1)] - [(j_2, k_2)] = [(j_1, k_1)] + (-1) \cdot [(j_2, k_2)].$$

One can verify, first of all, that this corresponds to the usual notions we have in mind for addition and multiplication. Then one can tediously verify that all commutativity, associativity, and distributivity properties we know about addition and multiplication are valid.

One can also order the integers in the usual way. This is done with our abstract definition by defining

$$\begin{aligned} [(j_1, k_1)] < [(j_2, k_2)] \quad &\iff \quad j_1 + k_2 < k_1 + j_2, \\ [(j_1, k_1)] \leq [(j_2, k_2)] \quad &\iff \quad j_1 + k_2 \leq k_1 + j_2. \end{aligned}$$

Again, one can work through the definitions to see that this corresponds to the expected notion of "less than" and "less than or equal to," and that it satisfies the usual properties we know about these relations.

Finally, we define the absolute value function on $\mathbb{Z}$ by

$$|k| = \begin{cases} k, & 0 < k, \\ 0, & k = 0, \\ -k, & k < 0. \end{cases}$$

This is the expected thing, and has all the expected properties, although, again, one must tediously prove them.

1.3 Construction of the rational numbers

We next turn to constructing the rational numbers, i.e., fractions of the form $\frac{j}{k}$. Since we do not know what "division" means, we need to make this definition using what we have at hand, which is the set of integers and all of the associated operations. We again work with equivalence of pairs, but now pairs of integers.

Given pairs $(j_1, k_1), (j_2, k_2) \in \mathbb{Z} \times \mathbb{Z}_{>0}$, we say they are ***equivalent*** if $j_1 \cdot k_2 = j_2 \cdot k_1$. Given a pair $(j, k) \in \mathbb{Z} \times \mathbb{Z}_{>0}$, we denote by $[(j, k)]$ the set of such pairs equivalent to (j, k). The set of ***rational numbers*** is then the set $\{[(j, k)] \mid j \in \mathbb{Z},\ k \in \mathbb{Z}_{>0}\}$ of equivalent pairs. We denote $[(j, k)] = \frac{j}{k}$.

Now one defines the expected operations on $\mathbb{Q}$. Thus we define

$$\begin{aligned} [(j_1, k_2)] + [(j_2, k_2)] &= [(j_1 \cdot k_2 + j_2 \cdot k_1, k_1 \cdot k_2)], \\ [(j_1, k_1)] \cdot [(j_2, k_2)] &= [(j_1 \cdot j_2, k_1 \cdot k_2)]. \end{aligned}$$

We can now also define division. If $[(j_1, k_1)], [(j_2, k_2)] \in \mathbb{Q}$ with $j_2 \neq 0$, we define

$$\frac{[(j_1, k_1)]}{[(j_2, k_2)]} = [(j_1 \cdot k_2, j_2 \cdot k_1)].$$

One can verify easily that these are the usual operations for rational numbers, and that they satisfy the expected properties of commutativity, associativity, and distributivity. We also define subtraction in the expected way:

$$[(j_1, k_1)] - [(j_2, k_2)] = [(j_1, k_1)] + [(-1, 1)] \cdot [(j_2, k_2)].$$

We also have the order relations on $\mathbb{Q}$ given by

$$\begin{aligned} [(j_1, k_1)] < [(j_2, k_2)] \quad &\iff \quad j_1 \cdot k_2 < j_2 \cdot k_1, \\ [(j_1, k_1)] \leq [(j_2, k_2)] \quad &\iff \quad j_1 \cdot k_2 \leq j_2 \cdot k_1. \end{aligned}$$

These are easily verified to be the usual "less than" and "less than or equal to," and they have the expected properties. We denote by $\mathbb{Q}_{>0}$ the set of rational numbers greater than 0.

Finally, we define the absolute value function on $\mathbb{Q}$ by

$$|[(j, k)]| = \begin{cases} [(j, k)], & [(0, 1)] < [(j, k)], \\ [(0, 1)], & [(j, k)] = [(0, 1)], \\ -[(j, k)], & [(j, k)] < [(0, 1)]. \end{cases}$$

Again, one can verify that the usual properties of absolute value are valid.

Now we have at hand the set $\mathbb{Q}$ of rational numbers, with its expected operations of addition, subtraction, multiplication, and division, with its expected relations $<$ and $\leq$, and with its absolute value function. Our next step is to build the real numbers. There are two common ways of building the real numbers, and we shall sketch both. They turn out to be equivalent, and this is something we will outline in Chapter 2.

1.4 The (lack of the) least upper bound property for $\mathbb{Q}$

The set of rational numbers has (at least) two defects, one of which we address in this section, and the other of which we address in the next section.

We begin with some easily understood definitions.

1.2 Definition: Let $A \subseteq \mathbb{Q}$.

(i) The set A is ***bounded*** if there exists $M \in \mathbb{Q}_{>0}$ such that $|q| \leq M$ for all $q \in A$.

(ii) A ***lower bound*** for A is a number $\ell \in \mathbb{Q}$ such that $\ell \leq q$ for all $q \in A$.

(iii) An ***upper bound*** for A is a number $u \in \mathbb{Q}$ such that $q \leq u$ for all $q \in A$.

(iv) A number $\ell \in \mathbb{Q}$ is a ***greatest lower bound*** for A if $\ell' \leq \ell$ for every lower bound ℓ' for A.

(v) A number $u \in \mathbb{Q}$ is a ***least upper bound*** for A if $u \leq u'$ for every upper bound u' for A. •

One can easily see that A is bounded if and only if it has both an upper and lower bound (Exercise E1.3).

The subjects of least upper bound and greatest lower bound are dominated by two important questions:

1. If a set A is bounded, does it always possess a greatest lower bound and a least upper bound?
2. If a set A possesses a least upper bound or greatest lower bound, is it unique?

The answer to these questions is, "No," and, "Yes." Just by using properties of $\mathbb{Q}$ as enumerated so far, and not presupposing the existence of the real numbers, it is not easy to prove that a bounded set might not possess, say, a least upper bound. However, it one presupposes the existence of the real numbers, this *is* easy to see. So let us consider this for the moment.

1.3 Example: Take $A = \{x \in \mathbb{Q} \mid x \leq \sqrt{2}\}$. We know that $\sqrt{2}$ is irrational by Example 0.4. Now suppose that u is any upper bound for A. Thus $x \leq u$ for every $x \in A$. Since $u > \sqrt{2}$, $u - \sqrt{2} > 0$. Thus $u - \frac{1}{2}(u - \sqrt{2}) > \sqrt{2}$ and so is also an upper bound. Moreover, it is smaller than u, and so there can no no least upper bound. •

It is possible to make the preceding example make sense without presupposing the existence of the real numbers, but that would take too much effort. And, in any case, it is something that is at least intuitively clear.

The main point, however, is that the quite intuitive property of existence of least upper bounds is not true in $\mathbb{Q}$. It *is* true, however, for the set of real numbers.

1.5 Cauchy sequences in $\mathbb{Q}$

Next we consider another way of characterising an important defect of the rational numbers. It will not be too difficult to imagine that this defect is rather related to the lack of existence of a least upper bound.

We consider sequences of rational numbers. In Definition 2.1 we carefully consider what we mean by a sequence of real numbers, and this definition carries over to sequences of real numbers. We will use this notion of sequences of rational numbers in the following definition.

1.4 Definition: Let (q_n) be a sequence in $\mathbb{Q}$.

(i) The sequence (q_n) is a ***Cauchy sequence*** if, for every $\epsilon \in \mathbb{Q}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|q_n - q_m| < \epsilon$ for $m, n \geq N$.

(ii) The sequence (q_n) ***converges*** to $L \in \mathbb{Q}$ if, for every $\epsilon \in \mathbb{Q}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|q_n - L| < \epsilon$ for $n \geq N$. •

Now, sequences that converge are Cauchy; we will prove this for Cauchy sequences of real numbers as one half of Theorem 2.23, and the proof for rational numbers is exactly the same. However, it is not the case that Cauchy sequences of rational numbers converge to some rational number. As with the least upper bound property in the preceding section, this is easy to imagine if one presupposes the real numbers, but is not so easy to prove otherwise. We thus consider an example of a nonconvergent Cauchy sequence in $\mathbb{Q}$.

1.5 Example: For $n \in \mathbb{Z}_{>0}$, consider the rational number $q_n = (1 + \frac{1}{n})^n$. Perhaps the reader has run into the sequence (q_n), and indeed it converges to the real number $\mathrm{e} \approx 2.718281828\ldots$, i.e., to Euler's constant. Since it is a convergent sequence of real numbers, it is also a convergent sequence of rational numbers (convergence for real numbers is defined in Definition 2.4, but, anyway, is exactly like for rational numbers). Therefore, it is a Cauchy sequence of rational numbers. However, since it's limit in the set of real numbers is e, which is not rational, it cannot converge in $\mathbb{Q}$. •

It is possible to come up with nonconvergent Cauchy sequences without making reference *a priori* to the set of real numbers, but it is simply difficult to do this.

This lack of necessity of Cauchy sequences converging is problematic, since Cauchy sequences "seem like" they want to converge., and it would be ever so convenient if they did.

1.6 The definition of the real numbers

In the previous two sections we gave two properties *not* possessed by the rational numbers, but which we would like to be possessed, and in fact are possessed, by the real numbers. Indeed, one can take these properties as the *definition* of the real numbers in some sense. In this section we sketch how to do this in both cases, but without going through the details.

1.6.1 Using Dedekind cuts

The method we outline here is to add least upper bounds to sets of rational numbers that may not have one. The idea is made somewhat precise in the following.

A ***Dedekind cut*** in $\mathbb{Q}$ is a partition of $\mathbb{Q}$ into two sets A and B with the following properties:

1. A and B are both nonempty;
2. $A \cup B = \mathbb{Q}$;
3. A is closed downwards, i.e., for all $q, r \in A$ with $q < r$, if $r \in A$ then $q \in A$;
4. B is closed upwards, i.e., for all $q, r \in B$ with $q < r$, if $q \in B$ then $r \in B$;
5. A contains no greatest element, i.e., there is no $q \in A$ such that $r \leq q$ for all $r \in A$.

The set $\mathbb{R}$ of ***real numbers*** is then the set of Dedekind cuts. Let us consider a few examples of this to see how the definitions work. We do not do the work to show that these are indeed Dedekind cuts.

1.6 Examples: 1. The real number we know as $\sqrt{2}$ is defined by the Dedekind cut

$$A = \{q \in \mathbb{Q} \mid \ q \cdot q < 2\}, \quad B = \mathbb{Q} \setminus A.$$

2. For $n \in \mathbb{Z}_{>0}$, let

$$A_n = \left\{ q \in \mathbb{Q} \mid \ q < \sum_{j=0}^{n} \frac{1}{j!} \right\}$$

and $A = \cup_{n=0}^{\infty} A_n$. The number e can be defined to be the Dedekind cut with defined by A and $B = \mathbb{Q} \setminus A$.

3. Rational numbers are defined by Dedekind cuts as follows. Let $q \in \mathbb{Q}$, and associate to q the Dedekind cut

$$A = \{r \in \mathbb{Q} \mid \ r < q\}, \quad B = \mathbb{Q} \setminus A.$$ •

Note that a Dedekind cut is effectively defined by one of the sets A and B, and in what follows we will work with the set A.

We need to define the standard constructions on $\mathbb{R}$ using this definition. First we define the order relations "less than" and "less than or equal to" as follows:

$$A_1 < A_2 \iff A_1 \subset A_2, \qquad A_1 \leq A_2 \iff A_1 \subseteq A_2.$$

Next we define addition and multiplication:

$$A_1 + A_2 = \{a_1 + a_2 \mid \ a_1 \in A_1,\ a_2 \in A_2\},$$

$$A_1 \cdot A_2 = \begin{cases} \{a_1 \cdot a_2 \mid \ a_1 \in A_1,\ a_2 \in A_2,\ a_1, a_2 \geq 0\} & A_1, A_2 \geq 0, \\ \qquad \cup \{q \in \mathbb{Q} \mid \ q \leq 0\}, & \\ -(A_1 \cdot (-A_2)), & A_1 \geq 0,\ A_2 < 0, \\ -((-A_1) \cdot A_2), & A_1 < 0,\ A_2 \geq 0, \\ (-A_1) \cdot (-A_2), & A_1, A_2 < 0. \end{cases}$$

Finally, we define the absolute value function by

$$|A| = \begin{cases} A, & A > 0, \\ 0, & A = 0, \\ -A, & A < 0. \end{cases}$$

One can show that all these operations satisfy the expected properties. We leave these rather tedious exercises to the reader.

1.6.2 Using Cauchy sequences

The construction we sketch here "fills in the gaps" between the convergent Cauchy sequences with the nonconvergent Cauchy sequences. We will be slightly precise about this.

Two Cauchy sequences (q_n) and (r_n) of rational numbers are ***equivalent*** if, for each $\epsilon \in \mathbb{Q}_{>0}$, there exists $N \in \mathbb{Z}_{>0}$ such that $|q_n - r_n| < \epsilon$ for every $n \geq N$. The idea is that equivalent Cauchy sequences have tails that get close to one another. Given a Cauchy sequence (q_n), we denote by $[(q_n)]$ the set of all Cauchy sequences equivalent to (q_n). We then define $\mathbb{R}$, the set of ***real numbers***, to be the set

$$\{[(q_n)] \mid [(q_n)] \text{ is a Cauchy sequence in } \mathbb{Q}\}.$$

This construction, then, includes in $\mathbb{R}$ *all* Cauchy sequences, even those that do not converge. In this way, it is a natural sort of definition.

One now has to show that $\mathbb{R}$ has all of the attributes one is familiar with. First, we indicate how we add and multiply in $\mathbb{R}$:

$$[(q_n)] + [(r_n)] = [(q_n + r_n)],$$
$$[(q_n)] \cdot [(r_n)] = [(q_n \cdot r_n)].$$

One has many things to verify here. First of all, the definitions need to be shown to be independent of the choice of equivalent Cauchy sequences. Second, one should show that the righthand sides of the preceding definitions are, in fact, Cauchy sequences. After this, one should show that addition and multiplication have the expected properties.

We can define the usual order relations on $\mathbb{R}$ as well. Specifically, we write $[(q_n)] < [(r_n)]$ if there exists $N \in \mathbb{Z}_{>0}$ such that $q_n < r_n$ for $n \geq N$. We also write $[(q_n)] \leq [(r_n)]$ if either $[(q_n)] < [(r_n)]$ or $[(q_n)] = [(r_n)]$. Of course, these order relations have all of the expected properties.

Finally, we define the absolute value function on $\mathbb{R}$. For this, we note that, if (q_n) is a Cauchy sequence, then $(|q_n|)$ is a Cauchy sequence. We then define

$$|[(q_n)]| = [(|q_n|)].$$

The absolute value function has all of the properties we expect. One of the most important of these is the following:

$$|[(q_n)] + [(r_n)]| \leq |[(q_n)]| + |[(r_n)]|,$$

which is called the ***triangle inequality***. If we eliminate the tedious equivalence class notation, the triangle inequality reads $|a + b| \leq |a| + |b|$ for $a, b \in \mathbb{R}$.

What we arrive at is the set $\mathbb{R}$ with its familiar algebraic, order properties, and valuation properties; but these properties are possessed by $\mathbb{Q}$. What $\mathbb{R}$ has that $\mathbb{Q}$ does not is the property that Cauchy sequences converge, an assertion we will prove as Theorem 2.23. This important property has a name: ***completeness***.

1.6.3 The punchline

In this section we have constructed the real numbers in two different ways. It turns out that the two constructions give equivalent outcomes. That is, the sets of real numbers we construct in each case amount to the same thing. The way one normally proceeds is by selecting one of the constructions as the starting point, and proves everything with this as a starting point. That is to say, one works with either (1) all bounded sets in $\mathbb{R}$ have a least upper bound or (2) all Cauchy sequences in $\mathbb{R}$ converge. In this text we will be assuming that all bounded sets have a least upper bound, and proving facts with this as a starting point.

1.7 The least upper bound property for $\mathbb{R}$

Based on our discussion above, we will assume that all bounded subsets of $\mathbb{R}$ have a least upper bound. From the point of view of analysis, the least upper bound property is one of the most fundamental properties of the real numbers. This property (or something equivalent to it) is usually stated as an assumption in a first-year calculus course, and the purpose of this section is to discuss and explain this property in more detail than we managed above in our rather hurried construction of the real numbers.

We begin with definitions we have already seen for rational numbers, but now in the setting of real numbers.

1.7 Definition: A nonempty subset A of $\mathbb{R}$ is said to be ***bounded below*** if there is a number s such that $s \leq a$ for all $a \in A$ and in this case any such number s is called a ***lower bound*** for A. Similarly, A is said to be ***bounded above*** if there is a number t such that $a \leq t$ for all $a \in A$, and any such number t is called an ***upper bound*** for A. Finally, A is said to be ***bounded*** if it is both bounded below and bounded above and to be ***unbounded*** if it is not bounded. •

In general, a set need not be bounded above, but if it is then it will necessarily have an infinite number of upper bounds. Indeed, if a set is bounded above then any number larger than an upper bound is also an upper bound.

1.8 Example: Consider the sets $A = [0, 1]$ and $B = (-\infty, 1)$. Here A is bounded below and bounded above, and hence bounded, while B is bounded above but is not bounded below. Moreover, the lower bounds for A are the numbers in the interval $(-\infty, 0]$, the upper bounds for A are the numbers in the interval $[1, \infty)$, and the sets A and B have the same upper bounds. Notice that the smallest of the upper bounds for A belongs to A but that the smallest of the upper bounds for B does not belong to B. •

Suppose that A is a nonempty subset of $\mathbb{R}$ that is bounded above. Then the set U of all upper bounds of A is an infinite set which is bounded below; in fact, U is bounded below since any number in A is a lower bound for U. One can now ask whether U contains a number which is smaller then every other number in U, i.e., whether there is an upper bound for A which is smaller than every other upper bound for A. It is clear that there can be at most one such smallest upper bound, i.e., if it exists then it is unique. But, as the reader may believe from our discussion in the first part of this chapter, it is not at all clear whether a smallest upper bound must always exist.

However, our construction of the reals by Dedekind cuts above is precisely designed so that the following theorem holds.

1.9 Theorem: *Every nonempty subset of $\mathbb{R}$ that is bounded above has a least upper bound, i.e., an upper bound that is smaller than any other upper bound. Similarly, every nonempty subset of $\mathbb{R}$ that is bounded below has a greatest lower bound, i.e., a lower bound that is greater than any other lower bound.*

This theorem justifies the following definition.

1.10 Definition: Suppose that A is a nonempty subset of $\mathbb{R}$.

(i) If A is bounded above then the smallest upper bound for A is called the ***least upper bound*** or ***supremum*** of A and is denoted by $\sup A$.

(ii) If A is bounded below then the largest lower bound for A is called the ***greatest lower bound*** or ***infimum*** of A and is denoted by $\inf A$. •

The proof of the following result is both easy and instructive.

1.11 Proposition: *If A and B are two nonempty subsets of $\mathbb{R}$ that are bounded above and if $A \subseteq B$ then $\sup A \leq \sup B$.*

Proof: If a is any element of A then $a \in B$ since $A \subseteq B$, and thus $a \leq \sup B$ since $\sup B$ is an upper bound for B. But this means that $\sup B$ is an upper bound for A, and therefore $\sup A \leq \sup B$ by the definition of $\sup A$ as the smallest upper bound of A. ∎

The next result gives a useful characterization of the supremum of those sets that are bounded above.

1.12 Theorem: *If A is a nonempty subset of $\mathbb{R}$ that is bounded above and if $s \in \mathbb{R}$ is an upper bound for A, then $s = \sup A$ if and only if*

(i) $a \leq s$ for all $a \in A$ and

(ii) $A \cap (s - \epsilon, s] \neq \emptyset$ for all $\epsilon > 0$.

Proof: Suppose first that $s = \sup A$. Then s is an upper bound for A and thus (i) holds. If (ii) were false there would be an $\epsilon > 0$ such that $A \cap (s - \epsilon, s] = \emptyset$. Then, since s is an upper bound for A and since $A \cap (s - \epsilon, s] = \emptyset$, $s - \epsilon$ would also be an upper bound for A. But this contradicts the fact that s is the least upper bound for A, and therefore (ii) must be true.

Now suppose that, conversely, (i) and (ii) are true. Then (i) implies that s is an upper bound for A. If s were not the least upper bound for A then there would be an upper bound t for A such that $t < s$. Thus $a \leq t$ for all $a \in A$ and so if $\epsilon = (s - t)/2$ then $t + \epsilon = s - \epsilon$ and $a < t + \epsilon$ for all $a \in A$, and thus $A \cap (s - \epsilon, s] = \emptyset$. This contradicts (ii) and shows that $s = \sup A$. ∎

Theorem 1.12 implies the following assertion that is sometimes useful.

1.13 Corollary: *If A is a nonempty subset of $\mathbb{R}$ that is bounded above and if $s = \sup A$ then either $s \in A$ or else $A \cap (s - \epsilon, s) \neq \emptyset$ for all $\epsilon > 0$.*

Exercises

E1.1 For $j_1, j_2, k \in \mathbb{Z}$, prove the distributive rule $(j_1 + j_2) \cdot k = j_1 \cdot k + j_2 \cdot k$.

E1.2 Show that the relations $<$ and $\leq$ on $\mathbb{Z}$ have the following properties:

1. $[(0, j)] < [(0, 0)]$ for all $j \in \mathbb{Z}_{>0}$;
2. $[(0, j)] < [(k, 0)]$ for all $j, k \in \mathbb{Z}_{>0}$;
3. $[(0, j)] < [(0, k)]$, $j, k, \in \mathbb{Z}_{\geq 0}$, if and only if $k < j$;
4. $[(0, 0)] < [(j, 0)]$ for all $j \in \mathbb{Z}_{>0}$;
5. $[(j, 0)] < [(k, 0)]$, $j, k \in \mathbb{Z}_{\geq 0}$, if and only if $j < k$;
6. $[(0, j)] \leq [(0, 0)]$ for all $j \in \mathbb{Z}_{\geq 0}$;
7. $[(0, j)] \leq [(k, 0)]$ for all $j, k \in \mathbb{Z}_{\geq 0}$;
8. $[(0, j)] \leq [(0, k)]$, $j, k, \in \mathbb{Z}_{\geq 0}$, if and only if $k \leq j$;
9. $[(0, 0)] \leq [(j, 0)]$ for all $j \in \mathbb{Z}_{\geq 0}$;
10. $[(j, 0)] \leq [(k, 0)]$, $j, k \in \mathbb{Z}_{\geq 0}$, if and only if $j \leq k$.

E1.3 Show that a subset $A \subseteq \mathbb{Q}$ is bounded if and only it is has a lower bound and an upper bound.

E1.4 For each of the following subsets of $\mathbb{R}$, find the least upper bound and greatest lower bound if they exist.

(a) $\{1, 3\}$

(b) $[0, 4]$

(c) $\left\{\frac{1}{n} \;\middle|\; \in \mathbb{Z}_{>0}\right\}$

(d) $\left\{\frac{n}{n+1} \;\middle|\; n \in \mathbb{Z}_{>0}\right\}$

(e) $\left\{n + \frac{(-1)^n}{n} \;\middle|\; n \in \mathbb{Z}_{>0}\right\}$

(f) $\bigcap_{n=1}^{\infty} \left(1 - \frac{1}{n}, 1 + \frac{1}{n}\right)$

E1.5 Find the least upper bound and the greatest lower bound of the sets below. For any given $\epsilon > 0$ find a number in the set that is greater than $\sup A - \epsilon$ and a number in the set that is smaller than $\inf A + \epsilon$.

(a) $A = \left\{\frac{4+x}{x} \;\middle|\; x \geq 1\right\}$

(b) $A = \left\{\frac{\sqrt{x-1}}{x} \;\middle|\; x \geq 2\right\}$

(c) $A = \{x \mid x^2 - x < 6\}$

E1.6 If $S = \{x \in \mathbb{R} \mid x^2 + x < 3\}$, find $\sup S$ and $\inf S$.
Hint: *Draw a graph.*

E1.7 Let $A = \left\{\frac{n}{n+1} \;\middle|\; n \in \mathbb{Z}_{>0}\right\}$ and notice that $\sup A = 1$. For $\epsilon = 0.01$ and for $\epsilon = 0.0001$ find an element of A larger that $1 - \epsilon$.

E1.8 Suppose A is a nonempty bounded subset of $\mathbb{R}$ and let $-A$ denote the set $\{-x \mid x \in A\}$. Show that $\sup(-A) = -\inf A$.

E1.9 For each nonempty subset S of $\mathbb{R}$ and each real number k, let $kS = \{ks \mid s \in S\}$.

(a) Suppose that $S = \{-1, 2, 4\}$ (so that $3S = \{-3, 6, 12\}$ and $-2S = \{2, -4, -8\}$). Find $\sup S$, $\sup 3S$, $\sup -2S$, $\inf S$, $\inf 3S$, and $\inf -2S$. Do you notice a pattern?

(b) If S is any nonempty bounded subset of $\mathbb{R}$ show that:

1. If $k \geq 0$ then $\sup kS = k \cdot \sup S$ and $\inf kS = k \cdot \inf S$.
2. If $k < 0$ then $\sup kS = k \cdot \inf S$ and $\inf kS = k \cdot \sup S$.

E1.10 Prove that, if a subset S of $\mathbb{R}$ has a maximal element s (that is, there exists an $s \in S$ such that $x \leq s$ for all x in S), then $s = \sup S$.

E1.11 Prove that $x = \sup\{q \in \mathbb{Q} \mid \ q < x\}$ for each $x \in \mathbb{R}$.

E1.12 Suppose that A and B are two nonempty bounded subsets of $\mathbb{R}$ and let

$$A + B = \{x + y \mid \ x \in A \text{ and } y \in B\}.$$

Show that

$$\sup(A + B) = \sup A + \sup B.$$

E1.13 Let A and B be two nonempty sets of real numbers that are bounded above and put $a = \sup A$, $b = \sup B$, and

$$AB = \{xy \mid \ x \in A \text{ and } y \in B\}.$$

Answer the following questions.

(a) Give an example to show that, in general, $\sup(AB) \neq ab$.

(b) Show that, if $a < 0$ and $b < 0$, then $\inf AB = ab$.

(c) Show that, if $\inf A > 0$ and $\inf B0$, then $\sup AB = ab$.

E1.14 For nonempty sets $A, B \subseteq \mathbb{R}$ determine which of the following statements are true and which are false. Prove those that are true and give counterexamples for those that are false.

(a) $\sup(A \cap B) \leq \min\{\sup A, \sup B\}$

(b) $\sup(A \cap B) = \min\{\sup A, \sup B\}$

(c) $\sup(A \cup B) \geq \max\{\sup A, \sup B\}$

(d) $\sup(A \cup B) = \max\{\sup A, \sup B\}$

E1.15 Let A be a nonempty subset of $\mathbb{R}$ that is bounded above and put $s = \sup A$.

(a) Show that if $s \notin A$ then the set $A \cap (s - \epsilon, s)$ is infinite for any $\epsilon > 0$.

(b) Give an example to show that, if $s \in A$, then the set $A \cap (s - \epsilon, s)$ may be finite for each $\epsilon > 0$.

E1.16 Let a and b be two numbers satisfying $a < b$.

(a) Show that the set $\{x \mid \ a < x < b\}$ contains neither its least upper bound not its greatest lower bound.

(b) Show that the set $\{x \mid \ a \leq x \leq b\}$ contains both its least upper bound and its greatest lower bound.

E1.17 Let S denote the set consisting of all those numbers with decimal expansions of the form $x = 0.a_1a_2a_3\ldots$, where all but finitely many of the digits $a_1, a_2, \ldots$ are 5 or 6. Find $\sup S$ and $\inf S$.

E1.18 Suppose that P is a subset of $[0, \infty)$ with the property that, for any integer k, there is an $x_k \in P$ such that $kx_k \leq 1$. Prove that $0 = \inf P$.

E1.19 If P and Q are two subsets of $\mathbb{R}$ such that $\sup P = \sup Q$ and $\inf P = \inf Q$ does it follow that $P = Q$? Support your answer with either a proof or a counterexample.

E1.20 For any two nonempty subsets of $\mathbb{R}$ let us write $P \leq Q$ if, for each $x \in P$, there is a $y \in Q$ satisfying $x \leq y$.

(a) Show that, if $P \leq Q$, then $\sup P \leq \sup Q$.

(b) Give an example to show that, if $P \leq Q$, then it does not follow that $\inf P \leq \inf Q$.

(c) Give an example to show that, if $P \leq Q$ and if $Q \leq P$, then it does not follow that $P = Q$.

E1.21 Let S be a nonempty subset of $\mathbb{R}$ that is bounded below. Show that $\inf S = \sup\{x \in \mathbb{R} \mid x \text{ is a lower bound for } S\}$.

E1.22 Prove that every nonnegative real number has a nonnegative square root, i.e., prove that, if x is any nonnegative real number, then there is nonnegative number y such that $y^2 = x$.
Hint: *Consider the set* $\sup\{u \in \mathbb{R} \mid u^2 \leq x\}$.

E1.23 One of the important consequences of Theorem 1.9 is the Archimedean property of $\mathbb{R}$: The set $\mathbb{Z}_{>0}$ of positive integers is not bounded above in $\mathbb{R}$. Although this property may seem obvious, its proof actually depends on Theorem 1.9. Prove the Archimedean property of $\mathbb{R}$.

E1.24 Suppose that D is a nonempty subset of $\mathbb{R}$ and that f and g are two functions from D to $\mathbb{R}$ such that $f(x) \leq g(y)$ for all $x, y \in D$. Show that $f(D)$ is bounded above, that $g(D)$ is bounded below, and that $\sup f(D) \leq \inf g(D)$. (Recall that $f(D) = \{f(x) \mid x \in D\}$.)

E1.25 Let D be a nonempty subset of $\mathbb{R}$ and suppose that f and g are two real-valued functions defined on D. Let $f + g$ denote the function on D defined by the formula $(f + g)(x) = f(x) + g(x)$ for all $x \in D$.

(a) If $f(D)$ and $g(D)$ are bounded above prove that $(f + g)(D)$ is bounded above and that $\sup(f + g)(D) \leq \sup f(D) + \sup g(D)$.

(b) Show by means of an example that the inequality in (a) may be strict.

(c) State and prove the analog of (a) for infima.

Chapter 2

Sequences of real numbers

Sequences arise in a variety of ways in real analysis. They show up, for example, when considering convergence of infinite series, such as we shall discuss in Chapter 5. What is less clear to a newcomer, however, is that sequences in some way characterize the very fabric of the set $\mathbb{R}$ of real numbers, and also its multi-dimensional version $\mathbb{R}^d$.

2.1 Definitions and examples

We get the ball rolling with the basic definitions and some elementary examples.

2.1 Definition: Let $\mathbb{Z}_{>0}$ denote the natural numbers, i.e., the positive integers. A ***sequence*** (of real numbers) is a function from $\mathbb{Z}_{>0}$ to $\mathbb{R}$. If a is a sequence, it is usual to write a_n in place of $a(n)$, to call a_n the ***nth term*** of the sequence, and to denote the sequence by (a_n), or by $(a_n)_{n\in\mathbb{Z}_{>0}}$ if we wish for the index set to be explicit. •

It is customary to think of terms of a sequence as being ordered, so that a_5 "occurs before" a_7 in the sequence, and so on. It is sometimes convenient to "start" a sequence with a_0 (or a_2 or even a_4) instead of a_1.

Sequences may be defined explicitly, e.g., by writing

$$a_n = \frac{n}{n+1},$$

or by ***recursion***, e.g., by the formulae

$$a_1 = 1, \quad a_{n+1} = \frac{a_n}{n+1}$$

or

$$a_0 = a_1 = 1, \quad a_{n+2} = a_{n+1} + a_n.$$

It is common (but not strictly correct) to describe a sequence by its first few terms, e.g.,

$$2, 4, 8, 16, \ldots.$$

The difficulty is that the terms are not really determined this way—there is an interesting sequence whose first five terms are $2, 4, 8, 16, 31$ (the number of regions obtained by joining every pair of points among n points in general position on the circumference of a circle for $n = 2, 3, 4, 5, \ldots$).

Beware of the difference between a *sequence* and the *range* of that sequence. The range of a sequence is analogous to the range of a function, i.e., the set of values attained by the sequence (or function). The following two examples illustrate this.

2.2 Examples: 1. The sequence $1, 2, 3, 1, 2, 3, 1, 2, 3, \ldots$ has an infinite number of terms, but its range contains only the three points $1, 2, 3$.

2. The range of the sequence $(\sin \frac{n\pi}{2})$ is the set $\{-1, 0, 1\}$. •

2.3 Definition: A sequence (a_n) is said to be ***bounded*** if there exist an $M > 0$ such that $|a_n| \leq M$ for all n. •

Notice that, according to this definition, a sequence is bounded if and only if its range is a bounded set in the sense of Definition 1.7.

2.4 Definition: A sequence (a_n) is said to have a ***limit*** L if, for any $\epsilon > 0$, there exists a number N (depending on ϵ) such that

$$|a_n - L| < \epsilon$$

for every integer $n \geq N$. If (a_n) has a limit L, then we say that the sequence is ***convergent*** and that it ***converges*** to L and we write $\lim_{n\to\infty} a_n = L$. If (a_n) is not convergent, it is said to ***diverge*** or to be ***divergent***. If, for every $c > 0$, there exists a number N such that $a_n > c$ for all $n \geq N$, we will write $\lim_{n\to\infty} a_n = \infty$ and say that (a_n) ***diverges to*** $\boldsymbol{\infty}$. Similarly, if, for every $c > 0$, there exists a number N such that $a_n < -c$ for all $n \geq N$, we will write $\lim_{n\to\infty} a_n = -\infty$ and say that (a_n) ***diverges to*** $\boldsymbol{-\infty}$. •

2.5 Examples: 1. If $a_n = n$, then $\lim_{n\to\infty} a_n = \infty$ or (a_n) diverges to ∞.

2. If $a_n = (-1)^n$ then $\lim_{n\to\infty} a_n$ does not exist, (a_n) is divergent, but it does not diverge to $\pm\infty$.
3. If $a_n = \frac{n}{n+1}$ then $\lim_{n\to\infty} a_n = 1$ or (a_n) converges to 1. •

In simple cases, such as Example 2.5–3, we can establish that the limit L exists by determining L (and then, if necessary, applying Definition 2.4 to prove that it is the limit). However, in many important cases it is not easy to find L or to apply the definition directly. For such cases we need general theorems to help us establish whether a limit exists. The two most important such results are Theorem 2.9 and Theorem 2.23. The next theorem is a first negative step in this direction (negative because it says "an unbounded sequence cannot converge").

2.6 Proposition: *Every convergent sequence of real numbers is bounded.*

Proof: Suppose that (a_n) is a convergent sequence that converges to L. Then, by definition, for any $\epsilon > 0$ there exists an $N > 0$ such that $|a_n - L| < \epsilon$ for all $n \geq N$. In particular, choosing $\epsilon = 1$ gives an $N > 0$ such that $L - 1 < a_n < L + 1$ for all $n \geq N$. The set $\{a_1, a_2, \ldots, a_{N-1}\}$ is bounded because it is finite and hence has a minimum and a maximum; let A and B be the minimum and maximum, respectively. So if $m = \min\{A, L-1\}$ and if $M = \max\{B, L+1\}$, then clearly $m \le a_n \le M$ for all n and so (a_n) is bounded. ■

Example 2.5–2 shows that the converse of Proposition 2.6 is false and so boundedness is necessary but not sufficient for convergence. To obtain a sufficient condition for convergence, we introduce the following definition.

2.7 Definition: A sequence (a_n) is said to be ***nondecreasing*** if $a_m \le a_n$ whenever $m < n$ and to be ***strictly increasing*** if $a_m < a_n$ whenever $m < n$. Similarly, the sequence is said to be ***nonincreasing*** if $a_m \ge a_n$ whenever $m < n$ and to be ***strictly decreasing*** if $a_m > a_n$ whenever $m < n$. Finally, (a_n) is said to be ***monotone*** if it is either nondecreasing or nonincreasing. •

Note that a strictly increasing sequence is nondecreasing and a strictly decreasing sequence is nonincreasing. The point of this definition is to have a formal way to avoid the ambiguous term "increasing," which sometimes means strictly increasing and, at other times, nondecreasing.

If (a_n) is a nondecreasing sequence and if $A = \{a_n \mid n \in \mathbb{Z}_{>0}\}$, then clearly $a_1 = \inf A$ and (a_n) is bounded if and only if A is bounded above.

The following example shows that a strictly increasing sequence may be bounded or unbounded. The terms in part 1 of this example are the partial sums of the exponential series while those in part 2 are the partial sums of the harmonic series, cf. Example 5.2–4.

2.8 Examples: 1. If

$$a_n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}$$

then $a_n < a_{n+1}$ for all n and, since $n! > 2^{n-1}$ if $n \ge 2$,

$$\begin{aligned} a_n &< 1 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{2^{n-1}} \\ &= 3 - 2^{1-n} < 3 \end{aligned}$$

by the familiar formula for the sum of a geometric series. The sequence (a_n) is therefore strictly increasing and bounded.

2. If

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

then $a_{n+1} > a_n$ for all n and

$$\begin{aligned} a_{2^n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \cdots + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^n}\right) \\ &> 1 + \frac{1}{2} + \underbrace{\left(\frac{1}{4} + \frac{1}{4}\right)}_{2 \text{ terms}} + \underbrace{\left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right)}_{4 \text{ terms}} + \cdots + \underbrace{\left(\frac{1}{2^n} + \cdots + \frac{1}{2^n}\right)}_{2^{n-1}\ textrmterms} \\ &> 1 + \frac{1}{2} + \cdots + \frac{1}{2} = 1 + \frac{n}{2}. \end{aligned}$$

The sequence (a_n) is thus strictly increasing and unbounded. •

The two series in this example illustrate the next theorem.

2.9 Theorem: *Every bounded monotone sequence is convergent. More precisely, if (a_n) is a bounded nondecreasing sequence then $\lim_{n\to\infty} a_n = \sup\{a_n \mid n \in \mathbb{Z}_{>0}\}$.*

Proof: Suppose that (a_n) is a bounded nondecreasing sequence. Then $\{a_n \mid n \in \mathbb{Z}_{>0}\}$ is a nonempty bounded set; let $L = \sup\{a_n \mid n \in \mathbb{Z}_{>0}\}$. Then, by Theorem 1.12, given any $\epsilon > 0$, there is an m such that $L - \epsilon < a_m \leq L$. Now, for any $n \geq m$, we have $a_m \leq a_n \leq L$ by the definition of L and thus $|L - a_n| < \epsilon$. This shows that (a_n) converges and that $\lim_{n\to\infty} a_n = L$.

The case of a nonincreasing sequence is similar. ■

Now looking back at Example 2.8, we see that the sequence in part 1 is convergent by Theorem 2.9, while the sequence in 2 is divergent by Proposition 2.6. Note that Proposition 2.6 and Theorem 2.9 together imply that a monotone sequence is convergent if and only if it is bounded.

While Theorem 2.9 is a very useful result, we require some additional ideas in order to discuss sequences that are not monotone and/or do not converge.

2.2 Subsequences

A subsequence of a sequence is, roughly speaking, a sequence that is obtained by taking only certain terms from the given sequence in the given order. Thus

$$2, 4, 6, 8, \ldots$$

is a subsequence of $1, 2, 3, \ldots$ whereas

$$2, 1, 8, 7, 3, 2, 31, \ldots$$

is not. It is possible for a subsequence of a divergent sequence to converge. For example, both of the sequences $1, 1, 1, \ldots$ and $2, 2, 2, \ldots$ are convergent subsequences of the divergent sequence $1, 2, 3, 1, 2, 3, 1, 2, 3, \ldots$.

Slightly more generally, if (a_n) is a sequence, then (a_{2n}) or $a_2, a_4, a_6, \ldots$ is a subsequence and (a_{3n+1}) or $a_1, a_4, a_7, \ldots$ is another. Here is a formal definition.

2.10 Definition: If (a_n) is a sequence and if (n_k) is a strictly increasing sequence of positive integers, then the sequence (a_{n_k}) is said to be a ***subsequence*** of (a_n). •

2.11 Example: Suppose that (a_n) is a sequence that is bounded above and put $t = \sup\{a_n \mid n \in \mathbb{Z}_{>0}\}$. We consider two cases.

1. If $a_n = t$ for some n, then t may or may not be the limit of a convergent subsequence whereas, if $a_n < t$ for all n, then t is the limit of a convergent subsequence. To see this, first notice that, if (a_n) is the sequence $1, 0, 0, \ldots$, then $t = 1$ and no subsequence of (a_n) converges to 1.

