



Department of Mathematics, London School of Economics

Introduction to Abstract Mathematics

(Second part of MA103)

Amol Sasane

Contents

1	Analysis	1
1.1	The real numbers	1
1.2	Sequences and limits	8
1.2.1	Sequences	8
1.2.2	Limit of a convergent sequence	10
1.2.3	Bounded and monotone sequences	14
1.2.4	Algebra of limits	17
1.2.5	Sandwich theorem	21
1.2.6	Subsequences	23
1.3	Continuity	26
1.3.1	Definition of continuity	27
1.3.2	Continuous functions preserve convergent sequences	30
1.3.3	Extreme value theorem	32
1.3.4	Intermediate value theorem	34
2	Algebra	39
2.1	Groups	39
2.1.1	Definition of a Group	39
2.1.2	Subgroups	45
2.1.3	Homomorphisms and isomorphisms	49
2.1.4	Cosets and Lagrange's theorem	52
2.2	Vector spaces	55
2.2.1	Definition of a vector space	55
2.2.2	Subspaces and linear combinations	59

2.2.3	Basis of a vector space	62
2.2.4	Linear transformations	65
	Solutions	69
	Bibliography	127
	Index	129

Chapter 1

Analysis

Analysis is the *theory* behind real numbers, sequences, and functions. The word ‘theory’ is important. You might, for example, have a good idea of what we mean by a ‘limit’ of a convergent sequence or the notion of a ‘continuous’ function, but in this part of the course we try to formalize such notions.

1.1 The real numbers

The rational number system is inadequate for many purposes. For instance, there is no rational number q such that $q^2 = 2$, and the set

$$S = \{q \in \mathbb{Q} \mid q^2 \leq 2\}$$

does not have a largest element in \mathbb{Q} . So we see that the rational number system has certain holes. The real number system \mathbb{R} fills these gaps. Thus the set

$$S = \{q \in \mathbb{R} \mid q^2 \leq 2\}$$

has a largest element. This is a consequence of a very important property of the real numbers, called the least upper bound property. But before we state this property of \mathbb{R} , we need a few definitions.

Definitions. Let S be a subset of \mathbb{R} .

1. An element $u \in \mathbb{R}$ is said to be an *upper bound of S* if for all $x \in S$, $x \leq u$. If the set of all upper bounds of S is not empty, then S is said to be *bounded above*.
2. An element $l \in \mathbb{R}$ is said to be a *lower bound of S* if for all $x \in S$, $l \leq x$. If the set of all lower bounds of S is not empty, then S is said to be *bounded below*.
3. The set S is said to be *bounded* if it is bounded above and it is bounded below.

Examples.

1. The set $S = \{x \in \mathbb{R} \mid 0 \leq x < 1\}$ is bounded. Any real number y satisfying $1 \leq y$ (for instance 1, π , 100) is an upper bound of S , and any real number z satisfying $z \leq 0$ (for instance 0, -1) is a lower bound of S .

2. The set $S = \{n \mid n \in \mathbb{N}\}$ is not bounded. Although it is bounded below (any real number $x \leq 1$ serves as a lower bound), it has no upper bound, and so it is not bounded above.
3. The set¹ $S = \{(-1)^n \mid n \in \mathbb{N}\}$ is bounded. It is bounded above by 1 and bounded below by -1 .
4. The set $S = \{\frac{1}{n} \mid n \in \mathbb{N}\}$ is bounded. Any real number x satisfying $1 \leq x$ is an upper bound, and 0 is a lower bound.
5. The sets \mathbb{Z} and \mathbb{R} are neither bounded above nor bounded below. Indeed, this is a consequence of the inequality $z < z + 1$.
6. The set \emptyset is bounded. (Why?) ◇

We now introduce the notions of a least upper bound (also called supremum) and a greatest lower bound (also called infimum) of a subset S of \mathbb{R} .

Definitions. Let S be a subset of \mathbb{R} .

1. An element $u_* \in \mathbb{R}$ is said to be a *least upper bound of S* (or a *supremum of S*) if
 - (a) u_* is an upper bound of S , and
 - (b) if u is an upper bound of S , then $u_* \leq u$.
2. An element $l_* \in \mathbb{R}$ is said to be a *greatest lower bound of S* (or an *infimum of S*) if
 - (a) l_* is a lower bound of S , and
 - (b) if l is a lower bound of S , then $l \leq l_*$.

Example. If $S = \{x \in \mathbb{R} \mid 0 \leq x < 1\}$, then the supremum of S is 1 and the infimum of S is 0.

Clearly 1 is an upper bound of S .

Now we show that if u is another upper bound, then $1 \leq u$. Suppose not, that is, $u < 1$. Then we have

$$0 \leq u < \frac{u+1}{2} < 1, \tag{1.1}$$

where the first inequality is a consequence of the facts that u is an upper bound of S and $0 \in S$, while the last two inequalities follow using $u < 1$. From (1.1), it follows that the number $\frac{u+1}{2}$ satisfies $0 < \frac{u+1}{2} < 1$, and so it belongs to S . The middle inequality in (1.1) above then shows that u cannot be an upper bound for S , a contradiction. Hence 1 is a supremum.

Next we show that this is the only supremum, since if u_* is another supremum, then in particular u_* is also an upper bound, and the above argument shows that $1 \leq u_*$. But $1 < u_*$ is not possible as 1 is an upper bound, and as u_* is a supremum, u_* must be less than or equal to 1. So it follows that $u_* = 1$.

Similarly one can show that the infimum of S is 0. ◇

In the above example, there was a unique supremum and infimum of the set S . In fact, this is always the case and we have the following result.

¹Note that this set is simply the finite set (that is, the set has finite cardinality) $\{-1, 1\}$. More generally, any finite set S is bounded.

Theorem 1.1.1 *If the least upper bound of a subset S of \mathbb{R} exists, then it is unique.*

Proof Suppose that u_* and u'_* are two least upper bounds of S . Then in particular u_* and u'_* are also upper bounds of S . Now since u_* is a least upper bound of S and u'_* is an upper bound of S , it follows that

$$u_* \leq u'_*. \quad (1.2)$$

Furthermore, since u'_* is a least upper bound of S and u_* is an upper bound of S , it follows that

$$u'_* \leq u_*. \quad (1.3)$$

From (1.2) and (1.3), we obtain $u_* = u'_*$. ■

Thus it makes sense to talk about *the* least upper bound of a set. The least upper bound of a set S (if it exists) is denoted by

$$\sup S$$

(the abbreviation of ‘supremum of S ’). Similarly, the infimum of a set S (if it exists) is also unique, and is denoted by

$$\inf S.$$

When the supremum and the infimum of a set belong to the set, then we give them special names:

Definitions.

1. If $\sup S \in S$, then $\sup S$ is called a *maximum of S* , denoted by $\max S$.
2. If $\inf S \in S$, then $\inf S$ is called a *minimum of S* , denoted by $\min S$.

Examples.

1. If $S = \{x \in \mathbb{R} \mid 0 \leq x < 1\}$, then $\sup S = 1 \notin S$ and so $\max S$ does not exist. But $\inf S = 0 \in S$, and so $\min S = 0$.
2. If $S = \{n \mid n \in \mathbb{N}\}$, then $\sup S$ does not exist, $\inf S = 1$, $\max S$ does not exist, and $\min S = 1$.
3. If $S = \{(-1)^n \mid n \in \mathbb{N}\}$, then $\sup S = 1$, $\inf S = -1$, $\max S = 1$, $\min S = -1$.
4. If $S = \{\frac{1}{n} \mid n \in \mathbb{N}\}$, then $\sup S = 1$ and $\max S = 1$. We show below (after Theorem 1.1.2) that $\inf S = 0$. So $\min S$ does not exist.
5. For the sets \mathbb{Z} and \mathbb{R} , \sup , \inf , \max , \min do not exist.
6. For the set \emptyset , \sup , \inf , \max , \min do not exist. ◇

In the above examples, we note that if S is nonempty and bounded above, then its supremum exists. In fact this is a fundamental property of the real numbers, called the *least upper bound property* of the real numbers, which we state below:

If S is a nonempty subset of \mathbb{R} having an upper bound, then $\sup S$ exists.

Remarks(*).

1. Note that the set of rational numbers do not possess this property. For instance, the set

$$S = \{q \in \mathbb{Q} \mid q^2 \leq 2\}$$

has an upper bound, say 2 (indeed, if $q > 2$, then $q^2 > 4 > 2$, and so $q \notin S$), but we now show that it does not have a supremum in \mathbb{Q} . Assume on the contrary that $u_* \in \mathbb{Q}$ is a supremum for S . Define

$$r = u_* - \frac{u_*^2 - 2}{u_* + 2}. \quad (1.4)$$

Then we can check that

$$r^2 - 2 = \frac{2(u_*^2 - 2)}{(u_* + 2)^2} \quad (1.5)$$

We have the following cases:

1° Suppose that $u_*^2 < 2$. From (1.5) we obtain $r \in S$, and from (1.4) it follows that $r > u_*$, which contradicts the fact that u_* is an upper bound of S .

2° Suppose that $u_*^2 = 2$. This is impossible, since u_* is a rational number.

3° Suppose that $u_*^2 > 2$. We see that r is an upper bound of S (indeed, if $q > r$, then $q^2 > r^2 > 2$, and so $q \notin S$). But (1.4) implies that $r < u_*$, contradicting the fact that u_* is the supremum.

2. In Exercise 6 on page 7 below, given a nonempty set S of \mathbb{R} , we define $-S = \{-x \mid x \in S\}$. One can show that if a nonempty subset S of \mathbb{R} is bounded below, then $-S$ is bounded above and so $\sup(-S)$ exists, by the least upper bound property. The negative of this supremum, namely $-\sup(-S)$, can then be shown to serve as the greatest lower bound of S (this is precisely the content of Exercise 6). Thus the real numbers also have the ‘greatest lower bound property’: If S is a nonempty subset of \mathbb{R} having an lower bound, then $\inf S$ exists.

We now prove the following theorem, which is called the *Archimedean property* of the real numbers.

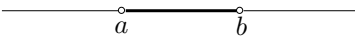
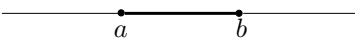
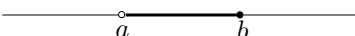
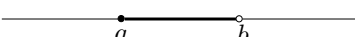
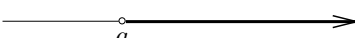
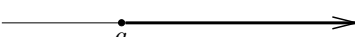
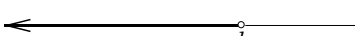


Theorem 1.1.2 (Archimedean property.) *If $x, y \in \mathbb{R}$ and $x > 0$, then there exists an $n \in \mathbb{N}$ such that $y < nx$.*

Proof If not, then the nonempty set $S = \{nx \mid n \in \mathbb{N}\}$ has an upper bound y , and so by the least upper bound property of the reals, it has a least upper bound u_* . But then $u_* - x < u_*$ (since x is positive) and so it follows that $u_* - x$ cannot be an upper bound of S . Hence there exists a natural number m such that $u_* - x < mx$, that is, $u_* < (m+1)x \in S$. This contradicts the fact that u_* is an upper bound of S . ■

Example. If $S = \{\frac{1}{n} \mid n \in \mathbb{N}\}$, then $\inf S = 0$. We know that 0 is a lower bound of S . Suppose that l is a lower bound of S such that $l > 0$. By the Archimedean property (with the real numbers x and y taken as $x = 1 (> 0)$ and $y = \frac{1}{l}$), there exists a $n \in \mathbb{N}$ such that $\frac{1}{l} = y < nx = n \cdot 1 = n$, and so $\frac{1}{n} < l$, contradicting the fact that l is a lower bound of S . Thus any lower bound of S must be less than or equal to 0. Hence 0 is the infimum of S . ◇

²the last inequality follows from (1.5)

Definition. An *interval* is a set consisting of all the real numbers between two given real numbers, or of all the real numbers on one side or the other of a given number. So an interval is a set of any of the following forms, where $a, b \in \mathbb{R}$:

$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$	
$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$	
$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$	
$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$	
$(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$	
$[a, \infty) = \{x \in \mathbb{R} \mid a \leq x\}$	
$(-\infty, b) = \{x \in \mathbb{R} \mid x < b\}$	
$(-\infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$	
$(-\infty, \infty) = \mathbb{R}$	

In the above notation for intervals, a parenthesis ‘(’ or ‘)’ means that the respective endpoint is not included, and a square bracket ‘[’ or ‘]’ means that the endpoint is included. Thus $[0, 1)$ means the set of all real numbers x such that $0 \leq x < 1$. (Note that the use of the symbol ∞ in the notation for intervals is simply a matter of convenience and is not to be taken as suggesting that there is a number ∞ .)

In analysis, in order to talk about notions such as *convergence* and *continuity*, we will need a notion of ‘closeness’ between real numbers. This is provided by the absolute value $|\cdot|$, and the distance between real numbers x and y is $|x - y|$. We give the definitions below.

Definitions.

1. The *absolute value* of a real number x is denoted by $|x|$, and it is defined as follows:

$$|x| = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

2. The *distance* between two real numbers x and y is the absolute value $|x - y|$ of their difference.

Thus $|1| = 1$, $|0| = 0$, $|-1| = 1$, and the distance between the real numbers -1 and 1 is equal to $|-1 - 1| = |-2| = 2$. The distance gives a notion of closeness of two points, which is crucial in the formalization of the notions of analysis. We can now specify regions comprising points close to a certain point $x_0 \in \mathbb{R}$ in terms of inequalities in absolute values, that is, by demanding that the distance of the points of the region, to the point x_0 , is less than a certain positive number δ , say $\delta = 0.01$ or $\delta = 0.0000001$, and so on. See Exercise 9 on page 8 and Figure 1.1.

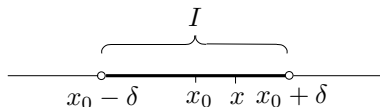


Figure 1.1: The interval $I = (x_0 - \delta, x_0 + \delta) = \{x \in \mathbb{R} \mid |x - x_0| < \delta\}$ is the set of all points in \mathbb{R} whose distance to the point x_0 is strictly less than δ (> 0).

The following properties of the absolute value will be useful in the sequel.

Theorem 1.1.3 *If x, y are real numbers, then*

$$|xy| = |x| |y| \quad \text{and} \quad (1.6)$$

$$|x + y| \leq |x| + |y|. \quad (1.7)$$

Proof We prove (1.6) by exhausting all possible cases:

1^o $x = 0$ or $y = 0$. Then $|x| = 0$ or $|y| = 0$, and so $|x| |y| = 0$. On the other hand, as $x = 0$ or $y = 0$, it follows that $xy = 0$ and so $|xy| = 0$.

2^o $x > 0$ and $y > 0$. Then $|x| = x$ and $|y| = y$, and so $|x| |y| = xy$. On the other hand, as $x > 0$ and $y > 0$, it follows that $xy > 0$ and so $|xy| = xy$.

3^o $x > 0$ and $y < 0$. Then $|x| = x$ and $|y| = -y$, and so $|x| |y| = x(-y) = -xy$. On the other hand, as $x > 0$ and $y < 0$, it follows that $xy < 0$ and so $|xy| = -xy$.

4^o $x < 0$ and $y > 0$. This follows from 3^o above by interchanging x and y .

5^o $x < 0$ and $y < 0$. Then $|x| = -x$ and $|y| = -y$, and so $|x| |y| = (-x)(-y) = xy$. On the other hand, as $x < 0$ and $y < 0$, it follows that $xy > 0$ and so $|xy| = xy$.

This proves (1.6).

Next we prove (1.7). First observe that from the definition of $|\cdot|$, it follows that for any real $x \in \mathbb{R}$, $|x| \geq x$: indeed if $x \geq 0$, then $|x| = x$, while if $x < 0$, then $-x > 0$, and so $|x| = -x > 0 > x$. From (1.6), we also have $|-x| = |-1 \cdot x| = |-1| |x| = 1|x| = |x|$, for all $x \in \mathbb{R}$, and so it follows that $|x| = |-x| \geq -x$ for all $x \in \mathbb{R}$. We have the following cases:

1^o $x + y \geq 0$. Then $|x + y| = x + y$. As $|x| \geq x$ and $|y| \geq y$, we obtain $|x| + |y| \geq x + y = |x + y|$.

2^o $x + y < 0$. Then $|x + y| = -(x + y)$. Since $|x| \geq -x$ and $|y| \geq -y$, it follows that $|x| + |y| \geq -x + (-y) = -(x + y) = |x + y|$.

This proves (1.7). ■

It is easy to check that the distance satisfies the following properties:

D1. (*Positive definiteness.*) For all $x, y \in \mathbb{R}$, $|x - y| \geq 0$. If $|x - y| = 0$ then $x = y$.

D2. (*Symmetry.*) For all $x, y \in \mathbb{R}$, $|x - y| = |y - x|$.

D3. (*Triangle inequality.*) For all $x, y, z \in \mathbb{R}$, $|x - z| \leq |x - y| + |y - z|$.

Exercises.

1. Provide the following information about the set S

An upper bound	A lower bound	Is S bounded?	$\sup S$	$\inf S$	If $\sup S$ exists, then is $\sup S$ in S ?	If $\inf S$ exists, then is $\inf S$ in S ?	$\max S$	$\min S$

where S is given by:

- (a) $(0, 1]$
- (b) $[0, 1]$
- (c) $(0, 1)$
- (d) $\{\frac{1}{n} \mid n \in \mathbb{Z} \setminus \{0\}\}$
- (e) $\{-\frac{1}{n} \mid n \in \mathbb{N}\}$
- (f) $\{\frac{n}{n+1} \mid n \in \mathbb{N}\}$
- (g) $\{x \in \mathbb{R} \mid x^2 \leq 2\}$
- (h) $\{0, 2, 5, 2005\}$
- (i) $\{(-1)^n (1 + \frac{1}{n}) \mid n \in \mathbb{N}\}$
- (j) $\{x^2 \mid x \in \mathbb{R}\}$
- (k) $\{\frac{x^2}{1+x^2} \mid x \in \mathbb{R}\}$.

2. Determine whether the following statements are TRUE or FALSE.

- (a) If u is an upper bound of a subset S of \mathbb{R} , and $u' < u$, then u' is not an upper bound for S .
- (b) If u_* is the least upper bound of a subset S of \mathbb{R} , and ϵ is any positive real number, then $u_* - \epsilon$ is not an upper bound of S .
- (c) Every subset of \mathbb{R} has a maximum.
- (d) Every subset of \mathbb{R} has a supremum.
- (e) Every bounded subset of \mathbb{R} has a maximum.
- (f) Every bounded subset of \mathbb{R} has a supremum.
- (g) Every bounded nonempty subset of \mathbb{R} has a supremum.
- (h) Every set that has a supremum is bounded above.
- (i) For every set that has a maximum, the maximum belongs to the set.
- (j) For every set that has a supremum, the supremum belongs to the set.
- (k) For every set S that is bounded above, $|S|$ defined by $\{|x| \mid x \in S\}$ is bounded.
- (l) For every set S that is bounded, $|S|$ defined by $\{|x| \mid x \in S\}$ is bounded.
- (m) For every bounded set S , if $\inf S < x < \sup S$, then $x \in S$.

3. For any nonempty bounded set S , prove that $\inf S \leq \sup S$, and that the equality holds iff³ S is a singleton set (that is a set with cardinality 1).

4. Let A and B be nonempty subsets of \mathbb{R} that are bounded above and such that $A \subset B$. Prove that $\sup A \leq \sup B$.

5. Let A and B be nonempty subsets of \mathbb{R} that are bounded above and define

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

Prove that $\sup(A + B)$ exists and that $\sup(A + B) \leq \sup A + \sup B$.

6. Let S be a nonempty subset of real numbers which is bounded below. Let $-S$ denote the set of all real numbers $-x$, where x belongs to S . Prove that $\inf S$ exists and $\inf S = -\sup(-S)$.

7. Let S be a nonempty set of positive real numbers, and define $S^{-1} = \{\frac{1}{x} \mid x \in S\}$. Show that S^{-1} is bounded above iff $\inf S > 0$. Furthermore, in case $\inf S > 0$, show that $\sup S^{-1} = \frac{1}{\inf S}$.

³The abbreviation 'iff' is standard in Mathematics, and it stands for 'if and only if'.

8. Let A_n , $n \in \mathbb{N}$, be a collection of sets.

The notation $\bigcap_{n \in \mathbb{N}} A_n$ denotes the intersection of the sets A_n , $n \in \mathbb{N}$, that is,

$$\bigcap_{n \in \mathbb{N}} A_n = \{x \mid \forall n \in \mathbb{N}, x \in A_n\},$$

and we use $\bigcup_{n \in \mathbb{N}} A_n$ to denote the union of the sets A_n , $n \in \mathbb{N}$, that is,

$$\bigcup_{n \in \mathbb{N}} A_n = \{x \mid \exists n \in \mathbb{N} \text{ such that } x \in A_n\}.$$

Prove that

$$(a) \emptyset = \bigcap_{n \in \mathbb{N}} \left(0, \frac{1}{n}\right).$$

$$(b) \{0\} = \bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right].$$

$$(c) (0, 1) = \bigcup_{n \in \mathbb{N}} \left[\frac{1}{n+2}, 1 - \frac{1}{n+2}\right].$$

$$(d) [0, 1] = \bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right).$$

9. Let $x_0 \in \mathbb{R}$ and $\delta > 0$. Prove that $(x_0 - \delta, x_0 + \delta) = \{x \in \mathbb{R} \mid |x - x_0| < \delta\}$.

10. Prove that if x, y are real numbers, then $||x| - |y|| \leq |x - y|$.

11. Show that a subset S of \mathbb{R} is bounded iff there exists a $M \in \mathbb{R}$ such that for all $x \in S$, $|x| \leq M$.

1.2 Sequences and limits

1.2.1 Sequences

The notion of a sequence occurs in ordinary conversation. An example is the phrase “an unfortunate sequence of events”. In this case, we envision one event causing another, which in turn causes another event and so on. We can identify a *first* event, a *second* event, etcetera.

A sequence of real numbers is a list

$$a_1, a_2, a_3, \dots$$

of real numbers, where there is the *first* number (namely a_1), the *second* number (namely a_2), and so on. For example,

$$1, \frac{1}{2}, \frac{1}{3}, \dots$$

is a sequence. The first number is 1, the second number is $\frac{1}{2}$ and so on. (There may not be a connection between the numbers appearing in a sequence.) If we think of a_1 as $f(1)$, a_2 as $f(2)$, and so on, then it becomes clear that a sequence is a special type of function, namely one with domain \mathbb{N} and co-domain \mathbb{R} .

Definition. A *sequence* is a function $f : \mathbb{N} \rightarrow \mathbb{R}$.

Only the notation is somewhat unusual. Instead of writing $f(n)$ for the value of f at a natural number n , we write a_n . The entire sequence is then written in any one of the following ways:

$$(a_n)_{n \in \mathbb{N}}, (a_n)_{n=1}^{\infty}, (a_n)_{n \geq 1}, (a_n).$$

In $(a_n)_{n=1}^{\infty}$, the ∞ symbol indicates that the assignment process $1 \mapsto a_1, 2 \mapsto a_2, \dots$ continues indefinitely. The n th term a_n of a sequence may be defined explicitly by a formula involving n , as in the example given above:

$$a_n = \frac{1}{n}, \quad n \in \mathbb{N}.$$

It might also sometimes be defined recursively. For example,

$$a_1 = 1, \quad a_{n+1} = \frac{n}{n+1}a_n \text{ for } n \in \mathbb{N}.$$

(Write down the first few terms of this sequence.)

Examples.

1. $(\frac{1}{n})_{n \in \mathbb{N}}$ is a sequence with the n th term given by $\frac{1}{n}$, for $n \in \mathbb{N}$. This is the sequence

$$1, \frac{1}{2}, \frac{1}{3}, \dots$$

2. $(1 + \frac{1}{n})_{n \in \mathbb{N}}$ is a sequence with the n th term given by $1 + \frac{1}{n}$, for $n \in \mathbb{N}$. This is the sequence

$$2, \frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \frac{6}{5}, \frac{7}{6}, \dots$$

3. $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$ is a sequence with the n th term given by $(-1)^n (1 + \frac{1}{n})$, for $n \in \mathbb{N}$. This is the sequence

$$-2, \frac{3}{2}, -\frac{4}{3}, \frac{5}{4}, -\frac{6}{5}, \frac{7}{6}, \dots$$

4. $((-1)^n)_{n \in \mathbb{N}}$ is a sequence with the n th term given by $(-1)^n$, for $n \in \mathbb{N}$. This sequence is simply

$$-1, 1, -1, 1, -1, 1, \dots$$

with the n th term equal to -1 if n is odd, and 1 if n is even.