2. Now suppose that $a_n < t$ for all n. We will show that, in this case, there is a subsequence (a_{n_k}) that converges to t.
   First we make an observation. For any number $\epsilon > 0$ and any positive integer m, consider the number
   $$\epsilon' = \min\{\epsilon/m, t - a_1, t - a_2, \ldots, t - a_m\}.$$
   Then $\epsilon' > 0$ and so, by Theorem 1.12, there will be an integer n such that $t - \epsilon'_m < a_n < t$, and $n > m$ by the definition of ϵ'. We shall use this observation repeatedly to construct the subsequence (n_k).
   Let $n_1 = 1$ and let $\epsilon_1 = t - a_1$. Then $\epsilon_1 > 0$ and there is an n_2 such that $t - \epsilon_1 < a_{n_2} < t$ and $n_2 > n_1$. Now, if
   $$\epsilon_2 = \min\{\epsilon_1/2, t - a_1, t - a_2, \ldots, t - a_{n_2}\},$$
   then there is an n_3 such that $t - \epsilon_2 < a_{n_3} < t$ and $n_3 > n_2$. Next, if
   $$\epsilon_3 = \min\{\epsilon_2/2, t - a_1, t - a_2, \ldots, t - a_{n_3}\},$$
   then there is an n_4 such that $t - \epsilon_3 < a_{n_3} < t$ and $n_4 > n_3$. Continuing in this manner will give a subsequence (a_{n_k}) that converges to t. •

The following lemma (whose proof is easy and will be omitted) asserts that every subsequence of a convergent sequence converges to the same limit. What turns out to be much more interesting is that, as several of the above examples show, a subsequence of a divergent sequence may well converge.

2.12 Lemma: *If (a_{n_k}) is a subsequence of a sequence (a_n) and if $\lim_{n\to\infty} a_n = L$, then $\lim_{k\to\infty} a_{n_k} = L$.*

The next theorem turns out to be crucial for the subsequent development of the theory and is even equivalent to the least upper bound property.

2.13 Theorem: (Bolzano–Weierstrass theorem) *Every bounded sequence has a convergent subsequence.*

Proof: Suppose that (a_n) is a bounded sequence and let m and M be numbers satisfying $m \le a_n \le M$ for all n. Put $c = (m + M)/2$, the midpoint of $[m, M]$. If the inequality $m \le a_n \le c$ holds for infinitely many n, let $m_1 = m$ and $M_1 = c$. On the other hand, if the inequality $m \le a_n \le c$ holds for only finitely many n, then the inequality $c \le a_n \le M$ must hold for infinitely many n, and in this case put $m_1 = c$ and $M_1 = M$. So, in either case $[m_1, M_1] \subseteq [m, M]$, $M_1 - m_1 = 2^{-1}(M - m)$, and the inequality $m_1 \le a_n \le M_1$ holds for infinitely many n. Now let n_1 be any integer for which $m_1 \le a_{n_1} \le M_1$.

The argument just described can be repeated. Let c_1 be the midpoint of $[m_1, M_1]$. Since the inequality $m_1 \le a_n \le M_1$ holds for infinitely many n, it follows that $m_1 \le a_n \le c_1$ for infinitely many n or $c_1 \le a_n \le M_1$ for infinitely many n or both.

So by letting either $m_2 = m_1$ and $M_2 = c_1$ or by letting $m_2 = c_1$ and $M_2 = M_1$ we can be sure that $[m_2, M_2] \subseteq [m_1, M_1]$, that

$$M_2 - m_2 = 2^{-1}(M_1 - m_1) = 2^{-2}(M - m),$$

and that $m_2 \leq a_n \leq M_2$ for infinitely many n. Now let n_2 be any integer satisfying $n_2 > n_1$ and $m_2 \leq a_{n_2} \leq M_2$.

We can evidently continue this construction (or more formally, proceed by recursion) and doing so will yield two sequences (m_k) and (M_k) of numbers and a sequence (n_k) of positive integers with the following properties:

$$\begin{gathered} m_1 \leq m_2 \leq \cdots ; \\ M_1 \geq M_2 \geq \cdots ; \\ n_1 < n_2 < \cdots ; \\ M_k - m_k = 2^{-k}(M - m); \\ m_k \leq a_{n_k} \leq M_k; \end{gathered}$$

for each k.

Now (m_k) is a bounded nondecreasing sequence and so, by Theorem 2.9, it converges to $L = \sup\{m_k \mid k \in \mathbb{Z}_{>0}\}$. It is easy to see that, if j and k are any two positive integers, then $m_j < M_k$. Indeed, if $j \leq k$ then $m_j \leq m_k < M_k$, whereas, if $j \geq k$, then $m_j < M_j \leq M_k$. This means that, for each $k \in \mathbb{Z}_{>0}$, M_k is an upper bound for the set $\{m_j \mid j \in \mathbb{Z}_{>0}\}$ and hence $L \leq M_k$. It now follows that both a_{n_k} and L belong to the interval $[m_k, M_k]$ and hence that

$$|L - a_{n_k}| \leq M_k - m_k = 2^{-k}(M - m)$$

for each k, and this clearly implies that $\lim_{k\to\infty} a_{n_k} = L$. This means that (a_{n_k}) is the required convergent subsequence. ■

Convergent sequences are in some sense easy to deal with, but we need to gain a better understanding of divergent sequences. Recall from Lemma 2.12 and Theorem 2.13 that, if a sequence converges to a limit L, then every subsequence converges to L. On the other hand, it is clear from Example 2.14 (below) that a subsequence of a divergent sequence may converge or diverge, and that the convergent subsequences need not all converge to the same limit. This suggests that a suitable "replacement" for the limit of a convergent sequence in the context of arbitrary sequences might be either

1. the set of all limits of convergent subsequences or
2. the largest and smallest numbers that are the limits of convergent subsequences.

Now 2 seems to be simpler than 2, but there is a potential problem with 2. Namely, even if the set of all limits of convergent subsequences is bounded, it is not clear that this set contains its least upper bound or, equivalently, that it has a largest element. The next lemma will show that this is, in fact, not a problem and that we really can use 2. But before giving this lemma it will be useful to first present some examples of sequences and, for each one, the set all limits of convergent subsequences.

2.14 Example: Here are six sequences and, for each one, the corresponding set of all limits of convergent subsequences:

Sequence	Set of limits of all convergent subsequences
$1, \frac{1}{2}, \frac{1}{3}, \ldots$	$\{0\}$
$1, -1, 1, -1, \ldots$	$\{-1, 1\}$
$1, 2, 3, \ldots$	$\emptyset$
$1, \frac{3}{2}, \frac{1}{3}, \frac{5}{4}, \frac{1}{5}, \frac{7}{6}, \ldots$	$\{0, 1\}$
$0, 1, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \ldots, \frac{7}{8}, \frac{1}{16}, \ldots$	$[0, 1]$
(a_n) is an enumeration of $\mathbb{Q}$	$\mathbb{R}$

Let us characterize the convergent subsequences of a sequence.

2.15 Lemma: *Suppose that (a_n) is a sequence and let T denote the set of all limits of convergent subsequences.*

(i) Suppose that $T \neq \emptyset$.

(a) If $\sup T < \infty$, then $\sup T \in T$ and, if $\inf T > -\infty$, then $\inf T \in T$.

(b) If $\sup T = \infty$, there is a subsequence of (a_n) diverging to ∞ and, if $\inf T = -\infty$, there is a subsequence of (a_n) diverging to $-\infty$.

(ii) If $T = \emptyset$ then no subsequence of (a_n) converges and there is a subsequence diverging to ∞ or to $-\infty$ (it is possible that some subsequences of (a_n) diverge to ∞ and others diverge to $-\infty$).

Proof: (i) Suppose first that T is not empty and that $\sup T < \infty$. Let $t = \sup T$ and let $\epsilon > 0$. Theorem 1.12 implies that $T \cap (t-\epsilon, t] \neq \emptyset$ and hence there is a convergent subsequence of (a_n) whose limit lies in the interval $(t-\epsilon, t]$. Now if (a_{n_k}) is such a subsequence then $t - \epsilon < a_{n_k} < t + \epsilon$ for infinitely many values of k and, therefore, a_n must lie in the interval $(t-\epsilon, t+\epsilon)$ for infinitely many values of n.

Since this is true for each $\epsilon > 0$ and it follows that, for each positive integer k, the inequality $t - 1/k < a_n < t + 1/k$ must hold for infinitely many values of n. We can use this to build a subsequence whose limit is t as follows. Let n_1 be any positive integer satisfying $t - 1 < a_{n_1} < t + 1$, let n_2 be any integer satisfying $n_2 > n_1$ and $t - 1/2 < a_{n_2} < t + 1/2$, let n_3 be any integer satisfying $n_3 > n_2$ and $t - 1/3 < a_{n_3} < t + 1/3$, etc. This construction gives a subsequence (a_{n_k}) that converges to t and therefore $t \in T$.

This proves the assertion in (i a) dealing with the case $\sup T < \infty$, and the proof of the one dealing with the case $\inf T > -\infty$ is similar and will be omitted.

Now suppose that T is not empty and that $\sup T = \infty$. Then, for any positive integer p, there is a convergent subsequence of (a_n) whose limit lies in the interval (p, ∞); if q is any other integer, there will be an integer $n > q$ such that $a_n > p$. This means that, given any two integers p and q, there is an integer n such that $n > q$ and $a_n > p$. We can use this observation to build a subsequence of (a_n) that diverges to ∞ as follows. Let n_1 be any positive integer satisfying $1 < a_{n_1}$, let n_2 be any integer satisfying $n_2 > n_1$ and $2 < a_{n_2}$, let n_3 be any integer satisfying $n_3 > n_2$ and $3 < a_{n_3}$, etc.

This proves the assertion in (i b) dealing with the case $\sup T = \infty$, and the proof of the one dealing with the case $\inf T = -\infty$ is similar and will be omitted.

(ii) Now suppose that $T = \emptyset$. The definition of T then implies that no subsequence of (a_n) can converge to a real number. Consider a positive integer k and the inequality $-k < a_n < k$. If this inequality holds for infinitely many n, then (a_n) would have a bounded subsequence and this subsequence would, by the Bolzano–Weierstrass theorem, have a convergent subsequence. But, since a subsequence of a subsequence is a subsequence, this is impossible by the assumption that $T = \emptyset$. This shows that, for each integer k, the inequality $-k < a_n < k$ holds for only finitely many integers n. So, for each integer k, the inequality $|a_n| \geq k$ must hold for infinitely many integers n, and hence at least one of the inequalities $a_n \geq k$ and $a_n \leq -k$ must hold for infinitely many values of n. We can use this observation to build a subsequence of (a_n) that diverges to ∞ or to $-\infty$ as follows.

First suppose that, for each positive integer k, the inequality $a_n \geq k$ holds for infinitely many values of n. Let n_1 be any integer satisfying $a_{n_1} \geq 1$, let n_2 be any integer satisfying $n_2 > n_1$ and $a_{n_2} \geq 2$, let n_3 be any integer satisfying $n_3 > n_2$ and $a_{n_3} \geq 3$, etc. Then (a_{n_k}) is clearly a subsequence diverging to ∞.

Now suppose that there is a positive integer k such that the inequality $a_n \geq k$ holds for only finitely many values of n. Then, for each positive integer k, the inequality $a_n \leq -k$ must hold for infinitely many values of n, and a similar argument will lead to a subsequence that diverges to $-\infty$. ■

Consider a sequence (a_n) and let T be as in the preceding lemma. We want to associate with this sequence two elements of the set $\mathbb{R} \cup \{-\infty, \infty\}$ to be denoted by $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$ with the following desiderata in mind. If the sequence converges to a real number, say L, then both $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$ are to be equal to L. More generally, $\limsup_{n\to\infty} a_n$ is to be the largest element of the set $\mathbb{R} \cup \{-\infty, \infty\}$ to which a subsequence of (a_n) converges or diverges, and $\liminf_{n\to\infty} a_n$ is to be the smallest element to which a subsequence converges or diverges (where, of course, we regard $-\infty < x < \infty$ for all $x \in \mathbb{R}$). The point of the above lemma is that this really does make sense. For example, if, say, $T \neq \emptyset$ and $\sup T < \infty$, then $\sup T$ belongs to T and so is the limit of a convergent subsequence. In this example, then, $\limsup_{n\to\infty} a_n$ will be ∞ if a subsequence of (a_n) diverges to ∞ and will be $\sup T$ otherwise. As a second example, if $T \neq \emptyset$ and $\sup T = \infty$, then there is a subsequence diverging to ∞ and $\limsup_{n\to\infty} a_n$ will be ∞. And as a final example, if $T = \emptyset$, then no subsequence converges to a real number and there are subsequences diverging to at least one of $\pm\infty$; so $\limsup_{n\to\infty} a_n$ will be ∞ if there is a subsequence diverging to ∞ and will be $-\infty$ otherwise. The next definition makes this intuitive discussion of the definitions of $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$ precise.

2.16 Definition: Suppose that (a_n) is a sequence and let T be the set of all limits of convergent subsequences.

(i) The ***limit superior*** of the sequence (a_n) is the element $\limsup_{n\to\infty} a_n$ of the set $\mathbb{R} \cup \{-\infty, \infty\}$ defined as follows:

   (a) If (a_n) is not bounded above, then $\limsup_{n\to\infty} a_n = \infty$.

   (b) If (a_n) is bounded above and if $T \neq \emptyset$, then $\limsup_{n\to\infty} a_n$ is the largest number that is the limit of a convergent subsequence of (a_n).

   (c) If (a_n) is bounded above and if $T = \emptyset$, then $\limsup_{n\to\infty} a_n = -\infty$.

(ii) The ***limit inferior*** of the sequence (a_n) is the element $\liminf_{n\to\infty} a_n$ of the set $\mathbb{R} \cup \{-\infty, \infty\}$ defined as follows:

   (a) If (a_n) is not bounded below, then $\liminf_{n\to\infty} a_n = -\infty$.

   (b) If (a_n) is bounded below and if $T \neq \emptyset$, then $\liminf_{n\to\infty} a_n$ is the smallest number that is the limit of a convergent subsequences of (a_n).

   (c) If (a_n) is bounded below and if $T = \emptyset$, then $\liminf_{n\to\infty} a_n = \infty$. •

The idea that the limit superior and limit inferior are in some sense replacements for the limit is borne out by the following result.

2.17 Proposition: *Suppose that (a_n) is a sequence.*

(i) If the sequence (a_n) converges to a number or diverges to $\pm\infty$ then

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n = \limsup_{n\to\infty} a_n. \tag{2.1}$$

(ii) If $\liminf_{n\to\infty} a_n = \lim_{n\to\infty} a_n$ then either the sequence (a_n) converges or diverges to $\pm\infty$; in either case (2.1) holds.

Proof: It is possible to prove this directly from the above definitions, but it is more easily deduced from the properties of the limit superior and inferior contained in Propositions 2.19 and 2.20 (cf. Exercise E2.44). ■

2.18 Example: This example is a continuation of Example 2.14.

Sequence	$\liminf_{n\to\infty} a_n$	$\limsup_{n\to\infty} a_n$
$1, \frac{1}{2}, \frac{1}{3}, \ldots$	0	0
$1, -1, 1, -1, \ldots$	-1	1
$1, 2, 3, \ldots$	∞	∞
$1, \frac{3}{2}, \frac{1}{3}, \frac{5}{4}, \frac{1}{5}, \frac{7}{6}, \ldots$	0	1
$0, 1, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \ldots, \frac{7}{8}, \frac{1}{16}, \ldots$	0	1
(a_n) is an enumeration of $\mathbb{Q}$	$-\infty$	∞

The next two propositions contain important properties of the limit superior; the corresponding properties of the limit inferior can be deduced from these propositions and the easily verified formula

$$\liminf_{n\to\infty} a_n = -\limsup_{n\to\infty}(-a_n) \tag{2.2}$$

(cf. Exercise E2.35).

2.19 Proposition: *Let (a_n) be a sequence of numbers and let L be a number.*

(i) If $L = \limsup_{n\to\infty} a_n$ then

(a) for each $\epsilon > 0$ the inequality $a_n > L+\epsilon$ holds for only finitely many n and

(b) for each $\epsilon > 0$ the inequality $a_n > L-\epsilon$ holds for infinitely many n.

(ii) If conditions (i a) and (i b) hold, then $L = \limsup_{n\to\infty} a_n$.

Proof: (i) Suppose that $\limsup_{n\to\infty} a_n = L$. Since the limit superior of (a_n) is a real number, it follows from Definition 2.16 that this sequence is bounded above (by M, say), that L is a limit of a convergent subsequence, and that no number larger than L is the limit of a convergent subsequence. If (i a) did not hold, then, for some $\epsilon > 0$, we would have

$$a_n > L + \epsilon \text{ for infinitely many } n.$$

But then there would be a subsequence of (a_n) all of whose terms belong to the interval $[L+\epsilon, M]$. By the Bolzano–Weierstrass theorem, the subsequence would itself have a subsequence converging to a point in $[L+\epsilon, M]$. But then this limit would also be the limit of a convergent subsequence of (a_n), contradicting the definition of L. If (i b) did not hold, then, for some $\epsilon > 0$, $a_n > L-\epsilon$ for only finitely many n, contradicting the fact that L is the limit of a convergent subsequence of (a_n).

(ii) Now suppose, conversely, that L satisfies (i a) and (i b) and let $L' = \limsup_{n\to\infty} a_n$. Consider first the case in which $L > L'$, and put $\epsilon = (L-L')/2$. Then the definition of L' and part (i) implies that the inequality

$$a_n > L' + \epsilon = L - \epsilon$$

holds for only finitely many integers n, contradicting the assumption that L satisfies (i b). Now consider the case in which $L' > L$ and put $\epsilon = (L'-L)/2$. Then (just as in the first case) the definition of L' and part (i) implies that the inequality

$$a_n > L' - \epsilon = L + \epsilon$$

holds for only infinitely many integers n, contradicting the assumption that L satisfies (i b).

This means that both of the assumptions $L > L'$ and $L' > L$ lead to contradictions, and thus $L = L'$. ■

Definition 2.16 together with Proposition 2.19 give us two different ways of characterizing the limit superior and limit inferior of a sequence. We get a third characterization by considering the limit of the sequence formed from (a_n) by taking the suprema and infima of its tails.

2.20 Proposition: *For any bounded sequence* (a_n) *we have*

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_k \mid k \geq n\}$$

and

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_k \mid k \geq n\}.$$

Before turning to the proof of this proposition it will be useful to consider an example to help unravel the notation.

2.21 Example: Let (a_n) denote the sequence

$$-2, \frac{1}{4}, \frac{-4}{3}, \frac{1}{16}, \frac{-5}{4}, \frac{1}{64}, \ldots$$

By considering subsequences of this sequence, it is clear that $\limsup_{n\to\infty} a_n = 0$ and that $\liminf_{n\to\infty} a_n = -1$. Put $M_n = \sup\{a_k \mid k \geq n\}$ and $m_n = \inf\{a_k \mid k \geq n\}$ for $n \in \mathbb{Z}_{>0}$. Then

$$\begin{aligned}
M_1 &= \sup\left\{-2, \frac{1}{4}, -\frac{4}{3}, \frac{1}{16}, \ldots\right\} = \frac{1}{4} \\
M_2 &= \sup\left\{\frac{1}{4}, -\frac{4}{3}, \frac{1}{16}, \ldots\right\} = \frac{1}{4} \\
M_3 &= \sup\left\{-\frac{4}{3}, \frac{1}{16}, \ldots\right\} = \frac{1}{16} \\
&\vdots
\end{aligned}$$

and so (M_n) is the sequence

$$\frac{1}{4}, \frac{1}{4}, \frac{1}{16}, \frac{1}{16}, \ldots$$

and thus $\lim_{n\to\infty} M_n = 0$. Similarly, (m_n) is the sequence

$$-2, -\frac{4}{3}, -\frac{4}{3}, -\frac{5}{4}, -\frac{5}{4} \ldots$$

and thus $\lim_{n\to\infty} m_n = -1$. •

Proof of Proposition 2.20: Suppose that A is a lower bound for $\{a_n \mid n \in \mathbb{Z}_{>0}\}$ and let $M_n = \sup\{a_k \mid k \geq n\}$ for each $n \in \mathbb{Z}_{>0}$. Since $\{a_k \mid k \geq n\} \supseteq \{a_k \mid k \geq n+1\}$, it follows from Proposition 1.11 that $M_{n+1} \leq M_n$. Hence (M_n) is a nonincreasing sequence, and, since $A \leq a_k \leq M_k$ for all $k \geq 1$, the sequence (M_n) is bounded below by A. It now follows from Theorem 2.9 that $\lim_{n\to\infty} M_n$ exists and that, if $L = \lim_{n\to\infty} M_n$, then $L = \inf\{M_n \mid n \in \mathbb{Z}_{>0}\}$. To show that $L = \limsup_{n\to\infty} a_n$ it is only necessary to show that L satisfies the two condition in Proposition 2.19.

Let ϵ be a positive number. It follows from Theorem 1.12 that there is an $n \in \mathbb{Z}_{>0}$ such that $M_n \leq L + \epsilon$. But then $a_k \leq L + \epsilon$ for all $k \geq n$, and hence the inequality

$a_k > L+\epsilon$ holds for only finitely many k. Now suppose that the inequality $a_k > L-\epsilon$ holds for only finitely many k. Then there would exist $n \in \mathbb{Z}_{>0}$ such that $a_k \le L-\epsilon$ for all $k \ge n$. But then $M_n \le L - \epsilon$, contradicting the fact that $L = \inf\{M_n \mid n \in \mathbb{Z}_{>0}\}$. So L satisfies conditions (i a) and (i b) of Proposition 2.19 and so equals $\limsup_{n\to\infty} a_n$.

The assertion dealing with the limit inferior can be proven in a similar manner or can be deduced from what was just proven and equation (2.2). ■

2.3 Cauchy sequences

The definition of the limit of a sequence can be used to verify whether or not a given number is the limit of a given sequence. But this definition cannot be used directly in answering the question of whether or not a given sequence converges. Now whether a sequence converges clearly depends only on the sequence itself and, at least in principle, we should be able to decide this solely by examining the sequence itself (and without looking outside the sequence at candidates for its limit). The purpose of this section is to identify a property that a given sequence may or may not satisfy and that is both intrinsic to the sequence and equivalent to its converging.

Consider a convergent sequence (a_n) and let $L = \lim_{n\to\infty} a_n$ and let ϵ be a a positive number. Then there is an integer N such that $|a_n - L| < \epsilon/2$ for all $n > N$. Now

$$|a_m - a_n| \le |a_m - L| + |L - a_n|$$

for any two integers m and n by the triangle inequality, and, therefore, $|a_m - a_n| < \epsilon$ for any two integers $m, n \ge N$. This turns out to be the desired condition.

2.22 Definition: A sequence (a_n) is said to be a ***Cauchy sequence*** if, for every $\epsilon > 0$, there is an integer N such that $|a_n - a_m| < \epsilon$ for all $n, m \ge N$. •

The following result now shows the importance of the notion of a Cauchy sequence.

2.23 Theorem: *A sequence of real numbers is convergent if and only if it is a Cauchy sequence.*

Proof: The discussion preceding the statement of the theorem shows that every convergent sequence is a Cauchy sequence. Now suppose that, conversely, (a_n) is a Cauchy sequence. Taking $\epsilon = 1$ in Definition 2.22 shows that there is an integer N such that $|a_n - a_m| < 1$ for all $m, n \ge N$. This means that all but a finite number of the terms of the sequence (a_n) are contained in the interval $(a_N - 1, a_N + 1)$ and, therefore, (a_n) itself must be a bounded sequence. The Bolzano–Weierstrass theorem (Theorem 2.13) now implies that (a_n) has a convergent subsequence; let (a_{n_k}) be such a subsequence and put $L = \lim_{k\to\infty} a_{n_k}$. The rest of the proof consists of showing that $L = \lim_{n\to\infty} a_n$.

Notice that

$$|a_n - L| = |a_n - a_{n_k} + a_{n_k} - L| \le |a_n - a_{n_k}| + |a_{n_k} - L|$$

for any two positive integers k and n. If ϵ is some positive number, then, since $L = \lim_{k\to\infty} a_{n_k}$, it follows that there will be an integer N_1 such that $|a_{n_k} - L| < \epsilon/2$ whenever $k \geq N_1$. Next, the assumption that (a_n) is a Cauchy sequence implies that there is an integer N_2 such that $|a_n - a_m| < \epsilon/2$ whenever $m, n \geq N_2$. Now let $N = \max\{N_1, N_2\}$ and recall that $n_k \geq k$ for all k by the definition of a subsequence. So, if $n \geq N$, then

$$|a_n - L| \leq |a_n - a_{n_N}| + |a_{n_N} - L| < \epsilon,$$

and this shows that $L = \lim_{n\to\infty} a_n$. ■

2.24 Example: We can use the preceding theorem to show that the sequence whose nth term is

$$a_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n-1}\frac{1}{n}$$

is convergent. To do this, we must examine the expression $|a_m - a_n|$ and, in doing so, we may as well assume that $m > n$. The fact that $\frac{1}{n} - \frac{1}{n+1} > 0$ for any positive integer n implies that, if $m - n$ is even, then

$$\begin{aligned}
|a_m - a_n| &= \left|(-1)^n \frac{1}{n+1} + (-1)^{n+1}\frac{1}{n+2} + \cdots + (-1)^{m+1}\frac{1}{m}\right| \\
&= \left|\frac{1}{n+1} - \frac{1}{n+2} + \frac{1}{n+3} - \frac{1}{n+4} + \cdots + \frac{1}{m-1} - \frac{1}{m}\right| \\
&= \frac{1}{n+1} - \frac{1}{n+2} + \frac{1}{n+3} - \frac{1}{n+4} + \cdots + \frac{1}{m-1} - \frac{1}{m} \\
&= \frac{1}{n+1} - \left(\frac{1}{n+2} - \frac{1}{n+3}\right) - \cdots - \left(\frac{1}{m-2} - \frac{1}{m-1}\right) - \frac{1}{m} \\
&< \frac{1}{n+1}
\end{aligned}$$

whereas, if $m - n$ is odd, then

$$\begin{aligned}
|a_m - a_n| &= \left|\frac{1}{n+1} - \frac{1}{n+2} + \cdots \frac{1}{m-2} - \frac{1}{m-1} + \frac{1}{m}\right| \\
&= \frac{1}{n+1} - \left(\frac{1}{n+2} - \frac{1}{n+3}\right) - \cdots - \left(\frac{1}{m-1} - \frac{1}{m}\right) \\
&< \frac{1}{n+1}.
\end{aligned}$$

Now suppose that an $\epsilon > 0$ is given and let N be an integer satisfying $N + 1 > \frac{1}{\epsilon}$. Then $|a_n - a_m| \leq \frac{1}{n+1} < \epsilon$ whenever $m \geq n \geq N$ and, therefore, (a_n) is a Cauchy, and hence a convergent, sequence.

The value of $\lim_{n\to\infty} a_n$ will be determined in Example 5.13–1 to be $\ln 2$. •

2.25 Example: Let (a_n) be as in Example 2.24. Recall from this example that the limit $\lim_{n\to\infty} a_n$ exists and that, for any integer $m \geq 5$, we have $|a_4 - a_m| < 1/5$. So, if $L = \lim_{n\to\infty} a_n$, then $|a_4 - L| \leq 1/5$, and since

$$a_4 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = \frac{7}{12},$$

we can conclude that $|\frac{7}{12} - L| < \frac{1}{4}$. This is consistent with the assertion made in the previous example that $L = \ln 2$. •

Exercises

E2.1 Determine $\lim_{n\to\infty} a_n$ if:

(a) $a_n = \dfrac{n+1}{3n}$

(b) $a_n = \dfrac{n^3+n}{n^2+n}$

(c) $a_n = \left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{4}\right)\cdots\left(1-\frac{1}{n+1}\right)$

(d) $a_n = \dfrac{1}{n^2} + \dfrac{2}{n^2} + \cdots \dfrac{n}{n^2}$

E2.2 Find $\lim_{n\to\infty} a_n$ in the following cases:

(a) $a_n = \dfrac{n^2}{n+1000}$, $n \geq 1$.

(b) $a_n = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n(n+1)}$, $n \geq 1$.

In each case, you must give a proof using the definition of $\lim_{n\to\infty} a_n$.
Hint: *For (b), express $\frac{1}{k(k+1)}$ as the difference of two fractions.*

E2.3 Using the definition of a limit, find $\lim_{n\to\infty} a_n$, where

$$a_n = \begin{cases} \frac{n^2+1}{\sin(n^2+1)}, & \text{if } n < 2^{1000}, \\ \frac{\sin(n^2+1)}{n^2+1}, & \text{if } n \geq 2^{1000}. \end{cases}$$

E2.4 It is easy to see that $\lim_{n\to\infty} \frac{n}{n+1} = 1$. For each number $\epsilon > 0$ find an integer N such that

$$\left|\frac{n}{n+1} - 1\right| < \epsilon$$

for all $n \geq N$.

E2.5 Prove, using the definition of convergence, that

$$\lim_{n\to\infty} \frac{n-1}{n+1} = 1.$$

E2.6 Define a sequence (a_n) by the two conditions $a_1 = \sqrt{2}$ and $a_{n+1} = \sqrt{2+a_n}$ for $n \geq 1$. Show that this sequence converges and find its limit.
Hint: *Show that the sequence is monotone and bounded.*

E2.7 Assuming that $\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = \mathrm{e}$, compute:

(a) $\lim_{n\to\infty} \left(1+\frac{1}{3n}\right)^n$;

(b) $\lim_{n\to\infty} \left(1-\frac{1}{n^2}\right)^{n^2}$.

E2.8 Show that the sequence $(n^n/(n!\mathrm{e}^n))$ has a limit.

E2.9 Recall that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Let $a_n = \binom{2n}{n}4^{-n}$, for all $n \geq 1$.

(a) Prove that the sequence (a_n) converges to a limit.

(b) Find $\lim_{n\to\infty} a_n$.

E2.10 Show that, if
$$a_n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!},$$
then $\lim_{n\to\infty} a_n$ exists by showing that the sequence is Cauchy.

E2.11 Suppose that $0 \le b < 1$ and let $a_n = (1+b)(1+b^2)\cdots(1+b^n)$ for $n \ge 1$. Show that $\lim_{n\to\infty} a_n$ exists.
Hint: *Show that* $1 + b < \mathrm{e}^b$.

E2.12 Let (a_n) and (b_n) be two nondecreasing sequences with the property that, for each positive integer n, there are integers p and q such that $a_n \le b_p$ and $b_n \le a_q$. Show that (a_n) and (b_n) either both converge or both diverge to ∞ and that, moreover, if they both converge they have the same limit.

E2.13 Prove the uniqueness of limits: If (x_n) is a sequence in $\mathbb{R}$ such that $\lim_{n\to\infty} x_n = a$ and $\lim_{n\to\infty} x_n = b$ then $a = b$. Include the cases where the sequence may diverge to $\pm\infty$.

E2.14 Suppose that S is a nonempty subset of $\mathbb{R}$ that is bounded above and put $s = \sup S$. Show that there is a sequence (x_n) such that $x_n \in S$ for all n and $\lim_{n\to\infty} x_n = s$.

E2.15 Show that, if a sequence (x_n) satisfies $|x_n - x_{n+1}| < 3^{-n}$ for all n, then it is convergent.

E2.16 Let (x_n) be a Cauchy sequence and suppose that, for every $\epsilon > 0$, there is an integer n satisfying $n > 1/\epsilon$ and $|x_n| < \epsilon$. Prove that $\lim_{n\to\infty} x_n = 0$.

E2.17 Show that any sequence (a_n) satisfying $|a_n - a_{n+1}| \le |a_{n-1} - a_n|/2$ for all $n \ge 2$ is convergent.

E2.18 Suppose that (a_n) and (b_n) are two sequences and assume that $|a_m - a_n| \le |b_m - b_n|$ for all integers m and n. Show that, if the sequence (b_n) is convergent, then so is the sequence (a_n).

E2.19 Suppose that (a_n) is a sequence of numbers and that a is a number.

(a) Show by means of examples that, if $\lim_{n\to\infty}|a_n| = |a|$, then the limit $\lim_{n\to\infty} a_n$ need not exist and that, if it does exist, it need not equal $|a|$.

(b) Prove that $\lim_{n\to\infty}|a_n| = 0$ if and only if $\lim_{n\to\infty} a_n = 0$.

E2.20 Suppose that (a_n) and (b_n) are two sequences of numbers.

(a) Prove that, if $\lim_{n\to\infty} a_n = 0$ and if (b_n) is bounded, then $\lim_{n\to\infty} a_n b_n = 0$.

(b) Show by example that, in part (a), it is necessary that the sequence (b_n) be bounded.

(c) Prove that, if $\lim_{n\to\infty} a_n = A$ and if $\lim_{n\to\infty} b_n = B$, then $\lim_{n\to\infty} a_n b_n = AB$.
Hint: *Subtract and add the same quantity to* $a_n b_n - AB$ *and use (a).*

E2.21 Let (a_n) and (b_n) be convergent sequences of real numbers such that $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$, for some $A, B \in \mathbb{R}$. Show (using the ϵ-N definition of a limit) that

(a) $\lim_{n\to\infty} a_n + b_n = A + B$

(b) $\lim_{n\to\infty} a_n b_n = AB$

Hint: *For part (b), see Exercise E2.20.*

E2.22 Show that, if (a_n) and (b_n) are two sequences such that $a_n \leq b_n$ for every $n \in \mathbb{Z}_{>0}$ and $\lim_{n\to\infty} a_n = \infty$, then $\lim_{n\to\infty} b_n = \infty$.

E2.23 Show that, if (a_n) is a sequence such that $\lim_{n\to\infty} a_n = \infty$, then $\lim_{n\to\infty} \frac{1}{a_n} = 0$.

E2.24 Suppose that $\lim_{n\to\infty} a_n = a$, with $a \neq 0$. Prove that $\lim_{n\to\infty} 1/a_n = 1/a$.

E2.25 (a) Show that $\sqrt{xy} \leq (x+y)/2$ for any two nonnegative numbers x and y. (The left and right sides of this inequality are the geometric and arithmetic means of x and y, respectively.)

(b) Let x_1 and y_1 be two numbers satisfying $0 < x_1 < y_1$ and put $x_{n+1} = \sqrt{x_n y_n}$ and $y_{n+1} = (x_n + y_n)/2$ for $n \geq 1$. Prove that the sequences (x_n) and (y_n) are convergent and that they have the same limit.

E2.26 Let (a_n) be a sequence and define a sequence (b_n) by putting

$$b_n = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

for $n \geq 1$.

(a) Show that, if $\lim_{n\to\infty} a_n = 0$, then $\lim_{n\to\infty} b_n = 0$.

(b) Show that, if $\lim_{n\to\infty} a_n = L$, then $\lim_{n\to\infty} b_n = L$, where L is a real number. (This result is important in the study of probability theory and Fourier series.)

(c) Give an example for which (a_n) diverges but (b_n) converges.

E2.27 (a) Use induction to prove Bernoulli's inequality: $(1+h)^n > 1 + nh$ or all integers $n > 1$ and all numbers $h > 0$.

(b) Use this inequality to prove that if $r > 1$ then $\lim_{n\to\infty} r^n = \infty$.

(c) Prove that if $|r| < 1$ then $\lim_{n\to\infty} r^n = 0$.

E2.28 In this exercise, we will prove the following statement, known as ***Fekete's Lemma***:

> *Let $a_1, a_2, a_3, \ldots$ be a sequence of nonnegative real numbers with the property that $a_{m+n} \leq a_m + a_n$ for all m, n. Then, the sequence $\left(\frac{a_n}{n}\right)$ converges to $\inf\{\frac{a_n}{n} \mid n \geq 1\}$.*

So, let (a_n) be a sequence as in the hypothesis of the above statement. The set $\{\frac{a_n}{n} \mid n \geq 1\}$ is bounded below by 0, so its infimum exists. Set $L = \inf\{\frac{a_n}{n} \mid n \geq 1\}$.

(a) For positive integers k, m, prove that $a_{km} \leq m a_k$.

(b) Show that, for any $\epsilon > 0$, there exists N such that, for all $n \geq N$, $\frac{a_n}{n} \leq L + \epsilon$.

Hint: *Let $\epsilon > 0$ be arbitrary. Pick a k such that $\frac{a_k}{k} < L + \epsilon/2$ (why can this be done?). Now, any n can be expressed as $n = km + r$, with*

$0 \le r < k$ (why?). Part (a) shows that $a_n \le m a_k + a_r$. Deduce from this that there exists N such that, for all $n > N$, $\frac{a_n}{n} < \frac{a_k}{k} + \epsilon/2$.

(c) Deduce from (b) that $\lim_{n\to\infty} \frac{a_n}{n} = L$.

E2.29 Suppose that (a_{n_k}) is a subsequence of a monotone sequence (a_n). Is it true that, if the limit $\lim_{k\to\infty} a_{n_k}$ exists, then so does $\lim_{n\to\infty} a_n$ and that these limits are equal? Why?

E2.30 Suppose that f is a continuous function whose domain is the real line and that (a_n) is a convergent sequence. Show that $\lim_{n\to\infty} f(a_n) = f(\lim_{n\to\infty} a_n)$.

E2.31 A sequence (a_n) of real numbers is said to be ***contractive*** if there exists a constant β with $0 < \beta < 1$ such that

$$|a_{n+1} - a_n| \le \beta, \ |a_n - a_{n-1}|$$

for all $n \ge 1$. Show that any contractive sequence is a Cauchy sequence, and, therefore, is convergent.

Hint: *Obtain an estimate for $|a_n - a_m|$ in terms of β and $|a_1 - a_0|$.*

E2.32 A function $f\colon \mathbb{R} \to \mathbb{R}$ is said to be a ***contraction*** on $\mathbb{R}$ if there exists a constant β with $0 < \beta < 1$ such that

$$|f(x) - f(y)| \le \beta |x - y|$$

for all $x, y \in \mathbb{R}$. In this problem, we will prove the following ***contraction mapping theorem***:

A contraction on $\mathbb{R}$ is continuous and has a unique fixed point.

(A fixed point of a function $f\colon \mathbb{R} \to \mathbb{R}$ is a point $x \in \mathbb{R}$ such that $f(x) = x$.)

Let $f\colon \mathbb{R} \to \mathbb{R}$ be a contraction on $\mathbb{R}$.

(a) Show that f is continuous.

(b) Show that f can have at most one fixed point.
Hint: *Suppose there were two.*

(c) Let x_0 be any point in $\mathbb{R}$. Let $x_1 = f(x_0)$, $x_2 = f(x_1)$, and in general, for $n \ge 1$, $x_n = f(x_{n-1})$. Show that (x_n) is a convergent sequence.
Hint: *Exercise E2.31.*

(d) Show that $x = \lim_{n\to\infty} x_n$ is a fixed point of f.

E2.33 Determine the limit superior $\limsup_{n\to\infty} a_n$ and the limit inferior $\liminf_{n\to\infty} a_n$ if:

(a) $a_n = 3 + (-1)^n(1 + 1/n)$;

(b) $a_n = 1 + \sin(n\pi/2)$;

(c) $a_n = (2 - 1/n)(-1)^n$;

(d) $a_n = -n/4 + [n/4] + (-1)^n$, where $[\cdot]$ denotes the greatest integer function;

(e) $a_n = \cos\left(\frac{n}{2}\pi\right) + \frac{1}{n}\sin\left(\frac{2n+1}{2}\pi\right)$.

E2.34 Suppose that (a_n) is a sequence such that $L = \limsup_{n\to\infty} a_n$ is a real number. Then we know that, for any number $M > L$, there are only finitely many integers n for which $a_n > M$. Show by means of an example that it is possible to have $a_n > L$ for infinitely many n.

E2.35 Prove equation (2.2).

E2.36 Suppose that (a_n) and (b_n) are two bounded sequences.

(a) Show that, if $a_n \le b_n$ for all n then $\limsup_{n\to\infty} a_n \le \limsup_{n\to\infty} b_n$.

(b) Show that

$$\limsup_{n\to\infty} (a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

(c) Give an example for which equality holds in (a) and an example for which strict inequality holds.

E2.37 Can you find two *distinct* sequences (a_n) and (b_n) such that each is a subsequence of the other? If yes, give an example; if no, give a proof.

E2.38 Find the limit superior, the limit inferior, and the set of limits of convergent subsequences of the sequence $1, 2, 1, 4, 1, 6, 1, 8, \ldots$.

E2.39 For each of the sequences (a_n) below, identify the set of all limits of convergent subsequences. Also, determine $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$.

(a) $a_n = \dfrac{n!}{n^n}$, for $n \ge 1$

(b) $a_n = (-1)^n \left(1 + \dfrac{(-1)^n}{n}\right)$, for $n \ge 1$

(c) $1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, 1, 2, \ldots$ (The first term is 1, the next two terms are $1, 2$, the next three terms are $1, 2, 3$, and so on.)

(d) $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots$ (The first term is $\frac{1}{2}$, the next two terms are $\frac{1}{3}, \frac{2}{3}$, the next three terms are $\frac{1}{4}, \frac{2}{4}, \frac{3}{4}$, and so on.)

E2.40 For each of the sequences (a_n) below, identify the set of all limits of convergent subsequences. Also, determine $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$.

(a) $a_n = \left(2 - \dfrac{1}{n}\right)(-1)^n$, for all $n \ge 1$

(b) $a_n = 1 + \dfrac{1}{2} + \dfrac{1}{3} + \cdots + \dfrac{1}{n}$, for all $n \ge 1$

(c) $1, 1, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{3}, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$ (The first term is 1, the next two terms are $1, \frac{1}{2}$, the next three terms are $1, \frac{1}{2}, \frac{1}{3}$, and so on.)

E2.41 Let

$$a_n = \begin{cases} -n, & n \equiv 0 \mod 3, \\ 1 + 1/n, & n \equiv 1 \mod 3, \\ -1 - 1/n, & n \equiv 2 \mod 3 \end{cases}$$

for $n \in \mathbb{Z}_{>0}$.

(a) Show that (a_n) is bounded above, is not bounded below, and that the set of limits of convergent subsequences is $\{-1, 1\}$.