5. $(1)_{n \in \mathbb{N}}$ is a sequence with the n th term given by 1 , for $n \in \mathbb{N}$. This is the constant sequence

$$1, 1, 1, \dots$$

6. $(n)_{n \in \mathbb{N}}$ is a sequence with the n th term given by n , for $n \in \mathbb{N}$. This is the increasing sequence

$$1, 2, 3, \dots$$

7. $(\frac{1}{1^1} + \frac{1}{2^2} + \frac{1}{3^3} + \dots + \frac{1}{n^n})_{n \in \mathbb{N}}$ is a sequence with the n th term given by $\frac{1}{1^1} + \frac{1}{2^2} + \frac{1}{3^3} + \dots + \frac{1}{n^n}$, for $n \in \mathbb{N}$. This is the sequence of ‘partial sums’

$$\frac{1}{1^1}, \frac{1}{1^1} + \frac{1}{2^2}, \frac{1}{1^1} + \frac{1}{2^2} + \frac{1}{3^3}, \dots$$

◇

1.2.2 Limit of a convergent sequence

A sequence can be graphed. For instance, the first 4 points of the graph of the sequence $(\frac{1}{n})_{n \in \mathbb{N}}$ are displayed in Figure 1.2. This portion of the graph suggests that the terms of the sequence

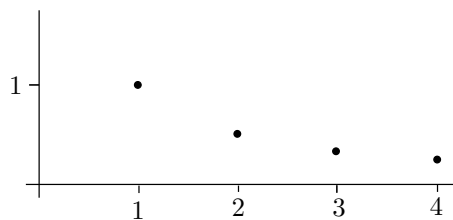


Figure 1.2: First four points of the graph of the sequence $(\frac{1}{n})_{n \in \mathbb{N}}$.

$(\frac{1}{n})_{n \in \mathbb{N}}$ tend toward 0 as n increases. This is consistent with the idea of convergence that you might have encountered before: a sequence $(a_n)_{n \in \mathbb{N}}$ converges to some real number L , if the terms a_n get “closer and closer” to L as n “increases without bound”. Symbolically, this is represented using the notation

$$\lim_{n \rightarrow \infty} a_n = L,$$

where L denotes the limit of the sequence. If there is no such finite number L to which the terms of the sequence get arbitrarily close, then the sequence is said to diverge.

The problem with this characterization is its imprecision. Exactly what does it mean for the terms of a sequence to get “closer and closer”, or “as close as we like”, or “arbitrarily close” to some number L ? Even if we accept this apparent ambiguity, how would one use the definition given in the preceding paragraph to prove theorems that involve sequences? Since sequences are used throughout analysis, the concepts of their convergence and divergence must be carefully defined.

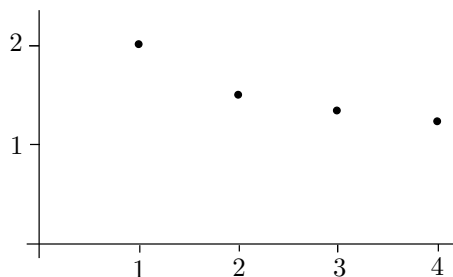


Figure 1.3: First four points of the graph of the sequence $(1 + \frac{1}{n})_{n \in \mathbb{N}}$.

For example, the terms of $(1 + \frac{1}{n})_{n \in \mathbb{N}}$ get “closer and closer” to 0 (indeed the distance to 0 keeps decreasing), but its limit is 1. See Figure 1.3.

The terms of $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$ get “as close as we like” or “arbitrarily close” to 1, but the sequence has no limit. See Figure 1.4.

Definition. The sequence $(a_n)_{n \in \mathbb{N}}$ is said to *converge to* L if for every $\epsilon > 0$, there exists⁴ an $N \in \mathbb{N}$ such that for all $n > N$,

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

⁴depending on ϵ

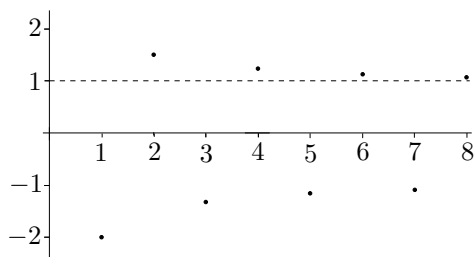


Figure 1.4: First eight points of the graph of the sequence $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$.

Then we say that $(a_n)_{n \in \mathbb{N}}$ is *convergent* (with limit L) and write

$$\lim_{n \rightarrow \infty} a_n = L.$$

If there does not exist a number L such that $\lim_{n \rightarrow \infty} a_n = L$, then the sequence $(a_n)_{n \in \mathbb{N}}$ is said to be *divergent*.

Note that $|a_n - L| < \epsilon$ iff $a_n \in (L - \epsilon, L + \epsilon)$. Hence pictorially, for a convergent sequence with limit L , this definition means the following, as illustrated in Figure 1.5: Pick any $\epsilon > 0$, and consider the shaded strip of width ϵ around the horizontal line passing through L . Then one can find a $N \in \mathbb{N}$, large enough, such that all the terms a_n of the sequence, for $n > N$ lie in the shaded strip.

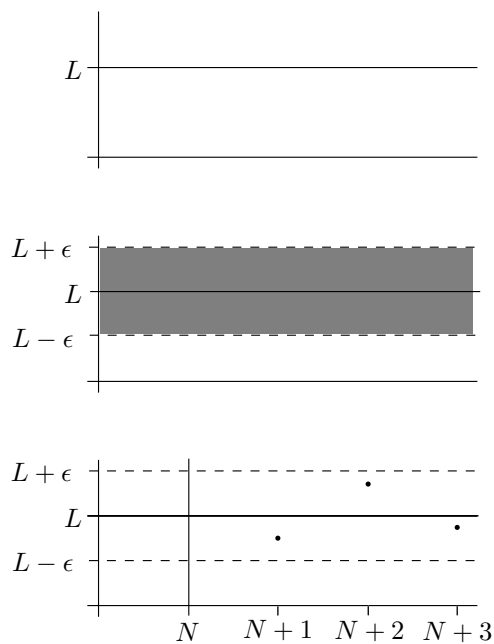


Figure 1.5: Convergence of a sequence with limit L .

Examples.

1. $(\frac{1}{n})_{n \in \mathbb{N}}$ is a convergent sequence with limit 0.

Given $\epsilon > 0$, we need to find a N such that for all $n > N$,

$$|a_n - L| = \left| \frac{1}{n} - 0 \right| = \frac{1}{n} < \epsilon.$$

If we choose $N \in \mathbb{N}$ such that $N > \frac{1}{\epsilon}$ (such a N exists by the Archimedean property!), then for $n > N$ ($\Leftrightarrow \frac{1}{n} < \frac{1}{N}$), we have

$$|a_n - L| = \left| \frac{1}{n} - 0 \right| = \frac{1}{n} < \frac{1}{N} < \epsilon.$$

Hence $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$.

2. $(1 + \frac{1}{n})_{n \in \mathbb{N}}$ is a convergent sequence with limit 1.

Given $\epsilon > 0$, we need to find a N such that for all $n > N$,

$$|a_n - L| = \left| 1 + \frac{1}{n} - 1 \right| = \frac{1}{n} < \epsilon.$$

Again we choose a $N \in \mathbb{N}$ such that $N > \frac{1}{\epsilon}$ and so for $n > N$ we have

$$|a_n - L| = \frac{1}{n} < \frac{1}{N} < \epsilon.$$

Hence $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right) = 1$.

3. $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$ is a divergent sequence.

Let $a_n = (-1)^n (1 + \frac{1}{n})$ for $n \in \mathbb{N}$. In order to prove that

$$(a_n)_{n \in \mathbb{N}} \text{ is divergent,}$$

we have to show that

$$\neg [(a_n)_{n \in \mathbb{N}} \text{ is convergent}],$$

that is,

$$\neg [\exists L \in \mathbb{R} \text{ such that } \forall \epsilon > 0 \exists N \in \mathbb{N} \text{ such that } \forall n > N, |a_n - L| < \epsilon],$$

that is,

$$\forall L \in \mathbb{R} \exists \epsilon > 0 \text{ such that } \forall N \in \mathbb{N} \exists n > N \text{ such that } |a_n - L| \geq \epsilon.$$

Let $L \in \mathbb{R}$. Now we need to prove

$$\exists \epsilon > 0 \text{ such that } \forall N \in \mathbb{N} \exists n > N \text{ such that } |a_n - L| \geq \epsilon.$$

Let $\epsilon = 1$. (It is not *obvious* that $\epsilon = 1$ would work, but it has been found by trial and error.) Now we will show that

$$\forall N \in \mathbb{N} \exists n > N \text{ such that } |a_n - L| \geq \epsilon.$$

So let $N \in \mathbb{N}$. If $L \geq 0$, then choose n to be any odd number $> N$. Then we have

$$|a_n - L| = \left| (-1)^n \left(1 + \frac{1}{n}\right) - L \right| = \left| -1 - \frac{1}{n} - L \right| = 1 + \frac{1}{n} + L > 1 = \epsilon.$$

If $L < 0$, then choose n to be any even number $> N$. Then we have

$$|a_n - L| = \left| (-1)^n \left(1 + \frac{1}{n}\right) - L \right| = \left| 1 + \frac{1}{n} - L \right| = 1 + \frac{1}{n} - L > 1 = \epsilon.$$

Thus we have shown that for all $L \in \mathbb{R}$, there exists a $\epsilon > 0$ (namely $\epsilon = 1$) such that for all $N \in \mathbb{N}$, there exists a $n > N$ (namely any odd number $> N$ if $L \geq 0$, and any even number $> N$ if $L < 0$) such that $|a_n - L| \geq \epsilon$. Thus the sequence is divergent. \diamond

The notation $\lim_{n \rightarrow \infty} a_n$ suggests that the limit of a convergent sequence is unique. Indeed this is the case, and we prove this below.

Theorem 1.2.1 *A convergent sequence has a unique limit.*

Proof Consider a convergent sequence $(a_n)_{n \in \mathbb{N}}$ and suppose that it has distinct limits L_1 and L_2 . Let

$$\epsilon = \frac{|L_1 - L_2|}{2} > 0,$$

where the positivity of the ϵ defined above follows from the fact that $L_1 \neq L_2$. Since L_1 is a limit, $\exists N_1 \in \mathbb{N}$ such that for all $n > N_1$,

$$|a_n - L_1| < \epsilon.$$

Since L_2 is a limit, $\exists N_2 \in \mathbb{N}$ such that for all $n > N_2$,

$$|a_n - L_2| < \epsilon.$$

Consequently for $n > \max\{N_1, N_2\}$,

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

a contradiction. ■

Exercises.

1. (a) Prove that the constant sequence $(1)_{n \in \mathbb{N}}$ is convergent.
 (b) Can the limit of a convergent sequence be one of the terms of the sequence?
 (c) If none of the terms of a convergent sequence equal its limit, then prove that the terms of the sequence cannot consist of a finite number of distinct values.
 (d) Prove that the sequence $((-1)^n)_{n \in \mathbb{N}}$ is divergent.
2. Prove that $\lim_{n \rightarrow \infty} \frac{1}{n} \neq 1$.
3. In each of the cases listed below, give an example of a divergent sequence $(a_n)_{n \in \mathbb{N}}$ that satisfies the given conditions. Suppose that $L = 1$.
 (a) For every $\epsilon > 0$, there exists an N such that for infinitely many $n > N$, $|a_n - L| < \epsilon$.
 (b) There exists an $\epsilon > 0$ and a $N \in \mathbb{N}$ such that for all $n > N$, $|a_n - L| < \epsilon$.
4. Let S be a nonempty subset of \mathbb{R} that is bounded above. Show that there exists a sequence $(a_n)_{n \in \mathbb{N}}$ contained in S (that is, $a_n \in S$ for all $n \in \mathbb{N}$) and which is convergent with limit equal to $\sup S$.
5. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence such that for all $n \in \mathbb{N}$, $a_n \geq 0$. Prove that if $(a_n)_{n \in \mathbb{N}}$ is convergent with limit L , then $L \geq 0$.
6. A sequence $(a_n)_{n \in \mathbb{N}}$ is said to be *Cauchy* if for every $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that for all $n, m > N$, $|a_n - a_m| < \epsilon$.

Show that every convergent sequence is Cauchy.

HINT: $|a_n - a_m| = |a_n - L + L - a_m| \leq |a_n - L| + |a_m - L|$.

1.2.3 Bounded and monotone sequences

It is cumbersome to check from the definition if a sequence is convergent or not. In this section, we will study a condition under which we can conclude that a sequence is convergent even without knowing its limit! We will prove that if a sequence is both ‘bounded’ as well as ‘monotone’, then it is always convergent.

Definition. A sequence $(a_n)_{n \in \mathbb{N}}$ is said to be *bounded* if there exists a $M > 0$ such that

$$\text{for all } n \in \mathbb{N}, \quad |a_n| \leq M. \quad (1.8)$$

Note that a sequence is bounded iff the set $S = \{a_n \mid n \in \mathbb{N}\}$ is bounded. (See Exercise 11 on page 8).

Examples.

1. $(1)_{n \in \mathbb{N}}$ is bounded, since $|1| = 1 \leq 1$ for all $n \in \mathbb{N}$.
2. $(\frac{1}{n})_{n \in \mathbb{N}}$ is bounded, since $|\frac{1}{n}| = \frac{1}{n} \leq 1$ for all $n \in \mathbb{N}$.
3. $(1 + \frac{1}{n})_{n \in \mathbb{N}}$ is bounded, since $|1 + \frac{1}{n}| = 1 + \frac{1}{n} \leq 2$ for all $n \in \mathbb{N}$.
4. $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$ is bounded, since $|(-1)^n (1 + \frac{1}{n})| = 1 + \frac{1}{n} \leq 2$ for all $n \in \mathbb{N}$.
5. The sequence $(a_n)_{n \in \mathbb{N}}$ defined by

$$a_n = \frac{1}{1^1} + \frac{1}{2^2} + \frac{1}{3^3} + \cdots + \frac{1}{n^n}, \quad n \in \mathbb{N}$$

is bounded. Indeed this can be seen as follows:

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

(Write down a detailed proof using induction on n .) So all the terms are bounded by $\frac{3}{2}$, and so the sequence is bounded.

6. The sequence $(a_n)_{n \in \mathbb{N}}$ given by $a_n = n$ for $n \in \mathbb{N}$, is not bounded. Indeed, given any $M > 0$, there exists an $N \in \mathbb{N}$ such that $M < N$ (Archimedean property with $y = M$ and $x = 1$). Thus

$$\neg[\exists M > 0 \text{ such that for all } n \in \mathbb{N}, |a_n| = |n| = n \leq M],$$

and so $(n)_{n \in \mathbb{N}}$ is not bounded. ◇

The sequences $(1)_{n \in \mathbb{N}}$, $(\frac{1}{n})_{n \in \mathbb{N}}$, $(1 + \frac{1}{n})_{n \in \mathbb{N}}$ are all convergent, and we have shown above that these are also bounded. This is not a coincidence, and in the next theorem we show that the set of all convergent sequences is contained in the set of all bounded sequences. See Figure 1.6.

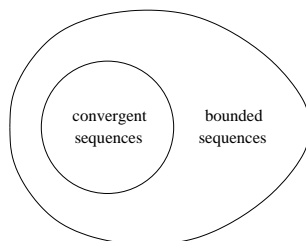


Figure 1.6: All convergent sequences are bounded.

Theorem 1.2.2 *If a sequence is convergent, then it is bounded.*

Proof Let $(a_n)_{n \in \mathbb{N}}$ be a convergent sequence with limit L . Let $\epsilon := 1 > 0$. Then $\exists N \in \mathbb{N}$ such that for all $n > N$

$$|a_n - L| < \epsilon = 1.$$

Hence for all $n > N$,

$$|a_n| = |a_n - L + L| \leq |a_n - L| + |L| < 1 + |L|.$$

Let $M = \max\{|a_1|, \dots, |a_N|, 1 + |L|\}$. Then for all $n \in \mathbb{N}$

$$|a_n| \leq M$$

and so $(a_n)_{n \in \mathbb{N}}$ is bounded. ■

Definitions. A sequence $(a_n)_{n \in \mathbb{N}}$ is said to be *increasing* if for all $n \in \mathbb{N}$, $a_n \leq a_{n+1}$. Thus $(a_n)_{n \in \mathbb{N}}$ is increasing if

$$a_1 \leq a_2 \leq a_3 \leq \dots$$

A sequence $(a_n)_{n \in \mathbb{N}}$ is said to be *decreasing* if for all $n \in \mathbb{N}$, $a_n \geq a_{n+1}$. Thus $(a_n)_{n \in \mathbb{N}}$ is decreasing if

$$a_1 \geq a_2 \geq a_3 \geq \dots$$

A sequence is said to be *monotone* if it is increasing or decreasing.

Examples.

Sequence	Is it increasing?	Is it decreasing?	Is it monotone?
$(\frac{1}{n})_{n \in \mathbb{N}}$	No	Yes	Yes
$(1 + \frac{1}{n})_{n \in \mathbb{N}}$	No	Yes	Yes
$((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$	No	No	No
$(1)_{n \in \mathbb{N}}$	Yes	Yes	Yes
$(n)_{n \in \mathbb{N}}$	Yes	No	Yes
$(\frac{1}{1^1} + \frac{1}{2^2} + \frac{1}{3^3} + \dots + \frac{1}{n^n})_{n \in \mathbb{N}}$	Yes	No	Yes

◇

The following theorem can be useful in showing that sequences converge when one does not know the limit beforehand.

Theorem 1.2.3 *If a sequence is monotone and bounded, then it is convergent.*

Proof

1° Let $(a_n)_{n \in \mathbb{N}}$ be an increasing sequence. Since $(a_n)_{n \in \mathbb{N}}$ is bounded, it follows that the set

$$S = \{a_n \mid n \in \mathbb{N}\}$$

has an upper bound and so $\sup S$ exists. We show that in fact $(a_n)_{n \in \mathbb{N}}$ converges to $\sup S$. Indeed given $\epsilon > 0$, then since $\sup S - \epsilon < \sup S$, it follows that $\sup S - \epsilon$ is not an upper bound for S and so $\exists a_N \in S$ such that $\sup S - \epsilon < a_N$, that is

$$\sup S - a_N < \epsilon.$$

Since $(a_n)_{n \in \mathbb{N}}$ is an increasing sequence, for $n > N$, we have $a_N \leq a_n$. Since $\sup S$ is an upper bound for S , $a_n \leq \sup S$ and so $|a_n - \sup S| = \sup S - a_n$. Thus for $n > N$ we obtain

$$|a_n - \sup S| = \sup S - a_n \leq \sup S - a_N < \epsilon.$$

2° If $(a_n)_{n \in \mathbb{N}}$ is a decreasing sequence, then clearly $(-a_n)_{n \in \mathbb{N}}$ is an increasing sequence. Furthermore if $(a_n)_{n \in \mathbb{N}}$ is bounded, then $(-a_n)_{n \in \mathbb{N}}$ is bounded as well ($|-a_n| = |a_n| \leq M$). Hence by the case considered above, it follows that $(-a_n)_{n \in \mathbb{N}}$ is a convergent sequence with limit

$$\sup\{-a_n \mid n \in \mathbb{N}\} = -\inf\{a_n \mid n \in \mathbb{N}\} = -\inf S,$$

where $S = \{a_n \mid n \in \mathbb{N}\}$ (see Exercise 6 on page 7). So given $\epsilon > 0$, $\exists N \in \mathbb{N}$ such that for all $n > N$, $|-a_n - (-\inf S)| < \epsilon$, that is, $|a_n - \inf S| < \epsilon$. Thus $(a_n)_{n \in \mathbb{N}}$ is convergent with limit $\inf S$. ■

Examples.

1. We have shown that the sequence $(a_n)_{n \in \mathbb{N}}$ defined by

$$a_n = \frac{1}{1^1} + \frac{1}{2^2} + \frac{1}{3^3} + \cdots + \frac{1}{n^n}, \quad n \in \mathbb{N}$$

is monotone (indeed, it is increasing since $a_{n+1} - a_n = \frac{1}{(n+1)^{(n+1)}} > 0$ for all $n \in \mathbb{N}$) and bounded (see item 5 on page 14). Thus it follows from Theorem 1.2.3 that this sequence⁵ is convergent.

2. The following table gives a summary of the valid implications, and gives counterexamples to implications which are not true.

Question	Answer	Reason/Counterexample
Is every convergent sequence bounded?	Yes	Theorem 1.2.2
Is every bounded sequence convergent?	No	$((-1)^n)_{n \in \mathbb{N}}$ is bounded, but not convergent.
Is every convergent sequence monotone?	No	$\left(\frac{(-1)^n}{n}\right)_{n \in \mathbb{N}}$ is convergent, but not monotone: $-1 < \frac{1}{2} > -\frac{1}{3}$.
Is every monotone sequence convergent?	No	$(n)_{n \in \mathbb{N}}$ is not convergent.
Is every bounded AND monotone sequence convergent?	Yes	Theorem 1.2.3

◇

⁵Although it is known that this sequence is convergent to some limit $L \in \mathbb{R}$, it is so far not even known if the limit L is rational or irrational, and this is still an open problem in mathematics! Also associated with this sequence

is the interesting identity $\sum_{n=1}^{\infty} \frac{1}{n^n} = \int_0^1 \frac{1}{x^x} dx$, the proof of which is beyond the scope of this course.

Exercises.

1. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence defined by

$$a_1 = 1 \text{ and } a_n = \frac{2n+1}{3n} a_{n-1} \text{ for } n \geq 2.$$

Prove that $(a_n)_{n \in \mathbb{N}}$ is convergent.

2. If $(b_n)_{n \in \mathbb{N}}$ is a bounded sequence, then prove that $(\frac{b_n}{n})_{n \in \mathbb{N}}$ is a convergent sequence with limit 0.
3. (a) (*) If $(a_n)_{n \in \mathbb{N}}$ is a convergent sequence with limit L , then prove that the sequence $(s_n)_{n \in \mathbb{N}}$, where

$$s_n = \frac{a_1 + \cdots + a_n}{n}, \quad n \in \mathbb{N},$$

is also convergent with limit L .

- (b) Give an example of a sequence $(a_n)_{n \in \mathbb{N}}$ such that $(s_n)_{n \in \mathbb{N}}$ is convergent but $(a_n)_{n \in \mathbb{N}}$ is divergent.
4. (*) Given a bounded sequence $(a_n)_{n \in \mathbb{N}}$, define

$$l_k = \inf\{a_n \mid n \geq k\} \text{ and } u_k = \sup\{a_n \mid n \geq k\}, \quad k \in \mathbb{N}.$$

Show that the sequences $(l_n)_{n \in \mathbb{N}}$, $(u_n)_{n \in \mathbb{N}}$ are bounded and monotone, and conclude that they are convergent. (Their respective limits are denoted by $\liminf_{n \rightarrow \infty} a_n$ and $\limsup_{n \rightarrow \infty} a_n$.)

5. (*) Recall the definition of a Cauchy sequence from Exercise 6 on page 13. Prove that every Cauchy sequence is bounded.

1.2.4 Algebra of limits

In this section we will learn that if we ‘algebraically’ combine the terms of convergent sequences, then the new sequence which is obtained, is again convergent, and moreover the limit of this sequence is the same algebraic combination of the limits. In this manner we can sometimes prove the convergence of complicated sequences by breaking them down and writing them as an algebraic combination of simple sequences. Thus, we conveniently apply arithmetic rules to compute the limits of sequences if the terms are the sum, product, quotient of terms of simpler sequences with a known limit. For instance, using the formal definition of a limit, one can show that the sequence $(a_n)_{n \in \mathbb{N}}$ defined by

$$a_n = \frac{4n^2 + 9}{3n^2 + 7n + 11}$$

converges to $\frac{4}{3}$. However, it is simpler to observe that

$$a_n = \frac{n^2 \left(4 + \frac{9}{n^2}\right)}{n^2 \left(3 + \frac{7}{n} + \frac{11}{n^2}\right)} = \frac{4 + \frac{9}{n^2}}{3 + \frac{7}{n} + \frac{11}{n^2}},$$

where the terms $\frac{9}{n^2}$, $\frac{7}{n}$, $\frac{11}{n^2}$ all have limit 0, and by a repeated application of Theorem 1.2.4 given below, we obtain

$$\lim_{n \rightarrow \infty} a_n = \frac{\lim_{n \rightarrow \infty} \left(4 + \frac{9}{n^2}\right)}{\lim_{n \rightarrow \infty} \left(3 + \frac{7}{n} + \frac{11}{n^2}\right)} = \frac{\lim_{n \rightarrow \infty} 4 + \lim_{n \rightarrow \infty} \frac{9}{n^2}}{\lim_{n \rightarrow \infty} 3 + \lim_{n \rightarrow \infty} \frac{7}{n} + \lim_{n \rightarrow \infty} \frac{11}{n^2}} = \frac{4 + 0}{3 + 0 + 0} = \frac{4}{3}.$$

Theorem 1.2.4 *If $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ are convergent sequences, then the following hold:*

1. For all $\alpha \in \mathbb{R}$, $(\alpha a_n)_{n \in \mathbb{N}}$ is a convergent sequence and $\lim_{n \rightarrow \infty} \alpha a_n = \alpha \lim_{n \rightarrow \infty} a_n$.
2. $(|a_n|)_{n \in \mathbb{N}}$ is a convergent sequence and $\lim_{n \rightarrow \infty} |a_n| = \left| \lim_{n \rightarrow \infty} a_n \right|$.
3. $(a_n + b_n)_{n \in \mathbb{N}}$ is a convergent sequence and $\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$.
4. $(a_n b_n)_{n \in \mathbb{N}}$ is a convergent sequence and $\lim_{n \rightarrow \infty} a_n b_n = \left(\lim_{n \rightarrow \infty} a_n \right) \left(\lim_{n \rightarrow \infty} b_n \right)$.
5. For all $k \in \mathbb{N}$, $(a_n^k)_{n \in \mathbb{N}}$ is a convergent sequence and $\lim_{n \rightarrow \infty} a_n^k = \left(\lim_{n \rightarrow \infty} a_n \right)^k$.
6. If for all $n \in \mathbb{N}$, $b_n \neq 0$ and $\lim_{n \rightarrow \infty} b_n \neq 0$, then $\left(\frac{1}{b_n} \right)_{n \in \mathbb{N}}$ is convergent and moreover,

$$\lim_{n \rightarrow \infty} \frac{1}{b_n} = \frac{1}{\lim_{n \rightarrow \infty} b_n}.$$

Proof Let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ converge to L_a and L_b , respectively.