(b) Prove that $\limsup_{n\to\infty} a_n = 1$ and $\liminf_{n\to\infty} a_n = -\infty$.

E2.42 (a) Suppose that (a_n) and (b_n) are two bounded sequences. Show that $\limsup_{n\to\infty}(a_n + b_n) \le \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$.

(b) Give an example in which the inequality above is strict, and another in which the inequality holds with equality.

E2.43 Prove that, if a sequence diverges to ∞, then it has no convergent subsequences.

E2.44 Prove that, if L is a real number and that (a_n) is a sequence, then $\lim_{n\to\infty} a_n = L$ if and only if $\liminf_{n\to\infty} a_n = L$ and $\limsup_{n\to\infty} a_n = L$ (cf. Proposition 2.17).

Chapter 3

Sequences in $\mathbb{R}^d$

Sequences and their limits can be defined in $\mathbb{R}^d$ for $d \geq 2$ almost exactly as they were defined in $\mathbb{R}$. (It will be convenient to use d for the dimension of the Euclidean space because of the ubiquitous use of n as a subscript.) As is customary in linear algebra, if $\boldsymbol{x}$ is a vector in $\mathbb{R}^d$ then x_k denotes the kth-component of $\boldsymbol{x}$ for $k = 1, \ldots, d$, so that $\boldsymbol{x} = (x_1, \ldots, x_d)$.

The ***norm*** of a vector $\boldsymbol{x} \in \mathbb{R}^d$ is the number

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + \cdots + x_d^2}.$$

Notice that, if $d = 1$, then the norm is just the absolute value. The norm has the following properties:

1. $\|\boldsymbol{x}\| \geq 0$ for all $\boldsymbol{x} \in \mathbb{R}^d$ with equality if only if $\boldsymbol{x} = 0$;
2. $\|a\boldsymbol{x}\| = |a|\|\boldsymbol{x}\|$ for all $a \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^d$;
3. $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$.

The first two of these properties are obvious, and the third, called the ***triangle inequality***, is most easily verified by first proving the Cauchy–Schwartz inequality

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|,$$

which is valid for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$. Both this inequality and property 3 above should be familiar from linear algebra.

If $(\boldsymbol{x}_n)$ is a sequence in $\mathbb{R}^d$ the kth-component of $\boldsymbol{x}_n$ is denoted by $x_{n,k}$ and so $\boldsymbol{x}_n$ itself is $(x_{n,1}, \ldots, x_{n,d})$.

3.1 Definition: A sequence $(\boldsymbol{x}_n)$ in $\mathbb{R}^d$ is said to have a ***limit*** $\boldsymbol{L}$, or to ***converge to*** $\boldsymbol{L}$ if, for any $\epsilon > 0$, there exists an integer N such that

$$\|\boldsymbol{x}_n - \boldsymbol{L}\| < \epsilon$$

for every $n \geq N$, and in this case we will write $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{L}$. •

The following theorem says that we find limits "component-wise" and means that many of the results for sequences in $\mathbb{R}$ can be carried over effortlessly to $\mathbb{R}^d$.

3.2 Theorem: *If $(\boldsymbol{x}_n)$ is a sequence in $\mathbb{R}^d$ and if $\boldsymbol{L}$ is a point in $\mathbb{R}^d$, then $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{L}$ if and only if $\lim_{n\to\infty} x_{n,i} = L_i$ for $i \in \{1,\ldots,d\}$.*

Proof: The theorem follows almost immediately from the identity

$$\|\boldsymbol{x}_n - \boldsymbol{L}\| = \sqrt{(x_{n,1} - L_1)^2 + \cdots + (x_{n,d} - L_d)^2}.$$

Indeed, suppose that $\epsilon > 0$ is given. If $\lim_{n\to\infty} x_{n,i} = L_i$ for $i \in \{1,\ldots,d\}$, then there is an N such that $|x_{n,i} - L_i| < \epsilon/\sqrt{d}$ for all $n \geq N$ and for $i \in \{1,\ldots,d\}$. But then it follows that

$$\|\boldsymbol{x}_n - \boldsymbol{L}\| < \sqrt{\epsilon^2/d + \cdots + \epsilon^2/d} = \epsilon$$

for all $n \geq N$. Conversely, if $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{L}$, there is an N such that $\|\boldsymbol{x}_n - \boldsymbol{L}\| < \epsilon$ for all $n \geq N$, and then

$$|x_{n,i} - L_i| \leq \sqrt{(x_{n,1} - L_1)^2 + \cdots + (x_{n,d} - L_d)^2} < \epsilon$$

for all $i \in \{1,\ldots,d\}$ and all $n \geq N$. ■

3.3 Definition: A sequence $(\boldsymbol{x}_n)$ in $\mathbb{R}^d$ is said to be ***bounded*** if there exists a number $M > 0$ such that $\|\boldsymbol{x_n}\| \leq M$ for all $n \in \mathbb{Z}_{>0}$. •

Since every ball in $\mathbb{R}^d$ lies inside some cube and every cube lies inside some ball, boundedness can also be defined by requiring that, for each $i \in \{1,\ldots,d\}$, there exist numbers m_i and M_i such that $m_i \leq x_i \leq M_i$ for all $n \in \mathbb{Z}_{>0}$. It is left as an exercise to prove that the two definitions are equivalent (cf. Exercise E3.1).

If $d \geq 2$, there is no useful definition of "greater than" in $\mathbb{R}^d$. (What would it mean to say that $\boldsymbol{x} < \boldsymbol{y}$ in $\mathbb{R}^2$, let alone $\mathbb{R}^d$?) Hence, there can be no results for $\mathbb{R}^d$ analogous to those for $\mathbb{R}$ which were phrased in terms of the order properties of $\mathbb{R}$, such as Theorem 2.9. Nevertheless, it is possible to generalize many of the convergence theorems for $\mathbb{R}$ to $\mathbb{R}^d$ by working component-wise.

Subsequences of sequences in $\mathbb{R}^d$ can be defined just as they were for sequences in $\mathbb{R}$ and a Bolzano–Weierstrass theorem for sequences in $\mathbb{R}^d$ can be proven by repeated applications of the theorem for $\mathbb{R}$.

3.4 Theorem: (Bolzano–Weierstrass theorem) *Every bounded sequence in $\mathbb{R}^d$ has a convergent subsequence.*

Proof: Suppose that $(\boldsymbol{x}_n)$ is a bounded sequence in $\mathbb{R}^d$. Then it is clear that $(x_{n,i})_{n\in\mathbb{Z}_{>0}}$ is a bounded sequence in $\mathbb{R}$ for $i \in \{1,\ldots,d\}$. We cannot simply apply the Bolzano–Weierstrass theorem (Theorem 2.13) to each of the d components of the sequence $(\boldsymbol{x}_n)$ since the resulting d convergent subsequences may be indexed on completely different subsets of $\mathbb{Z}_{>0}$. Instead, we apply the Bolzano–Weierstrass theorem to the sequence $(x_{n,1})$ to find a convergent subsequence $(x_{n_k,1})$. Next, we apply the Bolzano–Weierstrass theorem to the sequence $(x_{n_k,2})$ to find a convergent subsequence of it and hence a convergent subsubsequence of the original sequence. This idea works but both the notation and the terminology is rapidly getting out of hand, and it is necessary to change notation and replace the original sequence $(\boldsymbol{x}_n)$ by the subsequence $(\boldsymbol{x}_{n_k})$. Now the sequence $(x_{n,1})$ converges, and we can apply the Bolzano–Weierstrass theorem to the sequence $(x_{n,2})$ to find a convergent subsequence, say, $(x_{n_k,2})$. The sequence $(x_{n_k,1})$ still converges by Lemma 2.12, and the sequence (x_{n_k}) is a subsequence of the original sequence. Let us now change notation once again, replacing the sequence $(\boldsymbol{x}_n)$ by the subsequence $(\boldsymbol{x}_{n_k})$; this gives a sequence which is a subsequence of the original sequence and with the property that both of the sequences $(x_{n,1})$ and $(x_{n,2})$ converge. We can now repeat the argument, applying the Bolzano–Weierstrass theorem to the sequence $(x_{n,3})$.

After carrying out this argument d times, we will get a sequence $(\boldsymbol{x}_n)$ which is a subsequence of the original sequence, and such that the sequence $(x_{n,i})$ converges for $i \in \{1, \ldots, d\}$, and hence, by Theorem 3.2, such that $\lim_{n\to\infty} \boldsymbol{x}_n$ exists. ■

The notion of Cauchy sequence in $\mathbb{R}^d$ can be made entirely analogously to the same notion in $\mathbb{R}$. Moreover, it is still true that a sequence in $\mathbb{R}^d$ converges if and only if it is a Cauchy sequence (cf. Exercise E3.3). The fact that this sort of equivalence is not true for infinite-dimensional $\mathbb{R}$-vector spaces with norms is a launching point for a great deal of interesting and useful analysis.

A ***contraction*** on $\mathbb{R}^d$ is a function f from $\mathbb{R}^d$ to $\mathbb{R}^d$ such that there is a number $K < 1$ with the property that

$$|f(\boldsymbol{u}) - f(\boldsymbol{v})| \leq K\|\boldsymbol{u} - \boldsymbol{v}\|$$

for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$. The important property of contractions is they have unique fixed points, i.e., if f is a contraction on $\mathbb{R}^d$, there is exactly one point $\boldsymbol{x} \in \mathbb{R}^d$ with the property that $f(\boldsymbol{x}) = \boldsymbol{x}$ (cf. the following theorem). This fact (albeit in a more general context) can be used to prove a number of important theorems, such as the differentiability of the inverse of a one-to-one differentiable function or the existence of solutions of ordinary differential equations.

3.5 Theorem: (Contraction mapping theorem) *A contraction on $\mathbb{R}^d$ is continuous and has a unique fixed point.*

Proof: The proof will be left as an exercise (cf. Exercise E3.9). ■

Exercises

E3.1 Prove the equivalence of the two possible definitions of boundedness in $\mathbb{R}^d$.

E3.2 Find the limits, if they exist, of these sequences in $\mathbb{R}^2$:

(a) $\left((-1)^n, \dfrac{1}{n}\right)$;

(b) $\left(1, \dfrac{1}{n}\right)$;

(c) $\left(\left(\dfrac{1}{n}\right)(\cos(n\pi)), \left(\dfrac{1}{n}\right)\left(\sin\left(n\pi + \dfrac{\pi}{2}\right)\right)\right)$;

(d) $\left(\dfrac{1}{n}, n^{-n}\right)$.

E3.3 (a) Give a definition of a Cauchy sequence in $\mathbb{R}^d$.

(b) Show that a sequence $(\boldsymbol{x}_n)$ in $\mathbb{R}^d$ is Cauchy if and only if $(x_{n,k})$ is a Cauchy sequence of real numbers for $k = 1, \ldots, d$.

(c) Prove that every convergent sequence in $\mathbb{R}^d$ is Cauchy.

(d) Prove that every Cauchy sequence in $\mathbb{R}^d$ converges.

E3.4 Suppose that $(\boldsymbol{x}_n)$ is a Cauchy sequence of points in $\mathbb{R}^d$.

(a) Prove that $(\boldsymbol{x}_n)$ is bounded.

(b) Prove that, if (a_n) is a sequence of real numbers such that $\lim_{n\to\infty} a_n = 0$ then $((1 + a_n)\boldsymbol{x}_n)$ is a Cauchy sequence in $\mathbb{R}^d$.

E3.5 (a) Prove that every subset of $\mathbb{R}$ that is not bounded above contains a sequence that diverges to ∞.

(b) Prove that every unbounded subset of $\mathbb{R}^d$ contains a sequence $(\boldsymbol{x}_n)$ with the property that $\lim_{n\to\infty}\|\boldsymbol{x}_n\| = \infty$.

E3.6 Let $S = \{(x, y) \in \mathbb{R}^2 \mid xy > 1\}$ and $B = \{\|(x, y)\| \mid (x, y) \in S\}$. Find $\inf B$.

E3.7 For $n \geq 1$, let $a_n = (-1)^n$, and let

$$b_n = \begin{cases} 1, & \text{if } n < 10^{100}, \\ 1/n & \text{otherwise.} \end{cases}$$

Let $(\boldsymbol{x}_n)$ be the sequence in $\mathbb{R}^2$ defined by $\boldsymbol{x}_n = (a_n, b_n)$. Determine the set of all limits of convergent subsequences of $(\boldsymbol{x}_n)$.

E3.8 Suppose that S is a subset of $\mathbb{R}^d$, that $\boldsymbol{y}$ is a point in S, and that f is a real-valued function with domain S. Show that the following two statements are equivalent:

(a) $\lim_{\boldsymbol{x}\to\boldsymbol{y}} f(\boldsymbol{x}) = f(\boldsymbol{y})$;

(b) $\lim_{n\to\infty} f(\boldsymbol{x}_n) = f(\boldsymbol{y})$ whenever $(\boldsymbol{x}_n)$ is a sequence in S such that $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{y}$.

E3.9 Suppose that f is a contraction on $\mathbb{R}^d$ and let K denote the associated constant (so that $0 < K < 1$). Prove Theorem 3.5 as follows. Let $\boldsymbol{x}_1$ be any point in $\mathbb{R}^d$ and let $\boldsymbol{x}_2 = f(\boldsymbol{x}_1), \boldsymbol{x}_3 = f(\boldsymbol{x}_2), \ldots$.

(a) Prove that f is continuous.

(b) Prove that f has at most one fixed point.
Hint: *Suppose there were two.*

(c) Obtain an estimate for $\|\boldsymbol{x}_{m+1} - \boldsymbol{x}_m\|$ in terms of K and $\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$ for any integer m.

(d) Obtain an estimate for $\|\boldsymbol{x}_n - \boldsymbol{x}_m\|$ in terms of K and $\|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$ for any two integers m, n.

(e) Show that $(\boldsymbol{x}_n)$ is a Cauchy sequence in $\mathbb{R}^d$.

(f) Show that the vector $\boldsymbol{x} = \lim_{n\to\infty} \boldsymbol{x}_n$ is a fixed point of f.

E3.10 For each number α put $f_\alpha(x) = \alpha x$ for all $x \in \mathbb{R}$.

(a) For which numbers α is the function f_α a contraction on $\mathbb{R}$?

(b) For those numbers α for which f_α is a contraction, what is the fixed point of f_α?

E3.11 For what intervals $[0, r]$, $r \leq 1$, is

$$\begin{aligned} f\colon [0, r] &\to [0, r] \\ x &\mapsto x^2 \end{aligned}$$

contraction?

Chapter 4

Some topology in $\mathbb{R}^d$

Topology is, roughly speaking, that part of mathematics that studies open and closed sets and continuous functions, and the purpose of this section is to give an introduction to these ideas and some related ones in the context of Euclidean space. In particular, we shall prove that a continuous function defined on a closed and bounded set has a bounded range and attains both a maximum and a minimum value. This result should be familiar from first-year calculus, where it is often stated but rarely proven.

4.1 Open and closed sets

The properties of the norm (from the beginning of Chapter 3) make it reasonable to regard $\|\boldsymbol{x}-\boldsymbol{y}\|$ as the distance between two points $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$. In fact, if $d=1$ this is just the usual distance and if $d\in\{2,3\}$ it is also the usual "physical" distance by Pythagoras's theorem.

For each point $\boldsymbol{x}\in\mathbb{R}^d$ and each $r>0$, put

$$\begin{aligned} B_r(\boldsymbol{x}) &= \{\boldsymbol{y}\in\mathbb{R}^d \mid \|\boldsymbol{y}-\boldsymbol{x}\|<r\},\\ \overline{B}_r(\boldsymbol{x}) &= \{\boldsymbol{y}\in\mathbb{R}^d \mid \|\boldsymbol{y}-\boldsymbol{x}\|\le r\},\\ \widehat{B}_r(\boldsymbol{x}) &= \{\boldsymbol{y}\in\mathbb{R}^d \mid 0<\|\boldsymbol{y}-\boldsymbol{x}\|<r\}.\end{aligned}$$

These are the ***open ball*** of radius r and center $\boldsymbol{x}$, the ***closed ball*** of radius r and center $\boldsymbol{x}$, and the ***punctured ball*** of radius r and center $\boldsymbol{x}$, respectively.

If $d=1$, then $B_r(\boldsymbol{x})$ and $\overline{B}_r(\boldsymbol{x})$ are just the familiar open and closed intervals $(x-r,x+r)$ and $[x-r,x+r]$, respectively, and $\widehat{B}_r(\boldsymbol{x})=(x-r,x)\cup(x,x+r)$.

If $d=2$, then $B_r(\boldsymbol{x})$ and $\overline{B}_r(\boldsymbol{x})$ are disks in $\mathbb{R}^2$ with center at $\boldsymbol{x}$ and radius r, with the former not including its rim and the latter including its rim. The set $\widehat{B}_r(\boldsymbol{x})$ is the "punctured" disk with center $\boldsymbol{x}$ and radius r that includes neither its center nor its rim.

If $d=3$, the three sets $B_r(\boldsymbol{x})$, $\overline{B}_r(\boldsymbol{x})$, and $\widehat{B}_r(\boldsymbol{x})$ are balls (in the familiar sense) with center $\boldsymbol{x}$ and radius r.

Balls will be essential to phrase many of the notions we will introduce to characterize subsets of $\mathbb{R}^d$. We begin this process now.

4.1 Definition: For any subset A of $\mathbb{R}^d$, let $A^\complement$ denote the ***complement*** of A in $\mathbb{R}^d$, i.e.,

$$A^\complement = \mathbb{R}^d \setminus A = \{\boldsymbol{x} \in \mathbb{R}^d \mid \boldsymbol{x} \notin A\}. \qquad \bullet$$

Notice that, for any two subsets A and B of $\mathbb{R}^d$, we have

$$A \subseteq B \text{ if and only if } B^\complement \subseteq A^\complement.$$

4.2 Definition: A subset $A \subseteq \mathbb{R}^d$ is said to be ***open*** if, for each point $\boldsymbol{x} \in A$, there is an $r > 0$ such that $B_r(\boldsymbol{x}) \subseteq A$, and to be ***closed*** if $A^\complement$ is open. •

Open and closed sets are important building blocks for the structure of $\mathbb{R}^d$. It is obvious that $\mathbb{R}^d$ is an open set and hence that the empty set is a closed set. The empty set is also an open subset of $\mathbb{R}^d$, and to see this it is best to argue by contradiction. Indeed, if a subset A of $\mathbb{R}^d$ is not open then, by definition, it must contain a point $\boldsymbol{x}$ with the property that $B_r(\boldsymbol{x}) \nsubseteq A$ for all numbers $r > 0$. But the empty set does not have this property since it does not contain any points. The fact that the empty set is open now implies that $\mathbb{R}^d$ is closed. So the empty set and $\mathbb{R}^d$ are both open and closed subsets of $\mathbb{R}^d$, and it is true (but not so easy to prove) that these are the only subsets of $\mathbb{R}^d$ that are both open and closed.

Points in open sets have a useful property.

4.3 Definition: A point $\boldsymbol{x} \in \mathbb{R}^d$ is said to be an ***interior point*** of a subset $A \subseteq \mathbb{R}^d$ if there is an $r > 0$ such that $B_r(\boldsymbol{x}) \subseteq A$. By $\mathrm{int}(A)$ we denote the set of interior points of A, which is the ***interior*** of A. •

Sometimes the notation A° is used for the interior.

4.4 Proposition: *A set A is open if and only if every point of A is an interior point of A. Equivalently, A is open if and only if* $\mathrm{int}(A) = A$.

Proof: This follows immediately from Definitions 4.2 and 4.3. ■

Informally speaking, we can say that a set is ***finite*** if it is empty, or contains 1 point, or contains 2 points, or ... and is ***infinite*** if it is not finite. More formally, a set is usually said to be finite if every one-to-one function from the set into itself is onto and to be infinite if there is a one-to-one function from the set into itself that is not onto. For example, the function that multiplies each integer by 2 is a one-to-one function from $\mathbb{Z}_{>0}$ into itself that is not onto, and so $\mathbb{Z}_{>0}$ is infinite. It is important to not confuse the adjectives "finite" and "bounded": Every finite set is bounded but a bounded set need not be finite.

Let us consider some examples of open and closed sets.

4.5 Examples: 1. Every nonempty finite subset of $\mathbb{R}^d$ is closed and not open. Indeed, suppose that A is a nonempty finite subset of $\mathbb{R}^d$. To show that A is closed it is necessary to show that $A^\complement$ is open, so consider a point $\boldsymbol{x} \in A^\complement$. Then the number

$$r = \min\{\|\boldsymbol{x} - \boldsymbol{y}\| \mid \boldsymbol{y} \in A\}$$

is positive since A is finite and $\boldsymbol{x} \notin A$, and it is clear that $B_r(\boldsymbol{x}) \subseteq A^{\complement}$. Finally, to see that A is not open, it is only necessary to observe that every nonempty open set is infinite since it contains a ball and since every ball is infinite.

2. It is evident that the graph of, say, $y = x$ or $= x^2$ is a closed subset of $\mathbb{R}^2$ and that the graph of, say, $z = 2x - y$ or $z = x^2 + y^2$ is a closed subset of $\mathbb{R}^3$. However, the graph of $y = \sin(1/x)$ is not a closed subset of $\mathbb{R}^2$ because any point of the form $(0, a)$ with $|a| \leq 1$ is not on the graph and is not an interior point of the complement of the graph.
3. For any $\boldsymbol{x} \in \mathbb{R}^d$ and any $r > 0$, the sets $B_r(\boldsymbol{x})$ and $\widehat{B}_r(\boldsymbol{x})$ are open but not closed subsets of $\mathbb{R}^d$, and the set $\overline{B}_r(\boldsymbol{x})$ is a closed but not an open subset of $\mathbb{R}^d$. To see that $B_r(\boldsymbol{x})$ is an open set, consider a point $\boldsymbol{y} \in B_r(\boldsymbol{x})$ and put $s = r - \|\boldsymbol{y} - \boldsymbol{x}\|$. Then $s > 0$ by the definition of $B_r(\boldsymbol{x})$ and we can show that $B_s(\boldsymbol{y}) \subseteq B_r(\boldsymbol{x})$ as follows. If $\boldsymbol{z} \in B_s(\boldsymbol{y})$, then

$$\|\boldsymbol{z} - \boldsymbol{x}\| \leq \|\boldsymbol{z} - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{x}\| < s + \|\boldsymbol{y} - \boldsymbol{x}\| = r,$$

and hence $\boldsymbol{z} \in B_r(\boldsymbol{x})$. This shows that $B_r(\boldsymbol{x})$ is open.

To see that $B_r(\boldsymbol{x})$ is not a closed set, it is necessary to show that its complement $B_r(\boldsymbol{x})^{\complement}$ is not open, and for this it is necessary to find a point that belongs to $B_r(\boldsymbol{x})^{\complement}$ but which is not an interior point of $B_r(\boldsymbol{x})^{\complement}$. Let $\boldsymbol{y}$ be a point satisfying $\|\boldsymbol{x} - \boldsymbol{y}\| = r$, so that $\boldsymbol{y} \in B_r(\boldsymbol{x})^{\complement}$ (in fact, $\boldsymbol{y}$ is on the "boundary" of $B_r(\boldsymbol{x})$, something we shall make sense of shortly). Intuitively, it seems clear that the ball $B_s(\boldsymbol{y})$ for any $s > 0$ contains points in $B_r(\boldsymbol{x})$, so that $B_s(\boldsymbol{y}) \nsubseteq B_r(\boldsymbol{x})^{\complement}$ as required. The following argument verifies this intuition. Let s be any number such that $0 < s < r$ and let $\boldsymbol{z} = \boldsymbol{y} + \frac{s}{2r}(\boldsymbol{x} - \boldsymbol{y})$ (so that $\boldsymbol{z}$ is a point on the line segment joining $\boldsymbol{y}$ and $\boldsymbol{x}$). Then $\boldsymbol{z} \in B_s(\boldsymbol{y})$ since

$$\|\boldsymbol{z} - \boldsymbol{y}\| = \frac{s}{2r}\|\boldsymbol{x} - \boldsymbol{y}\| = \frac{s}{2},$$

and $\boldsymbol{z} \notin B_r(\boldsymbol{x})^{\complement}$ since

$$\|\boldsymbol{z} - \boldsymbol{x}\| = \left(1 - \frac{s}{2r}\right)\|\boldsymbol{x} - \boldsymbol{y}\| = r - \frac{s}{2} < r.$$

This proves the two assertions about $B_r(\boldsymbol{x})$, and those dealing with $\overline{B}_r(\boldsymbol{x})$ and $\widehat{B}_r(\boldsymbol{x})$ will be left as exercises. •

It follows from Definition 4.2 that every nonempty open set is a union of open balls. (Indeed, let A be a nonempty open set and, for each point $\boldsymbol{x} \in A$, let $r_{\boldsymbol{x}}$ be any positive number such that $B_{r_{\boldsymbol{x}}}(\boldsymbol{x}) \subseteq A$. Then $A = \bigcup_{\boldsymbol{x} \in A} B_{r_{\boldsymbol{x}}}(\boldsymbol{x})$.) The converse of this—that every union of open balls is an open set—is a consequence of the next result.

4.6 Proposition: *(i) The union of any number of open subsets of $\mathbb{R}^d$ is open.*

(ii) The intersection of any finite number of open subsets of $\mathbb{R}^d$ is open.

(iii) The union of any finite number of closed subsets of $\mathbb{R}^d$ is closed.

(iv) The intersection of any number of closed subsets of $\mathbb{R}^d$ is closed.

Proof: (i) Suppose that a set A is the union of open sets and let $\boldsymbol{x}$ be a point in A. Then $\boldsymbol{x}$ must belong to one of the open sets, say B, whose union is A. But then $B_r(\boldsymbol{x}) \subseteq B$ for some $r > 0$ since B is open, and thus $B_r(\boldsymbol{x}) \subseteq A$ since $B \subseteq A$. This shows that A is open.

(ii) Suppose that $A = \bigcap_{i=1}^m B_i$, where each of the sets $B_1, \ldots, B_m$ is open, and let $\boldsymbol{x}$ be a point in A. Then, for each $i \in \{1, \ldots, m\}$, there is an $r_i > 0$ such that $B_{r_i}(\boldsymbol{x}) \subseteq B_i$. If $r = \min\{r_1, \ldots, r_m\}$, then $r > 0$ and

$$B_r(\boldsymbol{x}) \subseteq B_{r_i}(\boldsymbol{x}) \subseteq B_i$$

for each $i \in \{1, \ldots, m\}$, and hence

$$B_r(\boldsymbol{x}) \subseteq \bigcap_{i=1}^m B_i = A.$$

This shows that A is open.

Finally, parts (iii) and (iv) are consequences of parts (i) and (ii) and the following two identities (which are known as ***de Morgan's laws***):

$$\left(\bigcup_{i\in I} A_i\right)^{\complement} = \bigcap_{i\in I} (A_i)^{\complement},$$
$$\left(\bigcap_{i\in I} A_i\right)^{\complement} = \bigcup_{i\in I} (A_i)^{\complement}. \qquad \blacksquare$$

The following example will show that one cannot omit the adjective "finite" in parts (ii) and (iii) of the above proposition.

4.7 Example: Consider a point $\boldsymbol{x} \in \mathbb{R}^d$. For any $k \in \mathbb{Z}_{>0}$ it is clear that $\frac{k-1}{k} < 1$ and hence that $\overline{B}_{(k-1)/k}(\boldsymbol{x}) \subseteq B_1(\boldsymbol{x})$, and, therefore,

$$\bigcup_{k=1}^{\infty} \overline{B}_{(k-1)/k}(\boldsymbol{x}) \subseteq B_1(\boldsymbol{x}).$$

On the other hand, if $\boldsymbol{y} \in B_1(\boldsymbol{x})$, then there is a $k \in \mathbb{Z}_{>0}$ such that $\|\boldsymbol{x}-\boldsymbol{y}\| \leq (k-1)/k$. Thus $\boldsymbol{y} \in \overline{B}_{(k-1)/k}(\boldsymbol{x})$ and this gives

$$\bigcup_{k=1}^{\infty} \overline{B}_{(k-1)/k}(\boldsymbol{x}) = B_1(\boldsymbol{x}),$$

and so the union of an infinite number of closed sets may well be open and not closed.

A similar analysis will show that

$$\bigcap_{k=1}^{\infty} B_{(k+1)/k}(\boldsymbol{x}) = \overline{B}_1(\boldsymbol{x}),$$

and so the intersection of an infinite number of open sets may be closed and not open. •

4.2 Cluster points, limit points, and boundary points

It will be useful to use sequences in discussing closed sets and continuous functions, and we shall need the "set version" of the Bolzano–Weierstrass Theorem (the "sequence version" was Theorems 2.13 and 3.4).

4.8 Theorem: (Bolzano–Weierstrass theorem) *If A is a bounded infinite subset of $\mathbb{R}^d$, then there is a point $\boldsymbol{x} \in \mathbb{R}^d$ with the property that the set $A \cap B_r(\boldsymbol{x})$ is infinite for each $r > 0$.*

Proof: Since A is infinite, there is a sequence $(\boldsymbol{x}_k)$ in A with $\boldsymbol{x}_k \neq \boldsymbol{x}_j$ whenever $k \neq j$. Then $(\boldsymbol{x}_k)$ is a bounded sequence in $\mathbb{R}^d$ since A is a bounded set in $\mathbb{R}^d$ and so, by the Bolzano–Weierstrass theorem for sequences, $(\boldsymbol{x}_k)$ has a convergent subsequence $(\boldsymbol{y}_j) = (\boldsymbol{x}_{k_j})$. Denote the limit of the convergent subsequence by $\boldsymbol{x}$. Then, for any $r > 0$, there is some integer N such that $\boldsymbol{y}_j \in B_r(\boldsymbol{x})$ for all $j \geq N$. But $\boldsymbol{y}_j = \boldsymbol{x}_{k_j} \in A$, so we have $\boldsymbol{y}_j \in A \cap B_r(\boldsymbol{x})$ for $j \geq N$. Thus $A \cap B_r(\boldsymbol{x})$ is an infinite set for any $r > 0$, and so $\boldsymbol{x}$ is the required point. ■

The following definition will be useful in studying closed sets.

4.9 Definition: A point $\boldsymbol{x} \in \mathbb{R}^d$ is said to be a ***cluster point*** of a subset $A \subseteq \mathbb{R}^d$ if $A \cap \widehat{B}_r(\boldsymbol{x}) \neq \emptyset$ for all $r > 0$; let $\operatorname{der}(A)$ denote the set of all cluster points of A, which is called the ***derived set*** of A. •

Sometimes the notation A' is used for the derived set.

The examples which follow will make it clear that a cluster point of A may belong to A or to $A^{\complement}$, and that points in A may or may not be cluster points of A.

4.10 Examples: 1. A finite set does not have any cluster points. To see this, suppose that A is a finite subset of $\mathbb{R}^d$. It is obvious that, if A is empty, then it has no cluster points, so suppose that A is nonempty and that $\boldsymbol{x}$ is a point in $\mathbb{R}^d$. If

$$r = \min\{\|\boldsymbol{x} - \boldsymbol{y}\| \mid \boldsymbol{y} \in A \text{ and } \boldsymbol{y} \neq \boldsymbol{x}\},$$

then $r > 0$ since A is finite, and $A \cap B_r(\boldsymbol{x})$ is either $\emptyset$ or $\{\boldsymbol{x}\}$. In either case, $A \cap \widehat{B}_r(\boldsymbol{x}) = \emptyset$, and thus $\boldsymbol{x}$ is not a cluster point of A.

2. It is easy to see that, for each $\boldsymbol{x} \in \mathbb{R}^d$ and any $r > 0$, each of the sets $\operatorname{der}(B_r(\boldsymbol{x}))$, $\operatorname{der}(\overline{B}_r(\boldsymbol{x}))$, and $\operatorname{der}(\widehat{B}_r(\boldsymbol{x}))$ is equal to $\overline{B}_r(\boldsymbol{x})$. •

4.11 Definition: A ***sequence in*** a set A is a sequence, all of whose terms belong to the set A. •

4.12 Proposition: *If $\boldsymbol{x} \in \mathbb{R}^d$ and if A is a subset of $\mathbb{R}^d$, the following three conditions are equivalent:*

(i) $\boldsymbol{x}$ is a cluster point of A;

(ii) $A \cap B_r(\boldsymbol{x})$ is infinite for each $r > 0$.

(iii) there is a sequence $(\boldsymbol{x}_n)$ in A such that $\boldsymbol{x}_m \neq \boldsymbol{x}_n$ whenever $m \neq n$ and $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x}$.

Proof: (i) $\Longrightarrow$ (iii) Let us assume that (i) is true. Then there is certainly a point $\boldsymbol{x}_1 \in A \cap \widehat{B}_1(\boldsymbol{x})$. Now, if $r = \min\{2^{-1}, \|\boldsymbol{x} - \boldsymbol{x}_1\|\}$, then $r > 0$ and so $A \cap \widehat{B}_r(\boldsymbol{x}) \neq \emptyset$ by (i); let $\boldsymbol{x}_2$ be any point in $A \cap \widehat{B}_r(\boldsymbol{x})$. Now let $r = \min\{2^{-2}, \|\boldsymbol{x} - \boldsymbol{x}_2\|\}$ and repeat the argument: $r > 0$ and $A \cap \widehat{B}_r(\boldsymbol{x}) \neq \emptyset$ by (i), and let $\boldsymbol{x}_3$ be any point in $A \cap \widehat{B}_r(\boldsymbol{x})$. Continuing in this manner will lead to a sequence $(\boldsymbol{x}_n)$ such that $\boldsymbol{x}_m \neq \boldsymbol{x}_n$ whenever $m \neq n$ and $\|\boldsymbol{x} - \boldsymbol{x}_n\| < 2^{-n}$ for all $n \in \mathbb{Z}_{>0}$. So the sequence $(\boldsymbol{x}_n)$ has all of the required properties.

(iii) $\Longrightarrow$ (ii) Let $(\boldsymbol{x}_n)$ be a sequence with the properties stated in (iii). Then, given an $r > 0$, there is an N such that $\|\boldsymbol{x}_n - \boldsymbol{x}\| < r$ for all $n \geq N$. But then $\{\boldsymbol{x}_n \mid n \geq N\}$ is an infinite subset of $B_r(\boldsymbol{x}) \cap A$, and thus $B_r(\boldsymbol{x}) \cap A$ is infinite.

(ii) $\Longrightarrow$ (i) This is obvious. ■

We can combine Theorem 4.8 and the preceding proposition to obtain the following succinct form of the Bolzano–Weierstrass theorem: Every bounded infinite set has a cluster point.

Now we turn to an idea related to, but different from, cluster points.

4.13 Definition: A point $\boldsymbol{x} \in \mathbb{R}^d$ is said to be a ***limit point*** of a subset $A \subseteq \mathbb{R}^d$ if $A \cap B_r(\boldsymbol{x}) \neq \emptyset$ for all $r > 0$; let $\mathrm{cl}(A)$ denote the set of all limit points of A, which is called the ***closure*** of A.

Sometimes the notation $\overline{A}$ is used for the closure.

It is clear from the definition that $\mathrm{der}(A) \subseteq \mathrm{cl}(A)$. It is also clear from the definition that $A \subseteq \mathrm{cl}(A)$, a relation that does not hold for the derived set. Let us illustrate by examples the difference between the two sets.

4.14 Examples: 1. Every point in a finite set is a limit point, merely because every point in a set is a limit point. Thus, for finite sets, we have $\mathrm{der}(A) \subset \mathrm{cl}(A)$.

2. As with the derived set, the closures $\mathrm{cl}(B_r(\boldsymbol{x}))$, $\mathrm{cl}(\overline{B}_r(\boldsymbol{x}))$, and $\mathrm{cl}(\widehat{B}_r(\boldsymbol{x}))$ are all equal to $\overline{B}_r(\boldsymbol{x})$. •

Let us give an alternative characterisations of limit points.

4.15 Proposition: *If $\boldsymbol{x} \in \mathbb{R}^d$ and if A is a subset of $\mathbb{R}^d$, the following two conditions are equivalent:*

(i) $\boldsymbol{x}$ is a limit point of A;

(ii) there is a sequence $(\boldsymbol{x}_n)$ in A such that $\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x}$.

Proof: (i) $\implies$ (ii) Assume $\boldsymbol{x}$ is a limit point of A. Let $n \in \mathbb{Z}_{>0}$ and let $\boldsymbol{x}_n \in A \cap B_{1/n}(\boldsymbol{x})$, this being possible by definition of limit point. The sequence $(\boldsymbol{x}_n)$ is in A, and converges to $\boldsymbol{x}$.

(ii) $\implies$ (i) Let $(\boldsymbol{x}_n)$ be a sequence in A converging to $\boldsymbol{x}$. Let $r > 0$ and let $N \in \mathbb{Z}_{>0}$ be such that $\|\boldsymbol{x}_n - \boldsymbol{x}\| < r$ for $n \geq N$, this being possible by definition of convergence. Now note that $\boldsymbol{x}_n \in A \cap B_r(\boldsymbol{x})$. Thus $\boldsymbol{x}$ is a limit point of A. ■

To fully characterize the difference between cluster points and boundary points, the following notion is useful.

4.16 Definition: A point $\boldsymbol{x} \in \mathbb{R}^d$ is said to be an ***isolated point*** of a subset $A \subseteq \mathbb{R}^d$ if there exists $r > 0$ such that $A \cap B_r(\boldsymbol{x}) = \{\boldsymbol{x}\}$. •

We now have the following result.

4.17 Proposition: *For a subset $A \subseteq \mathbb{R}^d$, a limit point $\boldsymbol{x} \in \mathbb{R}^d$ is a cluster point for A if and only if it is not an isolated point of A.*

Proof: We leave this as Exercise E4.23. ■

Next we define boundary points.

4.18 Definition: A point $\boldsymbol{x} \in \mathbb{R}^d$ is said to be a ***boundary point*** of a subset A of $\mathbb{R}^d$ if, for each $r > 0$, the ball $B_r(\boldsymbol{x})$ contains at least one point in A and at least one point in $A^{\complement}$; let $\mathrm{bd}(A)$ denote the set of boundary points of the set A. •

Sometimes the notation ∂A is used for the boundary.

Note that $\mathrm{bd}(A) = \mathrm{bd}(A^{\complement})$ because of the symmetry of A and $A^{\complement}$ in the definition of boundary point. It is obvious that $\mathrm{bd}(\emptyset) = \mathrm{bd}(\mathbb{R}^d) = \emptyset$.

The following characterisation of boundary is useful.

4.19 Proposition: *For a subset $A \subseteq \mathbb{R}^d$ we have $\mathrm{cl}(A) = A \cup \mathrm{bd}(A)$.*

Proof: Let $\boldsymbol{x} \in \mathrm{cl}(A)$. If $\boldsymbol{x} \in A$, then clearly $\boldsymbol{x} \in A \cup \mathrm{bd}(A)$. If $\boldsymbol{x} \in A^{\complement}$, then $\boldsymbol{x} \in A^{\complement} \cap B_r(\boldsymbol{x})$ for every $r > 0$. Also, since $\boldsymbol{x} \in \mathrm{cl}(A)$, $A \cap B_r(\boldsymbol{x}) \neq \emptyset$ for every $r > 0$. Thus $\boldsymbol{x} \in \mathrm{bd}(A)$ and so $\boldsymbol{x} \in A \cup \mathrm{bd}(A)$.

Next let $\boldsymbol{x} \in A \cup \mathrm{bd}(A)$. If $\boldsymbol{x} \in A$, then $\boldsymbol{x} \in \mathrm{cl}(A)$. If $\boldsymbol{x} \in \mathrm{bd}(A)$, then $A \cap B_r(\boldsymbol{x}) \neq \emptyset$ for each $r > 0$, in which case $\boldsymbol{x} \in \mathrm{cl}(A)$. ■

4.3 Properties of closed sets

It is necessary study closed sets in order to better understand cluster points, boundary points, and continuous functions.

4.20 Theorem: *The following conditions on a subset A of $\mathbb{R}^d$ are equivalent:*

(i) *A is closed;*

(ii) $\operatorname{der}(A) \subseteq A$, *i.e., A contains all of its cluster points;*

(iii) $\operatorname{bd}(A) \subseteq A$, *i.e., A contains all of its boundary points;*

(iv) $\operatorname{cl}(A) \subseteq A$, *i.e., A contains all of its limit points;*

(v) *if $(\boldsymbol{x}_n)$ is any convergent sequence in A, then $\lim_{n\to\infty} \boldsymbol{x}_n \in A$.*

Proof: (i) $\Longrightarrow$ (ii) Suppose that A is closed and consider a point $\boldsymbol{x} \in A^{\complement}$. Since $A^{\complement}$ is open, there is an $r > 0$ such that $B_r(\boldsymbol{x}) \subseteq A^{\complement}$, and thus $\boldsymbol{x} \notin \operatorname{der}(A)$. This argument shows that $A^{\complement} \subseteq \operatorname{der}(A)^{\complement}$, and, therefore, $\operatorname{der}(A) \subseteq A$ (cf. the remark following Definition 4.1).

(ii) $\Longrightarrow$ (iii) Suppose that $\operatorname{der}(A) \subseteq A$ and consider a point $\boldsymbol{x} \in A^{\complement}$. Then $\boldsymbol{x} \in \operatorname{der}(A)^{\complement}$ and so there is an $r > 0$ such that $A \cap \widehat{B}_r(\boldsymbol{x}) = \emptyset$. Now, since $\boldsymbol{x} \notin A$, this actually means that $B_r(\boldsymbol{x}) \cap A = \emptyset$, and, therefore, $\boldsymbol{x} \notin \operatorname{bd}(A)$. This argument shows that $A^{\complement} \subseteq \operatorname{bd}(A)^{\complement}$, and, therefore, $\operatorname{bd}(A) \subseteq A$.