1. If $\alpha = 0$, then $\alpha a_n = 0$ for all $n \in \mathbb{N}$ and clearly $(0)_{n \in \mathbb{N}}$ is a convergent sequence with limit 0. Thus

$$\lim_{n \rightarrow \infty} \alpha a_n = 0 = 0L_a = \alpha \lim_{n \rightarrow \infty} a_n.$$

If $\alpha \neq 0$, then given $\epsilon > 0$, let $N \in \mathbb{N}$ be such that for all $n > N$,

$$|a_n - L_a| < \frac{\epsilon}{|\alpha|},$$

that is

$$|\alpha a_n - \alpha L_a| = |\alpha| |a_n - L_a| < |\alpha| \frac{\epsilon}{|\alpha|} = \epsilon.$$

Hence $(\alpha a_n)_{n \in \mathbb{N}}$ is convergent with limit αL_a , that is,

$$\lim_{n \rightarrow \infty} \alpha a_n = \alpha L_a = \alpha \lim_{n \rightarrow \infty} a_n.$$

2. Given $\epsilon > 0$, let $N \in \mathbb{N}$ be such that for all $n > N$,

$$|a_n - L_a| < \epsilon.$$

Then we have for all $n > N$:

$$||a_n| - |L_a|| \leq |a_n - L_a| < \epsilon.$$

Hence $(|a_n|)_{n \in \mathbb{N}}$ is convergent with limit $|L_a|$, that is,

$$\lim_{n \rightarrow \infty} |a_n| = |L_a| = \left| \lim_{n \rightarrow \infty} a_n \right|.$$

3. Given $\epsilon > 0$, let $N_1 \in \mathbb{N}$ be such that for all $n > N_1$,

$$|a_n - L_a| < \frac{\epsilon}{2}.$$

Let $N_2 \in \mathbb{N}$ be such that for all $n > N_2$,

$$|b_n - L_b| < \frac{\epsilon}{2}.$$

Then for all $n > N := \max\{N_1, N_2\}$, we have

$$|a_n + b_n - (L_a + L_b)| = |a_n - L_a + b_n - L_b| \leq |a_n - L_a| + |b_n - L_b| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Hence $(a_n + b_n)_{n \in \mathbb{N}}$ is convergent with limit $L_a + L_b$, that is,

$$\lim_{n \rightarrow \infty} (a_n + b_n) = L_a + L_b = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n.$$

4. Note that

$$\begin{aligned} |a_n b_n - L_a L_b| &= |a_n b_n - L_a b_n + L_a b_n - L_a L_b| \\ &\leq |a_n b_n - L_a b_n| + |L_a b_n - L_a L_b| \\ &= |a_n - L_a| |b_n| + |L_a| |b_n - L_b|. \end{aligned} \tag{1.9}$$

Given $\epsilon > 0$, we need to find a N such that for all $n > N$,

$$|a_n b_n - L_a L_b| < \epsilon.$$

This can be achieved by finding a N such that each of the summands in (1.9) is less than $\frac{\epsilon}{2}$ for $n > N$. This can be done as follows.

STEP 1. Since $(b_n)_{n \in \mathbb{N}}$ is convergent, by Theorem 1.2.2 it follows that it is bounded: $\exists M > 0$ such that for all $n \in \mathbb{N}$, $|b_n| \leq M$. Let $N_1 \in \mathbb{N}$ be such that for all $n > N_1$,

$$|a_n - L_a| < \frac{\epsilon}{2M}.$$

STEP 2. Let $N_2 \in \mathbb{N}$ be such that for all $n > N_2$,

$$|b_n - L_b| < \frac{\epsilon}{2(|L_a| + 1)}.$$

Thus for $n > N := \max\{N_1, N_2\}$, we have

$$\begin{aligned} |a_n b_n - L_a L_b| &\leq |a_n - L_a| |b_n| + |L_a| |b_n - L_b| \\ &< \frac{\epsilon}{2M} M + |L_a| \frac{\epsilon}{2(|L_a| + 1)} \\ &= \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon. \end{aligned}$$

So $(a_n b_n)_{n \in \mathbb{N}}$ is a convergent sequence with limit $L_a L_b$, that is,

$$\lim_{n \rightarrow \infty} a_n b_n = L_a L_b = \left(\lim_{n \rightarrow \infty} a_n \right) \left(\lim_{n \rightarrow \infty} b_n \right).$$

5. This can be shown by using induction on k and from part 4 above. It is trivially true with $k = 1$. Suppose that it holds for some k , then $(a_n^k)_{n \in \mathbb{N}}$ is convergent and

$$\lim_{n \rightarrow \infty} a_n^k = \left(\lim_{n \rightarrow \infty} a_n \right)^k.$$

Hence by part 4 above applied to the sequences $(a_n)_{n \in \mathbb{N}}$ and $(a_n^k)_{n \in \mathbb{N}}$, we obtain that the sequence $(a_n \cdot a_n^k)_{n \in \mathbb{N}}$ is convergent and

$$\lim_{n \rightarrow \infty} a_n a_n^k = \left(\lim_{n \rightarrow \infty} a_n \right) \left(\lim_{n \rightarrow \infty} a_n^k \right) = \left(\lim_{n \rightarrow \infty} a_n \right) \left(\lim_{n \rightarrow \infty} a_n \right)^k = \left(\lim_{n \rightarrow \infty} a_n \right)^{k+1}.$$

Thus $(a_n^{k+1})_{n \in \mathbb{N}}$ is convergent and

$$\lim_{n \rightarrow \infty} a_n^{k+1} = \left(\lim_{n \rightarrow \infty} a_n \right)^{k+1}.$$

6. Let $N_1 \in \mathbb{N}$ be such that for all $n > N_1$,

$$|b_n - L_b| < \frac{|L_b|}{2}.$$

Thus for all $n > N_1$,

$$|L_b| - |b_n| \leq ||L_b| - |b_n|| \leq |b_n - L_b| < \frac{|L_b|}{2},$$

and so $|b_n| \geq \frac{|L_b|}{2}$. Let $\epsilon > 0$, and let $N_2 \in \mathbb{N}$ be such that for all $n > N_2$,

$$|b_n - L_b| < \frac{\epsilon |L_b|^2}{2}.$$

Hence for $n > N := \max\{N_1, N_2\}$, we have

$$\left| \frac{1}{b_n} - \frac{1}{L_b} \right| = \frac{|b_n - L_b|}{|b_n| |L_b|} < \frac{\epsilon |L_b|^2}{2} \frac{2}{|L_b|} \frac{1}{|L_b|} = \epsilon.$$

So $\left(\frac{1}{b_n} \right)_{n \in \mathbb{N}}$ is convergent and $\lim_{n \rightarrow \infty} \frac{1}{b_n} = \frac{1}{L_b} = \frac{1}{\lim_{n \rightarrow \infty} b_n}$.

■

Exercises.

1. Recall the convergent sequence $(a_n)_{n \in \mathbb{N}}$ from Exercise 1 on page 17 defined by

$$a_1 = 1 \text{ and } a_n = \frac{2n+1}{3n} a_{n-1} \text{ for } n \geq 2.$$

What is its limit?

HINT: If $(a_n)_{n \in \mathbb{N}}$ is a convergent sequence with limit L , then $(a_{n+1})_{n \in \mathbb{N}}$ is also a convergent sequence with limit L .

2. Suppose that the sequence $(a_n)_{n \in \mathbb{N}}$ is convergent, and assume that the sequence $(b_n)_{n \in \mathbb{N}}$ is bounded. Prove that the sequence $(c_n)_{n \in \mathbb{N}}$ defined by

$$c_n = \frac{a_n b_n + 5n}{a_n^2 + n}$$

is convergent, and find its limit.

3. (a) Let $(a_n)_{n \in \mathbb{N}}$ be a convergent sequence with limit L and suppose that $a_n \geq 0$ for all $n \in \mathbb{N}$. Prove that the sequence $(\sqrt{a_n})_{n \in \mathbb{N}}$ is also convergent, with limit \sqrt{L} .

HINT: First show that $L \geq 0$. Let $\epsilon > 0$. If $L = 0$, then choose $N \in \mathbb{N}$ large enough so that for $n > N$, $|a_n - L| = a_n < \epsilon^2$. If $L > 0$, then choose $N \in \mathbb{N}$ large enough so that for $n > N$, $|\sqrt{a_n} - \sqrt{L}| |\sqrt{a_n} + \sqrt{L}| = |a_n - L| < \epsilon \sqrt{L}$.

(b) Show that $(\sqrt{n^2 + n} - n)_{n \in \mathbb{N}}$ is a convergent sequence and find its limit.

HINT: 'Rationalize the numerator' by using $\sqrt{n^2 + n} + n$.

4. Prove that if $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ are convergent sequences such that for all $n \in \mathbb{N}$, $a_n \leq b_n$, then

$$\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n.$$

HINT: Use Exercise 5 on page 13.

1.2.5 Sandwich theorem

Another useful theorem that is useful in proving that sequences are convergent and in determining their limits is the so-called sandwich theorem. Roughly speaking, it says that if a sequence is sandwiched between two convergent limits with the *same* limit, then the sandwiched sequence is also convergent with the same limit.

Theorem 1.2.5 (Sandwich theorem.) *Let $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$ be convergent sequences with the same limit, that is,*

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n.$$

If $(c_n)_{n \in \mathbb{N}}$ is a third sequence such that

$$\text{for all } n \in \mathbb{N}, \quad a_n \leq c_n \leq b_n,$$

then $(c_n)_{n \in \mathbb{N}}$ is also convergent with the same limit, that is,

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = \lim_{n \rightarrow \infty} b_n.$$

Proof Let L denote the common limit of $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$:

$$\lim_{n \rightarrow \infty} a_n = L = \lim_{n \rightarrow \infty} b_n.$$

Given $\epsilon > 0$, let $N_1 \in \mathbb{N}$ be such that for all $n > N_1$, $|a_n - L| < \epsilon$. Hence for $n > N_1$,

$$L - a_n \leq |L - a_n| = |a_n - L| < \epsilon,$$

and so $L - a_n < \epsilon$, that is,

$$L - \epsilon < a_n.$$

Let $N_2 \in \mathbb{N}$ be such that for all $n > N_2$, $|b_n - L| < \epsilon$. So for $n > N_2$, $b_n - L < \epsilon$, that is,

$$b_n < L + \epsilon.$$

Thus for $n > N := \max\{N_1, N_2\}$, we have

$$L - \epsilon < a_n \leq c_n \leq b_n < L + \epsilon,$$

and so $L - \epsilon < c_n < L + \epsilon$. Consequently, $c_n - L < \epsilon$ and $-(c_n - L) < \epsilon$, and so

$$|c_n - L| < \epsilon.$$

This proves that $(c_n)_{n \in \mathbb{N}}$ is convergent with limit L . ■

Examples.

1. $\lim_{n \rightarrow \infty} \frac{n}{10^n} = 0.$

It can be shown by induction that for all $n \in \mathbb{N}$, $n^2 < 10^n$.

Consequently, we have

$$0 \leq \frac{n}{10^n} \leq \frac{n}{n^2} = \frac{1}{n}.$$

Since $\lim_{n \rightarrow \infty} 0 = 0 = \lim_{n \rightarrow \infty} \frac{1}{n}$, from the Sandwich theorem it follows that the sequence $(\frac{n}{10^n})_{n \in \mathbb{N}}$ is convergent and

$$\lim_{n \rightarrow \infty} \frac{n}{10^n} = 0.$$

Thus the sequence $\frac{1}{10}, \frac{2}{100}, \frac{3}{1000}, \frac{4}{10000}, \dots$ is convergent with limit 0.