(iii) $\Longrightarrow$ (iv) If $\operatorname{bd}(A) \subseteq A$, then $\operatorname{cl}(A) \subseteq A$ by Proposition 4.19.

(iv) $\Longrightarrow$ (v) Suppose that A is contains all of its limit points and let $(\boldsymbol{x}_n)$ be a sequence in A that converges to $\boldsymbol{x} \in \mathbb{R}^d$. Thus $\boldsymbol{x}$ is a limit point of A by Proposition 4.15, and thus $\boldsymbol{x} \in \operatorname{cl}(A)$ and so $\boldsymbol{x} \in A$.

(v) $\Longrightarrow$ (i) We prove the contrapositive. Suppose that A is not closed or, equivalently, that $A^{\complement}$ is not open. Then there is an $\boldsymbol{x} \in A^{\complement}$ with the property that $B_r(\boldsymbol{x}) \nsubseteq A^{\complement}$ for every $r > 0$. This means that, for each $k \in \mathbb{Z}_{>0}$, there will be a point $\boldsymbol{x}_k \in B_{1/k}(\boldsymbol{x}) \cap A$. But then $(\boldsymbol{x}_k)$ is a sequence in A converging to $\boldsymbol{x}$ and $\boldsymbol{x} \notin A$, and so (v) does not hold. ■

The following theorem gives us, for the first time, something interesting about sets that are both closed and bounded. Perhaps not obviously, such sets are extremely important in analysis, and we shall revisit these below in Theorem 4.27.

4.21 Theorem: *The following two conditions are equivalent for a subset $A \subseteq \mathbb{R}^d$:*

(i) *A is closed and bounded;*

(ii) *every sequence in A has a subsequence that converges to a point in A.*

Proof: (i) $\Longrightarrow$ (ii) Suppose first that A is closed and bounded and let $(\boldsymbol{x}_n)$ be a sequence in A. Then $(\boldsymbol{x}_n)$ is a bounded sequence and so has a convergent subsequence by the Bolzano–Weierstrass theorem and the limit of this subsequence is in A by Theorem 4.20.

(ii) $\Longrightarrow$ (i) Now suppose that A satisfies (ii). If A were not bounded, there would be a sequence $(\boldsymbol{x}_k)$ in A such that $\|\boldsymbol{x}_k\| > k$ for each integer k, and such a sequence clearly cannot have a convergent subsequence. But this contradicts (ii) and so A must be bounded. On the other hand, if A were not closed, there would be a boundary point $\boldsymbol{b}$ of A that is not in A by Theorem 4.20. Then, for each positive integer k, the ball $B_{1/k}(\boldsymbol{b})$ would contain a point, say $\boldsymbol{x}_k$, of A and so there would be a sequence $(\boldsymbol{x}_k)$ in A that converges to $\boldsymbol{b}$. But every subsequence of this sequence also converges to $\boldsymbol{b}$ by Lemma 2.12. This too contradicts (ii) and so A must be closed. ■

Closed and bounded subsets of $\mathbb{R}^d$ are sufficiently interesting to warrant a special name.

4.22 Definition: A subset $K \subseteq \mathbb{R}^d$ is ***compact*** if it is closed and bounded. •

For those of you who grow up and do more advanced analysis, you will see that "compact" and "closed and bounded" do not agree in more general spaces; but they do agree in Euclidean space. We shall have more to say about this below, specifically regarding the characterization of Theorem 4.27 for compact sets.

4.4 Continuous functions

We want to prove a theorem about the ranges of continuous functions that says, among other things, that a continuous function whose domain is a closed and bounded interval attains a minimum and a maximum value. As a first step to doing this it will be useful to formally reconcile the two common ways of thinking about continuity of functions: in terms of sequences and in terms of ϵ and δ.

4.23 Theorem: *For any nonempty set A, any point $\boldsymbol{a} \in A$, and any real-valued function f with domain A, the following are equivalent:*

(i) for each point $\boldsymbol{a} \in A$ and each $\epsilon > 0$, there is a $\delta > 0$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{a})| < \epsilon$ for all $\boldsymbol{x} \in A \cap B_\delta(\boldsymbol{a})$;

(ii) if $(\boldsymbol{a_n})$ is a sequence in A that converges to $\boldsymbol{a}$, then $\lim_{n\to\infty} f(\boldsymbol{a_n}) = f(\boldsymbol{a})$.

Proof: (i) $\implies$ (ii) Suppose that (i) holds and let $(\boldsymbol{a_n})$ be a sequence in A that converges to $\boldsymbol{a}$ and let ϵ be a positive number. Then (by (i)) there is a $\delta > 0$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{a})| < \epsilon$ for all $\boldsymbol{x} \in A \cap B_\delta(\boldsymbol{a})$. But, since $\boldsymbol{a} = \lim_{n\to\infty} \boldsymbol{a}_n$, there will be an integer N such that $\|\boldsymbol{a} - \boldsymbol{a}_n\| < \delta$ for all $n \geq N$. So, for any $n \geq N$, we have $\boldsymbol{a_n} \in A \cap B_\delta(\boldsymbol{a})$ and hence $|f(\boldsymbol{a_n}) - f(\boldsymbol{a})| < \epsilon$. As ϵ was an arbitrary positive number, this shows that $\lim_{n\to\infty} f(\boldsymbol{a}_n) = f(\boldsymbol{a})$ and hence (ii) holds.

(ii) $\implies$ (i) We will prove the contrapositive, so suppose that (i) does not hold. Then there must be an $\epsilon > 0$ such that, for each $\delta > 0$, there is a point $\boldsymbol{x} \in A \cap B_\delta(\boldsymbol{a})$ satisfying $|f(\boldsymbol{x}) - f(\boldsymbol{a})| \geq \epsilon$. In particular, for each positive integer n, we can take $\delta = 1/n$ and obtain a point $\boldsymbol{a}_n \in A \cap B_{1/n}(\boldsymbol{a})$ satisfying $|f(\boldsymbol{a}_n) - f(\boldsymbol{a})| \geq \epsilon$. But then $\lim_{n\to\infty} \boldsymbol{a}_n = \boldsymbol{a}$, and so we see that (ii) does not hold. ■

Now consider a nonempty subset A of $\mathbb{R}^d$ and a real-valued function f whose domain is A. As is well-known from first-year calculus, if the two equivalent conditions of the previous theorem hold for a point $\boldsymbol{a} \in A$, then f is said to be ***continuous at a*** and, if these conditions hold for every point $\boldsymbol{a} \in A$, then f is said to be ***continuous***. This definition makes every function continuous at every isolated point in its domain.

This discussion of continuous functions has an obvious analogue for $\mathbb{R}^m$-valued functions, i.e., for functions whose values are points in $\mathbb{R}^d$. In particular, such a function is continuous if and only if each of its m real-valued components are continuous (cf. Exercise E4.33 and the discussion of sequences in $\mathbb{R}^d$ in Chapter 3).

We are now ready to prove one of the fundamental theorems about continuous functions, one that has to do with functions with a compact domain.

4.24 Theorem: *Suppose that A is a compact subset of $\mathbb{R}^d$ and that $\boldsymbol{f}$ is a continuous $\mathbb{R}^m$-valued function with domain A.*

(i) The image $\operatorname{image}(\boldsymbol{f}) = \{\boldsymbol{f}(\boldsymbol{x}) \mid \boldsymbol{x} \in A\}$ *of $\boldsymbol{f}$ is a compact subset of $\mathbb{R}^m$.*

(ii) If $m = 1$, then f attains a minimum value and a maximum value on A, i.e., there are points $\boldsymbol{u}$ and $\boldsymbol{v}$ in A such that

$$f(\boldsymbol{u}) \leq f(\boldsymbol{x}) \leq f(\boldsymbol{v})$$

for all $\boldsymbol{x} \in A$.

Proof: (i) To show that $\operatorname{image}(\boldsymbol{f})$ is compact it is only necessary to show that it satisfies condition (ii) of Theorem 4.21. Thus suppose that $(\boldsymbol{y}_n)$ is a sequence in $\operatorname{image}(\boldsymbol{f})$. Then the definition of $\operatorname{image}(\boldsymbol{f})$ implies that there must be a sequence $(\boldsymbol{x}_n)$ in A such that $\boldsymbol{f}(\boldsymbol{x}_n) = \boldsymbol{y}_n$ for all $n \in \mathbb{Z}_{>0}$. Theorem 4.21 now implies that there is a subsequence $(\boldsymbol{x}_{n_k})$ of $(\boldsymbol{x}_n)$ that converges to a point, say $\boldsymbol{x}$, in A. But then Theorem 4.23 implies that $\lim_{k\to\infty} \boldsymbol{y}_{n_k} = \lim_{k\to\infty} \boldsymbol{f}(\boldsymbol{x}_{n_k}) = \boldsymbol{f}(\boldsymbol{x})$, and so $(\boldsymbol{y}_n)$ does indeed have a subsequence that converges to a point in $\operatorname{image}(\boldsymbol{f})$.

(ii) Now suppose that $m = 1$. Since $\operatorname{image}(f)$ is compact by (i), both $\inf \operatorname{image}(f)$ and $\sup \operatorname{image}(f)$ belong to $\operatorname{image}(f)$. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are points in A such that $f(\boldsymbol{u}) = \inf \operatorname{image}(f)$ and $f(\boldsymbol{v}) = \sup \operatorname{image}(f)$, then it is clear that f attains a minimum value at $\boldsymbol{u}$ and a maximum value at $\boldsymbol{v}$. ■

The proof of Theorem 4.24 used the Bolzano–Weierstrass theorem and thus depended on the completeness of $\mathbb{R}$, i.e., the fact that all Cauchy sequences in $\mathbb{R}$ converge. The intermediate value theorem is another theorem that is familiar from first-year calculus and that depends on the completeness of $\mathbb{R}$ (cf. Exercise E4.29).

4.5 Compact sets and open coverings

In this section we provide some alternative characterizations of compact sets. In particular, we shall see that compactness can be described more abstractly in terms of coverings by open sets; it is this more abstract characterization that carries over to more general settings than compact subsets of Euclidean space, where “compact” is no longer equivalent to “closed and bounded.”

4.25 Definition: Suppose that A is a subset of $\mathbb{R}^d$. A set $\mathcal{U}$ of subsets of $\mathbb{R}^d$ is said

(i) to be an ***open covering*** of A if each set in $\mathcal{U}$ is open and if $A \subseteq \bigcup_{U\in\mathcal{U}} U$, and

(ii) to have a ***finite subcover*** if there are $U_1, \ldots, U_n$ in $\mathcal{U}$ such that $A \subseteq \bigcup_{i=1}^{n} U_i$. •

4.26 Example: The set $\{(1/n, 1) \mid n \in \mathbb{Z}_{>0}\}$ of open intervals is an open covering of the interval $(0, 1)$ that has no finite subcover. •

The point of introducing open coverings and finite subcovers is that they can be used to give an third condition equivalent to the two conditions in Theorem 4.21, and to give another proof of Theorem 4.24. Why is it desirable to do this since we already know these theorems are true? The reason is that, in more advanced analysis, it is sometimes necessary, when proving properties of closed and bounded sets, to think of them in terms of open coverings and finite subcovers. There are also many situations in analysis and topology where the three conditions are no longer equivalent and the correct (or at least the most useful) one is the one involving open coverings and finite subcovers.

4.27 Theorem: (Heine–Borel theorem) *For a subset $A \subseteq \mathbb{R}^d$, the following three conditions are equivalent:*

(i) A is closed and bounded;

(ii) every sequence in A has a subsequence converging to a point in A;

(iii) every open covering of A has a finite subcover.

Proof: We already know that (i) and (ii) are equivalent.

(i) $\Longrightarrow$ (iii) For this, the more difficult part of the proof, we first prove a couple of lemmata.

1 Lemma: *If $K_1 \subseteq \mathbb{R}^m$ is compact and if $K_2 \subseteq \mathbb{R}^n$ is compact then $K_1 \times K_2 \subseteq \mathbb{R}^{m+n}$ is compact.*

Proof: Let us denote points in $\mathbb{R}^{m+n}$ by $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^m \times \mathbb{R}^n$. For $\boldsymbol{x} \in \mathbb{R}^m$, denote

$$K_{2,\boldsymbol{x}} = \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{y} \in K_2\}.$$

Let $\mathscr{U}$ be an open cover of $K_1 \times K_2$. For $\boldsymbol{x} \in K_1$, denote

$$\mathscr{U}_{\boldsymbol{x}} = \{U \in \mathscr{U} \mid U \cap K_{2,\boldsymbol{x}} \neq \emptyset\}.$$

For $U \in \mathscr{U}_{\boldsymbol{x}}$, define

$$V_U = \{\boldsymbol{y} \in U \mid (\boldsymbol{x}, \boldsymbol{y}) \in K_{2,\boldsymbol{x}}\}.$$

We claim that V_U is open. Indeed, let $\boldsymbol{y} \in V_U$ so that $(\boldsymbol{x}, \boldsymbol{y}) \in U$. Since U is open, there exists $r > 0$ such that $B_r(\boldsymbol{x}, \boldsymbol{y}) \subseteq U$. Therefore, $B_r(\boldsymbol{y}) \subseteq V_U$, and so V_U is open as claimed. Therefore, $(V_U)_{U \in \mathscr{U}_{\boldsymbol{x}}}$ is an open cover of K_2. Thus there exists $U_{\boldsymbol{x},1}, \dots, U_{\boldsymbol{x},k_{\boldsymbol{x}}} \in \mathscr{U}_{\boldsymbol{x}}$ such that $K_2 \subseteq \cup_{j=1}^{k_{\boldsymbol{x}}} V_{U_{\boldsymbol{x},j}}$.

Now, for $U \in \mathscr{U}$, denote

$$W_U = \{\boldsymbol{x} \in \mathbb{R}^m \mid (\boldsymbol{x}, \boldsymbol{y}) \in U \text{ for some } \boldsymbol{y} \in \mathbb{R}^n\}.$$

We claim that W_U is open. To see this, let $\boldsymbol{x} \in W_U$ and let $\boldsymbol{y} \in \mathbb{R}^n$ be such that $(\boldsymbol{x}, \boldsymbol{y}) \in U$. Since U is open, there exists $r > 0$ such that $B_r(\boldsymbol{x}, \boldsymbol{y}) \subseteq U$. Therefore, $B_r(\boldsymbol{x}) \subseteq W_U$, giving W_U as open, as desired. Now define $W_{\boldsymbol{x}} = \cap_{j=1}^{k_{\boldsymbol{x}}} W_{U_{\boldsymbol{x},j}}$ and note that, by Proposition 4.6, it follows that $W_{\boldsymbol{x}}$ is open. Thus $(W_{\boldsymbol{x}})_{\boldsymbol{x}\in K_1}$ is an open cover for K_1. By compactness of K_1, there exists $\boldsymbol{x}_1,\dots,\boldsymbol{x}_m \in K_1$ such that $K_1 = \cup_{l=1}^m W_{\boldsymbol{x}_l}$. Therefore,

$$K_1 \times K_2 = \cup_{\boldsymbol{x}\in K_1} K_{2,\boldsymbol{x}} = \cup_{\boldsymbol{x}\in K_1} \cup_{j=1}^{k_{\boldsymbol{x}}} U_{\boldsymbol{x},j} = \cup_{l=1}^m \cup_{j=1}^{k_{\boldsymbol{x}_l}} U_{\boldsymbol{x}_l,j},$$

so giving a finite subcover of $K_1 \times K_2$. ▼

2 Lemma: *The set $[a,b] \subseteq \mathbb{R}$ satisfies condition (iii).*

Proof: Let $\mathscr{U}$ be an open cover for $[a,b]$ and let

$$S_{[a,b]} = \{x \in \mathbb{R} \mid \; x \le b \text{ and } [a,x] \text{ has a finite subcover in } \mathscr{U}\}.$$

Note that $S_{[a,b]} \neq \emptyset$ since $a \in S_{[a,b]}$. Let $c = \sup S_{[a,b]}$. We claim that $c = b$. Suppose that $c < b$. Since $c \in [a,b]$ there is some $U \in \mathscr{U}$ such that $c \in U$. As U is open, there is some $r > 0$ sufficiently small that $B_r(c) \subseteq U$. By definition of c, there exists some $x \in (c-r,c)$ for which $x \in S_{[a,b]}$. By definition of $S_{[a,b]}$, there is a finite collection of open sets $U_1,\dots,U_m$ from $\mathscr{U}$ which cover $[a,x]$. Therefore, the finite collection $U_1,\dots,U_m,U$ of open sets covers $[a,c+\epsilon)$. This then contradicts the fact that $c = \sup S_{[a,b]}$, so showing that $b = \sup S_{[a,b]}$. The result follows by definition of $S_{[a,b]}$. ▼

3 Lemma: *If $A \subseteq \mathbb{R}^d$ satisfies condition (iii) and if $B \subseteq A$ is closed, then B satisfies (iii).*

Proof: Let $\mathscr{U}$ be an open cover for B and define $V = \mathbb{R}^d \setminus B$. Since B is closed, $\mathscr{U} \cup (V)$ is an open cover for A. Since A satisfies (iii), there exists $U_1,\dots,U_k \in \mathscr{U}$ such that $A \subseteq \cup_{j=1}^k U_j \cup V$. Therefore, $B \subseteq \cup_{j=1}^k U_j$, giving a finite subcover of B. ▼

Now we proceed with the proof of this part of the theorem. Suppose that A is closed and bounded. Let $R > 0$ be sufficiently large that $A \subseteq [-R,R]\times\cdots\times[-R,R]$. By Lemma 2 it follows that $[-R,R]$ satisfies condition (iii). By induction, using Lemma 1, it follows that $[-R,R]\times\cdots\times[-R,R]$ is compact. By Lemma 3 it follows that K is compact.

(iii) $\implies$ (i) Suppose that A satisfies (iii). Consider the set $\{B_n(\mathbf{0}) \mid n \in \mathbb{Z}_{>0}\}$ of open balls in $\mathbb{R}^d$; this set is certainly an open covering of A (it is even an open covering of all of $\mathbb{R}^d$). Then (iii) implies that there are integers $n_1,\dots,n_k$ such that $A \subseteq \bigcup_{i=1}^k B_{n_i}(\mathbf{0})$. Now, if $m = \max\{n_1,\dots,n_k\}$, then $B_m(\mathbf{0}) = \bigcup_{i=1}^k B_{n_i}(\mathbf{0})$ and so $A \subseteq B_m(\mathbf{0})$ and, therefore, A is bounded.

To show that A is also closed, it is sufficient to show that its complement $A^{\complement}$ is open. Let $\boldsymbol{x} \in A^{\complement}$ and consider the set

$$\mathscr{U} = \{\overline{B}_r(\boldsymbol{x})^{\complement} \mid \; r > 0\}$$

of open sets. Note that

$$\{\boldsymbol{x}\} = \bigcap_{r>0} \overline{B}_r(\boldsymbol{x}) \quad \implies \quad \{\boldsymbol{x}\}^{\complement} = \bigcup_{r>0} \overline{B}_r(\boldsymbol{x})^{\complement},$$

using de Morgan's Laws. Then, if $\boldsymbol{a} \in A$ then $\boldsymbol{a} \neq \boldsymbol{x}$ and so

$$\boldsymbol{a} \in \{\boldsymbol{x}\}^{\complement} \quad \Longrightarrow \quad \boldsymbol{a} \in \bigcup_{r>0} \overline{B}_r(\boldsymbol{x})^{\complement},$$

and so $\mathcal{U}$ is an open covering of A. But then (iii) implies that there are positive numbers $r_1, \ldots, r_n$ such that $A \subseteq \bigcup_{j=1}^{n} \overline{B}_{r_j}(\boldsymbol{x})^{\complement}$. Now, if $r = \min\{r_1, \ldots, r_n\}$, then $r > 0$ and

$$\overline{B}_r(\boldsymbol{x}) \subseteq \bigcap_{j=1}^{n} \overline{B}_{r_j}(\boldsymbol{x}) \quad \Longrightarrow \quad \overline{B}_r(\boldsymbol{x})^{\complement} = \bigcup_{j=1}^{n} \overline{B}_{r_j}(\boldsymbol{x})^{\complement}$$

and, therefore, $A \subseteq \overline{B}_r(\boldsymbol{x})^{\complement}$. But this means that $\overline{B}_r(\boldsymbol{x}) \subseteq A^{\complement}$ and so $B_r(\boldsymbol{x}) \subseteq A^{\complement}$. Since $\boldsymbol{x}$ was an arbitrary point in $A^{\complement}$, it follows that $A^{\complement}$ is open. ■

The more difficult part of Theorem 4.27 is the implication (i) $\Longrightarrow$ (iii). One part of the proof we pull out as the following example, since this may help us to understand this implication.

4.28 Example: We shall show that any open covering of $[0, 1]$ has a finite subcover. Indeed, suppose that $\mathcal{U}$ is an open covering of $[0, 1]$ and consider the set

$$C = \{x \in (0, 1] \mid \text{ there is finite subcover of } [0, x]\}.$$

The first step is to show that C is not empty. There is a set U in $\mathcal{U}$ with $0 \in U$ and hence with $(-r, r) \subseteq U$ for some $r > 0$. But then $[0, r/2] \subseteq U$ and so there is finite subcover (consisting of just one set, in fact) of $[0, r/2]$ and thus $r/2 \in C$.

To prove that $\mathcal{U}$ contains a finite subcover of $[0, 1]$, it is enough to show that $1 \in C$ and we will do this by putting $s = \sup C$ and showing that $s = 1$ and that $s \in C$.

Since $s \in [0, 1]$ there will be a set $U \in \mathcal{U}$ with $s \in U$ and hence with $(s-\epsilon, s+\epsilon) \subseteq U$ for some $\epsilon > 0$. Now, by Theorem 1.12, the interval $(s - \epsilon, s]$ must contain a point in C, say x. Then there are a finite number of sets $U_1, \ldots, U_n$ in $\mathcal{U}$ such that $[0, x] \subseteq \bigcup_{i=1}^{n} U_i$. But then the sets $U, U_1, \ldots, U_n$ are a finite subcover of $[0, s]$ and so $s \in C$. Moreover, if $s < 1$ and if $r = \min\{1 - s, \epsilon/2\}$, then it is even true that $s + r \leq 1$ and these same $n + 1$ sets cover $[0, s + r]$, and hence $s + r \in C$. But this contradicts the choice of s and therefore $s = 1$. •

Let us give a couple of examples of the use of the finite subcover characterisation of compactness. For the first result, we shall make use of the following definition.

4.29 Definition: Let $A \subseteq \mathbb{R}^d$ and let f be a real-valued function on A. The function f is ***uniformly continuous*** if, for each $\epsilon > 0$, there exists $\delta > 0$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{y})| < \epsilon$ if $\boldsymbol{x}, \boldsymbol{y} \in A$ satisfy $\|\boldsymbol{x} - \boldsymbol{y}\| < \delta$. •

Note carefully the difference between "continuous" and "uniformly continuous." The next theorem shows that these notions align for compact domains.

4.30 Theorem: **(Heine–Cantor Theorem)** *If $K \subseteq \mathbb{R}^d$ is compact, then a real-valued function f on K is continuous if and only if it is uniformly continuous.*

Proof: It is clear that, if f is uniformly continuous, then it is continuous. To show that f is uniformly continuous when it is continuous and when its domain K is compact, let $\epsilon > 0$. Continuity of f implies that, for $\boldsymbol{x} \in K$, there exists $\delta_{\boldsymbol{x}} > 0$ such that $|f(\boldsymbol{y}) - f(\boldsymbol{x})| < \epsilon/2$ for $\boldsymbol{y} \in B_{\delta_{\boldsymbol{x}}}(\boldsymbol{x}) \cap K$. Note that the balls

$$\{B_{\delta_{\boldsymbol{x}}/2}(\boldsymbol{x}) \mid \boldsymbol{x} \in K\}$$

cover K. Compactness of K implies that there exist $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in K$ such that $K \subseteq \cup_{j=1}^k B_{\delta_{\boldsymbol{x}_j}/2}(\boldsymbol{x}_j)$. Let

$$\delta = \min\{\delta_{\boldsymbol{x}_1}/2, \dots, \delta_{\boldsymbol{x}_k}/2\}.$$

Now let $\boldsymbol{x}, \boldsymbol{y} \in K$ satisfy $\|\boldsymbol{x} - \boldsymbol{y}\| < \delta$. We must have $\boldsymbol{x} \in B_{\delta_{\boldsymbol{x}_j/2}}(\boldsymbol{x}_j)$ for some $j \in \{1, \dots, k\}$. This, in turn, gives

$$\|\boldsymbol{x}_j - \boldsymbol{y}\| \leq \|\boldsymbol{x}_j - \boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\| < \tfrac{1}{2}\delta_{\boldsymbol{x}_j} + \delta < \delta_{\boldsymbol{x}_j}.$$

Thus $\boldsymbol{x}, \boldsymbol{y} \in B_{\delta_{\boldsymbol{x}_j}}(\boldsymbol{x}_j)$. Then

$$|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq |f(\boldsymbol{x}) - f(\boldsymbol{x}_j)| + |f(\boldsymbol{x}_j) - f(\boldsymbol{y})| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

giving uniform continuity. ■

As another application of the open cover characterisation of a compact set, let us give an alternative proof of the fact, proved as Theorem 4.24(i) above, that the image of a compact set under a continuous function is compact.

4.31 Theorem: *If $A \subseteq \mathbb{R}^d$ is compact and if $\boldsymbol{f}$ is a continuous $\mathbb{R}^m$-valued function on A, then* $\operatorname{image}(\boldsymbol{f})$ *is compact.*

Proof: Let $\mathscr{U}$ be an open cover of $\operatorname{image}(\boldsymbol{f})$. Let $\boldsymbol{x} \in A$ and let $U_{\boldsymbol{x}} \in \mathscr{U}$ be such that $\boldsymbol{f}(\boldsymbol{x}) \in U_{\boldsymbol{x}}$. Openness of $U_{\boldsymbol{x}}$ implies that there exists $\epsilon_{\boldsymbol{x}} > 0$ such that $B_{\epsilon_{\boldsymbol{x}}}(\boldsymbol{f}(\boldsymbol{x})) \subseteq U_{\boldsymbol{x}}$. By continuity of $\boldsymbol{f}$, there exists $\delta_{\boldsymbol{x}} > 0$ such that, if $\boldsymbol{y} \in B_{\delta_{\boldsymbol{x}}}(\boldsymbol{x}) \cap A$, then $f(\boldsymbol{y}) \in B_{\epsilon_{\boldsymbol{x}}}(\boldsymbol{f}(\boldsymbol{x}))$. Note that

$$\{B_{\delta_{\boldsymbol{x}}}(\boldsymbol{x}) \mid \boldsymbol{x} \in A\}$$

is an open cover of A. Compactness of A means that there are $\boldsymbol{x}_1, \dots, \boldsymbol{x}_k \in A$ such that $A \subseteq \cup_{j=1}^k B_{\delta_{\boldsymbol{x}_j}}(\boldsymbol{x}_j)$. We claim that $\operatorname{image}(\boldsymbol{f}) \subseteq \cup_{j=1}^k U_{\boldsymbol{x}_j}$. Indeed, let $\boldsymbol{f}(\boldsymbol{x}) \in \operatorname{image}(\boldsymbol{f})$. Then $\boldsymbol{x} \in B_{\delta_{\boldsymbol{x}_j}}(\boldsymbol{x}_j)$ for some $j \in \{1, \dots, k\}$. But then

$$\boldsymbol{f}(\boldsymbol{x}) \in B_{\epsilon_{\boldsymbol{x}_j}}(\boldsymbol{x}_j) \subseteq U_{\boldsymbol{x}_j},$$

giving the desired conclusion. ■

Exercises

E4.1 Determine the interior points, the cluster points, the limit points, the boundary points, and the isolated points of each of the following sets (note that the sets in (a)–(c) are subsets of $\mathbb{R}^2$ whereas the set in (d) is a subset of $\mathbb{R}$):

(a) $\{(q,r) \mid q \text{ and } r \text{ are rational and } 0 \leq q, r < 1\}$

(b) $\{(x,y) \mid x \in \mathbb{R} \text{ and } -1 < y < 1\}$

(c) $\left\{ \left(\frac{1}{n}, \frac{1}{m}\right) \;\middle|\; n, m \in \mathbb{Z}_{>0} \right\}$

(d) $\displaystyle\bigcup_{n=1}^{\infty} \left(\frac{1}{n+1}, \frac{1}{n}\right)$

E4.2 Determine the interior points, the cluster points, the limit points, the isolated points, and the boundary points of each of the following sets. Also, classify each of the sets as open, closed, neither or both.

(a) $\left\{ \frac{1}{n} \;\middle|\; n \in \mathbb{Z}_{>0},\ n \geq 1 \right\}$

(b) $\{\boldsymbol{x} \in \mathbb{R}^d \mid \|\boldsymbol{x}\| = 1\}$

E4.3 Determine the interior points, the boundary points, the cluster points, the limit points, and the isolated points of each of the following subsets of $\mathbb{R}^2$. Also, classify each of the sets as open, closed, neither or both.

(a) $\{(m,n) \mid m, n \in \mathbb{Z}\}$

(b) $\{(x,y) \in \mathbb{R}^2 \mid 0 < x \leq 1\}$

E4.4 Determine the interior points, the cluster points, the limit points, the boundary points, and the isolated points for each of the following sets. Also, state whether the sets are open or closed or neither. (Note that the sets in (a)–(e) are subsets of $\mathbb{R}$.)

(a) $(1,2)$

(b) $[2,3]$

(c) $(1,4]$

(d) $\displaystyle\bigcap_{n=1}^{\infty} [-1, 1/n)$

(e) $(0,1) \cap \mathbb{Q}$

(f) $\mathbb{R}^d$

(g) A hyperplane in $\mathbb{R}^d$.

(h) $\{(x,y) \in \mathbb{R}^2 \mid 0 < x \leq 1\}$

(i) $\{\boldsymbol{x} \in \mathbb{R}^d \mid \|\boldsymbol{x}\| = 1\}$

E4.5 Let $S = \{(x,y) \in \mathbb{R}^2 \mid |x| \leq 1 \text{ and } |y| < 1\}$.

(a) Show that S is not closed by finding a sequence in S that converges to a point not in S.

(b) Show S is not closed by showing that $\mathbb{R}^2 \setminus S$ is not open.

(c) Determine the interior points of S.

E4.6 Consider the following subset S of $\mathbb{R}^2$:

$$S = \{(x, y) \in \mathbb{R}^2 \mid y > x^2\}.$$

Answer the following questions with complete justifications.

(a) Is S open?

(b) If possible, find a point in S that is not an interior point.

(c) Is S closed?

(d) If possible, find a point not in S that is the limit of a sequence of points in S.

(e) Is S bounded?

(f) Is S compact?

E4.7 Let (a_n) be a sequence of real numbers, and let T be the set of all limits of its convergent subsequences.

(a) Show that T contains all of its cluster points, i.e., $\operatorname{der}(T) \subseteq T$.

(b) Give an example of a sequence (a_n) for which $\operatorname{der}(T) = T$, and an example for which $\operatorname{der}(T) \neq T$.

E4.8 Give examples in $\mathbb{R}^d$ of the following or else explain why there are no such examples.

(a) A boundary point of a set S that is not a cluster point of S.

(b) A cluster point of a set S that is not a boundary point of S.

(c) An isolated point of a set S that is not a boundary point of S.

E4.9 Suppose that $A \subseteq \mathbb{R}$ is such that $A \neq \emptyset$ and $A^{\complement} \neq \emptyset$. The aim of this exercise is to show that $\operatorname{bd}(A) \neq \emptyset$.

Pick any $x \in A$ and $y \in A^{\complement}$. Without loss of generality, we may assume $x < y$. Define $P = \{p \in A \mid p < y\}$.

(a) Justify the fact that $\sup P$ exists.

Set $a = \sup P$, and define $Q = \{q \in A^{\complement} \mid q \geq a\}$.

(b) Prove that $a = \inf Q$.

(c) Deduce from part (b) that $a \in \operatorname{bd}(A)$ and, therefore, $\operatorname{bd}(A) \neq \emptyset$.

E4.10 Give examples of:

(a) An infinite set in $\mathbb{R}$ with no cluster points.

(b) A nonempty subset of $\mathbb{R}$ that is contained in its set of cluster points.

(c) A subset of $\mathbb{R}$ that has infinitely many cluster points but contains none of them.

(d) A set S such $\operatorname{bd}(S) = S \cup \operatorname{der}(S)$.

E4.11 Suppose that (A_n) is a sequence of nonempty subsets of $\mathbb{R}$, and assume that A_1 is bounded and that $A_1 \supseteq A_2 \supseteq A_3 \ldots,$, and let $a_n = \inf A_n$ for each n.

(a) Show that $a_1 \leq a_2 \leq a_3 \leq \ldots$.

(b) Show that, if each A_n is closed, then

$$\sup\{a_1, a_2, a_3, \ldots\} \in \bigcap_{n=1}^{\infty} A_n.$$

(c) Give an example to show that it is possible for $\cap_{n=1}^{\infty} A_n$ to be empty.

E4.12 Conditions (i)–(iv) in the statement of Theorem 4.20 are equivalent and in the proof of this theorem it was shown that (i) implies (ii), that (ii) implies (iii), that (iii) implies (iv), and that (iv) implies (v).

(a) Give a direct argument showing that (iv) implies (iii).

(b) Give a direct argument showing that (iii) implies (ii).

(c) Give a direct argument showing that (ii) implies (i).

E4.13 (a) Prove that, if $A \subseteq \mathbb{R}^d$ is both open and closed, then $\mathrm{bd}(A) = \emptyset$. (You may *not* use the as-yet-unproved fact that $\emptyset$ and $\mathbb{R}^d$ are the only subsets of $\mathbb{R}^d$ that are both open and closed.)

(b) Prove that, if $A \subseteq \mathbb{R}$ is both open and closed, then either $A = \emptyset$ or $A = \mathbb{R}$.
Hint: *Use Exercise E4.9.*

E4.14 Suppose that $\boldsymbol{x} \in \mathbb{R}^d$ and that S is a nonempty subset of $\mathbb{R}^d$. Show that, if $\boldsymbol{x}$ is a boundary point but not an isolated point of S, then $\boldsymbol{x}$ is a cluster point of S.

E4.15 Let A be a subset of $\mathbb{R}^d$. Show that $\mathrm{cl}(A) = A \cup \mathrm{der}(A)$.

E4.16 Suppose that $(\boldsymbol{a}_n)$ is a sequence in $\mathbb{R}^d$ and let $R = \{\boldsymbol{a}_n \mid n \in \mathbb{Z}_{>0}\}$ be its range, let C be the set of cluster points of R, and let L be the set of all limit points of R.

(a) Show that $C \subseteq L$.

(b) Give an example of a sequence in $\mathbb{R}$ such that $C \neq L$.

E4.17 Suppose that S is a subset of $\mathbb{R}^d$.

(a) Show that the set of cluster points of S is a closed subset of $\mathbb{R}^d$.

(b) Show that the set of limit points of S is a closed subset of $\mathbb{R}^d$.

(c) Show that the set of boundary points of S is a closed subset of $\mathbb{R}^d$.

(d) Show that the set of isolated points of S is a closed subset of $\mathbb{R}^d$.

E4.18 Suppose that A and B are two nonempty subsets of $\mathbb{R}^d$ and assume that A is open. Is the set

$$\{\boldsymbol{x} + \boldsymbol{y} \mid \boldsymbol{x} \in A \text{ and } \boldsymbol{y} \in B\}$$

open? Why?

E4.19 (a) Show that $\mathrm{der}(A \cup B) = \mathrm{der}(A) \cup \mathrm{der}(B)$ for any two nonempty subsets A and B of $\mathbb{R}^d$.

(b) Is it true that

$$\operatorname{der}\left(\bigcup_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n} \operatorname{der}(A_i)$$

whenever $A_1, \ldots, A_n$ are nonempty subsets of $\mathbb{R}^d$?

(c) Is it true that

$$\operatorname{der}\left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} \operatorname{der}(A_i)$$

for any sequence (A_i) of nonempty subsets of $\mathbb{R}^d$?

E4.20 (a) Show that, if $A \subseteq \mathbb{R}^d$, then $\operatorname{bd}(A \cup \operatorname{bd}(A)) \subseteq \operatorname{bd}(A)$.

(b) Show that, if $A \subseteq \mathbb{R}^d$, then $A \cup \operatorname{bd}(A)$ is a closed set.

(c) Give an example of a set $A \subseteq \mathbb{R}$ such that $\operatorname{bd}(A \cup \operatorname{bd}(A)) \neq \operatorname{bd}(A)$.

E4.21 Prove directly that the intersection of two closed sets is closed.

E4.22 If A is any subset of $\mathbb{R}^d$, then $\operatorname{der}(\operatorname{der}(A))$ is the set of all cluster points of $\operatorname{der}(A)$.

(a) Give an example of a subset A of $\mathbb{R}$ for which $\operatorname{der}(\operatorname{der}(A)) = \emptyset$ and $\operatorname{der}(A) \neq \emptyset$.

(b) Show that, if A is a nonempty subset of $\mathbb{R}^d$, then $\operatorname{der}(\operatorname{der}(A)) \subseteq \operatorname{der}(A)$.

(c) Give an example of a subset A of $\mathbb{R}$ for which $\operatorname{der}(\operatorname{der}(A)) = \operatorname{der}(A)$.

(d) Give an example of a subset A of $\mathbb{R}$ for which $\emptyset \neq \operatorname{der}(\operatorname{der}(A)) \neq \operatorname{der}(A)$.

E4.23 Prove Proposition 4.17.

E4.24 Suppose S is a compact subset of $\mathbb{R}^d$ and that $(\boldsymbol{x}_n)$ is a sequence in S. Show that every convergent subsequence of $(\boldsymbol{x}_n)$ converges to a point in S.

E4.25 Suppose that f is a real-valued function with domain $\mathbb{R}$ that is nondecreasing. Prove that the limit $\lim_{x \to 0+} f(x)$ exists, $\lim_{x \to 0+}$ meaning the limit as x approaches zero from the right.
Hint: *Let* $A = f(\{x \in \mathbb{R} \mid x > 0\})$ *and consider* $\inf A$.

E4.26 Suppose that S is a closed subset of $\mathbb{R}^d$ and that f is a continuous real-valued function with domain S. Show that, for every real number a, the set

$$\{\boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leq a\}$$

is a closed subset of $\mathbb{R}^d$.

E4.27 Let $\boldsymbol{f} \colon \mathbb{R}^d \to \mathbb{R}^m$ be a function. For any $B \subseteq \mathbb{R}^m$, the *preimage* under $\boldsymbol{f}$, of B is defined to be the set

$$\boldsymbol{f}^{-1}(B) \triangleq \{\boldsymbol{x} \in \mathbb{R}^d \mid \boldsymbol{f}(\boldsymbol{x}) \in B\}.$$

Show that $\boldsymbol{f}$ is a continuous function if and only if, for each open set $U \subseteq \mathbb{R}^m$, the preimage $\boldsymbol{f}^{-1}(U)$ is an open set in $\mathbb{R}^d$.

E4.28 Let S be a nonempty subset of $\mathbb{R}$.

(a) Give an example to show that, if S is not closed, then there is a continuous real-valued function with domain S that does not attain a maximum value.

(b) Give an example to show that, if S is not bounded, then there is a continuous real-valued function with domain S that does not attain a maximum value.

(c) Give an example of a continuous bounded function with domain $\mathbb{R}$ that attains neither a maximum or minimum value.

E4.29 (***Intermediate value theorem***) Suppose that f is a continuous real-valued function whose domain includes a closed interval $[a, b]$. Show that, if $f(a) < 0$ and $f(b) > 0$, then there is a point $c \in [a, b]$ such that $f(c) = 0$.
Hint: *Consider the set* $\{x \in [a, b] \mid f(x) < 0\}$.

E4.30 Show that every cubic polynomial has a root.
Hint: *Use Exercise E4.29.*

E4.31 Show that every continuous function from the unit interval to itself has a fixed point.
Hint: *Apply the intermediate value theorem to the function* $x \mapsto f(x) - x$.

E4.32 Suppose that f is a continuous function whose domain is $[1, 2]$ and that satisfies $0 \leq f(x) \leq 3$ for all $x \in [1, 2]$. Show that f has a fixed point, i.e., that there is a point $x \in [1, 2]$ satisfying $f(x) = x$.

E4.33 Let $\boldsymbol{f}$ be a function from $\mathbb{R}^m$ to $\mathbb{R}^n$ and recall that we can write $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x}))$ for $\boldsymbol{x} \in \mathbb{R}^m$, where $f_1, \ldots, f_n$ are real-valued functions on $\mathbb{R}^m$. Show that $\boldsymbol{f}$ is continuous if and only if f_i is continuous for $i \in \{1, \ldots, n\}$.

Chapter 5

Infinite series

In Chapter 2 we considered various attributes of sequences in $\mathbb{R}$. We saw in that section that many interesting examples of sequences arise as partial sums of infinite series. In this section we carefully consider infinite series, and their convergence.

5.1 Definitions and examples

The problem which will be considered in this section is that of assigning a meaning to an infinite sum of real numbers, i.e., to an expression of the form

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots$$

for a sequence (a_k). The natural way to do this would be to consider the limit (as n tends to ∞) of the sum of the first n terms, and this is, in fact, what will be done.

5.1 Definition: Consider a sequence (a_k) of real numbers defined for $k \geq 1$. The formal expression

$$\sum_{k=1}^{\infty} a_k \tag{5.1}$$

is called the ***infinite series*** corresponding to the sequence (a_k). For $n \in \mathbb{Z}_{>0}$ the sum

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

is called the n ***partial sum*** of (5.1) and the sequence (S_n) is called the ***sequence of partial sums*** corresponding to (5.1).

If the sequence (S_n) converges then the series (5.1) is said to ***converge*** and one says that the value of $\sum_{k=1}^{\infty} a_k$ is the limit $\lim_{n\to\infty} S_n$, and one writes $\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} S_n$. On the other hand, if the sequence (S_n) diverges, then the series (5.1) is said to ***diverge*** or to ***not exist*** and no meaning is assigned to the symbol $\sum_{k=1}^{\infty} a_k$. •

The point of this definition is that, for some sequences (a_k), the infinite series (5.1) is defined to be a real number, while for other sequences (5.1) is not assigned a meaning. Namely, if the limit $\lim_{n\to\infty} S_n$ exists, then (5.1) is defined to be the value of this limit, whereas if this limit does not exist then the expression (5.1) is not assigned a value. In the latter case, then, we cannot "talk" about and manipulate the "sum" as though it were a number—it is merely a "formal expression.[1]" Students sometimes ignore the importance of the sequence (S_n) of partial sums and thereby miss the essential feature of infinite series. Partial sums are finite sums, so they can be dealt with and manipulated according to all the familiar rules for addition; *the corresponding statement for infinite series is not true.* All statements about infinite series must be understood as abbreviations for statements about limits of partial sums.

Notice that, if two sequences are identical except for a finite number of terms, then the two corresponding infinite series will either both converge or both diverge. Indeed, for n sufficiently large the two nth partial sums will have a constant difference, and hence will either both converge or both diverge. A particularly important case of this arises when we "neglect the first few terms of a series:" If (a_k) is a sequence defined for $k \in \mathbb{Z}_{>0}$ and if $N \in \mathbb{Z}_{>0}$, then the two series $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} a_k$ will either both converge or both diverge. To see this, let S_n denote the nth partial sum $\sum_{k=1}^{n} a_k$, and let T_n denote the nth partial sum $\sum_{k=N}^{n} a_k$ for $n \geq N$. Then $S_n = S_{N-1} + T_n$ for $n \geq N$ and, since S_{N-1} is fixed, $\lim_{n\to\infty} S_n$ exists if and only if $\lim_{n\to\infty} T_n$ exists.

It seems somewhat ironic that, while it is generally difficult to calculate the value of a convergent infinite series, it is often relatively easy to decide whether an infinite series converges or diverges. What this really means, when rephrased in terms of partial sums, is that, while it is generally difficult to calculate the limit of a convergent sequence, it is often relatively easy to decide whether a sequence converges or diverges. (Of course, if a series is known to converge one can, in principle, calculate its sum as accurately as desired by simply calculating its nth partial sum for large values of n. But this is not the same as finding the value of the infinite series.) Before going on to describe these so-called "convergence tests," it will be useful to give a few examples.

5.2 Examples: 1. It is easy to see that both of the infinite series

$$\sum_{k=1}^{\infty} 1 \qquad \text{and} \qquad \sum_{k=1}^{\infty} (-1)^k$$

diverge. Indeed, for the first series the nth partial sum is n and so the sequence of partial sums diverges to ∞. And for the second the sequence of partial sums is $-1, 0 - 1, 0, \ldots$ and hence is divergent.

2. Let r be a real number and consider the infinite series $\sum_{k=0}^{\infty} r^k$. This series is called a ***geometric series***; let (S_n) be its sequence of partial sums. It is obvious that, if $r = 1$, then $S_n = n + 1$ and hence the geometric series diverges in this case. Now suppose that $r \neq 1$. Then

$$(1 - r)(1 + r + \cdots + r^n) = 1 - r^{n+1},$$

[1]Something which can be made precise, but which will definitely *not* be made precise here.

and hence

$$S_n = \frac{1 - r^{n+1}}{1 - r}.$$

This makes it clear that, if $|r| < 1$, the geometric series converges with

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r},$$

and that, if $|r| \geq 1$, the geometric series diverges.

3. Consider the series

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)}.$$

Notice that the nth partial sum is

$$\begin{aligned}\sum_{k=1}^{n} \frac{1}{k(k+1)} &= \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1}\end{aligned}$$

and, therefore,

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

Here the partial sums are said to be "telescoping sums" because of the way they simplified.

4. The series

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

is called the ***harmonic series***. Recall from Example 2.8–2 that the 2^nth partial sum S_{2^n} exceeds $n/2$, and hence the harmonic series diverges to ∞.

We now begin considering theorems that will allow us to decide questions about convergence without obtaining explicit formulae for the partial sums. The first theorem is very simple.

5.3 Theorem: *Let (a_k) be a sequence of numbers.*

(i) If the infinite series $\sum_{k=1}^{\infty} a_k$ converges, then $\lim_{k\to\infty} a_k = 0$.

(ii) If the sequence (a_k) does not converge to 0, then the infinite series $\sum_{k=1}^{\infty} a_k$ diverges.

Proof: The two parts of this theorem are, of course, logically equivalent and so it is only necessary to prove (i). Let (S_n) be the sequence of partial sums of the series $\sum_{k=1}^{\infty} a_k$ and let ϵ be a given positive number. Then $\lim_{n\to\infty} S_n$ exists and, if $L = \lim_{n\to\infty} S_n$, there is an integer N such that $|L - S_n| < \frac{\epsilon}{2}$ whenever $n \geq N$. But then, for any integer $n \geq N$, we have

$$\begin{aligned}|a_{n+1}| &= |S_{n+1} - S_n| \\ &\leq |S_{n+1} - L| + |L - S_n| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

This shows that $\lim_{n\to\infty} a_n = 0$. ■

Notice that Example 5.2–4 shows that the converse of this theorem is false. It is very important to realize that this theorem gives a *necessary*, but not a *sufficient*, condition for convergence of a series. This theorem can sometimes be used to establish the divergence of a series but cannot be used to establish convergence.

One way to understand the preceding example and theorem is that, if a series $\sum_{k=1}^{\infty} a_k$ is to converge, then it is necessary that $a_k \to 0$ quite quickly as $k \to \infty$, and that the terms of the harmonic series simply approach zero too slowly.

5.2 Convergence of nonnegative series

It turns out that series with nonnegative terms are easier to discuss than those in which the terms may be positive or negative. To see why, consider a sequence (a_k) of nonnegative numbers and let (S_n) be the sequence of partial sums of the infinite series $\sum_{k=1}^{\infty} a_k$. Then $S_{n+1} = S_n + a_{n+1} \geq S_n$ for each n and, therefore, (S_n) is a nondecreasing sequence. So, by Theorem 2.9, the sequence (S_n) converges if and only if it is bounded. It will be useful to record this observation as a theorem.

5.4 Theorem: *If (a_k) is a sequence of nonnegative terms, then the infinite series $\sum_{k=1}^{\infty} a_k$ converges if and only if the corresponding sequence of partial sums is bounded.*

Next we give four "convergence tests" that are useful in deciding whether a given series with nonnegative terms is convergent or divergent.

5.5 Theorem: (Comparison test) *Suppose that (a_k) and (b_k) are two sequences.*

(i) If $0 \leq a_k \leq b_k$ for all k and if the series $\sum_{k=1}^{\infty} b_k$ converges, then so does the series $\sum_{k=1}^{\infty} a_k$.

(ii) If $0 \leq a_k \leq b_k$ for all k and if the series $\sum_{k=1}^{\infty} a_k$ diverges, then so does the series $\sum_{k=1}^{\infty} b_k$.

(iii) If $b_k > 0$ for all k, if $0 \leq \lim_{k\to\infty} a_k/b_k < \infty$, and if $\sum_{k=1}^{\infty} b_k$ converges, then so does $\sum_{k=1}^{\infty} a_k$.

(iv) If $b_k > 0$ for all k, if $0 < \lim_{k\to\infty} a_k/b_k \leq \infty$, and if $\sum_{k=1}^{\infty} b_k$ diverges, then so does $\sum_{k=1}^{\infty} a_k$.

Proof: (i) Let $S_n = \sum_{k=1}^n a_k$ and $T_n = \sum_{k=1}^n b_k$ for $n \in \mathbb{Z}_{>0}$. Then the assumption that $0 \le a_k \le b_k$ for all k implies that $0 \le S_n \le T_n$ for all n, and the assumption that $\sum_{k=1}^\infty b_k$ converges implies that (T_n) is a bounded sequence. But then (S_n) too is a bounded sequence, and so $\sum_{k=1}^\infty a_k$ converges by Theorem 5.4.

(ii) This assertion is logically equivalent to (i).

(iii) The hypotheses implies that, if $L = \lim_{k\to\infty} a_k/b_k$, then there exists an integer N such that $a_k/b_k \le L+1$ for all $k \ge N$, and hence such that $a_k \le (L+1)b_k$ for all $k \ge N$. The desired conclusion now follows from (i), together with our earlier observation that we may "neglect the first few terms" in showing convergence or divergence (cf. page 70).

(iv) In this case, let $0 < L = \lim_{k\to\infty} a_k/b_k$. Let $\epsilon > 0$ be such that $L - \epsilon > 0$. Then there exists $N \in \mathbb{Z}_{>0}$ such that $L - \epsilon < a_k/b_k$ for all $k \ge N$. Thus, if (S_n) are the partial sums for the sum $\sum_{k=1}^\infty a_n$, $a_k > (L-\epsilon)b_k$ for $k \ge N$ and so, for $n > N$,

$$S_n - S_N = \sum_{k=N+1}^n a_n > (L-\epsilon)\sum_{k=N+1}^n b_n.$$

Therefore

$$\lim_{n\to\infty} S_n > \lim_{n\to\infty}\Big((L-\epsilon)\sum_{k=N+1}^n b_n + S_N\Big) = \infty,$$

as desired. ■

In applying parts (i) and (ii) of this theorem, it is sufficient that there be an n_0 such that the inequalities $0 \le a_k \le b_k$ hold for all $k \ge N$ (and not for all k as in the statement of the theorem). This follows from our earlier remark (cf. page 70) about "neglecting the first few terms" in proving convergence or divergence.

5.6 Examples: 1. Since $k^{1/2} \le k$ for $k \ge 1$, it follows that $1/k^{1/2} \ge 1/k$ for $k \ge 1$. The comparison test and Example 5.2–4 then imply that the series $\sum_{k=1}^\infty \frac{1}{k^{1/2}}$ diverges.

2. Since
$$\frac{1}{2n+1} > \frac{1}{2(n+1)}$$
for $n \in \mathbb{Z}_{>0}$, the comparison test and Example 5.2–4 imply that the series $\sum_{n=1}^\infty \frac{1}{2n+1}$ diverges. Notice that this series consists of odd terms of the harmonic sequence.

3. Since
$$\frac{1}{n^2} \le \frac{2}{n(n+1)}$$
for $n \in \mathbb{Z}_{>0}$, the comparison test and Example 5.2–3 imply that the series $\sum_{n=1}^\infty \frac{1}{n^2}$ converges.

4. Does the series $\sum_{n=1}^{\infty} \frac{3}{2n^2 - n^{1/2}}$ converge? Since the terms are positive and

$$\lim_{n\to\infty} \left(\frac{3}{2n^2 - n^{1/2}} \Big/ \frac{1}{n^2} \right) = \frac{3}{2},$$

the series converges by 3 and the comparison test. Note that it would be more awkward to use part (i) of the comparison test for this particular example. •

5.7 Theorem: (Integral test for positive series) *Suppose that* (a_k) *is a sequence and that* f *is a function which is continuous,*[2] *nonnegative-valued, and nonincreasing on the half-line* $[1, \infty)$*, and satisfies* $a_k = f(k)$ *for all* $k \in \mathbb{Z}_{>0}$*. Then the infinite series* $\sum_{k=1}^{\infty} a_k$ *and the improper integral*

$$\int_1^{\infty} f(x)\,\mathrm{d}x$$

either both converge or both diverge.

Proof: Since f is nonincreasing and nonnegative-valued, it is clear from thinking about its graph that

$$a_{k+1} = f(k+1) \le \int_k^{k+1} f(x)\,\mathrm{d}x \le f(k) = a_k \tag{5.2}$$

for all $k \in \mathbb{Z}_{>0}$ and, therefore, that

$$\sum_{k=1}^{n} a_{k+1} \le \sum_{k=1}^{n} \int_k^{k+1} f(x)\,\mathrm{d}x \le \sum_{k=1}^{n} a_k$$

or

$$\sum_{k=2}^{n+1} a_k \le \int_1^{n+1} f(x)\,\mathrm{d}x \le \sum_{k=1}^{n} a_k \tag{5.3}$$

for each positive integer n.

Suppose first that the series $\sum_{k=1}^{\infty} a_k$ converges. Then there is a number, say M, such that $\sum_{k=1}^{n} a_k \le M$ for all $n \ge 1$ (by Theorem 5.4) and thus

$$\int_1^{n+1} f(x)\,\mathrm{d}x \le M$$

for all $n \ge 1$ by (5.3). But

$$n \mapsto \int_1^{n} f(x)\,\mathrm{d}x$$

is a nondecreasing function of n since f is nonnegative valued and, therefore, the limit $\lim_{n\to\infty} \int_1^n f(x)\,\mathrm{d}x$ exists by Theorem 2.9.

[2]The assumption that f is continuous is made merely to ensure that it is integrable over any interval of finite length.

Now suppose that, conversely, the improper integral $\int_1^\infty f(x)\,\mathrm{d}x$ converges. Then

$$\sum_{k=2}^{n+1} a_k \le \int_1^{n+1} f(x)\,\mathrm{d}x \le \int_1^\infty f(x)\,\mathrm{d}x < \infty$$

for all $n \in \mathbb{Z}_{>0}$ by (5.3), and so the series $\sum_{k=1}^\infty a_k$ converges by 2.9. ■

Under the hypothesis of the integral test, it is sufficient to assume that there is a positive integer N such that $a_k = f(k)$ for all $k \ge N$, and f is nonincreasing, nonnegative-valued, and continuous on $[N, \infty)$. This remark is similar to one made following the proof of the comparison test and follows from our earlier remark (cf. page 70) about "neglecting the first few terms" in proving convergence or divergence.

Let us consider some applications of the integral test.

5.8 Example: The two infinite series

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

both converge for $p > 1$ and diverge for $p \le 1$. To see this, it is enough to apply the integral test with the functions $f(x) = \frac{1}{x^p}$ and $f(x) = \frac{1}{x(\log x)^p}$. In carrying out this integration is is necessary to distinguish between $p = 1$ and $p \ne 1$. •

The infinite series $\sum_{k=1}^\infty \frac{1}{k^p}$ is called the ***p-series***. Notice that, if $p = 1$, then the p-series is just the harmonic series.

Let us turn to another convergence test.

5.9 Theorem: (Ratio test for positive series) *Suppose that (a_k) is a sequence of positive terms and assume that $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = L$ exists.*

(i) If $L < 1$, then $\sum_{k=1}^\infty a_k$ converges.

(ii) If $L > 1$, then $\sum_{k=1}^\infty a_k$ diverges.

(iii) If $L = 1$, then this test is inconclusive.

Proof: (i) Suppose that $L < 1$ and let r be any number satisfying $L < r < 1$. Then there must be $N \in \mathbb{Z}_{>0}$ such that $0 < \frac{a_{k+1}}{a_k} < r$ for all $k \ge N$. Thus $a_{k+1} < ra_k$ for all $k \ge N$ and, therefore,

$$\begin{aligned}
a_{N+1} &\le a_N r \\
a_{N+2} &\le a_{N+1} r \le a_N r^2 \\
a_{N+3} &\le a_{N+2} r \le a_N r^3 \\
&\ \vdots \\
a_{N+k} &\le a_N r^k
\end{aligned}$$

for $k \geq 1$. Now $\sum_{k=1}^{\infty} r^k$ is a convergent geometric series and, therefore, $\sum_{k=1}^{\infty} a_k$ converges by the comparison test (Theorem 5.5(i)).

(ii) Suppose that $L > 1$. Then there exists an $N \in \mathbb{Z}_{>0}$ such that $\frac{a_{k+1}}{a_k} \geq 1$ for all $k \geq N$. This means that $a_{k+1} \geq a_k$ for all $k \geq N$ and, therefore,

$$\begin{aligned} a_{N+1} &\geq a_N \\ a_{N+2} &\geq a_{N+1} \geq a_N \\ a_{N+3} &\geq a_{N+2} \geq a_N \\ &\vdots \end{aligned}$$

This shows that $\lim_{k\to\infty} a_k \neq 0$ and therefore $\sum_{k=1}^{\infty} a_k$ diverges by Theorem 5.3.

(iii) If $a_k = k$ for all k or if $a_k = k^{-2}$ for all k, then $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$, yet the first of these series diverges and the second converges. ■

The following test sometimes works when the ratio test does not apply. Note that it uses the limit superior and so may be applicable even in cases where the terms are quite irregular and $\lim_{n\to\infty} \sqrt[n]{a_n}$ does not exist.

5.10 Theorem: (Root test for positive series) *Suppose that (a_n) is a sequence of positive terms and let $L = \limsup_{n\to\infty} \sqrt[n]{a_n}$.*

(i) If $L < 1$ then $\sum_{n=1}^{\infty} a_n$ converges.

(ii) If $L > 1$ then $\sum_{n=1}^{\infty} a_n$ diverges.

(iii) If $L = 1$ this test is inconclusive.

Proof: (i) Say $L < 1$ and let r be a number satisfying $L < r < 1$. Then (by Proposition 2.19) the inequality $\sqrt[n]{a_n} > r$ holds for only finitely many n and, therefore, there is an integer N such that $\sqrt[n]{a_n} \leq r$ for all $n \geq N$. Thus $a_n \leq r^n$ for all $n \geq N$ and, since $\sum_{n=1}^{\infty} r^n$ is a convergent geometric series, the comparison test implies that the series $\sum_{n=1}^{\infty} a_n$ converges.

(ii) Now suppose that $L > 1$. Then Proposition 2.19 implies that the inequality $\sqrt[n]{a_n} > 1$ holds for infinitely many n, and thus $a_n > 1$ for infinitely many n. But then it cannot be true that $\lim_{n\to\infty} a_n = 0$ and, therefore, the series $\sum_{n=1}^{\infty} a_n$ diverges by Theorem 5.3.

(iii) The same examples which were used in Theorem 5.9(iii) can be used here. ■

At this point some general comments on the four convergence tests in this section are probably in order. The comparison test is often the subtlest and most sensitive as well as the hardest of the four tests to use. In using the comparison test to decide whether a given series $\sum_{k=1}^{\infty} a_k$ converges or diverges, we must employ a second series whose behavior is understood. In fact, if we suspect that the given series converges, we must find a series $\sum_{k=1}^{\infty} b_k$ which we know converges and which satisfies $a_k \leq b_k$ for all sufficiently large values of k. On the other hand, if we suspect that the given series diverges, we must find a series $\sum_{k=1}^{\infty} b_k$ which we know diverges and which satisfies $b_k \leq a_k$ for all sufficiently large values of k. So, if we have a large store of series whose behaviors we understand, then it is possible to make good use of the comparison test.

The other three tests are more straightforward to use. If either the ratio or root tests predicts the convergence of a series, then that series converges at least as rapidly as a geometric series (this is clear from the proofs of these two tests). For series that converge relatively slowly (such as the p-series), the ratio and root tests are necessarily inconclusive and there is no alternative but to use the comparison or integral test, or possibly some other device.

In applying the ratio and root tests the following simple lemma is sometimes useful.

5.11 Lemma: $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Proof: If $f(x) = x^{1/x}$ then

$$\lim_{x\to\infty} \ln f(x) = \lim_{x\to\infty} \frac{\ln x}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0$$

by l'Hôspital's rule. Thus $\lim_{x\to\infty} f(x) = \mathrm{e}^0 = 1$ and so $\lim_{n\to\infty} \sqrt[n]{n} = 1$. ■

Let us consider an example that makes use of this fact.

5.12 Example: Consider the series $\sum_{n=0}^{\infty} nr^n$, where $r \geq 0$. Here

$$\lim_{n\to\infty} \sqrt[n]{nr^n} = \lim_{n\to\infty} r\sqrt[n]{n} = r$$

by Lemma 5.11, and so the root test implies that the series converges if $r < 1$ and diverges if $r > 1$. The root test is inconclusive if $r = 1$ but it is clear that this series diverges in this case. •

5.3 Convergence of general series

Consider a series (a_n) whose terms may be positive, negative, or zero, and let (S_n) be the sequence of partial sums of the series $\sum_{n=1}^{\infty} a_n$. Define two new sequences (p_n) and (q_n) as follows:

$$p_n = \begin{cases} a_n, & a_n \geq 0, \\ 0, & a_n < 0. \end{cases}$$

and

$$q_n = \begin{cases} 0, & a_n \geq 0, \\ -a_n, & a_n < 0. \end{cases}$$

Then $p_n, q_n \geq 0$, $a_n = p_n - q_n$, and $|a_n| = p_n + q_n$ for all n. Now put $U_n = \sum_{k=1}^{n} p_k$ and $V_n = \sum_{k=1}^{n} q_k$ for $n \in \mathbb{Z}_{>0}$, so that $S_n = U_n - V_n$ for $n \geq 1$ and $U_1 \leq U_2 \leq \cdots$, and $V_1 \leq V_2 \leq \cdots$. The sequences (U_n) and (V_n) thus either converge or diverge to ∞ by Theorem 2.9; put $U = \lim_{n\to\infty} U_n$ and $V = \lim_{n\to\infty} V_n$. There are four possibilities to be considered:

1. $U < \infty$ and $V < \infty$, in which case $\lim_{n\to\infty} S_n = U - V$;
2. $U < \infty$ and $V = \infty$, in which case $\lim_{n\to\infty} S_n = -\infty$;
3. $U = \infty$ and $V < \infty$, in which case $\lim_{n\to\infty} S_n = \infty$;
4. $U = \infty$ and $V = \infty$, in which case it is not clear what is $\lim_{n\to\infty} S_n$.

In case 4, the series $\sum_{n=1}^{\infty} a_n$ is very unstable and, by permuting its terms, can be made to do any of the following: converge to any specified number; diverge to ∞; diverge to $-\infty$; diverge without diverging to $\pm\infty$. Such series should be avoided since, if they converge, they "just barely" converge and do not obey the usual rules of arithmetic. In contrast, in cases 1–3 the series is stable and its value is unaffected by permutations of its terms.

The next two examples deal with the so-called ***alternating harmonic series***:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots .$$

These examples will show that this series belongs to case 4 and that it is very much affected by permutations of its terms.

5.13 Examples: 1. The key observation (or trick) in studying the convergence of the alternating harmonic series is to notice that $1/n = \int_0^1 x^{n-1}\,\mathrm{d}x$ for all $n \in \mathbb{Z}_{>0}$ and thus that

$$\begin{aligned}
\sum_{n=1}^{m} \frac{(-1)^{n+1}}{n} &= \sum_{n=1}^{m} (-1)^{n+1} \int_0^1 x^{n-1}\,\mathrm{d}x \\
&= \int_0^1 \sum_{n=1}^{m} (-1)^{n+1} x^{n-1}\,\mathrm{d}x \\
&= \int_0^1 \sum_{n=1}^{m} (-x)^{n-1}\,\mathrm{d}x \\
&= \int_0^1 \frac{1-(-x)^m}{1+x}\,\mathrm{d}x \\
&= \int_0^1 \frac{\mathrm{d}x}{1+x} - \int_0^1 \frac{(-x)^m}{1+x}\,\mathrm{d}x \\
&= \ln 2 - (-1)^m \int_0^1 \frac{x^m}{1+x}\,\mathrm{d}x.
\end{aligned}$$

Now, for $0 \le x \le 1$, we have

$$\int_0^1 \frac{x^m}{1+x}\,\mathrm{d}x \le \int_0^1 x^m\,\mathrm{d}x = \frac{1}{m+1}$$

since $1 + x \ge 1$ and, therefore,

$$\lim_{m\to\infty} \int_0^1 \frac{x^m}{1+x}\,\mathrm{d}x = 0.$$

This implies that

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n} &= \lim_{m\to\infty}\sum_{n=1}^{m}\frac{(-1)^{n+1}}{n}\\ &= \lim_{m\to\infty}\left(\ln 2-(-1)^m\int_0^1\frac{x^m}{1+x}\,\mathrm{d}x\right)\\ &= \ln 2.\end{aligned}$$

2. This example continues the analysis of the alternating harmonic series in part 1. In the notation introduced at the beginning of this section, we have

$$p_n = \begin{cases}1/n, & n \text{ odd},\\ 0, & n \text{ even}\end{cases}$$

and

$$q_n = \begin{cases}1/n, & n \text{ even},\\ 0, & n \text{ odd}\end{cases}$$

and thus

$$\sum_{n=1}^{\infty} p_n = 1+0+\frac{1}{3}+0+\frac{1}{5}+0+\cdots$$

and

$$\sum_{n=1}^{\infty} q_n = \frac{1}{2}+0+\frac{1}{4}+0+\frac{1}{6}+0+\cdots.$$

So each of $\sum_{n=1}^{\infty} p_n$ and $\sum_{n=1}^{\infty} q_n$ is more or less one half the sum of the harmonic series, and hence they both diverge.

Now let us consider the series

$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\frac{1}{9}+\frac{1}{11}-\frac{1}{6}+\frac{1}{13}-\cdots. \tag{5.4}$$

Notice that this series is a permutation of the alternating harmonic series and that, if (S_n) denotes its sequence of partial sums, then

$$S_{3n+1} = 1+\left(\frac{1}{3}-\frac{1}{2}+\frac{1}{5}\right)+\cdots+\left(\frac{1}{4n-1}-\frac{1}{2n}+\frac{1}{4n+1}\right) > 1$$

since

$$\frac{1}{2n-1}-\frac{1}{2n}+\frac{1}{2n+2} > 0$$

for all $n \geq 1$. This inequality means that, if the series (5.4) converges, then its limit exceeds 1 and hence does not converge to $\ln 2$. The alternating harmonic series and the series (5.4), therefore, behave very differently. •

The next theorem is very important because it sometimes allows us to use the comparison, integral, ratio, or root test to establish *convergence* for series with positive and negative terms. Note, however, that the converse of this theorem is false (cf. Example 5.13–1) so that tests for series with positive terms cannot be used to prove *divergence* of series with terms of both signs.

5.14 Theorem: *If $\sum_{n=1}^{\infty}|a_n|$ converges, then so does $\sum_{n=1}^{\infty} a_n$.*

Proof: Put $S_n = \sum_{k=1}^{n} a_k$ and $T_n = \sum_{k=1}^{n}|a_k|$ for $n \geq 1$. The hypothesis is that the limit $\lim_{n\to\infty} T_n$ exists and hence (T_n) is a Cauchy sequence by Theorem 2.23. So, given $\epsilon > 0$, there will be an integer $N \in \mathbb{Z}_{>0}$ such that $|T_n - T_m| < \epsilon$ for all $m, n \geq N$. But then, for any two integers m and n satisfying $n > m > N$, we have

$$\begin{aligned}|S_n - S_m| &= |a_{m+1} + a_{m+2} + \cdots + a_n| \\ &\leq |a_{m+1}| + |a_{m+2}| + \cdots + |a_n| \\ &= T_n - T_m < \epsilon.\end{aligned}$$

This means that (S_n) is a Cauchy sequence and then 2.23 implies that it converges. ■

5.15 Definition: The series $\sum_{n=1}^{\infty} a_n$ is said

(i) to ***converge absolutely*** if the series $\sum_{n=1}^{\infty}|a_n|$ converges and

(ii) to ***converge conditionally*** if $\sum_{n=1}^{\infty} a_n$ converges but $\sum_{n=1}^{\infty}|a_n|$ diverges. •

This is, at first sight, a very strange definition. It is, in fact, one of the (few?) situations in mathematics where terminology is a hindrance rather than an aid to understanding. The first thing to note is that a conditionally convergent series is convergent, the second is that (by Theorem 5.14) an absolutely convergent series is convergent, and the third is that there are series that converge conditionally but not absolutely (the alternating harmonic series, for example). When Theorem 5.14, in conjunction with the comparison, integral, ratio or root test is used to establish the convergence of a series with positive and negative terms, it is actually absolute convergence which is being established. To emphasise this, we state the corresponding versions of these theorems for general series.

5.16 Theorem: (Integral test for general series) *Suppose that (a_k) is a sequence and that f is a function which is continuous,*[3] *nonnegative-valued, and nonincreasing on the half-line $[1, \infty)$, and satisfies $|a_k| = f(k)$ for all $k \in \mathbb{Z}_{>0}$. Then the infinite series $\sum_{k=1}^{\infty} a_k$ converges if the improper integral*

$$\int_1^{\infty} f(x)\,\mathrm{d}x$$

converges.

Proof: By applying Theorem 5.7 to the series $\sum_{k=1}^{\infty}|a_k|$, we see that the series $\sum_{k=1}^{\infty} a_k$ is absolutely convergent, and so convergent by Theorem 5.14. ■

[3] The assumption that f is continuous is made merely to ensure that it is integrable over any interval of finite length.

5.17 Theorem: (Ratio test for general series) *Suppose that* (a_k) *is a sequence of nonzero terms and assume that* $\lim_{k\to\infty}|\frac{a_{k+1}}{a_k}| = L$ *exists.*

(i) If $L < 1$*, then* $\sum_{k=1}^{\infty} a_k$ *converges.*

(ii) If $L > 1$*, then* $\sum_{k=1}^{\infty} a_k$ *diverges.*

(iii) If $L = 1$*, then this test is inconclusive.*

Proof: (i) If $L < 1$ then the series $\sum_{k=1}^{\infty}|a_k|$ converges. Thus $\sum_{k=1}^{\infty} a_k$ converges absolutely and so converges by Theorem 5.14.

(ii) Referring to the proof of the corresponding part of Theorem 5.9, we see that, if $L > 1$, then there exists $N \in \mathbb{Z}_{>0}$ such that $|a_k| \geq |a_N|$ for all $k \geq N$. This prohibits $\lim_{k\to\infty} a_k = 0$, and so prohibits convergence of the series by Theorem 5.3.

(iii) This follows from the corresponding part of Theorem 5.9. ■

5.18 Theorem: (Root test for general series) *Suppose that* (a_n) *is a sequence of nonzero terms and let* $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.

(i) If $L < 1$ *then* $\sum_{n=1}^{\infty} a_n$ *converges.*

(ii) If $L > 1$ *then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

(iii) If $L = 1$ *this test is inconclusive.*

Proof: This follows from Theorem 5.10 in exactly the same way as Theorem 5.17 follows from Theorem 5.9. ■

The only test in these notes that is capable of establishing the convergence of a series which does not converge absolutely is the following one.

5.19 Theorem: (Alternating series test) *If* (a_n) *is a sequence such that*

(i) $\lim_{n\to\infty} a_n = 0$,

(ii) $|a_{n+1}| < |a_n|$ *for all* n*, and*

(iii) $a_{n+1}/a_n < 0$ *for all* n,

then the series $\sum_{n=1}^{\infty} a_n$ *converges.*

Proof: Suppose, to be definite, that $a_1 > 0$. Then $a_{2n} + a_{2n+1} < 0$, whereas $a_{2n-1} + a_{2n} > 0$ for all n. Let (S_n) denote the sequence of partial sums of the series in question. Then

$$\begin{aligned} S_{2n} &= a_1 + \cdots + a_{2n} \\ &= a_1 + \cdots + a_{2n-2} + a_{2n-1} + a_{2n} \\ &= S_{2n-2} + a_{2n-1} + a_{2n} \\ &> S_{2n-2} \end{aligned}$$

for all $n \in \mathbb{Z}_{>0}$, and so the sequence (S_{2n}) is strictly increasing. Moreover,

$$\begin{aligned} S_{2n} &= a_1 + (a_2 + a_3) + \cdots + (a_{2n-2} + a_{2n-1}) + a_{2n} \\ &< a_1 + (a_2 + a_3) + \cdots + (a_{2n-2} + a_{2n-1}) + (a_{2n} + a_{2n+1}) < a_1, \end{aligned}$$

and so the sequence (S_{2n}) is bounded above by a_1. But then (S_{2n}) must be a convergent sequence by Theorem 2.9; let S denote its limit. Now $S_{2n+1} = S_{2n} + a_{2n+1}$ for all $n \in \mathbb{Z}_{>0}$, and, since $\lim_{n\to\infty} a_n = 0$, it follows that

$$\lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} (S_{2n} + a_{2n+1}) = S$$

and, therefore, $\lim_{n\to\infty} S_n = S$ and so the series $\sum_{n=1}^{\infty} a_n$ converges. ■

Let us point out the significance of the three hypothesis. Condition (iii) says that the terms are nonzero and alternate in sign and condition (ii) says that, in magnitude, the terms are strictly decreasing. Of course, (i) is necessary for convergence by Theorem 5.3.

We have already remarked that a series converges conditionally if the sum of its positive terms diverges to ∞ and the series of its negative terms diverges to $-\infty$, yet the sum itself converges and has a finite value. This phenomenon can occur because of a careful "balancing" or "near-cancellation" of the positive and negative contributions within any particular partial sum. A special case of such "near-cancellation" occurs in Example 5.13–2 and in the proof of Theorem 5.19. We might expect that permuting the terms in a conditionally convergent series (a "rearrangement" of the series) would alter this balancing act and hence change the sum of the series. This is indeed the case and is exactly the point of the second part of Example 5.13–2.

5.20 Theorem: *By permuting the terms of a conditionally convergent series we can*

(i) make it converge to any number we wish or diverge to $\pm\infty$ or

(ii) make its partial sums oscillate between any two numbers.

Proof: We will only show how to rearrange the terms of a conditionally convergent series so as to make it converge to a given positive number S; rearranging the terms to make the series converge to a negative number or diverge to $\pm\infty$ is similar or to make its partial sums oscillate is similar.

First we can reorder the sequence, which we denote by (a_n), so that $(|a_n|)$ is nonincreasing, this being possible since conditional convergence ensures that $\lim_{n\to\infty} a_n = 0$. First add together the positive terms from the beginning of the sequence to make a sum that is larger than S (this is possible since the sum of the positive terms diverges). Now add enough negative terms from the beginning of the sequence to make the sum less than S. Next, add more positive terms, until the sum is again larger than S, and then add negative terms to make the sum less than S. Since the magnitudes of the terms is decreasing (this was the point of initially reordering so that $(|a_n|)$ is nonincreasing), the oscillations about S are becoming smaller and smaller and so the rearranged series converges to S. ■

5.4 Estimating remainders

If a series $\sum_{n=1}^{\infty} a_n$ converges to, say, S how well is S approximated by the nth partial sum? This is clearly an important question if we wish to estimate the sum of a series numerically. Now the error in the approximation is

$$|S - S_n| = \left|\sum_{k=1}^{\infty} a_k - \sum_{k=1}^{n} a_k\right| = \left|\sum_{k=n+1}^{\infty} a_k\right|.$$

We shall refer to $\sum_{k=n+1}^{\infty} a_k$ as the "remainder" or "tail" after the nth-term. Two of the tests for convergence already considered can be made to yield estimates for these tails.

5.21 Theorem: (Alternating series test remainder) *Suppose the series (a_n) satisfies the hypothesis of Theorem 5.19. Then*

$$|a_{n+1} + a_{n+2}| < \left|\sum_{k=n+1}^{\infty} a_k\right| < |a_{n+1}|$$

for any $n \geq 1$.

Proof: First notice that $a_n + a_{n+1}$ has the same sign as a_n and the opposite sign as a_{n-1} by assumptions (ii) and (iii). Thus

$$\begin{aligned}\left|\sum_{k=n+1}^{\infty} a_k\right| &= |(a_{n+1} + a_{n+2}) + (a_{n+3} + a_{n+4}) + \cdots| \\ &= |a_{n+1} + a_{n+2}| + |a_{n+3} + a_{n+4}| + \cdots \\ &> |a_{n+1} + a_{n+2}|\end{aligned}$$

and

$$\begin{aligned}\left|\sum_{k=n+1}^{\infty} a_k\right| &= |a_{n+1} + (a_{n+2} + a_{n+3}) + (a_{n+4} + a_{n+5}) + \cdots| \\ &= |a_{n+1}| - |a_{n+2} + a_{n+3}| - |a_{n+4} + a_{n+5}| - \cdots \\ &< |a_{n+1}|.\end{aligned}$$

The two inequalities in the theorem clearly follow from these two calculations. ■

5.22 Example: The alternating harmonic series satisfies the hypothesis of Theorem 5.19 and converges to $\ln 2$ by Example 5.13–1. It, therefore, follows from Theorem 5.21 that the 1000th partial sum of the alternating harmonic series will approximate $\ln 2$ with an error less than $1/1001$. (Rather disappointing accuracy!). •

5.23 Theorem: (Integral test remainder) *If the sequence (a_n) and the function f satisfies the hypotheses of the integral test (Theorem 5.7) and if $\sum_{n=1}^{\infty} a_n$ converges, then*

$$\int_{n+1}^{\infty} f(x)\,\mathrm{d}x \leq \sum_{k=n+1}^{\infty} a_k \leq \int_{n}^{\infty} f(x)\,\mathrm{d}x$$

for all $n \geq 1$.

Proof: If m and n are any two integers satisfying $m > n \geq 1$, then the inequality (5.2) implies

$$\begin{aligned}\int_{n+1}^{m} f(x)\,\mathrm{d}x &= \sum_{k=n+1}^{m-1} \int_{k}^{k+1} f(x)\,\mathrm{d}x \\ &\leq \sum_{k=n+1}^{m-1} a_k = \sum_{k=n}^{m-2} a_{k+1} \\ &\leq \sum_{k=n}^{m-2} \int_{k}^{k+1} f(x)\,\mathrm{d}x = \int_{n}^{m-1} f(x)\,\mathrm{d}x.\end{aligned}$$

The desired inequality now follows by taking the limit as m tends to infinity. ■

5.24 Example: What can we say about the error when the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is approximated by the sum $\sum_{n=1}^{10} \frac{1}{n^2}$? The error here is $\sum_{n=11}^{\infty} \frac{1}{n^2}$ and Theorem 5.23 implies that

$$\frac{1}{11} = \int_{11}^{\infty} \frac{1}{x^2}\,\mathrm{d}x \leq \sum_{n=11}^{\infty} \frac{1}{n^2} \leq \int_{10}^{\infty} \frac{1}{x^2}\,\mathrm{d}x = \frac{1}{10}.$$

We thus have

$$\left(\sum_{n=1}^{10} \frac{1}{n^2}\right) + \frac{1}{11} \leq \sum_{n=1}^{\infty} \frac{1}{n^2} \leq \left(\sum_{n=1}^{10} \frac{1}{n^2}\right) + \frac{1}{10},$$

and, since

$$1.5497 < \sum_{n=1}^{10} \frac{1}{n^2} < 1.5498,$$

it follows that

$$1.6406 \approx 1.5497 + .0909 < \sum_{k=1}^{\infty} \frac{1}{n^2} < 1.5498 + 0.1000 \approx 1.6498.$$

We can, therefore, say that $\sum_{n=1}^{\infty} \frac{1}{n^2} = 1.6452$ with an error of no more than ± 0.0046. In fact, the true value of $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is $\frac{\pi^2}{6} \approx 1.6449$. •

The final example in this chapter will illustrate that the ratio test can sometimes be used to estimate errors.

5.25 Example: How great is the error if $\sum \frac{n}{3^n}$ is approximated by $\sum_{n=1}^{10} \frac{n}{3^n}$? Observe that, if $n \geq 10$, then

$$\frac{n+1}{3^{n+1}} \Big/ \frac{n}{3^n} = \frac{(n+1)}{n}\frac{1}{3} < \frac{1.1}{3}$$

and hence

$$\frac{n+1}{3^{n+1}} < \frac{1.1}{3}\frac{n}{3^n}.$$

This implies that

$$\frac{10+k}{3^{10+k}} < \left(\frac{1.1}{3}\right)^k \frac{10}{3^{10}}$$

for all $k \geq 1$ (cf. the proof of the ratio test) and, therefore,

$$\begin{aligned}\sum_{n=11}^{\infty} \frac{n}{3^n} &= \sum_{k=1}^{\infty} \frac{10+k}{3^{10+k}} \\ &< \frac{1.1}{3}\frac{10}{3^{10}} + \left(\frac{1.1}{3}\right)^2 \frac{10}{3^{10}} + \cdots \\ &= \frac{10}{3^{10}}\frac{1.1}{3}\left(\frac{1}{1-1.1/3}\right) \\ &< 10^{-4}.\end{aligned}$$

•

The last two examples illustrate something which was said earlier, following Theorem 5.10. The series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges slowly in that the ratio and root tests are inconclusive and, correspondingly, the 1000th partial sum is not a good approximation. On the other hand, the two series $\sum_{n=1}^{\infty} \frac{1}{2^n}$ and $\sum_{n=1}^{\infty} \frac{n}{3^n}$ converge rapidly in that the ratio and root tests do predict convergence and, correspondingly, even the 10th partial sum is a good approximation.