2. $\lim_{n \rightarrow \infty} 2^{\frac{1}{n}} = 1.$

$2 > 1$ and so $2^{\frac{1}{n}} > 1$ (for otherwise $2 = (2^{\frac{1}{n}})^n \leq 1$, a contradiction). Let $a_n := 2^{\frac{1}{n}} - 1 \geq 0$. Then $2 = (1 + a_n)^n \geq 1 + na_n$ (It can be shown using induction that for all real $x \geq -1$ and for all $n \in \mathbb{N}$, $(1 + x)^n \geq 1 + nx$.) Hence

$$0 \leq a_n \leq \frac{1}{n}$$

and so using the Sandwich theorem it follows that $\lim_{n \rightarrow \infty} a_n = 0$. Consequently, $\lim_{n \rightarrow \infty} 2^{\frac{1}{n}} = 1$, that is, the sequence $2, \sqrt{2}, \sqrt[3]{2}, \sqrt[4]{2}, \sqrt[5]{2}, \dots$ is convergent with limit 1.

3. For any $a, b \in \mathbb{R}$, $\lim_{n \rightarrow \infty} (|a|^n + |b|^n)^{\frac{1}{n}} = \max\{|a|, |b|\}.$

Clearly,

$$(\max\{|a|, |b|\})^n \leq |a|^n + |b|^n \leq (\max\{|a|, |b|\})^n + (\max\{|a|, |b|\})^n$$

and so

$$\max\{|a|, |b|\} \leq (|a|^n + |b|^n)^{\frac{1}{n}} \leq 2^{\frac{1}{n}} \max\{|a|, |b|\}.$$

So using the Sandwich theorem, it follows that $\lim_{n \rightarrow \infty} (|a|^n + |b|^n)^{\frac{1}{n}} = \max\{|a|, |b|\}.$

In particular, with $a = 24$ and $b = 2005$, we have that $\lim_{n \rightarrow \infty} (24^n + 2005^n)^{\frac{1}{n}} = 2005$, that is, the sequence $2028, 2005.1436, 2005.001146260873, \dots$ is convergent with limit 2005.

4. If $a \in (0, 1)$, then $\lim_{n \rightarrow \infty} a^n = 0.$

Since $0 < a < 1$, it follows that $1 < \frac{1}{a}$ and so $h := \frac{1}{a} - 1 > 0$. Then we have

$$\frac{1}{a^n} = (1 + h)^n \geq 1 + nh \geq nh$$

and so

$$0 \leq a^n \leq \frac{1}{nh}.$$

Hence from the Sandwich theorem, it follows that $\lim_{n \rightarrow \infty} a^n = 0.$

5. $\lim_{n \rightarrow \infty} \left(\frac{1}{n^2 + 1} + \frac{1}{n^2 + 2} + \dots + \frac{1}{n^2 + n} \right) = 0.$

For all $n \in \mathbb{N}$, we have

$$\frac{n}{n^2 + n} \leq \frac{1}{n^2 + 1} + \frac{1}{n^2 + 2} + \dots + \frac{1}{n^2 + n} \leq \frac{n}{n^2 + 1},$$

and since

$$\lim_{n \rightarrow \infty} \frac{n}{n^2 + n} = 0 = \lim_{n \rightarrow \infty} \frac{n}{n^2 + 1}.$$

it follows from the Sandwich theorem that $\lim_{n \rightarrow \infty} \left(\frac{1}{n^2 + 1} + \frac{1}{n^2 + 2} + \dots + \frac{1}{n^2 + n} \right) = 0. \diamond$

Exercises.

1. Prove that the sequence $\left(\frac{n!}{n^n} \right)_{n \in \mathbb{N}}$ is convergent and $\lim_{n \rightarrow \infty} \frac{n!}{n^n} = 0.$

HINT: Observe that $0 \leq \frac{n!}{n^n} = \frac{1}{n} \cdot \frac{2}{n} \cdot \dots \cdot \frac{n}{n} \leq \frac{1}{n} \cdot 1 \cdot \dots \cdot 1 \leq \frac{1}{n}.$

2. Prove that for all $k \in \mathbb{N}$, the sequence

$$\left(\frac{1^k + 2^k + 3^k + \cdots + n^k}{n^{k+2}} \right)_{n \in \mathbb{N}}$$

is convergent and

$$\lim_{n \rightarrow \infty} \frac{1^k + 2^k + 3^k + \cdots + n^k}{n^{k+2}} = 0.$$

3. (a) Using induction, prove that if $x \geq -1$ and $n \in \mathbb{N}$, then

$$(1+x)^n \geq 1+nx. \quad (1.10)$$

(b) Show that for all $n \in \mathbb{N}$,

$$1 \leq n^{\frac{1}{n}} < (1 + \sqrt{n})^{\frac{2}{n}} \leq \left(1 + \frac{1}{\sqrt{n}}\right)^2.$$

HINT: Take $x = \frac{1}{\sqrt{n}}$ in the inequality (1.10).

(c) Prove that $(n^{\frac{1}{n}})_{n \in \mathbb{N}}$ is convergent and find its limit.

4. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence contained in the interval (a, b) (that is, for all $n \in \mathbb{N}$, $a < a_n < b$). If $(a_n)_{n \in \mathbb{N}}$ is convergent with limit L , then prove that $L \in [a, b]$.

HINT: Use Exercise 5 on page 13.

Give an example to show that L needn't belong to (a, b) .

5. Let $(a_n)_{n \in \mathbb{N}}$ be a convergent sequence, and let $(b_n)_{n \in \mathbb{N}}$ satisfy $|b_n - a_n| < \frac{1}{n}$ for all $n \in \mathbb{N}$. Show that $(b_n)_{n \in \mathbb{N}}$ is also convergent. What is its limit?

HINT: Observe that $-\frac{1}{n} + a_n < b_n < a_n + \frac{1}{n}$ for all $n \in \mathbb{N}$.

1.2.6 Subsequences

In this section we prove an important result in analysis, known as the Bolzano-Weierstrass theorem, which says that every bounded sequence has a convergent 'subsequence'. We begin this section by defining what we mean by a subsequence of a sequence.

Definition. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence and let $(n_k)_{k \in \mathbb{N}}$ be a sequence of natural numbers such that $n_1 < n_2 < n_3 < \dots$. Then $(a_{n_k})_{k \in \mathbb{N}}$ is called a *subsequence* of $(a_n)_{n \in \mathbb{N}}$.

Examples.

1. $(\frac{1}{2n})_{n \in \mathbb{N}}$, $(\frac{1}{n^2})_{n \in \mathbb{N}}$, $(\frac{1}{n!})_{n \in \mathbb{N}}$ and $(\frac{1}{n^n})_{n \in \mathbb{N}}$ are all subsequences of $(\frac{1}{n})_{n \in \mathbb{N}}$. Also the sequence $(a_n)_{n \in \mathbb{N}}$ defined by

$$a_n = \frac{1}{p}, \text{ where } p \text{ is the } n\text{th prime in the infinite sequence of increasing primes } 2, 3, 5, 7, 11, \dots$$

is a subsequence of $(\frac{1}{n})_{n \in \mathbb{N}}$. The sequence

$$\frac{1}{2}, 1, \frac{1}{3}, \frac{1}{4}, \dots$$

is not a subsequence of $(\frac{1}{n})_{n \in \mathbb{N}}$.

2. The sequences $((-1)^{2n})_{n \in \mathbb{N}}$ (that is, the constant sequence $1, 1, 1, \dots$) and $((-1)^{2n-1})_{n \in \mathbb{N}}$ (that is the constant sequence $-1, -1, -1, \dots$) are subsequences of $((-1)^n)_{n \in \mathbb{N}}$. \diamond

Exercise(*). Beginning with 2 and 7, the sequence $2, 7, 1, 4, 7, 4, 2, 8, 2, 8, \dots$ is constructed by multiplying successive pairs of its terms and adjoining the result as the next one or two members of the sequence depending on whether the product is a one- or two-digit number. Thus we start with 2 and 7, giving the product 14, and so the next two terms are 1, 4. Proceeding in this manner, we get subsequent terms as follows:

$$\begin{array}{c} \underline{2, 7} \\ 2, 7, 1, 4 \\ 2, \underline{7, 1}, 4 \\ 2, 7, 1, 4, 7 \\ 2, 7, \underline{1, 4}, 7 \\ 2, 7, 1, 4, 7, 4 \\ 2, 7, 1, 4, \underline{7}, 4 \\ 2, 7, 1, 4, 7, 4, 2, 8 \\ 2, 7, 1, 4, \underline{7, 4}, 2, 8 \\ 2, 7, 1, 4, 7, 4, 2, 8, 2, 8 \\ \vdots \end{array}$$

Prove that this sequence has the constant subsequence $6, 6, 6, \dots$

HINT: Show that 6 appears an infinite number of times as follows. Since the terms 2, 8, 2, 8 are adjacent, they give rise to the adjacent terms 1, 6, 1, 6 at some point, which in turn give rise to the adjacent terms 6, 6, 6 eventually, and so on. Proceeding in this way, find out if you get a loop containing the term 6.

Theorem 1.2.6 *If $(a_n)_{n \in \mathbb{N}}$ is a convergent sequence with limit L , then any subsequence of $(a_n)_{n \in \mathbb{N}}$ is also convergent with the limit L .*

Proof Let $(a_{n_k})_{k \in \mathbb{N}}$ be a subsequence of $(a_n)_{n \in \mathbb{N}}$. Given $\epsilon > 0$, let $N \in \mathbb{N}$ be such that for all $n > N$, $|a_n - L| < \epsilon$. Since the sequence $n_1 < n_2 < n_3 < \dots$, it follows that there exists a $K \in \mathbb{N}$ such that $n_K > N$. Then for all $k > K$, $n_k > n_K > N$. Hence for $k > K$, $|a_{n_k} - L| < \epsilon$, and so $(a_{n_k})_{k \in \mathbb{N}}$ is convergent with limit L . \blacksquare

Examples.

1. $(\frac{1}{2n})_{n \in \mathbb{N}}$, $(\frac{1}{n^2})_{n \in \mathbb{N}}$, $(\frac{1}{n!})_{n \in \mathbb{N}}$ and $(\frac{1}{n^n})_{n \in \mathbb{N}}$ are convergent sequences with limit 0.
2. The sequence $((-1)^n)_{n \in \mathbb{N}}$ is divergent since the subsequence $1, 1, 1, \dots$ has limit 1, while the subsequence $-1, -1, -1, \dots$ has limit -1 . \diamond

Theorem 1.2.7 *Every sequence has a monotone subsequence.*

We first give an illustration of the idea behind this proof. Assume that $(a_n)_{n \in \mathbb{N}}$ is the given sequence. Imagine that a_n is the height of the hotel with number n , which is followed by hotel $n + 1$, and so on, along an infinite line, where at infinity there is the sea. A hotel is said to have the *seaview property* if it is higher than all hotels following it. See Figure 1.7. Now there are only

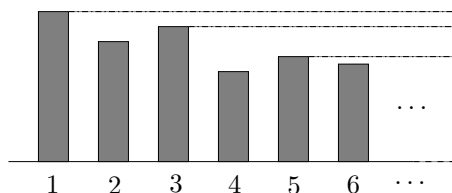


Figure 1.7: The seaview property.

two possibilities:

1° There are infinitely many hotels with the seaview property. Then their heights form a decreasing subsequence.

2° There is only a finite number of hotels with the seaview property. Then after the last hotel with the seaview property, one can start with any hotel and then always find one that is at least as high, which is taken as the next hotel, and then finding yet another that is at least as high as that one, and so on. The heights of these hotels form an increasing subsequence.

Proof Let

$$S = \{m \in \mathbb{N} \mid \text{for all } n > m, a_n < a_m\}.$$

Then we have the following two cases.

1° S is infinite. Arrange the elements of S in increasing order: $n_1 < n_2 < n_3 < \dots$. Then $(a_{n_k})_{k \in \mathbb{N}}$ is a decreasing subsequence of $(a_n)_{n \in \mathbb{N}}$.

2° S is finite. If S empty, then define $n_1 = 1$, and otherwise let $n_1 = \max S + 1$. Define inductively

$$n_{k+1} = \min\{m \in \mathbb{N} \mid m > n_k \text{ and } a_m \geq a_{n_k}\}.$$

(The minimum exists since the set $\{m \in \mathbb{N} \mid m > n_k \text{ and } a_m \geq a_{n_k}\}$ is a nonempty subset of \mathbb{N} : indeed otherwise if it were empty, then $n_k \in S$, and this is not possible if S was empty, and also impossible if S was not empty, since $n_k > \max S$.) Then $(a_{n_k})_{k \in \mathbb{N}}$ is an increasing subsequence of $(a_n)_{n \in \mathbb{N}}$. ■

An important consequence of the above theorem is the following result.