Exercises

E5.1 Determine whether the following series converge or diverge:

(a) $\sum_{n=1}^{\infty} \frac{2n^2+3}{\sqrt{n^2+3n+2}}$

(b) $\sum_{n=1}^{\infty} \frac{1\cdot 2\cdot 3\cdot \ldots \cdot n}{3\cdot 5\cdot 7\cdot \ldots \cdot (2n+1)}$

(c) $\sum_{n=1}^{\infty} n\mathrm{e}^{-n^2}$

(d) $\sum_{n=1}^{\infty} \frac{(n!)^2 1000^n}{2^{n^2}}$

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

(f) $\sum_{n=1}^{\infty} \frac{n^2(n+1)^n}{(2n)^n}$

(g) $\sum_{n=1}^{\infty} \frac{\mathrm{e}^{1/n}}{n^2}$

E5.2 Determine whether or not each of the following series is convergent:

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

(b) $\sum_{n=1}^{\infty} \frac{n\cos(n\pi)}{n+1}$

(c) $\sum_{n=1}^{\infty} \left(n+\frac{1}{n}\right)\left(n^6+\ln n\right)^{-1/2}$

E5.3 Determine whether each of the following series diverge or converge.

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{n}{(n^2+1)(n+2)}$

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n+6.022\times 10^{23}}$

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

E5.4 Determine whether or not the series

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

is convergent.

Hint: *Use the identity* $a-b=\frac{a^2-b^2}{a+b}$.

E5.5 Determine whether the series $\sum_{n=1}^{\infty}\left(k+\frac{1}{k}\right)\left(k^5+\ln k\right)^{-1/2}$ converges.

E5.6 Determine whether or not the series $\sum_{n=1}^{\infty}(1-\cos(1/n))$ is convergent.

E5.7 Find both the nth partial sum as well as the sum of each of the following series:

(a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(2n-1)(2n+1)}$

(b) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)}$

E5.8 Determine all the real numbers p for which the series

$$\sum_{k=2}^{\infty} \frac{1}{k(\ln k)^p}$$

converges.

E5.9 In each of the following cases, determine all the values of $x \in \mathbb{R}$ for which the series converges.

(a) $\sum_{n=0}^{\infty} \frac{n^2}{n+1} x^n$

(b) $\sum_{n=1}^{\infty} \left(1 + \frac{x}{n}\right)^{n^2}$

E5.10 Determine all the values of x for which the following series converge:

(a) $\sum_{n=1}^{\infty} \frac{(n+3)x^{2n}}{n^n}$

(b) $\sum_{n=1}^{\infty} \frac{(n+1)(2x-1)^n}{2^n}$

E5.11 Determine all the values of x for which the following series converge:

(a) $\sum_{n=1}^{\infty} \frac{(n+3)x^{2n}}{n^n}$

(b) $\sum_{n=1}^{\infty} \frac{(n+1)(2x-1)^n}{2^n}$

E5.12 Determine all the values of x for which the infinite series $\sum_{n=1}^{\infty} \left(1 + \frac{x}{n}\right)^{n^2}$ converges.

E5.13 Recall that the binomial coefficient $\binom{n}{k}$ is defined as $\frac{n!}{k!(n-k)!}$. Determine all values of $x \geq 0$ for which the series $\sum_{n=0}^{\infty} \binom{2n}{n} x^n$ converges. What can you say about the series when $x < 0$?

Hint: *You may find the following bounds on* $\binom{2n}{n}$ *to be useful. Use the grouping*

$$\begin{aligned}(2n)! &= (1 \cdot 2) \times (3 \cdot 4) \times \cdots \times ((2n-1)(2n)) \\ &\leq 2^2 \times 4^2 \times \cdots \times (2n)^2\end{aligned}$$

to deduce that $(2n)! \leq 4^n (n!)^2$*, and so,* $\binom{2n}{n} \leq 4^n$*. Similarly, use the grouping*

$$(2n)! = (2 \cdot 3) \times (4 \cdot 5) \times \cdots \times ((2n-2)(2n-1)) \times (2n)$$

to deduce that $\binom{2n}{n} \geq (\frac{1}{2n})4^n$*.*

E5.14 If $x > y > 0$, does the series $\sum_{n=1}^{\infty} \frac{1}{x^n - y^n}$ converge or diverge? Why?

E5.15 (a) Prove that, if (a_n) is a sequence of positive numbers such that the series $\sum_{n=1}^{\infty} a_n$ converges, then so does the series $\sum_{n=1}^{\infty} a_n^2$.

(b) Show by example that the converse of (a) is false.

(c) Give an example of a sequence (a_n) such that the series $\sum_{n=1}^{\infty} a_n$ converges but the series $\sum_{n=1}^{\infty} a_n^2$ diverges.

E5.16 Let (a_k) be a sequence of nonnegative real numbers.

(a) Show that, if the series $\sum_{k=1}^{\infty} a_k^2$ converges, then so does the series $\sum_{k=1}^{\infty} \frac{a_k}{k}$.

Hint: *Apply the Cauchy–Schwartz inequality*

$$\left| \sum_{k=1}^{n} x_k y_k \right| \leq \left(\sqrt{\sum_{k=1}^{n} x_k^2} \right) \left(\sqrt{\sum_{k=1}^{n} y_k^2} \right)$$

to the partial sums of $\sum_{k=1}^{\infty} \frac{a_k}{k}$*.*

(b) Is the converse also true in general?

E5.17 Show that, if $0 < r < 1$, then it is not true that $n^{-2} \leq r^n$ for infinitely many positive integers n. Explain why, in view of this fact, neither the ratio test nor the root test could possibly be used to show that the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges.

E5.18 Consider the series

$$\sum_{n=1}^{\infty} a_k = \frac{1}{10} + 2^2 + \frac{1}{10^3} + 2^4 + \frac{1}{10^5} + 2^6 + \cdots .$$

(a) Show that $\lim_{n\to\infty} \sqrt[n]{a_n}$ does not exist.

(b) Find $\limsup_{n\to\infty} \sqrt[n]{a_n}$.

(c) Does the series converge or diverge?

E5.19 (a) Show that, if (a_n) is a sequence of positive numbers, then

$$\limsup_{n\to\infty} \sqrt[n]{a_n} \leq \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

Hint: *Show that, if* $L > \limsup_{n\to\infty} \frac{a_{n+1}}{a_n}$, *then* $L \geq \limsup_{n\to\infty} \sqrt[n]{a_n}$.

(b) Explain the significance of this inequality with respect to the relative strengths of the ratio and root tests.

E5.20 If $a_n > 0$ for all $n \in \mathbb{Z}_{>0}$ and if the series $\sum_{n=1}^{\infty} a_n$, diverges prove that the series $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$ diverges.

E5.21 Let (a_k) be a sequence of positive real numbers that is nonincreasing. Show that the alternating series $\sum_{k=1}^{\infty}(-1)^{k-1}a_k$ is convergent if and only if $\lim_{k\to\infty} a_k = 0$.
Hint: *For the* $\Leftarrow$ *direction, show that the sequence of partial sums* (S_n) *is Cauchy by generalizing the argument of Example 2.24.*

E5.22 The integral test shows that the harmonic series $\sum_{k=1}^{\infty} \frac{1}{k}$ is divergent. Consider the sequence defined by

$$a_n = \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) - \ln n.$$

Show that (a_n) is a bounded monotone sequence, and hence is convergent.
Hint: *It may be useful to draw a figure illustrating the integral test. Express* a_n *in terms of areas in this figure.*

(The limit $\lim_{n\to\infty} a_n$ is called Euler's constant (or sometimes, the Euler–Mascheroni constant), and is usually denoted by γ. Its value is approximately $0.5772\ldots$. It is not known whether γ is rational or irrational!)

E5.23 Determine whether the following series converge absolutely, converge conditionally, or diverge:

(a) $\sum_{n=1}^{\infty}(-1)^n(1 - \cos(1/n))$

(b) $\sum_{n=1}^{\infty}(-1)^n \left(n^{1/n}-1\right)^n$
(c) $\sum_{n=1}^{\infty} \frac{(n+3)\cos n\pi}{1003n}$

E5.24 Determine whether the following series converge absolutely, converge conditionally, or diverge.

(a) $\displaystyle\sum_{n=1}^{\infty}(-1)^n\,(n\sin(1/n)-1)$ (c) $\displaystyle\sum_{n=1}^{\infty}\frac{\cos n}{n^2}$

(b) $\displaystyle\sum_{n=1}^{\infty}\frac{\cos n}{2^n}$ (d) $\displaystyle\sum_{n=2}^{\infty}\frac{(-1)^n}{\ln n}$

E5.25 Determine whether each of the following series converges absolutely, converges conditionally, or diverges.

(a) $\sum_{k=1}^{\infty}(-1)^k\frac{\ln k}{k}$
(b) $\frac{1}{2}-\frac{1}{3}+\frac{1}{4}-\frac{1}{3^2}+\frac{1}{6}-\frac{1}{3^3}+\frac{1}{8}-\frac{1}{3^4}+\ldots$
Hint: *Consider the partial sums* S_{2n}

E5.26 Determine whether each of the following series converges absolutely, converges conditionally, or diverges.

(a) $\sum_{k=2}^{\infty}(-1)^k\frac{1}{k(\ln k)}$
(b) $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{2^2}+\frac{1}{5}-\frac{1}{2^3}+\frac{1}{7}-\frac{1}{2^4}+\cdots$
Hint: *Consider the partial sums* S_{2n}.

E5.27 Prove that, if the series $\sum_{n=1}^{\infty}a_n$ is absolutely convergent, then so is the series $\sum_{n=1}^{\infty}\left(\frac{n+1}{n}\right)a_n$.

E5.28 Suppose that a series $\sum_{n=1}^{\infty}a_n$ converges absolutely and that a series $\sum_{n=1}^{\infty}b_n$ converges conditionally. Prove that the series $\sum_{n=1}^{\infty}a_nb_n$ converges. Does this series converge conditionally or absolutely?

E5.29 Consider two sequences (a_n) and (b_n) and suppose that $a_n \neq 0$ and that $b_n > 0$ for all n. Show that if the series $\sum_{n=1}^{\infty}b_n$ converges and if

$$\left|\frac{a_{n+1}}{a_n}\right| \leq \frac{b_{n+1}}{b_n}$$

for all n then the series $\sum_{n=1}^{\infty}a_n$ converges absolutely.

E5.30 Evaluate $\sum_{n=10}^{\infty}\mathrm{e}^{-n}$ and compare its value with the lower and upper estimates obtained from the integral test.

E5.31 Verify that each of the following two series converge, and obtain lower and upper estimates for the sums.

(a) $\sum_{n=1}^{\infty}\frac{n}{2^n(n+3)}$
(b) $\sum_{n=1}^{\infty}n\mathrm{e}^{-n}$

E5.32 (a) Give an example of a continuous function f with domain $[1,\infty)$ and with the property that the limit

$$\lim_{n\to\infty}\int_1^n f(x)\,\mathrm{d}x$$

exists, but the improper integral

$$\int_1^\infty f(x)\,\mathrm{d}x$$

does not exist.

(b) Prove that, if f is a nonnegative-valued continuous function with domain $[1,\infty)$ such that the limit

$$\lim_{n\to\infty}\int_1^n f(x)\,\mathrm{d}x$$

exists, then the improper integral

$$\int_1^\infty f(x)\,\mathrm{d}x$$

converges and

$$\lim_{n\to\infty}\int_1^n f(x)\,\mathrm{d}x = \int_1^\infty f(x)\,\mathrm{d}x.$$

E5.33 Suppose we have an inexhaustible supply of uniform rectangular cards. Show that, given any distance d, a finite number of these cards can be stacked on a table so that the outside edge of the top card projects a distance d beyond the edge of the table.

Chapter 6

Sequences of functions

In this chapter we are going to examine what it means for a sequence of functions to converge to a function. This turns out not to be as straightforward as one might think but, at the same time, is a very useful and powerful concept in analysis, applied mathematics, and geometry. We will begin with an example that illustrates both the use of convergence of functions and points out some of subtleties involved in a discussion of such convergence.

6.1 Motivating example

In the study of ordinary differential equations, one needs to know that, under appropriate conditions, the initial value problem

$$\begin{aligned} y' &= f(x, y) \\ y(x_0) &= y_0 \end{aligned}$$

has a unique solution $y(x)$ defined for x near x_0.

The first step towards a proof of such an existence and uniqueness theorem is to observe that a function y is a solution of the initial value problem if and only if it satisfies the integral equation

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\, \mathrm{d}t. \tag{6.1}$$

(This equivalence is easy to check using the fundamental theorem of calculus.) The point of replacing the original initial value problem by the integral equation is simply that integration is much easier to deal with than differentiation.

We will solve this integral equation by finding a sequence of approximate solutions that converges to the actual solution. The initial approximation is the constant function y_0 satisfying $y_0(x) = y_0$ for all x, and the successive approximations $y_1, y_2, \ldots$

are then defined as follows:

$$\begin{aligned}
y_1(x) &= y_0 + \int_{x_0}^{x} f(t, y_0(t))\,\mathrm{d}t,\\
y_2(x) &= y_0 + \int_{x_0}^{x} f(t, y_1(t))\,\mathrm{d}t,\\
&\vdots\\
y_n(x) &= y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t))\,\mathrm{d}t.
\end{aligned}$$

The question now, of course, is whether this sequence of functions converge (in some sense) to a function y and, if so, whether this limit function is a solution of (6.1). If we believe that the sequence (y_n) converges to a function y, it is tempting to argue as follows:

$$\begin{aligned}
y(x) &= \lim_{n\to\infty} y_n(x)\\
&= \lim_{n\to\infty} \left(y_0 + \int_{x_0}^{x} f(t, y_{n-1}(t))\,\mathrm{d}t \right)\\
&= y_0 + \lim_{n\to\infty} \int_{x_0}^{x} f(t, y_{n-1}(t))\,\mathrm{d}t\\
&= y_0 + \int_{x_0}^{x} \lim_{n\to\infty} f(t, y_{n-1}(t))\,\mathrm{d}t\\
&= y_0 + \int_{x_0}^{x} f(t, \lim_{n\to\infty} y_{n-1}(t))\,\mathrm{d}t\\
&= y_0 + \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t.
\end{aligned}$$

There are two questionable steps here, in the fourth and fifth lines. The fourth line involves interchanging the limit and the integral and is not always valid. The conditions under which this is valid is one of the topics to be studied in this chapter (cf. Theorem 6.9). The fifth line is also contentious although it is clearly related to the continuity of f. But, at any rate, if this calculation can be justified, then y is a solution of the given initial value problem.

Many of the classical methods of solving ordinary and partial differential equations involve series of functions (for example, power, Fourier, and Bessel series). To understand and justify the use of such methods, it is necessary to understand convergence of sequences and series of functions.

6.2 Pointwise convergence

Suppose that I is an interval in $\mathbb{R}$ and that $f_1, f_2, \ldots$ are real-valued functions defined on I. (The interval may be open, closed, or neither, and may be of finite or infinite length.) Then, for each point $x \in I$, we have a sequence $(f_n(x))$ of numbers. In fact, we can think of $f_n(x)$ as a function of the two variables $n \in \mathbb{Z}_{>0}$ and $x \in I$, and so examining this sequence for a fixed value of x is not unlike examining a function of two real variables along, say, a vertical line in the plane. For each fixed value of $x \in I$, we can ask whether the limit $\lim_{n\to\infty} f_n(x)$ exists or whether the series $\sum_{n=1}^{\infty} f_n(x)$ converges. The point here is that, since x is fixed, the discussion of convergence of sequences in Chapter 2 and of series in Chapter 5 can be brought to bear on these questions. Since this sequence and series are being considered for a fixed value of x, the adjective "pointwise" is attached to these questions and their answers to distinguish them from a related situation we will consider later in this chapter.

6.1 Definition: Let $I \subseteq \mathbb{R}$ be an interval and suppose that (f_n) is a sequence of real-valued functions defined on I. We will say that

(i) the sequence (f_n) ***converges pointwise*** on I if the limit $\lim_{n\to\infty} f_n(x)$ exists for each point $x \in I$ and that

(ii) the series $\sum_{n=1}^{\infty} f_n$ ***converges pointwise*** on I if the series $\sum_{n=1}^{\infty} f_n(x)$ converges for each point $x \in I$. •

Suppose that (f_n) is a sequence of functions defined on an interval $I \subseteq \mathbb{R}$. If this sequence converges pointwise on I, we can define a function f on I by letting $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x \in I$, and, in this case, we will write "$\lim_{n\to\infty} f_n = f$ pointwise". Similarly, if the series $\sum_{n=1}^{\infty} f_n$ converges pointwise on I, we can define a function f on I by letting $f(x) = \sum_{n=1}^{\infty} f_n(x)$ for all $x \in I$, and in this case, we will write "$\sum_{n=1}^{\infty} f_n = f$ pointwise". Notice that, if we let $g_n = \sum_{k=1}^{n} f_k$ for $n \in \mathbb{Z}_{>0}$, then $f = \sum_{n=1}^{\infty} f_n$ pointwise if and only if $f = \lim_{n\to\infty} g_n$ pointwise.

According to Definition 2.4, this means that the following condition holds.

6.2 Lemma: *A sequence (f_n) of functions on an interval $I \subseteq \mathbb{R}$ converges pointwise if and only if the following condition holds:*

for each $\epsilon > 0$ and for each $x \in I$, there is an $N \in \mathbb{Z}_{>0}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \geq N$.

Note that here N is chosen *after* ϵ and x are selected, and hence that N depends on ϵ and x, i.e., $N = N(\epsilon, x)$.

Let us apply these definitions in an example.

6.3 Example: It is easy to see that

$$\lim_{n\to\infty} x^n = \begin{cases} 1, & x = 1, \\ 0, & 0 \leq x < 1. \end{cases}$$

So, if $f_n(x) = x^n$, for $x \in [0,1]$ and if

$$f(x) = \begin{cases} 1, & x = 1, \\ 0 & 0 \leq x < 1, \end{cases}$$

then $\lim_{n\to\infty} f_n = f$ pointwise on $[0,1]$. Notice that each f_n is continuous and that the limit function f is discontinuous.

If ϵ and x are two numbers in the open interval $(0,1)$, then $|f_n(x) - f(x)| = x^n$, and thus $|f_n(x) - f(x)| < \epsilon$ if and only if $n > \frac{\ln\epsilon}{\ln x}$. So the smallest possible value of N satisfying Lemma 6.2 is the smallest integer larger than $\frac{\ln\epsilon}{\ln x}$. Thus N does indeed depend on both ϵ and x and, for a fixed $\epsilon \in (0,1)$, $\lim_{x\to 1-} N(\epsilon, x) = \infty$. •

Pointwise convergence is a very weak notion of convergence and consequently not a very useful one. That is to say, one can draw very few conclusions from the assertion that $\lim_{n\to\infty} f_n = f$ pointwise on I. For example, the preceding example shows that

1. f need not be continuous, even if each of the f_n is C^∞,

and the next few examples will show that

2. f need not be integrable, even if each of the f_n is integrable,
3. if f and each of the f_n is continuous, then it is not necessarily true that
$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x,$$
where $I = [a,b]$, and
4. if f and each of the f_n is C^1, it is not necessarily true that
$$\lim_{n\to\infty} f_n' = f' \text{ pointwise.}$$

6.4 Example: For each $n \in \mathbb{Z}_{>0}$, let a_n be a positive number and let f_n be the function with domain $[0,1]$ which is zero on $[1/n, 1]$ and whose graph over $[0, 1/n]$ is an isosceles triangle with base $[0, 1/n]$ and height a_n. Formally,

$$f_n(x) = \begin{cases} 2na_n x, & 0 \le x \le 1/(2n), \\ 2a_n - 2na_n x, & 1/(2n) \le x \le 1/n, \\ 0, & 1/n \le x \le 1. \end{cases}$$

The sequence (f_n) converges pointwise to the zero function on $[0,1]$, regardless of the choice of the sequence (a_n). Indeed, it is certainly true that $f_n(0) = 0$ for all n. And, if $x \in (0,1]$ and if $N \in \mathbb{Z}_{>0}$ such that $1/N < x$, then $f_n(x) = 0$ for all $n \ge N$, and, therefore, $\lim_{n\to\infty} f_n(x) = 0$. Notice that, just as in the preceding example, the N of Lemma 6.2 appears to depend on both x and ϵ.

It is easy to see that, by choosing the sequence (a_n) appropriately, the limit

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \lim_{n\to\infty} \frac{a_n}{2n}$$

can be made to equal 0, 1, or ∞ by taking, for example, $a_n = 1$, $a_n = 2n$, or $a_n = n^2$, respectively. At any rate, it is not necessarily true that

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 \Big(\lim_{n\to\infty} f_n(x)\Big)\,\mathrm{d}x.$$ •

The example just presented, where pointwise convergence of functions does not imply the convergence of their integrals, is perhaps artificial in that the functions are piecewise linear, but not C^∞. The next example exhibits this same phenomenon with C^∞ functions.

6.5 Examples: 1. If $f_n(x) = nx(1-x^2)^n$ and $g_n(x) = nf_n(x)$ for $x \in [0,1]$, and $n \in \mathbb{Z}_{>0}$, then

$$\lim_{n\to\infty} f_n(x) = 0, \quad \lim_{n\to\infty} g_n(x) = 0$$

pointwise on $[0,1]$ and

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \frac{1}{2} \quad \lim_{n\to\infty} \int_0^1 g_n(x)\,\mathrm{d}x = \infty$$

by l'Hôspital's rule. The verification of these claims is left to the reader.

2. If $r_1, r_2, \ldots$ is an enumeration of $\mathbb{Q}$ (that is, a list of all the rational numbers) and if

$$f_n(x) = \begin{cases} 1, & x \in \{r_1, r_2, \ldots, r_n\} \\ 0, & \text{otherwise,} \end{cases}$$

then $\lim_{n\to\infty} f_n = f$ pointwise, where

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

Here each of the f_n is integrable and satisfies

$$\int_0^1 f_n(x)\,\mathrm{d}x = 0,$$

but the limit f is not even Riemann integrable.

3. The series $\sum_{n=0}^\infty \frac{x^2}{(1+x^2)^n}$ obviously converges to 0 for $x = 0$, and for $x \neq 0$ is a multiple of a convergent geometric series whose value can be calculated as follows:

$$\begin{aligned} \sum_{n=0}^\infty \frac{x^2}{(1+x^2)^n} &= x^2 \sum_{n=0}^\infty (1+x^2)^{-n} \\ &= x^2 \frac{1}{1-(1+x^2)^{-1}} \\ &= 1+x^2. \end{aligned}$$

So the series $\sum_{n=0}^\infty \frac{x^2}{(1+x^2)^n}$ converges pointwise for all $x \in \mathbb{R}$ to the function f defined by

$$f(x) = \begin{cases} 0, & x = 0, \\ 1+x^2, & x \neq 0. \end{cases}$$

4. Notice that, in the preceding example, the partial sums of the series are C^∞ functions, but that the pointwise limit of the series is not even continuous. Let's explore this a little more deeply.

   Let $f_n(x) = \frac{\sin nx}{\sqrt{n}}$ for $n \in \mathbb{Z}_{>0}$ and $x \in \mathbb{R}$. Then $\lim_{n\to\infty} f_n = 0$ pointwise on $\mathbb{R}$ since $|\sin nx| \le 1$. In fact, if $\epsilon > 0$ is given and if N is an integer with $1/N < \epsilon$, then $|f_n(x)| < \epsilon$ for all $n \ge N$ and all x. So, in this example, the N of Lemma 6.2 depends only on ϵ, in contrast to several earlier examples. Here $f_n'(x) = \sqrt{n}\cos nx$, and thus $\lim_{n\to\infty} f_n'(x) = \infty$ for every x, and hence the sequence (f_n') does not converge pointwise. •

6.3 Uniform convergence

Example 6.3 and the Examples 6.5 illustrate the fact that pointwise convergence of a sequence (f_n) of functions does not guarantee that the limit function inherits desirable properties from the terms f_n. Example 6.4 clearly illustrates the weakness of the concept of pointwise convergence: if $a_n = 1$ for all n, then $\lim_{n\to\infty} f_n = 0$ pointwise on $[0,1]$, but there is no integer n for which the graph of f_n is close to that of the limit function 0 on the whole interval $[0,1]$. The definition of uniform convergence captures the idea that, for each large value of n, the number $f_n(x)$ should be close to $f(x)$ for *all* $x \in [0,1]$.

6.6 Definition: Let I be an interval and suppose that $f, f_1, f_2, \dots$ are real-valued functions defined on I. (The interval may again be open, closed, or neither, and may be of finite or infinite length.)

(i) We will say that the sequence (f_n) ***converges uniformly*** to f on I and write "$\lim_{n\to\infty} f_n = f$ uniformly on I" if, for each $\epsilon > 0$, there is an integer $N \in \mathbb{Z}_{>0}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n \ge N$ and all $x \in I$.

(ii) We will say that the series $\sum_{n=1}^\infty f_n$ ***converges uniformly*** to f on I and write "$\sum_{n=1}^\infty f_n = f$ uniformly on I" if $\lim_{n\to\infty} \sum_{k=1}^n f_k = f$ uniformly on I (as defined in (i)), i.e., if, for each $\epsilon > 0$, there is an integer $N \in \mathbb{Z}_{>0}$ such that $|\sum_{k=1}^n f_k(x) - f(x)| < \epsilon$ for all $n \ge N$ and all $x \in I$. •

If a sequence of functions converges uniformly on an interval, then it certainly converges pointwise on that interval.

Note the subtle difference between this definition and that of pointwise convergence. In the definition of pointwise convergence, N was selected after x and ϵ were specified, and so N can depend on both x and ϵ (cf. Lemma 6.2). In the definition of uniform convergence, N is selected after only ϵ is specified, and so N can depend on ϵ but must be independent of x.

There are two helpful ways to think about Definition 6.6. Firstly, the condition that $|f_n(x) - f(x)| < \epsilon$ for all $x \in I$ can be rewritten as

$$f(x) - \epsilon < f_n(x) < f(x) + \epsilon, \qquad x \in I,$$

and interpreted as saying that the graph of f_n lies in a "tubular" neighbourhood of radius ϵ centered along the graph of f. And, secondly, the condition that $|f_n(x) - f(x)| < \epsilon$ for all $x \in I$ means that the maximum value of $f_n - f$ on I is less than ϵ and that the minimum value of $f_n - f$ on I is more than $-\epsilon$. It is, therefore, possible to verify the condition in Definition 6.6 by either graphing functions or solving one-variable max/min problems.

6.7 Examples: 1. If (a_n) and (f_n) are as in Example 6.4, then $\lim_{n\to\infty} f_n = 0$ uniformly on $[0,1]$ if and only if $\lim_{n\to\infty} a_n = 0$. Indeed, this follows from the above discussion and the fact that a_n is the maximum value of $f_n = f_n - 0$ on $[0,1]$.

2. The sequence considered in Example 6.3 does not converge uniformly on $[0,1]$. This can be seen either by examining the graphs of the functions in question or by using Theorem 6.8 below.

3. Let (f_n) be the sequence considered in Example 6.5–4. Then

$$|f_n(x) - 0| = \left|\frac{\sin nx}{\sqrt{n}}\right| \le \frac{1}{\sqrt{n}} < \epsilon$$

for all $x \in \mathbb{R}$ and all $n > \epsilon^{-2}$ and, therefore, $\lim_{n\to\infty} \frac{\sin nx}{\sqrt{n}} = 0$ uniformly on $\mathbb{R}$ and hence on any interval.

4. This example is a continuation of Example 6.5–3. If f is as in that example and if I is a closed interval, then $\sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n} = f(x)$ uniformly on I if and only if $0 \notin I$. To see that this it is necessary, in view of the definition of uniform convergence of a series, to estimate the partial sums. First of all, if $x \neq 0$ we have

$$\begin{aligned}\sum_{k=0}^{n} \frac{x^2}{(1+x^2)^k} &= \left(\sum_{k=0}^{n} (1+x^2)^{-k}\right) x^2 \\ &= \left(\frac{1-(1+x^2)^{-n-1}}{1-(1+x^2)^{-1}}\right) x^2 \\ &= 1 + x^2 - (1+x^2)^{-n}\end{aligned}$$

and, therefore,

$$\left|\sum_{k=0}^{n} \frac{x^2}{(1+x^2)^k} - f(x)\right| = \begin{cases} 0, & x = 0, \\ (1+x^2)^{-n}, & x \neq 0. \end{cases}$$

Suppose first that $0 \in I$ and assume that 0 is not the right end point of I. Then $I = [a,b]$, where $a \le 0 < b$, and the number $x_n = \sqrt{2^{1/n} - 1}$ belongs to I at least for sufficiently large values of n. Now

$$\left|\sum_{k=0}^{n} \frac{x_n^2}{(1+x_n^2)^k} - f(x_n)\right| = 1 + x_n^2 - \left(1+x_n^2\right)^{-n} - (1+x_n^2) = \frac{1}{2}$$

for all sufficiently large values of n by the above calculation, and the formula for f contained in Example 6.5–3. A similar argument will show that the series does not converge uniformly if 0 is not the left end point of I.

Now suppose that $0 \notin I$ and assume, to be definite, that I consists of positive numbers. Then $I = [a, b]$, where $0 < a < b$, and

$$\left| \sum_{k=0}^{n} \frac{x^2}{(1+x^2)^k} - f(x) \right| = (1+x^2)^{-n} \le (1+a^2)^{-n}$$

for all $x \in I$ and all $n \in \mathbb{Z}_{>0}$. Now $\lim_{n\to\infty}(1+a^2)^{-n} = 0$ and so, given an $\epsilon > 0$, there will be an integer $N \in \mathbb{Z}_{>0}$ such that $(1+a^2)^{-N} < \epsilon$. But this means that

$$\left| \sum_{k=0}^{n} \frac{x^2}{(1+x^2)^k} - f(x) \right| < \epsilon$$

for all $n \ge N$ and all $x \in I$ and, therefore, the series converges uniformly on I. A similar argument will evidently show that the series converges uniformly on I if I consists of negative numbers.

Finally, note that, if $0 \in I$, then it also follows from Theorem 6.8 (below), and the fact that f is discontinuous at 0, that the convergence is not uniform on I. •

6.4 Properties of uniform convergence

The examples of Section 6.2 show that pointwise convergence is a very weak property in that there cannot be any theorems relating pointwise convergence to continuity, integration, and differentiation. The point of this section is to show that uniform convergence is a stronger and much more useful assumption than pointwise convergence; we will do this by proving some nice theorems relating uniform convergence to continuity, integration, and differentiation.

Examples 6.3 and 6.5–3 show that a pointwise limit of continuous functions need not be continuous. The first theorem of this section shows that a uniform limit of continuous functions is necessarily continuous.

6.8 Theorem: *Suppose that (f_n) is a sequence of continuous functions defined on an interval I. If there is a function f on I such that $\lim_{n\to\infty} f_n = f$ uniformly on I then f is continuous on I.*

Proof: Recall that, by definition, f is continuous on I if it is continuous at each point of I. So, to prove the theorem, we must show that, corresponding to each point $x_0 \in I$ and each number $\epsilon > 0$, there is a number $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ for all $x \in I$ satisfying $|x - x_0| < \delta$. First note that

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_n(x) + f_n(x) - f_n(x_0) + f_n(x_0) - f(x_0)| \\ &\le |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \end{aligned}$$

for all $x \in I$ and all $n \in \mathbb{Z}_{>0}$. Now, since $\lim_{n\to\infty} f_n = f$ uniformly on I, there is an integer $N \in \mathbb{Z}_{>0}$ such that $|f(x) - f_n(x)| < \epsilon$ for all $x \in I$ and all $n \geq N$ and then, since f_N is continuous, there is a number $\delta > 0$ such that $|f_N(x) - f_N(x_0)| < \epsilon/3$ for all $x \in I$ satisfying $|x - x_0| < \delta$. But then, if x is any point in I satisfying $|x - x_0| < \delta$ and if $n = N$, each of the three summands in the last line of the above estimate is less than $\epsilon/3$ and, therefore,

$$|f(x) - f(x_0)| < \epsilon$$

for all such x. ■

While Example 6.5–1 showed that we cannot interchange a pointwise limit and an integral, the next result shows that we can interchange a uniform limit and an integral.

6.9 Theorem: *Suppose that (f_n) is a sequence of continuous functions defined on an interval $[a, b]$ and assume that there is a function f on I such that $\lim_{n\to\infty} f_n = f$ uniformly on $[a, b]$. Then*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Proof: Recall from Theorem 6.8 that f is continuous and hence that all of the integrals in the statement of the theorem exist. It is necessary to show that, given an $\epsilon > 0$, there is an integer $N \in \mathbb{Z}_{>0}$ such that

$$\left| \int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x \right| < \epsilon$$

for all $n \geq N$. Now, since $\lim_{n\to\infty} f_n = f$ uniformly on $[a, b]$, it follows that there is an integer $N \in \mathbb{Z}_{>0}$ such that $|f_n(x) - f(x)| < \frac{\epsilon}{2(b-a)}$ for all $x \in [a, b]$ and all $n \geq N$. But then

$$\begin{aligned} \left| \int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x \right| &\leq \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x \\ &\leq \int_a^b \frac{\epsilon}{2(b-a)}\,\mathrm{d}x \\ &= \frac{\epsilon}{2} < \epsilon \end{aligned}$$

for all $n \geq N$. ■

We already know from Example 6.5–4 that, if $\lim_{n\to\infty} f_n = f$ pointwise, then it does not follow that $\lim_{n\to\infty} f_n' = f'$ pointwise and, in analogy with the preceding theorem, we might hope to be able to prove that, if $\lim_{n\to\infty} f_n = f$ uniformly, then $\lim_{n\to\infty} f_n' = f'$ uniformly. But Exercise 6.7 and the next example both show that this is false in general, although Theorem 6.11 shows it is true under certain hypotheses. The point here is that

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

is a more smooth function than f, but that f' is a less smooth function. (For instance, if f is, say, C^4, then g is C^5, whereas f' is C^3.) The effect of this is that, in the above theorem on integration, it was not necessary to make assumptions beyond those involving the functions themselves, whereas in the theorem on differentiation (see below) it is necessary to make assumptions about both the functions and their derivatives.

6.10 Example: Let $f_n(x) = \dfrac{x}{1+n^2x^2}$ for $n \in \mathbb{Z}_{>0}$ and $x \in [-1,1]$. It follows easily from elementary calculus[1] that

$$-\frac{1}{2n} \leq f_n(x) \leq \frac{1}{2n}$$

for all $n \in \mathbb{Z}_{>0}$ and all $x \in [-1,1]$, and, therefore, $\lim_{n\to\infty} f_n = 0$ uniformly on $[-1,1]$. On the other hand,

$$f_n'(x) = \frac{1-n^2x^2}{(1+n^2x^2)^2}$$

and so

$$\lim_{n\to\infty} f_n'(x) = \begin{cases} 1, & x = 0, \\ 0, & 0 < |x| \leq 1. \end{cases}$$

So here $\lim_{n\to\infty} f_n' \neq (\lim_{n\to\infty} f_n)'$ and, moreover, since $\lim_{n\to\infty} f_n'$ is not continuous, the sequence (f_n') does not converge uniformly by Theorem 6.8. •

The derivative of a function defined on an interval is generally only considered at interior points of that interval and, for this reason, a function is usually only described as being C^1 if its domain is an open interval. For the purpose of the next theorem and corollary, it will be convenient to modify this terminology a little. Namely, suppose that h is a function defined on a closed interval $[a,b]$. Then, for this theorem and corollary only, we will let

$$h'(x) = \begin{cases} \text{the right derivative of } h \text{ at } a, & x = a, \\ \text{the derivative of } h \text{ at } x, & a < x < b \\ \text{the left derivative of } h \text{ at } b, & x = b. \end{cases}$$

In addition, we will say that h is C^1 on $[a,b]$ if the function h' as just defined is continuous on $[a,b]$.

6.11 Theorem: *Suppose that (f_n) is a sequence of functions defined on an interval $[a,b]$ and assume that*

(i) f_n is C^1 on $[a,b]$ for each $n \in \mathbb{Z}_{>0}$,

(ii) there is a function g defined on $[a,b]$ such that $\lim_{n\to\infty} f_n' = g$ uniformly on $[a,b]$, and

[1] One computes (1) that the derivative of f_n vanishes at x if and only if $x = \pm 1/n$ (2) that $x = -1/n$ is a minimum and $1/n$ is a maximum of f_n, and (3) $f_n(-1/n) = -1/2n$ and $f_n(1/n) = 1/2n$.

(iii) there is a point $c \in [a, b]$ for which the limit $\lim_{n\to\infty} f_n(c)$ exists.

Then there is a function f on $[a, b]$ such that $\lim_{n\to\infty} f_n = f$ uniformly on $[a, b]$ and $f' = g$.

Proof: The function f that is asserted to exist must satisfy the equation

$$f(x) = f(a) + \int_a^x g(t)\,\mathrm{d}t = \lim_{n\to\infty} f_n(a) + \int_a^x g(t)\,\mathrm{d}t$$

by the fundamental theorem of calculus. (Notice that g is continuous by assumptions (i) and (ii), and Theorem 6.8 and hence is integrable.) This suggests that we attempt to carry out the proof by

1. first showing that the $\lim_{n\to\infty} f_n(a)$ exists and defining a function f on $[a, b]$ by the formula
$$f(x) = \lim_{n\to\infty} f_n(a) + \int_a^x g(t)\,\mathrm{d}t, \tag{6.2}$$
2. showing that $\lim_{n\to\infty} f_n = f$ uniformly on $[a, b]$, and
3. showing $f' = g$.

The proof of 1 is easy. We know that

$$f_n(c) - f_n(a) = \int_a^c f_n'(t)\,\mathrm{d}t$$

for all $n \in \mathbb{Z}_{>0}$ by the fundamental theorem of calculus and that $\lim_{n\to\infty} f_n' = g$ uniformly on $[a, c]$, and so

$$\begin{aligned}\lim_{n\to\infty} f_n(a) &= \lim_{n\to\infty}\left(f_n(c) - \int_a^c f_n'(t)\,\mathrm{d}t\right)\\ &= \lim_{n\to\infty} f_n(c) - \int_a^c g(t)\,\mathrm{d}t\end{aligned}$$

by Theorem 6.9.

We can now use (6.2) to define a function f on $[a, b]$.

Since g is continuous, the definition of f and the fundamental theorem of calculus imply that f satisfies 3. To show that f also satisfies 2, suppose that $\epsilon > 0$ is given. Then (by (ii)) there will be an integer $N \in \mathbb{Z}_{>0}$ such that $|f_n'(x) - g(x)| < \frac{\epsilon}{2(b-a)}$ for all $x \in [a, b]$ and all $n \geq N$ and such that $|f_m(a) - \lim_{n\to\infty} f_n(a)| < \epsilon/2$ for all $m \geq N$. It now follows from the fundamental theorem of calculus and the definition of f that

$$\begin{aligned}|f_m(x) - f(x)| &= \left|\left(f_m(a) + \int_a^x f_m'(t)\,\mathrm{d}t\right) - \left(\lim_{n\to\infty} f_n(a) + \int_a^x g(t)\,\mathrm{d}t\right)\right|\\ &\leq \left|f_m(a) - \lim_{n\to\infty} f_n(a)\right| + \left|\int_a^x (f_m'(t) - g(t))\,\mathrm{d}t\right|\\ &< \frac{\epsilon}{2} + \int_a^x |f_m'(t) - g(t)|\,\mathrm{d}t\\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

for all $x \in [a, b]$ and all $m \geq N$. ■

The hypotheses in this theorem seem somewhat technical but they are often easy to verify and so the theorem is useful. The next corollary is really just a restatement of a special case of the theorem and so does not require a proof.

6.12 Corollary: *Suppose that (f_n) is a sequence of C^1 functions defined on an interval $[a, b]$ and assume that there are two functions f and g, also defined on $[a, b]$, such that $\lim_{n\to\infty} f_n = f$ uniformly on $[a, b]$ and $\lim_{n\to\infty} f_n' = g$ uniformly on $[a, b]$. Then $f' = g$.*

The condition (iii) in the above theorem is not superfluous in that there are sequences of functions which satisfy (i) and (ii), but not the conclusion of the theorem. The simplest such sequence is the one in which $f_n(x) = n$ for all x and all n. This sequence certainly satisfies (i) and (ii), but not (iii) and does not converge pointwise.

This concludes the discussion of the properties of uniform convergence of sequences of functions. But what about uniform convergence of series of functions? In principle, there is nothing new here since a series is just the limit of its sequence of partial sums. But, in applying the three preceding theorems, it would be desirable to have a criteria for deciding when a series of functions converges uniformly, and there is a such a criteria.

6.13 Theorem: (Weierstrass M-test) *Suppose that (f_n) is a sequence of functions defined on an interval I and assume that there is a sequence (M_n) of numbers such that*

(i) $|f_n(x)| \leq M_n$ for all $x \in I$ and all $n \in \mathbb{Z}_{>0}$ and

(ii) the series $\sum_{n=1}^{\infty} M_n$ converges.

Then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on I.

Proof: Put $S_n(x) = \sum_{k=1}^{n} f_k(x)$ and $T_n = \sum_{k=1}^{n} M_k$ for $n \in \mathbb{Z}_{>0}$ and $x \in I$. Then the sequence (T_n) is clearly nondecreasing and it converges by assumption (ii), and hence is bounded by Proposition 2.6. Moreover, if $1 \leq n < m$, then

$$\begin{aligned}
|S_m(x) - S_n(x)| &= \left| \sum_{k=n+1}^{m} f_k(x) \right| \\
&\leq \sum_{k=n+1}^{m} |f_k(x)| \\
&\leq \sum_{k=n+1}^{m} M_k = T_m - T_n
\end{aligned}$$

for all $x \in I$. Now, since (T_n) is a Cauchy sequence, this calculation implies that $(S_n(x))$ is a Cauchy sequence for each $x \in I$ and, therefore, that the sequence (S_n) converges pointwise by Theorem 2.23. So, if we put $f(x) = \lim_{n\to\infty} S_n(x)$ for each

$x \in I$, then $f = \sum_{n=1}^{\infty} f_n$ pointwise on I, and it remains to show that this convergence is actually uniform. Now, if x is any point in I, then

$$\begin{aligned}\left|f(x) - \sum_{k=1}^{n} f_n(x)\right| &= |f(x) - S_m(x) + S_m(x) - S_n(x)| \\ &\leq |f(x) - S_m(x)| + |S_m(x) - S_n(x)| \\ &\leq |f(x) - S_m(x)| + T_m - T_n\end{aligned}$$

for any $m \geq n$ and, therefore,

$$\begin{aligned}\left|f(x) - \sum_{k=1}^{n} f_n(x)\right| &\leq \lim_{m\to\infty} (|f(x) - S_m(x)| + T_m - T_n) \\ &= \lim_{m\to\infty} T_m - T_n.\end{aligned}$$

Now, if a number $\epsilon > 0$ is given, there will be an integer $N \in \mathbb{Z}_{>0}$ such that $\lim_{m\to\infty} T_m - T_n < \epsilon$ whenever $n \geq N$, and, therefore, $|f(x) - \sum_{k=1}^{n} f_n(x)| < \epsilon$ for all $n \geq N$ and all $x \in I$. This makes it clear that $\sum_{n=1}^{\infty} f_n = f$ uniformly on I. ■

6.14 Example: Consider the series $\sum_{n=1}^{\infty} \frac{1}{n^3} \sin(nx)$ (this is an example of a Fourier series). Put $u_n(x) = \frac{\sin nx}{n^3}$ and $S_n(x) = \sum_{k=1}^{n} u_k(x)$ for $n \in \mathbb{Z}_{>0}$ and $x \in \mathbb{R}$. The series

$$\sum_{n=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{n^3}\sin(nx)\right) = \sum_{n=1}^{\infty} \frac{1}{n^2}\cos(nx)$$

converges uniformly on $\mathbb{R}$ by the Weierstrass M-test since

$$\left|\frac{1}{n^2}\cos(nx)\right| \leq \frac{1}{n^2}$$

and since the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges by Example 5.8.

Now $\sum_{n=1}^{\infty} \frac{1}{n^3} \sin(nx)$ converges to 0 when $x = 0$. Hence, by Theorem 6.11, we conclude that the series may be differentiated term-by-term, and

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=1}^{\infty} \frac{1}{n^3}\sin(nx) = \sum_{n=1}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{n^3}\sin(nx)\right).$$

•

6.5 The sup-norm

Suppose that I is an interval (which may be open, closed. or neither, and of finite or infinite length) in $\mathbb{R}$. Recall that a function f with domain I is said to be bounded if there is a number M such that $|f(x)| \leq M$ for all $x \in I$. Let $\mathrm{B}(I)$ denote the set consisting of all the bounded functions with domain I and $\mathrm{C}(I)$ the set consisting of all the continuous functions with domain I. In general, neither of these sets of functions is a subset of the other, although if I is closed and bounded then $\mathrm{C}(I) \subseteq \mathrm{B}(I)$ by Theorem 4.24(i). Both of these sets of functions are (infinite-dimensional) vector spaces in the sense that, if f and g are functions in one of these sets and if a is a number, then the functions $f+g$ and af defined by the equations

$$(f+g)(x) = f(x) + g(x)$$

and

$$(af)(x) = af(x)$$

also belong to that set. This is known in the case of $\mathrm{C}(I)$ (since linear combinations of continuous functions are again continuous) and in the case of $\mathrm{B}(I)$ can be shown as follows. If f and g are in $\mathrm{B}(I)$ then there are numbers M and N satisfying $|f(x)| \leq M$ and $|g(x)| \leq N$ for all $x \in I$, and consequently

$$|(f+g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq M + N$$

and

$$|(af)(x)| = |af(x)| = |a||f(x)| \leq |a|M$$

for all $x \in I$. The origin in both of these vector spaces is the zero-function.

For any function $f \in \mathrm{B}(I)$, the quantity

$$\|f\| = \sup\{|f(x)| \mid x \in I\}$$

is a finite number called the **sup-*norm*** of f on I. (Notice that the sup-norm $\|\cdot\|$ depends on I. but that I is not included in the notation.) Thus $\|\cdot\|$ is a function on $\mathrm{B}(I)$ and it has the three properties:

1. $\|f\| \geq 0$ for all $f \in \mathrm{B}(I)$ with equality if and only if $f = 0$;
2. $\|af\| = |a| \, \|f\|$ for all $a \in \mathbb{R}$ and all $f \in \mathrm{B}(I)$;
3. $\|f+g\| \leq \|f\| + \|g\|$ for all $f, g \in \mathrm{B}(I)$.

Only the third of these properties is not obvious and it can be proven as follows. If f and g are two functions in $\mathrm{B}(I)$ and if x is a point in I, then

$$|(f+g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|$$

and, therefore,

$$\|f+g\| = \sup\{|(f+g)(x)| \mid x \in I\} \leq \|f\| + \|g\|.$$

The sup-norm on $\mathrm{B}(I)$, therefore, has the same basic properties as the usual norm has on $\mathbb{R}^d$ (cf. the beginning of Chapter 4). Now a cursory examination will reveal that all of the ideas presented in Chapter 4 depended ultimately on the balls $B_r(\boldsymbol{x})$ and that the definitions of these balls depended only on the norm. This suggests that it should be possible to carry over to $\mathrm{B}(I)$ all of the definitions that were made for $\mathbb{R}^d$ in Chapter 4 and most (but definitely not all) of the results that were proven there. In doing this we must replace $\mathbb{R}^d$ by $\mathrm{B}(I)$ and regard this as the set we are studying and we must replace the vectors in $\mathbb{R}^d$ by the functions in $\mathrm{B}(I)$ and regard these as the points in the set we are studying. Moreover, the functions that were studied in Section 4.4 (which all had subsets of $\mathbb{R}^d$ as their domains) must now be replaced by functions defined on subsets of $\mathrm{B}(I)$, i.e., functions whose argument is itself a function. (We already know many examples of such functions: evaluation, differentiation, and integration are just three such examples).

The definitions of distance between points in $\mathrm{B}(I)$ and convergence of sequences in $\mathrm{B}(I)$ are completely analogous to those in $\mathbb{R}^d$.

6.15 Definition: Let I be an interval in $\mathbb{R}$.

(i) The ***distance*** between two functions f and g in $\mathrm{B}(I)$ is the number $|f - g|$.

(ii) A sequence (f_n) in $\mathrm{B}(I)$ is said to ***converge*** to a function $f \in \mathrm{B}(I)$ if, for each $\epsilon > 0$, there is an integer $N \in \mathbb{Z}_{>0}$ such that $\|f_n - f\| < \epsilon$ whenever $n \geq N$. •

Consider two functions f and g in $\mathrm{B}(I)$ and an $\epsilon > 0$. Now

$$\|f - g\| = \sup\{|f(x) - g(x)| \mid x \in I\}$$

and so, if $\|f - g\| < \epsilon$, then $|f(x) - g(x)| < \epsilon$ for all $x \in I$ and, conversely, if $|f(x) - g(x)| < \epsilon$ for all $x \in I$, then $\|f - g\| \leq \epsilon$. In particular, if (f_n) is a sequence in $\mathrm{B}(I)$, this observation takes the following form: If $\|f_n - f\| < \epsilon$ for all $n \geq N$, then $|f_n(x) - f(x)| < \epsilon$ for all $x \in I$ and all $n \geq N$ and, conversely, if $|f_n(x) - g(x)| < \epsilon$ for all $x \in I$ and all $n \in N$, then $\|f_n - f\| \leq \epsilon$ for all $n \in \mathbb{Z}_{>0}$. This just means that convergence in $\mathrm{B}(I)$ in the sense of Definition 6.15 is the same as uniform convergence on I. This observation can be summed up as follows:

6.16 Proposition: *A sequence (f_n) in $\mathrm{B}(I)$ converges to a function $f \in \mathrm{B}(I)$ (in the sense of Definition 6.15) if and only if $\lim_{n\to\infty} f_n = f$ uniformly on I.*

Consider again an interval I in $\mathbb{R}$. A sequence (f_n) of functions in $\mathrm{B}(I)$ is said to be ***pointwise Cauchy*** if the numerical sequence $(f_n(x))$ is Cauchy for all $x \in I$. With this definition, it follows from Theorem 2.23 that the sequence (f_n) converges pointwise on I if and only if it is pointwise Cauchy. The following definition and theorem capture the corresponding definition and characterization for uniform convergence.

6.17 Definition: Suppose that I is an interval in $\mathbb{R}$. A sequence (f_n) of functions in $\mathrm{B}(I)$ is said to be ***uniformly Cauchy*** on I if, for each $\epsilon > 0$, there exists an integer N such that $\|f_n - f_m\| < \epsilon$ for all $m, n > N$. •

6.18 Theorem: *Suppose that I is an interval in $\mathbb{R}$. A sequence (f_n) of functions in $\mathrm{B}(I)$ converges uniformly if and only if it is uniformly Cauchy.*

Proof: Suppose first that (f_n) converges uniformly on I to a function f. Then, given $\epsilon > 0$, there will be an integer N such that $|f_n(x) - f(x)| < \epsilon/2$ for all $x \in I$ and all

$n \geq N$. But this implies that

$$\begin{aligned}|f_n(x) - f_m(x)| &\leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

for all $x \in I$ and all $m, n \geq N$, hence $\|f_n - f_m\| \leq \epsilon$ for all $m, n \geq N$, and, therefore, (f_n) is a uniformly Cauchy.

Now suppose, conversely, that (f_n) is a uniformly Cauchy sequence on I. Then $|f_m(x) - f_n(x)| \leq \|f_m - f_n\|$ for all $x \in I$ and all $m, n \in \mathbb{Z}_{>0}$ and, therefore, $(f_n(x))$ is a Cauchy and hence (by Theorem 2.23) a convergent sequence of numbers for each $x \in I$. This means that the formula $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in I$ defines a function f on I such that $\lim_{n\to\infty} f_n = f$ pointwise on I. To complete the proof it is sufficient to show that this convergence is actually uniform on I.

Let ϵ be a positive number and let x be a point in I. Then, as (f_n) is uniformly Cauchy, there is an integer N such that $\|f_m - f_n\| < \epsilon/2$ for all $m, n \geq N$. And, as $\lim_{n\to\infty} f_n = f$ pointwise on I, there is an integer M such that $|f_n(x) - f(x)| < \epsilon/2$ for all $n \geq M$. So, if $n \geq N$ and if $m = \max\{M, N\}$, then

$$\begin{aligned}|f_n(x) - f(x)| &\leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| \\ &\leq \|f_n - f_m\| + |f_m(x) - f(x)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

Since this is true for any point $x \in I$, it follows that $\|f_n - f\| \leq \epsilon$ whenever $n \geq N$, and this just means that $\lim_{n\to\infty} f_n = f$ uniformly on I. ■

It is possible to use Theorem 6.18 to give a very short proof of the Weierstrass M-test (Theorem 6.13).

6.19 Example: Let f and f_n for $n \in \mathbb{Z}_{>0}$ be as in Example 6.3. Then $\|f_n - f\| = 1$ for all $n \in \mathbb{Z}_{>0}$ since, if x is a little less than 1, then $f_n(x) = x^n$ is almost 1 and $f(x) = 0$. More precisely, if we let

$$x_n = \left(\frac{n}{n+1}\right)^{1/n}$$

for $n \in \mathbb{Z}_{>0}$, then $0 < x_n < 1$ for each n and

$$\|f_n - f\| \geq |f_n(x_n) - f(x_n)| = x_n^n = \frac{n}{n+1}$$

and so $\|f_n - f\| = 1$ for each $n \in \mathbb{Z}_{>0}$. This shows once more that $\lim_{n\to\infty} f_n = f$ pointwise, but not uniformly, on $[0, 1]$. •

6.20 Corollary: *If I is a closed and bounded interval in $\mathbb{R}$, then every sequence in $\mathrm{C}(I)$ that is uniformly Cauchy converges uniformly to a function in $\mathrm{C}(I)$.*

Proof: Suppose that (f_n) is a uniformly Cauchy sequence in $\mathrm{C}(I)$. Now, since I is compact, we know from Theorem 4.24pl:cptcont1 that $\mathrm{C}(I) \subseteq \mathrm{B}(I)$ and hence (f_n) is a uniformly Cauchy sequence in $\mathrm{B}(I)$. The preceding theorem now asserts that there is a function f in $\mathrm{B}(I)$ such that $\lim_{n\to\infty} f_n = f$ uniformly on I, and then Theorem 6.8 implies that f is continuous, i.e., belongs to $\mathrm{C}(I)$. ■

The emphasis on the completeness of the real line is the major difference between the discussion contained in these notes and that which could or would be given in a first-year calculus course. As has already been remarked, the following assertions are equivalent and any one of them can be taken as the statement of the completeness of the real line:

1. every nonempty subset of $\mathbb{R}$ that is bounded above has a least upper bound (Theorem 1.9);
2. the Bolzano–Weierstrass theorem (Theorem 2.13);
3. every Cauchy sequence in $\mathbb{R}$ converges (Theorem 2.23).

In the context of $\mathrm{B}(I)$, conditions 1 and 2 no longer make sense, since there is no suitable order on this set of functions, but the analogue of condition 3 is true by Theorem 6.18. And, if I is a closed and bounded interval in $\mathbb{R}$, the analogue of 3 is also true for $\mathrm{C}(I)$ by Corollary 6.20.

Exercises

E6.1 Show that each of the following sequences of functions converges pointwise on $[0,1]$ to the zero function and determine whether the convergence is uniform on that interval.

(a) $f_n(x) = nx\mathrm{e}^{-nx}$
(b) $f_n(x) = nx\mathrm{e}^{-n^2x}$
(c) $f_n(x) = nx\mathrm{e}^{-nx^2}$

E6.2 Determine if the following sequences (f_n) of functions on $[0,1]$ converge pointwise and/or uniformly.

(a) $f_n(x) = \begin{cases} 2nx, & x \in [0, \frac{1}{2n}], \\ 2 - 2nx, & x \in (\frac{1}{2n}, \frac{1}{n}], \\ 0, & x \in (\frac{1}{n}, 1] \end{cases}$

(b) $f_n(x) = \begin{cases} \frac{\sin(1/n^2x)}{\sqrt{n}}, & x \in (0,1], \\ 0, & x = 0 \end{cases}$

(c) $f_n(x) = \displaystyle\sum_{k=1}^{n} \frac{\cos(kx)}{k}$

E6.3 Let $f_n(x) = \frac{1}{x} + \frac{1}{n}$ and let $g_n(x) = 1 + \frac{1}{n}$ for $n \in \mathbb{Z}_{>0}$ and $x \in (0,1)$.

(a) Show that $\lim_{n\to\infty} f_n = 1/x$ uniformly on $(0,1)$.
(b) Show that $\lim_{n\to\infty} g_n = 1$ uniformly on $(0,1)$.
(c) Show that $\lim_{n\to\infty} f_n g_n = 1/x$ pointwise, but not uniformly, on $(0,1)$.

E6.4 Determine the pointwise limit $\lim_{n\to\infty} f_n$ of the given sequence of functions on the interval $[0,1]$, and also determine whether the convergence is uniform on that interval.

(a) $f_n(x) = 1/(1+nx)$
(b) $f_n(x) = x/(1+nx)$

E6.5 Compute the pointwise limits of the following sequences of functions on the indicated interval and explain why the convergence is not uniform.

(a) $S_n(x) = \dfrac{n^2x^n}{1+n^2x^n}$ on $[0,2]$
(b) $S_n(x) = \dfrac{1-n^2x^2}{1+n^2x^2}$ on $[-1,1]$

E6.6 Let $f_n(x) = \frac{nx^2}{1+nx^2}$ for $x \in \mathbb{R}$ and $n \in \mathbb{Z}_{>0}$.

(a) Show that the sequence (f_n) converges pointwise on $\mathbb{R}$.
(b) Show that the sequence (f_n) does not converge uniformly on any closed interval containing 0.
(c) Show that the sequence (f_n) converges uniformly on any closed interval not containing 0.

E6.7 Show that $\sum_{n=0}^{\infty} x^n$ converges uniformly on $[-a, a]$ for every a satisfying $0 < a < 1$, but does not converge uniformly on $(-1, 1)$.

E6.8 Prove that $\sum_{n=1}^{\infty} \frac{x^n}{n!}$ converges uniformly on $[-a, a]$ for each positive real number a, but does not converge uniformly on $\mathbb{R}$.

E6.9 (a) Show that $\lim_{n\to\infty} nx^p e^{-nx} = 0$ pointwise on $[0, 1]$ for each positive number p.

(b) Determine all the positive numbers p for which the convergence in (a) is uniform on $[0, 1]$.

E6.10 Consider the sequence of functions (f_n), where

$$f_n(x) = \sin\left(\frac{\pi x}{2n}\right)$$

for $n \geq 1$ and $x \in \mathbb{R}$.

(a) Determine the pointwise limit of this sequence on $\mathbb{R}$.

(b) Show that the sequence converges uniformly on the closed interval $[-a, a]$ for any $a > 0$.

(c) Show that the sequence does not converge uniformly on $\mathbb{R}$.

E6.11 Suppose that a sequence of functions (f_n) defined on an interval I satisfies

1. $|f_n(x) - f_n(y)| \leq |x - y|$ for any $x, y \in I$ and any $n \geq 1$ and
2. (f_n) converges pointwise on I to a function f.

Show that f satisfies $|f(x) - f(y)| \leq 2|x - y|$ for all $x, y \in I$.

E6.12 Determine all the values of x for which the series $\sum_{n=1}^{\infty} \frac{1}{(x(2-x))^n}$ converges. Also, determine interval(s) on which this series converges uniformly. On what set is this series a continuous function?

E6.13 Find a sequence of functions (f_n) defined on $[0, 1]$ such that each f_n is discontinuous at each point in $[0, 1]$, but (f_n) converges uniformly to a function f that is continuous on $[0, 1]$.

Hint: *Define $f_1\colon [0, 1] \to \mathbb{R}$ by setting $f_1(x)$ to be 1 for $x \in \mathbb{Q}$, and 0 for $x \notin$ rational.*

E6.14 Prove that, if a sequence (f_n) of functions converges uniformly to a function f on an interval I, and if a second sequence (g_n) converges uniformly to a second function g on I then the sequence $(f_n + g_n)$ converges uniformly to $(f + g)$ on I.

E6.15 (a) Prove that, if (f_n) converges uniformly to f on an interval I, if (g_n) converges uniformly to g on I, and if (f_n) and (g_n) are uniformly bounded on I, then $(f_n g_n)$ converges uniformly to fg on I. (Use the following definition: A sequence (f_n) of functions on I is said to be ***uniformly bounded*** on I if there exists a constant M such that $|f_n(x)| < M$ for all n and for all $x \in I$.)

(b) Give an example to show that, if one of (f_n) or (g_n) is uniformly bounded and the other is not, then the result of (a) is false.

E6.16 Suppose that f is a real-valued function with domain $\mathbb{R}$ and with the property that $|f(x) - f(y)| \leq K|x - y|$ for all $x, y \in \mathbb{R}$ and some $K > 0$. Suppose also that (g_n) is a sequence of functions from $\mathbb{R}$ to $\mathbb{R}$ which converges uniformly on $\mathbb{R}$ to a function g. Prove that the sequence $(f \circ g_n)$ converges uniformly to $f \circ g$ on $\mathbb{R}$.

E6.17 Suppose that (f_n) is a sequence of differentiable functions on $[0, 1]$ that converges uniformly to some function. Must the sequence f_n' converge uniformly? ***Hint:*** *Consider the functions* $f_n(x) = (\sin(n^2 x))/n$ *or* $g_n(x) = x^{n+1}/(n + 1)$ *on* $[0, 1]$.

E6.18 Recall the series $\sum_{n=0}^{\infty} \binom{2n}{n} x^n$ from Exercise E5.13. It was determined there that the series converges pointwise on $(-1/4, 1/4)$.

(a) Show that the series $\sum_{n=1}^{\infty} \binom{2n}{n} x^n$ converges uniformly on the interval $[-a, a]$ for any real number a satisfying $0 < a < 1/4$.

(b) Set $f(x) = \sum_{n=1}^{\infty} \binom{2n}{n} x^n$ for $x \in (-1/4, 1/4)$. Explain why the derivative of $f(x)$ is given by $f'(x) = \sum_{n=1}^{\infty} n\binom{2n}{n} x^{n-1}$ for any $x \in (-1/4, 1/4)$.

E6.19 Show that, if the series $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly on an interval I, then $\lim_{n\to\infty} u_n = 0$ uniformly on I.

E6.20 (a) Use the Weierstrass M-test to show that the series $\sum_{n=1}^{\infty} n^{-3/2} \cos(x/n)$ converges uniformly on $\mathbb{R}$. Let S denote its sum.

(b) Deduce from (a) and Theorem 6.11 that the series $\sum_{n=1}^{\infty} n^{-1/2} \sin(x/n)$ converges uniformly on all closed intervals of finite length, and that its derivative is the function $S(x)$ of (a). This theorem does not imply that the series $\sum_{n=1}^{\infty} n^{-1/2} \sin(x/n)$ converges uniformly on $\mathbb{R}$. Why?

(c) Use the Weierstrass M-test and the inequality $|\sin u| \leq |u|$ (which is valid for all $u \in \mathbb{R}$) to prove that the series $\sum_{n=1}^{\infty} n^{-1/2} \sin(x/n)$ converges uniformly on all closed intervals of finite length.

(d) Does the series $\sum_{n=1}^{\infty} n^{-1/2} \sin(x/n)$ converges uniformly on $\mathbb{R}$?

E6.21 Let $S_n(x) = nx/(nx + 1)$ for $n \geq 1$ and $x \in [0, 1]$.

(a) Compute the pointwise limit $\lim_{n\to\infty} S_n(x)$ on $[0, 1]$. Is the convergence uniform on this interval?

(b) Compute $\lim_{n\to\infty} \int_0^1 S_n(x)\, \mathrm{d}x$ and

$$\int_0^1 \lim_{n\to\infty} S_n(x)\, \mathrm{d}x.$$

E6.22 (a) Consider functions defined on a subset S of $\mathbb{R}^2$. How should the definition of "sup-norm of a function with respect to S" be defined?

(b) Suppose that (g_n) is a sequence of functions defined a compact subset S of $\mathbb{R}^2$ which converges uniformly on S to a function g. Prove that

$$\lim_{n\to\infty} \iint_S g_n(x, y)\, \mathrm{d}A = \iint_S g(x, y)\, \mathrm{d}A.$$

E6.23 If $f(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n^3}$, show that

$$\int_0^{\pi} f(x)\,\mathrm{d}x = 2\sum_{n=1}^{\infty} \frac{1}{(2n-1)^4}.$$

Justify each step in your proof by quoting an appropriate theorem.

E6.24 (a) Prove that the series

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

converges uniformly on the interval $[a, \infty)$ for any $a > 0$.

(b) Show that if we put

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{1+n^2x}$$

for $x > 0$, then f is a continuous function on $(0, \infty)$.

(c) Explain why the derivative f' of f is given by the formula

$$f'(x) = -\sum_{n=1}^{\infty} \frac{n^2}{(1+n^2x)^2}$$

for all $x > 0$ and deduce that f is strictly decreasing on $(0, \infty)$.

(d) Let $f_n(x) = \sum_{k=1}^{n} \frac{1}{1+k^2x}$ be the nth partial sum of f. Use the fact that $\lim_{x\to\infty} f_n(x) = 0$ for each fixed $n \in \mathbb{Z}_{>0}$ to deduce that $\lim_{x\to\infty} f(x) = 0$.

(e) If f_n is as in (d), show that $f_n(1/n^2) > n/2$ and deduce that $\lim_{x\to 0+} f(x) = \infty$.

(f) Use the information in parts (a)–(e) to make a rough sketch of the graph of f.

E6.25 Show that the series $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{x^2+n}$ is uniformly convergent on $\mathbb{R}$, but is not absolutely convergent for any $x \in \mathbb{R}$.

Hint: *Use the Cauchy criterion for uniform convergence.*

E6.26 For each $n \in \mathbb{Z}_{>0}$, let f_n be the function with domain $\mathbb{R}$ defined as follows:

$$f_n(x) = \begin{cases} x/n, & |x| \le n \\ 1, & |x| > n. \end{cases}$$

(a) Determine all the values of x for which the limit $\lim_{n\to\infty} f_n(x)$ exists.

(b) Determine the sup-norm $\|f_n\|$ of f_n on the interval $[-a, a]$ for each $n \in \mathbb{Z}_{>0}$ and each $a > 0$.

Hint: *Consider separately the two cases $n > a$ and $n \le a$.*

(c) For which numbers $a > 0$ does the sequence (f_n) converge uniformly on the interval $[-a, a]$?

(d) Does the sequence (f_n) converge uniformly on $\mathbb{R}$?

E6.27 Consider the functions $f_n(x) = nx^n(1-x)$ for $x \in [0,1]$ and $n \in \mathbb{Z}_{>0}$.

(a) Find $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in [0,1]$.

(b) Find the sup-norm $\|f_n - f\|$ of $f_n - f$ on $[0,1]$.

(c) Does $\lim_{n\to\infty} f_n = f$ uniformly $[0,1]$?

(d) Does $\displaystyle\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 f(x)\,\mathrm{d}x$?

E6.28 Decide whether each of the following statements is true or false and justify your answers with either a proof or an example.

(a) If $\lim_{n\to\infty} f_n = f$ uniformly on an interval I, then $\lim_{n\to\infty} f_n = f$ pointwise on that interval.

(b) If $\lim_{n\to\infty} f_n = f$ pointwise on an interval I and if each f_n, is of class C^2 on I, then $\lim_{n\to\infty} f_n = f$ uniformly on I.

(c) If $\lim_{n\to\infty} f_n = f$ pointwise on $[0,1]$ and if

$$\lim_{n\to\infty} \int_0^1 f_n(t)\,\mathrm{d}t = \int_0^1 f(t)\,\mathrm{d}t,$$

then $\lim_{n\to\infty} f_n = f$ uniformly on $[0,1]$.

E6.29 Consider the functions f_n defined as follows:

$$f_n(x) = \begin{cases} 1/n, & 0 \le x \le n^2, \\ 0, & x > n^2. \end{cases}$$

Show that $\lim_{n\to\infty} f_n = 0$ uniformly on $[0,\infty)$, that the improper integral

$$\int_0^\infty f_n(x)\,\mathrm{d}x$$

exists for each $n \ge 1$, but that $\displaystyle\lim_{n\to\infty} \int_0^\infty f_n(x)\,\mathrm{d}x = \infty$. This seems to contradict Theorem 6.9; explain why there is, in fact, no contradiction.

E6.30 When answering the following problems, state clearly either (1) how you use results from this course or (2) what assumptions that you had to make about things that *do not* follow from results in this course.

(a) Show that if (C_n) is a bounded sequence of real numbers then the series

$$u(x,t) = \sum_{n=1}^{\infty} C_n \mathrm{e}^{-n^2 t} \sin(nx)$$

is a solution of the boundary value problem

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}, \quad u(0,t) = u(\pi,t) = 0$$

for $t > 0$ and $x \in [0,\pi]$.

(b) Show that if, for $t = 0$, the series in (a) is uniformly convergent for $x \in [0, \pi]$, then $u(x, t)$ is continuous on the set

$$\{(x, t) \in \mathbb{R}^2 \mid 0 \le t < \infty \text{ and } 0 \le x \le \pi\}.$$

Remark: The partial differential equation $u_{xx} = u_t$ is the ***heat equation*** and can be derived from conservation of heat and Gauss's theorem. The function u is the expansion in a Fourier series of the solution of the heat equation.

Chapter 7

Power series and Taylor series

An important class of series of functions are the power series, such as arise in the theory of Taylor series. The theory of Taylor series, then, plays an important role in complex analysis.[1] In this chapter we shall consider the special properties of power series.

7.1 Definitions and examples

An infinite series of the form

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

where x is a real variable, c is a number, and (a_n) is a sequence of numbers, is called a ***power series***. Such series play important roles in numerous areas in analysis: they are used in approximation theory, they provide an important method of solving differential equations, and they are indispensable in the study of functions of a complex variable.

There are several questions about power series which must be dealt with before they can be understood. Firstly, for which values of x does a power series converge? And, secondly, on the set consisting of those x for which the series converges, is the convergence pointwise or uniform, is the power series a continuous function, and can we differentiate and integrate the series formally term-by-term? These questions will be answered systematically in the next two sections; the present section will simply give some examples to motivate the general discussion in those sections.

7.1 Examples: 1. Consider the power series $\sum_{n=1}^{\infty} \frac{x^n}{n}$. Here $c = 0$, $a_0 = 0$, and $a_n = 1/n$ for $n \geq 1$. Now

$$\lim_{n\to\infty} \left|\frac{x^{n+1}}{n+1}\right| \Big/ \left|\frac{x^n}{n}\right| = |x| \lim_{n\to\infty} \frac{n}{n+1} = |x|$$

and so, by the ratio test (Theorem 5.9), the power series converges for $|x| < 1$ and diverges for $x > 1$. As we shall see in the proof of Theorem 7.3 below, the

[1] While we consider here only real-valued functions of a real variable, the theory for complex-valued functions of a complex-variable is hardly any different.

series also diverges for $x < -1$, although this does not follow immediately from the ratio test as stated in Theorem 5.9 since the series is alternating. In addition, the series clearly diverges for $x = 1$ and converges for $x = -1$, since then it is just the alternating harmonic series.

For any $a \in (0,1)$, the series $\sum_{n=1}^{\infty} \frac{a^n}{n}$ converges and it follows from this, the Weierstrass M-test, and the inequality $\frac{|x|^n}{n} \leq \frac{a^n}{n}$, which is valid for $|x| \leq a$, that the power series in question converges uniformly on the interval $[-a, a]$. On the other hand, it is not hard to show that the power series does not converge uniformly on the interval $(-1, 1)$. In fact, if it did, then (taking $\epsilon = 1$ in the definition of uniform convergence) there would be an integer N with the property that $\sum_{n=N}^{\infty} \frac{x^n}{n} < 1$ for all $x \in (-1,1)$, and hence with the property that $\sum_{n=N}^{M} \frac{x^n}{n} < 1$ for all integers $M > N$. Now the series $\sum_{n=N}^{\infty} \frac{1}{n}$ diverges to ∞, and so there will be an integer M satisfying $M > N$ and $\sum_{n=N}^{M} \frac{1}{n} > 1$. Since this is a finite sum, it follows by continuity that there will be an $x \in (0,1)$ so close to 1 that $\sum_{n=N}^{M} \frac{x^n}{n} > 1$, and this is a contradiction.

2. Consider the power series $\sum_{n=0}^{\infty} \frac{n^2 x^{2n}}{2^n}$. This is a power series with $c = 0$ and with every second coefficient equal to zero or with every second term missing. We can use either the ratio or the root test to determine the x for which this series converges. In fact,

$$\lim_{n\to\infty} \left(\frac{(n+1)^2 x^{2(n+1)}}{2^{n+1}} \Big/ \frac{n^2 x^{2n}}{2^n} \right) = \lim_{n\to\infty} \left(\frac{n+1}{n} \right)^2 \frac{x^2}{2} = \frac{x^2}{2}$$

and

$$\lim_{n\to\infty} \left(\frac{n^2 x^{2n}}{2^n} \right)^{1/n} = \lim_{n\to\infty} \frac{n^{2/n} x^2}{2} = \frac{x^2}{2}$$

by Example 6.5–1, and both the ratio and the root tests assert that the series $\sum_{n=0}^{\infty} \frac{n^2 x^{2n}}{2^n}$ converges for $|x| < \sqrt{2}$ and diverges for $|x| > \sqrt{2}$. It is obvious that this power series diverges for $x = \pm\sqrt{2}$. Finally, the argument using the Weierstrass M-test given in 1 will work here, and leads to the conclusion that this series converging uniformly on the interval $[-a, a]$ for any $a \in [0, \sqrt{2})$. •

7.2 Convergence of power series

The two examples in the preceding section suggest that, for any power series of the form $\sum_{n=0}^{\infty} a_n x^n$, there should be a number R such that the series converges pointwise for $|x| < R$, converges uniformly on the interval $[-r, r]$ for each $r \in (0, R)$, and diverges for $|x| > R$. There is indeed such a number and (as the next theorem will show) it is given by the formula

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

with the understanding that $\frac{1}{0} = \infty$ and that $\frac{1}{\infty} = 0$. Because of these properties, R is known as the ***radius of convergence*** of the power series.

7.2 Example: For the power series considered in Example 7.1–1, we have

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} n^{-1/n} = 1$$

by Lemma 5.11 and so $R = 1$. For the power series considered in Example 7.1–2, we have

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} \left(\frac{n^2}{2^n}\right)^{1/(2n)} = \frac{1}{\sqrt{2}}$$

and so $R = \sqrt{2}$. •

7.3 Theorem: *Consider the power series*

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

and the quantity $R = 1 \Big/ \left(\limsup_{n\to\infty} \sqrt[n]{|a_n|}\right)$, *where* $1/0 = \infty$ *and* $1/\infty = 0$.

(i) If $R = 0$*, then the power series converges only for* $x = c$*.*

(ii) If $0 < R < \infty$*, then the power series converges absolutely for any* x *satisfying* $|x-c| < R$*, diverges for any* x *satisfying* $|x-c| > R$*, and converges uniformly on* $[c-r, c+r]$ *for any* $r \in (0, R)$*.*

(iii) If $R = \infty$*, then the power series converges absolutely for any* x *and uniformly on any interval of finite length.*

Proof: Suppose first that $R = 0$. Then $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \infty$ and so some subsequence of the sequence $(\sqrt[n]{|a_n|})$ must diverge to ∞. So, if $x \neq c$, then some subsequence of the sequence $(\sqrt[n]{|a_n|}|x-c|)$ will diverge to ∞, and hence some subsequence of the sequence $(|a_n(x-c)|^n)$ will also diverge to ∞. But this implies that $\lim_{n\to\infty} a_n(x-c)^n \neq 0$, and so it follows from Theorem 5.3 that the power series does not converge. This proves (i).

Now suppose that $R > 0$ and consider a fixed number x and the quantity

$$L = \limsup_{n\to\infty} \sqrt[n]{|a_n(x-c)^n|}.$$

Then

$$L = |x-c| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x-c|}{R}$$

and so, if $|x-c| < R$, then $L < 1$ and the series converges absolutely by the root test (Theorem 5.10). On the other hand, if $|x-c| > R$, then $L > 1$, and we conclude divergence of the series by the general version of the root test (Theorem 5.18).

Now consider a number $r \in (0, R)$. Then $|a_n(x-c)^n| \leq |a_n|r^n$ for all $x \in [c-r, c+r]$ and all n, and the series $\sum_{n=0}^{\infty}|a_n r^n|$ converges by the first part of the proof, and, therefore, the power series converges uniformly on the interval $[c-r, c+r]$ by the Weierstrass M-test (Theorem 6.13). This completes the proof of the theorem. ■

Let us consider a few interesting facts about radii of convergence of power series. To do this, let us begin by proving a generalization of Lemma 5.11.

7.4 Lemma: *If P is a polynomial function, then*

$$\lim_{n\to\infty} \sqrt[n]{|P(n)|} = 1.$$

Proof: Let us write

$$P(n) = p_k n^k + p_{k-1}n^{k-1} + \cdots + p_1 n + p_0$$

for $p_0, p_1, \dots, p_k \in \mathbb{R}$ with $p_k \neq 0$. We then have

$$\begin{aligned}\lim_{n\to\infty} \sqrt[n]{|P(n)|} &= \lim_{n\to\infty} \sqrt[n]{|p_k n^k + p_{n-1}n^{k-1} + \cdots + p_1 n + p_0|} \\ &= \lim_{n\to\infty} \sqrt[n]{|p_k|}\,\sqrt[n]{n^k}\,\sqrt[n]{\left|1 + \frac{p_{n-1}}{p_k}\frac{1}{n} + \cdots + \frac{p_1}{p_k}\frac{1}{n^{k-1}} + \frac{p_0}{p_k}\frac{1}{n^k}\right|}.\end{aligned}$$

We have

$$\begin{aligned}\lim_{n\to\infty} \sqrt[n]{|p_k|} &= \lim_{n\to\infty} \exp\left(\frac{\ln(|p_k|)}{n}\right) \\ &= \exp\left(\lim_{n\to\infty} \frac{\ln(|p_k|)}{n}\right) = 1.\end{aligned}$$

By Lemma 5.11 we have

$$\lim_{n\to\infty} \sqrt[n]{n^k} = \lim_{n\to\infty} \left(\sqrt[n]{n}\right)^k = \left(\lim_{n\to\infty} \sqrt[n]{n}\right)^k = 1.$$

By combining the preceding two calculations with this one,

$$\begin{aligned}\lim_{n\to\infty} &\sqrt[n]{\left|1 + \frac{p_{n-1}}{p_k}\frac{1}{n} + \cdots + \frac{p_1}{p_k}\frac{1}{n^{k-1}} + \frac{p_0}{p_k}\frac{1}{n^k}\right|} \\ &= \lim_{n\to\infty} \exp\left(\frac{\ln(1 + \frac{p_{n-1}}{p_k}\frac{1}{n} + \cdots + \frac{p_1}{p_k}\frac{1}{n^{k-1}} + \frac{p_0}{p_k}\frac{1}{n^k})}{n}\right) \\ &= \exp\left(\lim_{n\to\infty} \frac{\ln(1 + \frac{p_{n-1}}{p_k}\frac{1}{n} + \cdots + \frac{p_1}{p_k}\frac{1}{n^{k-1}} + \frac{p_0}{p_k}\frac{1}{n^k})}{n}\right) = 1,\end{aligned}$$

one arrives at the statement of the lemma. ■

With the lemma at hand, one easily proves the following useful result.

7.5 Proposition: *Let the power series*

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

have radius of convergence R. Then the following statements hold:

(i) if P is a polynomial function, then the power series

$$\sum_{n=0}^{\infty} P(n)a_n(x-c)^n$$

has radius of convergence R;

(ii) if Q is a polynomial function for which $Q(n) \neq 0$ for every $n \in \mathbb{Z}_{\geq 0}$, then the power series

$$\sum_{n=0}^{\infty} \frac{a_n}{Q(n)}(x-c)^n$$

has radius of convergence R;

(iii) if $b \in \mathbb{R}$, then the power series

$$\sum_{n=0}^{\infty} a_n b^n (x-c)^n$$

has radius of convergence $\dfrac{R}{|b|}$.

Proof: (i) and (ii) We shall prove that the power series

$$\sum_{n=0}^{\infty} \frac{P(n)}{Q(n)}(x-c)^n$$

has radius of convergence R. This, however, is straightforward, given Lemma 7.4. Indeed, we have

$$\begin{aligned}\limsup_{n\to\infty} \sqrt[n]{\left|\frac{P(n)a_n}{Q(n)}\right|} &= \frac{\lim_{n\to\infty} \sqrt[n]{|P(n)|} \limsup_{n\to\infty} \sqrt[n]{|a_n|}}{\lim_{n\to\infty} \sqrt[n]{|Q(n)|}} \\ &= \limsup_{n\to\infty} \sqrt[n]{|a_n|},\end{aligned}$$

from which the conclusion follows immediately.

(iii) In this case, we directly compute

$$\limsup_{n\to\infty} \sqrt[n]{|a_n b^n|} = |b| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|b|}{R}.$$

Thus the radius of convergence of the power series

$$\sum_{n=0}^{\infty} a_n b^n (x-c)^n$$

is $\dfrac{R}{|b|}$, as stated. ■

7.3 Properties of power series

A power series with a nonzero radius of convergence defines a function on an open interval or on the whole real line. What are the properties of this function? For instance, is it continuous or differentiable, and (assuming that they exist) what are its integral and its derivative?

7.6 Theorem: *A power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ whose radius of convergence R is positive is a continuous function on $(c-R, c+R)$.*

Proof: Consider a point $a \in (c-R, c+R)$ and let r be a number satisfying $|a-c| < r < R$. For each integer n, the partial sum $\sum_{k=0}^{n} a_k(x-c)^k$ is a polynomial, and hence a continuous function and (by Theorem 7.3(iii)) these partial sums converge uniformly to $\sum_{n=0}^{\infty} a_n(x-c)^n$ on the interval $[c-r, c+r]$. It thus follows from Theorem 6.8 that $\sum_{n=0}^{\infty} a_n(x-c)^n$ is a continuous function on the interval $[c-r, c+r]$, and hence a continuous function at a. Since this is true for each $a \in (c-R, c+R)$, it means that $\sum_{n=0}^{\infty} a_n(x-c)^n$ is continuous at each point of $(c-R, c+R)$, and hence on $(c-R, c+R)$. ■

The formal term-by-term derivative or integral of a power series are themselves power series. Are they actually the derivative and integral of the power series itself? If this were the case it would make the calculus of power series extremely easy. The first step in investigating this question is to determine the radius of convergence of these two associated power series.