Theorem 1.2.8 (Bolzano-Weierstrass theorem.) *Every bounded sequence has a convergent subsequence.*

Proof Let $(a_n)_{n \in \mathbb{N}}$ be a bounded sequence. Then there exists a $M > 0$ such that for all $n \in \mathbb{N}$, $|a_n| \leq M$. From Theorem 1.2.7 above, it follows that the sequence $(a_n)_{n \in \mathbb{N}}$ has a monotone subsequence $(a_{n_k})_{k \in \mathbb{N}}$. Then clearly for all $k \in \mathbb{N}$, $|a_{n_k}| \leq M$ and so the sequence $(a_{n_k})_{k \in \mathbb{N}}$ is also bounded. Since $(a_{n_k})_{k \in \mathbb{N}}$ is monotone and bounded, it follows from Theorem 1.2.3 that it is convergent. ■

Example. Consider the sequence $(a_n)_{n \in \mathbb{N}}$ of fractional parts of integral multiples of $\sqrt{2}$, defined by

$$a_n = n\sqrt{2} - \lfloor n\sqrt{2} \rfloor, \text{ for } n \in \mathbb{N},$$

where for $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the *greatest integer part of x* , which is defined as the largest integer less than or equal to x :

$$\lfloor x \rfloor = \min\{z \in \mathbb{Z} \mid x < z\} - 1.$$

(By the Archimedean property, there exist natural numbers n_1 and n_2 such that $x < n_1$ and $-x < n_2$, that is, $-n_2 < x < n_1$. Hence the set $\{z \in \mathbb{Z} \mid x < z\}$ is a nonempty subset of \mathbb{Z} (indeed, n_1 belongs to it!), and it is bounded below (by $-n_2 \in \mathbb{Z}$), and so the minimum of the set $\{z \in \mathbb{Z} \mid x < z\}$ exists in \mathbb{Z} . Consequently $\lfloor \cdot \rfloor$ is a well-defined function.)

The terms of the sequence $(a_n)_{n \in \mathbb{N}}$ are as follows:

$$\begin{array}{ll} \sqrt{2} = 1.414213\dots & a_1 = 0.414213\dots \\ 2\sqrt{2} = 2.828427\dots & a_2 = 0.828427\dots \\ 3\sqrt{2} = 4.242640\dots & a_3 = 0.242640\dots \\ 4\sqrt{2} = 5.656854\dots & a_4 = 0.656854\dots \\ 5\sqrt{2} = 7.071067\dots & a_5 = 0.071067\dots \\ 6\sqrt{2} = 8.485281\dots & a_6 = 0.485281\dots \\ & \vdots \end{array}$$

The sequence $(a_n)_{n \in \mathbb{N}}$ is bounded: indeed, $0 \leq a_n < 1$. So by the Bolzano-Weierstrass theorem it has a convergent subsequence⁶. \diamond

Exercise. (*) Recall the definition of a Cauchy sequence from Exercise 6 on page 13, where we had already seen that every convergent sequence is Cauchy. Using Bolzano-Weierstrass theorem, we can prove the converse. Show that if a sequence is Cauchy, then it is convergent.

HINT: Proceed as follows. Let $(a_n)_{n \in \mathbb{N}}$ be a Cauchy sequence. From Exercise 5 on page 17, it follows that $(a_n)_{n \in \mathbb{N}}$ is bounded. By the Bolzano Weierstrass theorem, it follows that $(a_n)_{n \in \mathbb{N}}$ has a convergent subsequence, say $(a_{n_k})_{k \in \mathbb{N}}$ with limit L . Prove (using the fact that $(a_n)_{n \in \mathbb{N}}$ is Cauchy), that then $(a_n)_{n \in \mathbb{N}}$ is itself convergent with limit L .

1.3 Continuity

A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is a rule of correspondence that assigns to each real number a unique real number. Many bizarre functions make appearances in analysis, and in order to avoid falling into pitfalls with simplistic thinking, we need definitions and hypothesis of theorems to be stated carefully and clearly.

Within the huge collection of functions, there is an important subset: the continuous functions. Continuous functions play a prominent role in analysis since they possess some useful properties.

In this section we give the formal definition of continuous functions and prove two of the most important properties: the extreme value theorem and the intermediate value theorem.

⁶In fact, it can be shown that these fractional parts a_n are “dense” in $(0, 1)$. Thus given any number $L \in (0, 1)$, there exists a subsequence of the sequence $(a_n)_{n \in \mathbb{N}}$ above that converges to L .

1.3.1 Definition of continuity

In everyday speech, a ‘continuous’ process is one that proceeds without gaps or interruptions or sudden changes. What does it mean for a function $f : \mathbb{R} \rightarrow \mathbb{R}$ to be continuous? The common informal definition of this concept states that a function f is continuous if one can sketch its graph without lifting the pencil. In other words, the graph of f has no breaks in it. If a break does occur in the graph, then this break will occur at some point. Thus (based on this visual view of continuity), we first give the formal definition of the continuity of a function *at a point* below. Next, if a function is continuous at *each* point, then it will be called continuous.

If a function has a break at a point, say c , then even if points x are close to c , the points $f(x)$ do not get close to $f(c)$. See Figure 1.8.

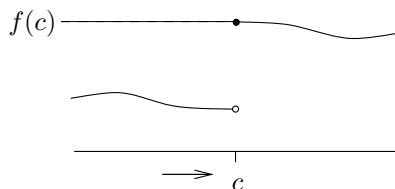


Figure 1.8: A function with a break at c . If x lies to the left of c , then $f(x)$ is not close to $f(c)$, no matter how close x comes to c .

This motivates the following definition of continuity, which guarantees that if a function is continuous at a point c , then we can make $f(x)$ as close as we like to $f(c)$, by choosing x sufficiently close to c . See Figure 1.9.

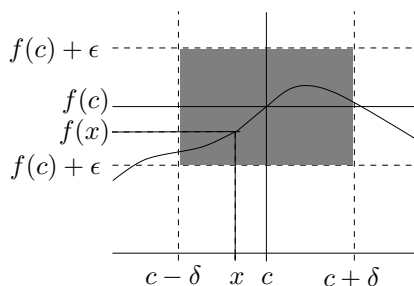


Figure 1.9: The definition of the continuity of a function at point c . If the function is continuous at c , then given any $\epsilon > 0$ (which determines a strip around the line $y = f(c)$ of width 2ϵ), there exists a $\delta > 0$ (which determines an interval of width 2δ around the point c) such that whenever x lies in this width (so that x satisfies $c - \delta < x < c + \delta$, that is, $|x - c| < \delta$), then $f(x)$ satisfies $f(c) - \epsilon < f(x) < f(c) + \epsilon$, that is, $|f(x) - f(c)| < \epsilon$.

Definitions. Let I be an interval in \mathbb{R} and let $c \in I$. A function $f : I \rightarrow \mathbb{R}$ is *continuous at c* if for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in I$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$.

A function $f : I \rightarrow \mathbb{R}$ is *continuous (on I)* if for every $c \in I$, f is continuous at c .

Examples.

1. $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = 1$ for all $x \in \mathbb{R}$ is continuous.

Let $c \in \mathbb{R} = (-\infty, \infty)$. Given $\epsilon > 0$, let δ be an arbitrary positive real number; for instance, let $\delta = 1$. Then if $x \in \mathbb{R}$ and $|x - c| < \delta = 1$, we have:

$$|f(x) - f(c)| = |1 - 1| = |0| = 0 < \epsilon.$$

So f is continuous at c . Since the choice of $c \in \mathbb{R}$ was arbitrary, it follows that f is continuous on \mathbb{R} . See Figure 1.10.

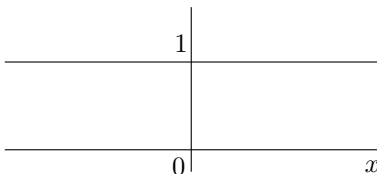


Figure 1.10: The continuous constant function 1.

2. $f : \mathbb{R} \rightarrow \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \in \mathbb{R} \setminus \{0\}. \end{cases}$$

is not continuous at 0.

Suppose that it is continuous at 0. Then given $\epsilon = \frac{1}{2} > 0$, let $\delta > 0$ be such that if $x \in \mathbb{R}$ and $|x| = |x - 0| < \delta$, then $|f(x) - f(0)| = |f(x) - 0| = |f(x)| < \epsilon = \frac{1}{2}$. But now take $x = \frac{\delta}{2} \in \mathbb{R}$, and so $|x| = |\frac{\delta}{2}| = \frac{\delta}{2} < \delta$. Thus $|f(x)| = |f(\frac{\delta}{2})| = |1| = 1 > \frac{1}{2} = \epsilon$, a contradiction. So f cannot be continuous at 0.

However, for all $c \in \mathbb{R} \setminus \{0\}$, f is continuous at c . This can be seen as follows. Given $\epsilon > 0$, let $\delta = \frac{|c|}{2} > 0$. Then if $x \in \mathbb{R}$ and $|x - c| < \delta$, we have

$$|c| - |x| \leq ||c| - |x|| \leq |c - x| = |x - c| < \delta = \frac{|c|}{2}$$

and so

$$|x| > \frac{|c|}{2} > 0.$$

Thus $x \neq 0$ and so $f(x) = 1$. Hence if $x \in \mathbb{R}$ and $|x - c| < \delta$, we obtain

$$|f(x) - f(c)| = |1 - 1| = |0| = 0 < \epsilon.$$

Consequently f is continuous at c . See Figure 1.11.

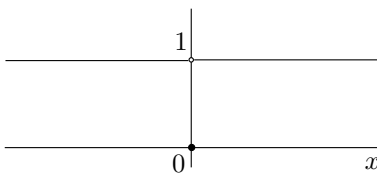


Figure 1.11: A function continuous everywhere except at 0.

3. $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = x$ for all $x \in \mathbb{R}$ is continuous.

Let $c \in \mathbb{R}$. Given $\epsilon > 0$, let $\delta = \epsilon$. Then if $x \in \mathbb{R}$ and $|x - c| < \delta$, we have:

$$|f(x) - f(c)| = |x - c| < \delta = \epsilon.$$

So f is continuous at c . Since the choice of $c \in \mathbb{R}$ was arbitrary, it follows that f is continuous on \mathbb{R} .

4. $f : (0, 1) \rightarrow \mathbb{R}$ given by $f(x) = \frac{1}{x}$ for all $x \in \mathbb{R}$ is continuous (on $(0, 1)$).

Let $c \in (0, 1)$. Given $\epsilon > 0$, let $\delta = \min \left\{ \frac{\epsilon}{2}, \frac{\epsilon c^2}{2} \right\}$ (> 0). Then if $x \in (0, 1)$ and $|x - c| < \delta$, we have

$$|c| - |x| \leq ||c| - |x|| \leq |c - x| = |x - c| < \delta \leq \frac{|c|}{2}$$

and so

$$\frac{|c|}{2} < |x|, \text{ that is, } \frac{1}{|x|} < \frac{2}{|c|}.$$

Consequently, if $x \in (0, 1)$ and $|x - c| < \delta$,

$$\left| \frac{1}{x} - \frac{1}{c} \right| = \frac{|c - x|}{|x| |c|} = \frac{|x - c|}{|x| |c|} < \delta \cdot \frac{2}{|c|} \cdot \frac{1}{|c|} = \frac{2\delta}{c^2} \leq \epsilon.$$

So f is continuous at c . Since the choice of $c \in (0, 1)$ was arbitrary, it follows that f is continuous (on $(0, 1)$). \diamond

Exercises.

1. Let the function $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $f(x) = x^2$.

- (a) Prove that f is continuous at 0.
 (b) (*) Suppose that c is a nonzero real number. Prove that f is continuous at c .

In Exercise 1 on page 31, we will give a slick proof of the fact that f is continuous on \mathbb{R} .

2. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function that satisfies $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$.

- (a) Suppose that f is continuous at some real number c . Prove that f is continuous on \mathbb{R} .
 HINT: Since f is continuous at c , given $\epsilon > 0$, $\exists \delta > 0$ such that for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$. Show that given any other point $c' \in \mathbb{R}$, the function f is continuous at c' by showing that the same δ works (for this ϵ).
 (b) Give an example of such a function.

3. Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ and there exists a $M > 0$ such that for all $x \in \mathbb{R}$, $|f(x)| \leq M|x|$. Prove that f is continuous at 0.

HINT: Find $f(0)$.

4. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be defined by

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

Prove that for every $c \in \mathbb{R}$, f is not continuous at c .