7.7 Lemma: *For any sequence (a_n) and any $c \in \mathbb{R}$, the three power series*

$$\sum_{n=0}^{\infty} a_n(x-c)^n, \quad \sum_{n=1}^{\infty} n a_n(x-c)^{n-1}, \quad \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-c)^{n+1}$$

have the same radius of convergence.

Proof: For perfect transparency, let us make a change of index for the second and third of these series:

$$\sum_{n=1}^{\infty} n a_n(x-c)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1}(x-c)^n,$$

$$\sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-c)^{n+1} = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n}(x-c)^n.$$

We next calculate

$$\begin{aligned}
\limsup_{n\to\infty} |a_{n-1}|^{n/(n-1)} &= \limsup_{n\to\infty} \exp\left(\frac{n}{n-1}\ln(|a_{n-1}|)\right) \\
&= \exp\left(\limsup_{n\to\infty} \frac{n}{n+1}\ln(|a_{n-1}|)\right) \\
&= \exp\left(\limsup_{n\to\infty} \ln(|a_{n-1}|)\right) \\
&= \limsup_{n\to\infty} \exp(\ln(|a_{n-1}|)) = \limsup_{n\to\infty} |a_{n-1}|.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\limsup_{n\to\infty} \sqrt[n]{|a_{n-1}|} &= \left(\limsup_{n\to\infty} |a_{n-1}|\right)^{1/n} \\
&= \left(\limsup_{n\to\infty} |a_{n-1}|^{n/(n-1)}\right) \\
&= \limsup_{n\to\infty} (|a_{n-1}|^{n/(n-1)})^{1/n} \\
&= \limsup_{n\to\infty} |a_{n-1}|^{1/(n-1)} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}.
\end{aligned}$$

In similar manner,

$$\limsup_{n\to\infty} \sqrt[n]{|a_{n+1}|} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

The result now follows from Proposition 7.5. ■

Note that the second and third power series in the preceding lemma are obtained from the first one by formal term-by-term differentiation and integration, respectively.

7.8 Theorem: *If $\sum_{n=0}^{\infty} a_n(x-c)^n$ is a power series whose radius of convergence R is positive, then*

$$\int_c^x \left(\sum_{n=0}^{\infty} a_n(t-c)^n\right) \mathrm{d}t = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-c)^{n+1}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=0}^{\infty} a_n(x-c)^n\right) = \sum_{n=1}^{\infty} n a_n (x-c)^{n-1}$$

for all $x \in (c-R, c+R)$.

Proof: Let $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ and $P_n(x) = \sum_{k=0}^{n} a_k(x-c)^k$ for $n \in \mathbb{Z}_{>0}$ and $x \in (c-R, c+R)$. To prove the first formula, consider an $x \in (c-R, c+R)$ and let r be a number satisfying $|x-c| < r < R$. Then $\lim_{n\to\infty} P_n = f$ uniformly on

$[c-r, c+r]$ by Theorem 7.3, and hence uniformly on the interval with endpoints c and x, regardless of whether $x > c$ or $x < c$, and, therefore,

$$\begin{aligned}\int_c^x f(t)\,\mathrm{d}t &= \int_c^x \lim_{n\to\infty} P_n(t)\,\mathrm{d}t \\ &= \lim_{n\to\infty} \int_c^x P_n(t)\,\mathrm{d}t \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{a_k}{k+1} a_k (x-c)^{k+1} \\ &= \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-c)^{n+1}.\end{aligned}$$

This proves the first formula.

To prove the second formula, let $g(x) = \sum_{n=1}^{\infty} n a_n (x-c)^{n-1}$ and recall from Lemma 7.7 that this power series has a radius of convergence of R. So, if r is a number satisfying $0 < r < R$, then the power series $\sum_{n=1}^{\infty} n a_n (x-c)^{n-1}$ converges uniformly on $[c-r, c+r]$. But this just means that $\lim_{n\to\infty} P_n' = g$ uniformly on $[c-r, c+r]$ and, therefore, Theorem 6.11 implies that $f' = g$ on $[c-r, c+r]$. Finally, since r was an arbitrary number in $(0, R)$, it follows that $f' = g$ on $(c-R, c+R)$. ■

This theorem has two corollaries but it will be useful to consider several examples before going on to these corollaries.

7.9 Examples: 1. From our knowledge of the geometric series, we know that

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots$$

with a radius of convergence of 1. The above theorem on differentiation then gives

$$\begin{aligned}\frac{1}{(1-x)^2} &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(1 + x + x^2 + x^3 + \cdots\right) \\ &= 1 + 2x + 3x^2 + \cdots\end{aligned}$$

and thus

$$\frac{1}{(1-x)^2} = \sum_{n=0}^{\infty}(n+1)x^n. \tag{7.1}$$

On the other hand, since $\int_0^x \frac{\mathrm{d}t}{1-t} = -\ln(1-x)$ for $|x| < 1$, the above theorem on integration gives

$$\begin{aligned}\ln(1-x) &= -\int_0^x (1 + t + t^2 + \cdots)\,\mathrm{d}t \\ &= -\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots\right),\end{aligned}$$

or

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{n} x^n. \tag{7.2}$$

The two power series (7.1) and (7.2) both have radius of convergence of 1.
We can combine the above power series for $\ln(1+x)$ and $\ln(1-x)$ to obtain

$$\begin{aligned}
\ln\left(\frac{1+x}{1-x}\right) &= \ln(1+x) - \ln(1-x) \\
&= \left(x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots\right) \\
&\quad + \left(x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots\right) \\
&= 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots\right) \\
&= 2\sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}.
\end{aligned}$$

Letting, say, $x = 1/3$ in this last expression gives

$$\ln 2 = \ln\left(\frac{1+1/3}{1-1/3}\right) = 2\left(\frac{1}{3} + \frac{1}{81} + \frac{1}{1215} + \cdots\right).$$

2. Substituting $-x^2$ in place of x in the geometric series

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots$$

gives the power series

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots ;$$

this is valid for $x^2 < 1$, and so the radius of convergence of this last series is 1. Since $\int_0^x \frac{\mathrm{d}t}{1+t^2} = \arctan x$, we can integrate this last power series to obtain

$$\arctan x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots ,$$

and this power series too has a radius of convergence of 1. In particular, letting $x = 1/\sqrt{3}$ in this series gives

$$\frac{\pi}{6} = \frac{1}{\sqrt{3}}\left(1 - \frac{1}{3\cdot 3} + \frac{1}{5\cdot 3^2} - \frac{1}{7\cdot 3^3} + \cdots\right). \qquad \bullet$$

The theorem on differentiation and integration of power series has the following two consequences.

7.10 Corollary: *A power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ whose radius of convergence R is positive is a C^∞-function on the interval interval$(c-R, c+R)$ and*

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\sum_{n=0}^{\infty} a_n(x-c)^n\right)\Bigg|_{x=c} = k!a_k$$

for all $k \in \mathbb{Z}_{>0}$.

Proof: Applying the previous theorem and lemma k times gives

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\sum_{n=0}^{\infty} a_n(x-c)^n\right) = k!a_k + k!a_{k+1}(x-c) + \cdots$$

and the corollary is now clear. ■

It is obvious that the coefficients of power series determine the power series as a function, but what about the converse? That is to say, if two power series with the same centers represent the same function, must their corresponding coefficients be equal? The next result asserts that this is indeed the case.

7.11 Corollary: *If $\sum_{n=0}^{\infty} a_n(x-c)^n$ and $\sum_{n=0}^{\infty} b_n(x-c)^n$ are two power series with positive radii of convergence R and S, respectively, and if there is a number r such that $0 < r \leq \min\{R, S\}$ and*

$$\sum_{n=0}^{\infty} a_n(x-c)^n = \sum_{n=0}^{\infty} b_n(x-c)^n$$

for all $x \in (c-r, c+r)$, then $a_n = b_n$ for all $n \in \mathbb{Z}_{>0}$.

Proof: This corollary is an immediate consequence of the preceding one. ■

7.4 Taylor series and Taylor polynomials

Consider a power series $\sum_{n=0}^{\infty} a_n(x-c)^n$ whose radius of convergence R is positive, and put $f(x) = \sum_{n=0}^{\infty} a_n(x-c)^n$ for $x \in (c-R, c+R)$. Then, from the preceding two sections, we know that

1. f is a C^∞-function on the interval $(c-R, c+R)$ (by Corollary 7.10),
2. $a_n = \frac{f^{(n)}(c)}{n!}$ for $n \in \mathbb{Z}_{\geq 0}$ (also by Corollary 7.10), and
3. if $P_n(x) = \sum_{k=0}^{n} a_k(x-c)^k$ for $n \in \mathbb{Z}_{>0}$ and $x \in (c-R, c+R)$, then $\lim_{n\to\infty} P_n = f$ pointwise on $(c-R, c+R)$ and uniformly on $[c-r, c+r]$ for each $r \in (0, R)$ (by Theorem 7.3).

This section is concerned with what one might think of as a converse of 1, namely, is every C^∞-function a power series?

To describe this converse more precisely, suppose that f is a C^∞-function defined on some open interval I and that c is a point in I. We can then form the power series

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

and ask whether it converges on some interval of positive radius and, if so, whether it converges to f on that interval. This power series is called the ***Taylor series*** of f about c. Now the question of the convergence of this power series is that of the convergence of its sequence of partial sums, and this suggests that we consider its partial sums, i.e., the polynomials

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k \tag{7.3}$$

for $n \in \mathbb{Z}_{>0}$. The polynomial (7.3) is called the nth ***Taylor polynomial*** of f at c and has the property that $f^{(k)}(c) = P_n^{(k)}(c)$ for $k \in \mathbb{Z}_{\geq 0}$. This property suggests that P_n should be a good approximation to f near c since the two functions f and P_n have the same value, the same slope (and hence the same tangent line), the same concavity, etc., at c. In fact, if we let $R_n = f - P_n$, then $f = P_n + R_n$, and R_n is just the error when we regard P_n as an approximation to f. The question of how good an approximation P_n is to f, therefore, becomes the question of how small R_n is and, in particular, the question of whether $\lim_{n\to\infty} P_n(x) = f(x)$ for a particular value of x is the question of whether $\lim_{n\to\infty} R_n(x) = 0$. To answer this question we evidently need an expression for R_n and the next theorem contains two such expressions. Notice that this theorem does not require that f be a C^∞-function, but merely of class C^{n+1} on I.

7.12 Theorem: *Suppose that f is a function defined on an open interval I, that c is a point in I, and that n is a positive integer. If $f^{(n+1)}$ exists and is continuous on I and if R_n is as above, then*

$$R_n(x) = \frac{1}{n!}\int_c^x (x-u)^n f^{(n+1)}(u)\,\mathrm{d}u \tag{7.4}$$

and there is a point $\xi \in I$ such that

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}. \tag{7.5}$$

Proof: The best way to prove this theorem is probably to first prove a lemma and then deduce the theorem from the lemma. The reason for is two-fold: the lemma is of independent interest and doing it this way will make the proof easier to comprehend. Part (ii) of the lemma is a special case of the generalized mean value theorem for integrals (cf. Exercise E7.35).

1 Lemma: *Let g be a continuously differentiable function defined on an open interval I and let c be a point in I.*

(i) If g' exists and is continuous on I, then

$$\int_c^x (x-u)^j g(u)\,\mathrm{d}u = \frac{g(c)}{j+1}(x-c)^{j+1} + \frac{1}{j+1}\int_c^x (x-u)^{j+1} g'(u)\,\mathrm{d}u$$

for all $j \in \mathbb{Z}_{\geq 0}$ and all $x \in I$.

(ii) Corresponding to each $x \in I \setminus \{c\}$ and each $n \in \mathbb{Z}_{>0}$, there is a point $\xi \in I$ satisfying $|\xi - c| \leq |x-c|$ and

$$g(\xi) = \frac{n+1}{(x-c)^{n+1}}\int_c^x (x-u)^n g(u)\,\mathrm{d}u.$$

Proof: (i) This follows easily from the formula for integration by parts:

$$\begin{aligned}\int_c^x (x-u)^j g(u)\,\mathrm{d}u &= \frac{-1}{j+1}\int_c^x \left(\frac{\mathrm{d}}{\mathrm{d}x}(x-u)^{j+1}\right) g(u)\,\mathrm{d}u \\ &= \frac{-1}{j+1}\left((x-u)^{j+1}g(u)\Big|_c^x \right. \\ &\quad \left. - \int_c^x (x-u)^{j+1} g'(u)\,\mathrm{d}u\right) \\ &= \frac{g(c)}{j+1}(x-c)^{j+1} + \frac{1}{j+1}\int_c^x (x-u)^{j+1} g'(u)\,\mathrm{d}u.\end{aligned}$$

(ii) Fix a point $x \in I \setminus \{c\}$ and assume that $x > c$ (the proof in the case $x < c$ is similar). If

$$m = \inf\{g(u) \mid c \leq u \leq x\}$$

and

$$M = \sup\{g(u) \mid c \leq u \leq x\},$$

then it is certainly true that $m \leq g(u) \leq M$ for all $u \in [c, x]$. Since $x - u \geq 0$ for all $u \in [c, x]$, it follows that

$$m(x-u)^n \leq (x-u)^n g(u) \leq M(x-u)^n$$

for all $u \in [c, x]$. Integrating this inequality gives

$$m\int_c^x (x-u)^n\,\mathrm{d}u \leq \int_c^x (x-u)^n g(u)\,\mathrm{d}u \leq M\int_c^x (x-u)^n\,\mathrm{d}u,$$

hence

$$\frac{m}{n+1}(x-c)^{n+1} \leq \int_c^x (x-u)^n g(u)\,\mathrm{d}u \leq \frac{M}{n+1}(x-c)^{n+1},$$

and, therefore,

$$m \le \frac{n+1}{(x-c)^{n+1}} \int_c^x (x-u)^n g(u)\,\mathrm{d}u \le M. \tag{7.6}$$

Now let R denote the range of the restriction of g to the closed interval $[c, x]$. Since g is continuous on $[c, x]$ it follows from Theorem 4.24(i) that $m, M \in R$, and then (7.6) and the intermediate value theorem (see Exercise E4.29) imply that

$$\frac{n+1}{(x-c)^{n+1}} \int_c^x (x-u)^n g(u)\,\mathrm{d}u$$

also belongs to R. But, if this number belongs to R, then, by the definition of R, it must be equal to $g(\xi)$ for some point $\xi \in [c, x]$. ▼

Now we proceed with the proof of the theorem. The proof of (7.4) consists of repeated applications of Lemma 1(i). First of all,

$$\begin{aligned} f(x) - f(c) &= \int_c^x (x-u)^0 f'(u)\,\mathrm{d}u \\ &= (x-c)f'(c) + \int_c^x (x-u)^1 f''(u)\,\mathrm{d}u, \end{aligned}$$

by the fundamental theorem of calculus and Lemma 1 and, therefore,

$$f(x) = f(c) + f'(c)(x-c) + \int_c^x (x-u)^1 f''(u)\,\mathrm{d}u.$$

Since

$$f(x) = P_1(x) + R_1(x), \quad P_1(x) = f(c) + f'(c)(x-c)$$

by the definition of P_1 and R_1, this verifies (7.4) for $n = 1$. Next,

$$\int_c^x (x-u)^1 f''(u)\,\mathrm{d}u = \frac{f''(c)}{2}(x-c)^2 + \frac{1}{2}\int_c^x (x-u)^2 f'''(u)\,\mathrm{d}u$$

by Lemma 1, and thus

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \int_c^x (x-u)^2 f'''(u)\,\mathrm{d}u.$$

The first three terms on the right side are just $P_2(x)$, and so the fourth term must be $R_2(x)$ and this verifies (7.4) for $n = 2$. Let's play this game one more time:

$$\int_c^x (x-u)^2 f'''(x)\,\mathrm{d}u = \frac{f'''(c)}{3} + \frac{1}{3}\int_c^x (x-u)^3 f^{(4)}\,\mathrm{d}u$$

and hence

$$\begin{aligned} f(x) &= f(c) + f'(c)(x-c) + \frac{f''(c)}{2}(x-c)^2 + \frac{f'''(c)}{3!}(x-c)^3 \\ &\quad + \frac{1}{3!}\int_c^x (x-u)^3 f^{(4)}(x)\,\mathrm{d}u \\ &= P_3(x) + \frac{1}{3!}\int_c^x (x-u)^3 f^{(4)}(x)\,\mathrm{d}u. \end{aligned}$$

Just as before, this verifies (7.4) for $n = 3$. Continuing in this manner (or, better still, starting over and proceeding by induction) will lead to (7.4) for a general value of n. In doing so, the assumption that $f^{(n+1)}$ is continuous on I is sufficient to justify the last application of integration by parts in this calculation.

Next we verify (7.5). According to Lemma 1(ii), there is a point $\xi \in I$ such that

$$f^{(n+1)}(\xi) = \frac{n+1}{(x-c)^{n+1}} \int_c^x (x-u)^n f^{(n+1)}\,\mathrm{d}u$$

and, therefore,

$$R_n(x) = \frac{1}{n!}\frac{(x-c)^{n+1}}{n+1} f^{(n+1)}(\xi) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1}$$

by (7.4). ■

We can use either (7.4) or (7.5) in trying to decide whether the Taylor series of a given function converges to that function.

7.13 Examples: 1. The derivatives of the sine function at the origin are easy to calculate and the Taylor series of the sine function at 0 is

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .$$

Now consider a number $x \neq 0$ and an integer n. Then, by (7.5), there is a number ξ such that

$$R_n(x) = \frac{\sin^{(n+1)}(\xi)}{(n+1)!}(x-0)^{n+1}$$

and, since all of the derivatives of sin are $\pm\sin$ or $\pm\cos$, it follows that

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Now the power series $\sum_{n=0}^{\infty} \frac{|x|^n}{n!}$ has a radius of convergence equal to ∞ (this is an easy consequence of the ratio test) and, therefore, $\lim_{n\to\infty} \frac{|x|^n}{n!} = 0$ by Theorem 5.3. This means that $\lim_{n\to\infty} R_n(x) = 0$ and thus $\lim_{n\to\infty} P_n(x) = \sin x$ for each number x.

The sine function is therefore equal to its Taylor series for all real numbers and (from our knowledge of power series) it follows that

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

uniformly on intervals of finite length.

2. An argument virtually identical to that of the preceding example will show that the Taylor series of the cosine function at the origin converges pointwise on $\mathbb{R}$ and uniformly on intervals of finite length to the cosine function. Thus

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

uniformly on intervals of finite length.

3. Since the derivative of the exponential function e^x is just e^x, it follows that the Taylor series of this function is

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots .$$

It is easy to verify (using the ratio test) that this power series has a radius of convergence equal to ∞. To show that it actually converges to e^x, it is sufficient to prove that $\lim_{n\to\infty} R_n(x) = 0$ for each number x.

Now, for each number x and positive integer n, there is a number ξ such that

$$R_n(x) = \frac{1}{(n+1)!} \left.\frac{\mathrm{d}^{n+1}\mathrm{e}^x}{\mathrm{d}x^{n+1}}\right|_{x=\xi} x^{n+1} = \frac{x^{n+1}\mathrm{e}^{\xi}}{(n+1)!}$$

and, therefore,

$$0 \le R_n(x) \le \frac{|x|^{n+1}\mathrm{e}^{|x|}}{(n+1)!}.$$

We saw in 1 that $\lim_{n\to\infty} \frac{|x|^n}{n!} = 0$, hence $\lim_{n\to\infty} R_n(x) = 0$, and, therefore,

$$\mathrm{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

for all real numbers x.

4. We want to identify the Taylor series of the function $f(x) = (1+x)^r$ at 0, where r is a real number, and show that it converges to f on the interval $(-1,1)$. Now, if r is a positive integer, then f is just a polynomial (as can be seen by expanding $(1+x)^r$ by the binomial formula) and there is nothing to prove, so we may as well assume that r is neither 0 nor a positive integer. It is easy to see that

$$f^{(k)}(x) = r(r-1)\cdots(r-k+1)(1+x)^{r-k},$$

and hence the coefficient of x^k in the Taylor series of f at 0 is

$$\frac{f^{(k)}(0)}{k!} = \frac{r(r-1)\cdots(r-k+1)}{k!}.$$

Since this expression is reminiscent of a binomial coefficient, it will be convenient to denote it by $\binom{r}{k}$, with the understanding that $\binom{r}{0} = 1$. Notice that $\binom{r}{k} \neq 0$ for all $k \in \mathbb{Z}_{>0}$ since r is not a nonnegative integer. So the Taylor series of f at 0 is

$$\sum_{n=0}^{\infty} \binom{r}{n} x^n = 1 + rx + \frac{r(r-1)}{2!}x^2 + \frac{r(r-1)(r-2)}{3!}x^3 + \cdots.$$

The radius of convergence of this power series is best calculated by the ratio test:

$$\begin{aligned} \lim_{n\to\infty} \left|\binom{r}{n+1}x^{n+1}\right| \Big/ \left|\binom{r}{n}x^n\right| &= |x| \lim_{n\to\infty} \frac{|r-k|}{(k+1)!} \\ &= |x|, \end{aligned}$$

and so the radius of convergence is 1.

For this particular function f, it is very difficult to estimate the remainder R_n as was done in the previous examples, so the proof of the convergence of the Taylor series for f to f will be accomplished by another method. If we let g denote the Taylor series of f at 0, i.e., if

$$g(x) = \sum_{n=0}^{\infty} \binom{r}{n} x^n$$

for $|x| < 1$, then we know that g is a C^∞-function defined on the interval $(-1, 1)$. To show that the Taylor series of f at 0 converges to f, it is sufficient to show that $f(x) = g(x)$ for $|x| < 1$, and for this it is, in turn, sufficient to show that

$$f(0) = g(0), \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{g(x)}{f(x)}\right) = 0, \qquad x \in (-1, 1). \tag{7.7}$$

Indeed, the second of these equations implies that $f(x)/g(x)$ is constant, and then the first equation implies that this constant is 1.

Now, it is obvious from the definitions of f and g that $f(0) = 1 = g(0)$ and the quotient rule gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{g(x)}{f(x)}\right) &= \frac{g'(x)f(x) - g(x)f'(x)}{f^2(x)} \\ &= \frac{(1+x)^r g'(x) - r(1+x)^{r-1}g(x)}{(1+x)^{2r}} \\ &= \frac{(1+x)g'(x) - rg(x)}{(1+x)^{r+1}}. \end{aligned}$$

To complete this example, it will be sufficient to show that $(1+x)g'(x) = rg(x)$ for $|x| < 1$. Since g is a power series, its derivative g' can be found by term-by-term differentiation (by Theorem 7.8), and hence

$$\begin{aligned}
(1+x)g'(x) &= (1+x)\sum_{n=1}^{\infty} n\binom{r}{n}x^{n-1} \\
&= \sum_{n=1}^{\infty} n\binom{r}{n}x^{n-1} + \sum_{n=1}^{\infty} n\binom{r}{n}x^{n} \\
&= \binom{r}{1} + \sum_{n=1}^{\infty}(n+1)\binom{r}{n+1}x^{n} + \sum_{n=1}^{\infty} n\binom{r}{n}x^{n} \\
&= r + \sum_{n=1}^{\infty}\left[(n+1)\binom{r}{n+1} + n\binom{r}{n}\right]x^{n}.
\end{aligned}$$

The expression in the square brackets can be simplified as follows:

$$\begin{aligned}
(n+1)\binom{r}{n+1} + n\binom{r}{n} &= (n+1)\frac{r(r-1)\cdots(r-n)}{(n+1)!} + \\
&\qquad n\frac{r(r-1)\cdots(r-n+1)}{n!} \\
&= r(r-1)\cdots(r-n+1)\left(\frac{r-n}{n!} + \frac{n}{n!}\right) \\
&= r\binom{r}{n},
\end{aligned}$$

and, therefore,

$$(1+x)g'(x) = r + \sum_{n=1}^{\infty} r\binom{r}{n}x^{n} = rg(x).$$

5. For two numbers v and c satisfying $0 < v < c$, we have

$$\begin{aligned}
\frac{1}{\sqrt{1-v^2/c^2}} &= \left(1-\frac{v^2}{c^2}\right)^{-1/2} \\
&= 1 + \left(-\frac{1}{2}\right)\left(-\frac{v^2}{c^2}\right) + \frac{\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)}{2!}\left(-\frac{v^2}{c^2}\right)^2 + \cdots \\
&= 1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \cdots
\end{aligned}$$

by the previous example. This expression leads to the approximation

$$\frac{1}{\sqrt{1-v^2/c^2}} \approx 1 + \frac{1}{2}\frac{v^2}{c^2},$$

which is frequently used in relativity theory when $\frac{v}{c} \ll 1$.

6. Although it is true for all of the common functions, it is not always the case that the Taylor series of a C^∞-function converges to that function. The simplest function for which this does not occur is probably the one in the following example. The function f in this example satisfies $f(0) = 0$, $0 \leq f(x) \leq 1$ for all x, is asymptotic to 1 in both directions, and is incredibly flat at the origin. (Try using plotting the graph of this function using some software package.)

   Consider the function f defined by

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

   Then $f'(x) = 2x^{-3}e^{-1/x^2}$ for $x \neq 0$ by the usual formulae for differentiation, whereas $f'(0)$ must be calculated directly from the definition of the derivative and l'Hôspital's rule:

$$\begin{aligned} f'(0) &= \lim_{h\to 0} \frac{f(h) - f(0)}{h} = \lim_{h\to 0} \frac{e^{-1/h^2}}{h} \\ &= \lim_{h\to 0} \frac{1/h}{e^{1/h^2}} = \lim_{k\to\pm\infty} \frac{k}{e^{k^2}} \\ &= \lim_{k\to\pm\infty} \frac{1}{2ke^{k^2}} = 0. \end{aligned}$$

   A similar calculation will show that

$$f''(x) = \begin{cases} p_2(1/x)e^{-1/x^2}, & x \neq 0 \\ 0, & x = 0, \end{cases}$$

   where p_2 is a polynomial of degree 6, and an inductive argument will even show that

$$f^{(k)}(x) = \begin{cases} p_k(1/x)e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

   for $k \in \mathbb{Z}_{>0}$, where p_k is a polynomial of degree $3k$. So the derivatives $f^{(k)}$ of f exists on $\mathbb{R}$ for all $k \in \mathbb{Z}_{>0}$ and f is C^∞, yet the Taylor series of f about 0 is the zero function and does not converge to or equal f except at 0 itself. •

Exercises

E7.1 Consider a power series $\sum_{n=0}^{\infty} a_n(x-c)^n$. Show that, if the sequence $(|\frac{a_n}{a_{n+1}}|)$ converges or diverges to ∞, then the radius of convergence of the power series is

$$R = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|.$$

E7.2 Find the radius of convergence of the power series $\sum a_n x^n$, where a_n is as given below, and determine whether the series converges absolutely, converges conditionally, or diverges at the endpoints of the interval of convergence.

(a) $a_n = 2^{-n}n^4(n^5+1)$

(b) $a_n = \dfrac{(n!)^2}{(2n)!}$

(c) $a_n = (3^n+4)^n 5^{-n}$

E7.3 Determine the radius of convergence, R, of the power series $\sum_{n=1}^{\infty} a_n x^n$, where a_n is as given below. Also, determine whether the series converges absolutely, converges conditionally, or diverges at the points $x = \pm R$.

(a) $a_n = 10^{-n} + 10^n$

(b) $a_n = 1/\binom{2n}{n} = \frac{(n!)^2}{(2n)!}$.

E7.4 Determine the radius of convergence of each of the following power series:

(a) $\sum_{n=1}^{\infty} \frac{n^{10}}{n!} x^n$

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

E7.5 Determine the interval of convergence of each of the following series. (Do not investigate convergence or divergence at the endpoints.)

(a) $\displaystyle\sum_{n=1}^{\infty}\left(\frac{a^n}{n}+\frac{b^n}{n^2}\right)(x-1)^{2n}$, where a and b are two positive numbers

(b) $\displaystyle\sum_{n=1}^{\infty}\frac{(2x+3)^n}{n}$

(c) $\displaystyle\sum_{n=1}^{\infty}\frac{\mathrm{e}^n}{n^3+1}\left(\frac{x}{2}-1\right)^n$

E7.6 Determine the interval of convergence of the power series

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

E7.7 Determine the interval of convergence of and an explicit formula for the power series

$$f(x) = 1 + 2x + x^2 + 2x^3 + x^4 + \cdots.$$

E7.8 Prove that the power series

$$f(x) = \sum_{n=1}^{\infty} \frac{x^n}{(n+1)(n+2)}$$

converges uniformly on the closed interval $[-1, 1]$.

E7.9 Determine all of the numbers x for which the power series $\sum_{n=0}^{\infty} c_n (x-1)^n$ converges, where

$$c_n = \begin{cases} 2^n/n, & n \text{ is odd,} \\ 1/n, & n \text{ even.} \end{cases}$$

E7.10 Consider the following power series: $\sum_{n=0}^{\infty} x^n$.

(a) What is the radius of convergence R of this power series?

(b) For $x \in (-R, R)$, explain why $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$.

(c) For what values of x can we write

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}.$$

(d) For what values of x can we write

$$\int_0^x \frac{1}{1+y^2}\,\mathrm{d}y = \arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}.$$

(e) Can you conclude the validity of the formula

$$\pi = 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.$$

E7.11 (a) Determine the interval of convergence, I, of the power series $\sum_{n=1}^{\infty} \frac{x^n}{n(n+1)}$.

(b) Show that $\sum_{n=1}^{\infty} \frac{x^n}{n(n+1)}$ converges uniformly on I.

(c) Deduce from part (b) that the function $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n(n+1)}$ is continuous on I.

(d) Starting from the fact that $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ for $|x| < 1$, derive an expression for $\sum_{n=1}^{\infty} \frac{x^n}{n(n+1)}$ that is valid for *all* $x \in I$. You must carefully justify all the steps in your derivation.

E7.12 Determine the radius of convergence, R, of the power series $\sum_{n=1}^{\infty} a_n x^n$, where a_n is as given below. If $R \in (0, \infty)$, determine whether the series converges at the points $x = \pm R$.

(a) $a_n = \left(\frac{-1}{2}\right)^{n+1} \frac{\ln(n+1)}{n+1}$.

(b) $a_n = \frac{2^n}{n!}$.

E7.13 Note that (at least formally)

$$\sum_{n=1}^{\infty} nx^n = x\sum_{n=1}^{\infty} nx^{n-1} = x\left(\frac{\mathrm{d}}{\mathrm{d}x}\sum_{n=0}^{\infty} x^n\right) \tag{E7.1}$$

(a) Find the sum of the series $\sum_{n=1}^{\infty} nx^n$.
(b) For what values of x is the calculation (E7.1) valid? Justify your answer by appealing to theorems about power series.

E7.14 Find the sum of the power series $\sum_{n=100}^{\infty} x^{4n+3}$.

E7.15 (a) Find a simple expression for the sum of the power series $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n(n+1)}$.
Hint: *Multiply by x and take the second derivative.*
(b) Use a procedure similar to that in (a) to find a simple expression for $\sum_{n=0}^{\infty} n^2 x^n$.

E7.16 Let r be a fixed non-negative integer. It is not hard to check that the radius of convergence of the power series

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

is $R = \infty$, so that the series converges absolutely for all $x \in \mathbb{R}$ (you don't need to show this). Let $J_r(x)$ denote the function represented by this series. This function is called the *r-th order Bessel function of the first kind.*

Verify that $J_r(x)$ is a solution $y(x)$ to *Bessel's differential equation*

$$x^2y'' + xy' + (x^2 - r^2)y = 0.$$

You must justify all the steps in your verification by quoting appropriate theorems wherever necessary.

(*Remark*: Bessel functions have many applications; they arise for instance when studying the propagation of electromagnetic waves through cylindrical waveguides.)

E7.17 It is shown in Example 7.13–4 of the course reader (read it!) that the Taylor series about 0 of $f(x) = (1+x)^r$, where r is a real number, is given by $\sum_{n=0}^{\infty} \binom{r}{n} x^n$. Since r can be any real number, not necessarily a positive integer, the binomial coefficient $\binom{r}{n}$ is to be interpreted as $\frac{r(r-1)(r-2)\ldots(r-n+1)}{n!}$, with the understanding that $\binom{r}{0} = 1$. It is also shown in the example that the Taylor series, in fact, converges to $f(x)$ for all $x \in (-1, 1)$.

(a) Use the above fact (and some algebraic manipulations) to deduce that the Taylor series about 0 of $\frac{1}{\sqrt{1-4x}}$ is the series $\sum_{n=0}^{\infty} \binom{2n}{n} x^n$.

Conclude from this that $\sum_{n=0}^{\infty} \binom{2n}{n} x^n$ converges to $\frac{1}{\sqrt{1-4x}}$ for $x \in (-1/4, 1/4)$.

(b) (This is hard, and need not be attempted, but you may wish to give it a shot.) A solution to Exercise E5.13 shows that the series $\sum_{n=0}^{\infty}\binom{2n}{n}x^n$ converges conditionally at $x=-1/4$. Show that

$$\sum_{n=0}^{\infty}\binom{2n}{n}(-1/4)^n = \frac{1}{\sqrt{1-4(-1/4)}} = \frac{1}{\sqrt{2}}.$$

E7.18 The nth Catalan number is defined as $C_n = \frac{1}{n+1}\binom{2n}{n}$, for all integers $n \geq 0$.

(a) Use the inequality

$$\binom{2n}{n}4^{-n} \leq \frac{1}{\sqrt{\pi n}}\mathrm{e}^{\frac{1}{24n}}$$

to deduce that $\sum_{n=0}^{\infty} C_n 4^{-n}$ converges.

(b) Determine the interval of convergence, I, of the power series $\sum_{n=0}^{\infty} C_n x^n$.

(c) Show that $\sum_{n=0}^{\infty} C_n x^n$ converges uniformly on I.

(d) Deduce from part (c) that the function $f(x) = \sum_{n=0}^{\infty} C_n x^n$ is continuous on I.

(e) Use the result of Exercise E7.17(a) to derive an expression for $\sum_{n=0}^{\infty} C_n x^n$ that is valid for *all* $x \in I$. You must carefully justify all the steps in your derivation.

(Catalan numbers occur in various counting problems; the question above derives a *generating function* for the Catalan numbers.)

E7.19 If the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$ is R, what is the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^{2n}$?

E7.20 Suppose that the power series $\sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $R \in (0,\infty)$. Find the radius of convergence of $\sum_{n=0}^{\infty} a_n x^{n^2}$.

E7.21 Suppose that the power series $f(x) = \displaystyle\sum_{n=0}^{\infty} a_n x^n$ has a nonzero radius of convergence.

(a) If f is an odd function, show that

$$a_0 = a_2 = a_4 = \cdots = 0.$$

(b) If f is an even function, show that

$$a_1 = a_3 = a_5 = \cdots = 0.$$

E7.22 Prove that if (a_n) is a sequence such that the sequence $(|a_n|)$ is bounded away from both 0 and ∞, then the power series $\sum_{n=0}^{\infty} a_n x^n$ has a radius of convergence equal to 1.

E7.23 Suppose that the sequence (a_n) is bounded, but that the series $\sum_{n=0}^{\infty} a_n$ diverges. Prove that the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n x^n$ is equal to 1.

E7.24 Suppose that $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ are two power series whose radii of convergence are R_1 and R_2,, respectively, what can you say (prove) about the radius of convergence of $\sum_{n=0}^{\infty}(a_n + b_n)x^n$??

E7.25 (a) Determine the radius of convergence of the power series $f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ and obtain expressions for the power series $f'(x)$ and $\int_0^x f(t)\,\mathrm{d}t$ and calculate their radii of convergence.

(b) Repeat this for the power series $f(x) = \sum_{n=1}^{\infty} \frac{\mathrm{e}^n}{n^5+1}\left(\frac{x}{2}\right)^n$.

E7.26 Let $S(x)$ be the sum of the power series $\sum_{n=0}^{\infty} \frac{x^n}{n+1}$. (This is not the power series of a familiar function. Except for part (c), you must answer the following questions by using only properties of power series.)

(a) What is the longest open interval I on which this power series converges?

(b) If S continuous on I? Is S differentiable on I? Why?

(c) Show that, if $|x| < 1$ then

$$\left|\frac{x^n}{n+1}\right| \geq \left|\frac{x^{n+1}}{n+2}\right|.$$

(d) Prove that $S(x) > 0$ for all $x \in I$.

(e) Calculate $\frac{\mathrm{d}}{\mathrm{d}x}(xS(x))$ and use this to deduce a simple expression for $S(x)$.

E7.27 The radius of convergence of the power series

$$f(x) = \sum_{n=1}^{\infty} \frac{x^{2n+1}}{4^n n^2}$$

is 2 (it is not necessary to prove this).

(a) Prove that this series converges uniformly on $[-2, 2]$.

(b) Explain how we know that $f'(x)$ is defined at least on $(-2, 2)$. Quote relevant theorems.

(c) Determine all the open intervals on which f is increasing and all the open intervals on which the graph of f is concave up. (Do not attempt to find a formula for f.)

(d) What symmetry (if any) does f have? Use the information determined so far to make as accurate a sketch as you can of the graph of f.

E7.28 Consider the power series

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^3 9^n}(x-2)^{2n}.$$

(a) Determine the radius of convergence of this series and the interval I consisting of all the numbers x for which the series converges.

(b) Show that it converges uniformly on the set I.

(c) Obtain expressions for $f'(x)$ and $f''(x)$.

(d) Determine all the open intervals on which f is increasing, is decreasing, is concave up, and is concave down. Also, does the graph of f have any symmetry and what are its intercepts? Finally, make a rough sketch of the graph of f.

E7.29 Determine the Taylor series for each of the following functions about the given point (use any convenient procedure):

(a) $x \sin x$ about $x = 0$

(b) $\cos x$ about $x = \pi/3$

(c) $\sqrt{x}$ about $x = 1$

(d) $\displaystyle\int_0^x e^{-t^2}\,dt$ about $x = 0$

(e) $\displaystyle\int_0^x \frac{\sin t^2 - t^2}{t^6}\,dt$ about $x = 0$

E7.30 Use the Taylor series for $f(x) = e^x$ about $x = 0$ and the methods described in Section 7.4 to obtain the Taylor series for each of the following functions about $x = 0$:

(a) e^{-x}

(b) $\cosh x = \dfrac{1}{2}\left(e^x + e^{-x}\right)$

(c) $\sinh x = \dfrac{1}{2}\left(e^x - e^{-x}\right)$

(d) e^{x^2}

(e) xe^{x^2}

(f) $\displaystyle\int_0^x te^{t^2}\,dt = \frac{1}{2}\left(e^{x^2} - 1\right)$

E7.31 Find enough of the Taylor series expansion of the function $f(x) = x^3 e^x$ to determine $f^{(5)}(0)$.

E7.32 Find the Taylor series for $f(x) = \sin x$ about $x = \pi/3$ and prove that this Taylor series really does converge to $\sin x$ for all $x \in \mathbb{R}$.

E7.33 (a) Find the Taylor polynomial of degree n about $x = 0$ for the function e^x.

(b) Use your answer to (a) to find the Taylor polynomial of degree n about $x = 0$ for the function

$$f(x) = \begin{cases} (e^x - 1)/x, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

(c) Use your answer to (b) to express the integral

$$\int_{-1/2}^{0} \frac{e^x - 1}{x}\,dx$$

as an infinite series, and estimate the error in approximating the integral by the first n terms in the series.

E7.34 Determine the Taylor series about $x = 0$ of the function $f(x) = \frac{1}{1+2x^3}$. What is its radius of convergence of this Taylor series?

E7.35 Prove the generalized mean value theorem for integrals: If f and g are two continuous function on an interval $[a, b]$ and if f is nonnegative-valued and positive at at least one point, then there is point $\xi \in [a, b]$ such that

$$f(\xi) \int_a^b g(t)\,\mathrm{d}t = \int_a^b f(t)g(t)\,\mathrm{d}t$$

(cf. the proof of Lemma 1(ii) from the proof of Theorem 7.12).

E7.36 The binomial theorem states that

$$(1+x)^p = 1 + px + \frac{p(p-1)}{2!}x^2 + \frac{p(p-1)(p-2)}{3!}x^3 + \cdots$$

for $|x| \leq 1$. Use the familiar formula

$$\sin^{-1} x = \int_0^x \frac{\mathrm{d}u}{\sqrt{1-u^2}}$$

to deduce that

$$\frac{\pi}{6} = \frac{1}{2} + \frac{1}{2 \cdot 3 \cdot 2^3} + \frac{1 \cdot 3}{2 \cdot 4 \cdot 5 \cdot 2^5} + \cdots .$$

Justify each of the steps in your work.

E7.37 Use the binomial theorem to obtain simple approximations for the following expressions valid for arguments indicated:

(a) $\dfrac{1}{1-u}$ for $|u| \ll 1$

(b) $\left(1 - \dfrac{v^2}{c^2}\right)^{-1/2}$ for $\dfrac{v^2}{c^2} \ll 1$

(c) $(1+x^2)^{3/2}$ for $x^2 \ll 1$

(d) $(1+3x)^{\pi}$ for $|x| \ll 1$

E7.38 Compute the first few terms in the power series expansion about $x = 0$ of the following functions:

(a) $x^3 \cos x$

(b) $\frac{\mathrm{e}^{-x}}{(1-x)}$

(c) $\mathrm{e}^x \sin x$

(d) $\tan x$

Index