HINT: Use the fact that there are irrational numbers arbitrarily close to any rational number and rational numbers arbitrarily close to any irrational number.

5. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Prove that if for some $c \in \mathbb{R}$, $f(c) > 0$, then there exists a $\delta > 0$ such that for all $x \in (c - \delta, c + \delta)$, $f(x) > 0$.

If f and g are functions on \mathbb{R} , then the composition of g with f , denoted by $g \circ f$, is defined by

$$(g \circ f)(x) = g(f(x)), \quad x \in \mathbb{R}.$$

The following theorem implies that the composition of continuous functions is continuous.

Theorem 1.3.1 *Let $c \in \mathbb{R}$. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous at c and $g : \mathbb{R} \rightarrow \mathbb{R}$ is continuous at $f(c)$, then $g \circ f$ is continuous at c .*

Proof Since g is continuous at $f(c)$, it follows that given $\epsilon > 0$, $\exists \delta > 0$ such that for all $y \in \mathbb{R}$ satisfying $|y - f(c)| < \delta$,

$$|g(y) - g(f(c))| < \epsilon.$$

Since f is continuous at c , $\exists \delta_1 > 0$ such that for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta_1$,

$$|f(x) - f(c)| < \delta.$$

Consequently, for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta_1$, we have $|f(x) - f(c)| < \delta$, and so

$$|g(f(x)) - g(f(c))| < \epsilon,$$

that is,

$$|(g \circ f)(x) - (g \circ f)(c)| < \epsilon.$$

Hence $g \circ f$ is continuous at c . ■

1.3.2 Continuous functions preserve convergent sequences

We now give an alternative characterization of continuity.

Theorem 1.3.2 *Let I be an interval in \mathbb{R} and let $c \in I$. Suppose that $f : I \rightarrow \mathbb{R}$ is a function. Then f is continuous at c iff*

$$\boxed{\text{for every convergent sequence } (x_n)_{n \in \mathbb{N}} \text{ contained in } I \text{ with limit } c, (f(x_n))_{n \in \mathbb{N}} \text{ is convergent and } \lim_{n \rightarrow \infty} f(x_n) = f(c).} \quad (1.11)$$

Proof

ONLY IF: Suppose that f is continuous at $c \in I$ and let $(x_n)_{n \in \mathbb{N}}$ be a convergent sequence contained in I with limit c .

Since f is continuous at $c \in I$, given $\epsilon > 0$, $\exists \delta > 0$ such that for all $x \in I$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$.

Since $(x_n)_{n \in \mathbb{N}}$ is convergent with limit c , $\exists N \in \mathbb{N}$ such that for all $n > N$, $|x_n - c| < \delta$.

Consequently for $n > N$, $|f(x_n) - f(c)| < \epsilon$. So $(f(x_n))_{n \in \mathbb{N}}$ is convergent with limit $f(c)$.

IF: Suppose that (1.11) holds. Then we need to show that f is continuous at c and we prove this by contradiction. Assume that f is not continuous at c , that is,

$$\neg [\forall \epsilon > 0 \exists \delta > 0 \text{ such that } \forall x \in I \text{ such that } |x - c| < \delta, |f(x) - f(c)| < \epsilon]$$

that is,

$$\exists \epsilon > 0 \text{ such that } \forall \delta > 0 \exists x \in I \text{ such that } |x - c| < \delta \text{ but } |f(x) - f(c)| \geq \epsilon.$$

Hence if $\delta = \frac{1}{n}$, then we can find $x_n \in I$ such that $|x_n - c| < \delta = \frac{1}{n}$, but $|f(x_n) - f(c)| \geq \epsilon$.

CLAIM 1: The sequence $(x_n)_{n \in \mathbb{N}}$ is contained in I and is convergent with limit c .

Indeed, we have for all $n \in \mathbb{N}$, $x_n \in I$. Furthermore, given any $\zeta > 0$, we can find $N \in \mathbb{N}$ such that $\frac{1}{\zeta} < N$ (Archimedean property), that is, $\frac{1}{N} < \zeta$. Hence for $n > N$, $|x_n - c| < \frac{1}{N} < \zeta$. So $(x_n)_{n \in \mathbb{N}}$ is convergent with limit c .

CLAIM 2: The sequence $(f(x_n))_{n \in \mathbb{N}}$ does not converge to $f(c)$.

Indeed for all $n \in \mathbb{N}$, we have $|f(x_n) - f(c)| \geq \epsilon$. Thus for instance $\frac{\epsilon}{2} > 0$, but it is not possible to find a large enough $N \in \mathbb{N}$ such that for all $n > N$, $|f(x_n) - f(c)| < \frac{\epsilon}{2}$ (for if this were possible, then we would arrive at the contradiction $\epsilon \leq |f(x_n) - f(c)| < \frac{\epsilon}{2}$).

The CLAIMS 1 and 2 show that (1.11) does not hold, a contradiction. Hence f is continuous at c . ■

Exercises.

1. Recall Exercise 1 on page 29: The function $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ is given by $f(x) = x^2$. Using the characterization of continuity provided in Theorem 1.3.2, prove that f is continuous on \mathbb{R} .

2. Prove that⁷ if $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and $f(x) = 0$ if x is rational, then $f(x) = 0$ for all $x \in \mathbb{R}$.

HINT: Given any real number c , there exists a sequence of rational numbers $(q_n)_{n \in \mathbb{N}}$ that converges to c .

3. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a function that preserves divergent sequences, that is, for every divergent sequence $(x_n)_{n \in \mathbb{N}}$, $(f(x_n))_{n \in \mathbb{N}}$ is divergent as well. Prove that f is one-to-one.

HINT: Let x_1, x_2 be distinct real numbers, and consider the sequence $x_1, x_2, x_1, x_2, \dots$

Using the above theorem, we obtain the following useful Theorem 1.3.3. But before we state this result, we introduce some convenient notation.

Let I be an interval in \mathbb{R} . Given functions $f : I \rightarrow \mathbb{R}$ and $g : I \rightarrow \mathbb{R}$, we define the following:

1. If $\alpha \in \mathbb{R}$, then we define the function $\alpha f : I \rightarrow \mathbb{R}$ by $(\alpha f)(x) = \alpha \cdot f(x)$, $x \in I$.
2. We define the *absolute value of f* , $|f| : I \rightarrow \mathbb{R}$ by $|f|(x) = |f(x)|$, $x \in I$.
3. The *sum of f and g* , $f + g : I \rightarrow \mathbb{R}$ is defined by $(f + g)(x) = f(x) + g(x)$, $x \in I$.
4. The *product of f and g* , $fg : I \rightarrow \mathbb{R}$ is defined by $(fg)(x) = f(x)g(x)$, $x \in I$.
5. If $k \in \mathbb{N}$, then we define the *k th power of f* , $f^k : I \rightarrow \mathbb{R}$ by $f^k(x) = (f(x))^k$, $x \in I$.
6. If for all $x \in I$, $g(x) \neq 0$, then we define $\frac{1}{g} : I \rightarrow \mathbb{R}$ by $\left(\frac{1}{g}\right)(x) = \frac{1}{g(x)}$, $x \in I$.

Theorem 1.3.3 *Let I be an interval in \mathbb{R} and let $c \in I$. Suppose that $f : I \rightarrow \mathbb{R}$ and $g : I \rightarrow \mathbb{R}$ are continuous at c . Then:*

1. *For all $\alpha \in \mathbb{R}$, αf is continuous at c .*
2. *$|f|$ is continuous at c .*
3. *$f + g$ is continuous at c .*

⁷See Exercise 4 on page 29.

4. fg is continuous at c .
5. For all $k \in \mathbb{N}$, f^k is continuous at c .
6. If for all $x \in I$, $g(x) \neq 0$, then $\frac{1}{g}$ is continuous at c .

Proof Suppose that $(x_n)_{n \in \mathbb{N}}$ is a convergent sequence contained in I , with limit c . Since f and g are continuous at c , from Theorem 1.3.2, it follows that $(f(x_n))_{n \in \mathbb{N}}$ and $(g(x_n))_{n \in \mathbb{N}}$ are convergent with limits $f(c)$ and $g(c)$, respectively. Hence from Theorem 1.2.4, it follows that:

1. $(\alpha \cdot f(x_n))_{n \in \mathbb{N}}$ is convergent with limit $\alpha \cdot f(c)$, that is, $((\alpha f)(x_n))_{n \in \mathbb{N}}$ is convergent with limit $(\alpha f)(c)$. So from Theorem 1.3.2, it follows that αf is continuous at c .
2. $(|f(x_n)|)_{n \in \mathbb{N}}$ is convergent with limit $|f(c)|$, that is, $(|f|(x_n))_{n \in \mathbb{N}}$ is convergent with limit $|f|(c)$. So from Theorem 1.3.2, it follows that $|f|$ is continuous at c .
3. $(f(x_n) + g(x_n))_{n \in \mathbb{N}}$ is convergent with limit $f(c) + g(c)$, that is, $((f+g)(x_n))_{n \in \mathbb{N}}$ is convergent with limit $(f+g)(c)$. So from Theorem 1.3.2, it follows that $f+g$ is continuous at c .
4. $(f(x_n)g(x_n))_{n \in \mathbb{N}}$ is convergent with limit $f(c)g(c)$, that is, $((fg)(x_n))_{n \in \mathbb{N}}$ is convergent with limit $(fg)(c)$. So from Theorem 1.3.2, it follows that fg is continuous at c .
5. $((f(x_n))^k)_{n \in \mathbb{N}}$ is convergent with limit $(f(c))^k$, that is, $(f^k(x_n))_{n \in \mathbb{N}}$ is convergent with limit $f^k(c)$. So from Theorem 1.3.2, it follows that f^k is continuous at c .
6. $(\frac{1}{g(x_n)})_{n \in \mathbb{N}}$ is convergent with limit $\frac{1}{g(c)}$ (since for all $x \in I$, $g(x) \neq 0$, in particular $g(x_n) \neq 0$ and $g(c) \neq 0$), that is, $(\frac{1}{g}(x_n))_{n \in \mathbb{N}}$ is convergent with limit $(\frac{1}{g})(c)$. So from Theorem 1.3.2, it follows that $\frac{1}{g}$ is continuous at c .

■

Example. Since $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = x$ for $x \in \mathbb{R}$ is continuous (see Example 3 on page 28), it follows that for all $k \in \mathbb{N}$, x^k is continuous. Thus given arbitrary scalars a_0, a_1, \dots, a_N in \mathbb{R} , it follows that the functions $a_0 \cdot 1, a_1 \cdot x, \dots, a_N \cdot x^N$ are continuous. Consequently the *polynomial function* $p : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$p(x) = a_0 + a_1x + \dots + a_Nx^N, \quad x \in \mathbb{R}$$

is continuous. ◇

Exercise. Show that the rational function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$f(x) = \frac{x^2}{1+x^2}, \quad x \in \mathbb{R},$$

is continuous on \mathbb{R} .

1.3.3 Extreme value theorem

Below we show that a continuous function on an interval $[a, b]$ attains its maximum and minimum values. See Figure 1.12.

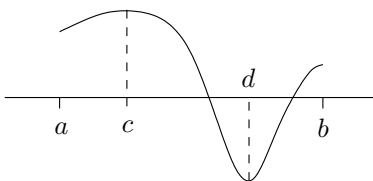


Figure 1.12: Extreme value theorem.

Theorem 1.3.4 (Extreme value theorem). *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, then there exist $c, d \in [a, b]$ such that*

$$f(c) = \sup\{f(x) \mid x \in [a, b]\}, \text{ and} \quad (1.12)$$

$$f(d) = \inf\{f(x) \mid x \in [a, b]\}. \quad (1.13)$$

Note that since $c, d \in [a, b]$, the supremum and infimum in (1.12) and (1.13) are in fact the maximum and minimum, respectively.

Proof

STEP 1. We first show that f is bounded, that is, the set

$$S = \{f(x) \mid x \in [a, b]\}$$

is bounded. Suppose that S is not bounded. Then given $n \in \mathbb{N}$, $\exists x_n \in [a, b]$ such that $|f(x_n)| > n$ (for if this fails for some $N \in \mathbb{N}$, then $|f(x)| \leq N$ for all $x \in [a, b]$, and so S would be bounded, a contradiction). In this way, we get a sequence $(x_n)_{n \in \mathbb{N}}$. Since

$$a \leq x_n \leq b \text{ for all } n \in \mathbb{N}, \quad (1.14)$$

$(x_n)_{n \in \mathbb{N}}$ is bounded, and so by the Bolzano-Weierstrass theorem (Theorem 1.2.8), it follows that it has a convergent subsequence $(x_{n_k})_{k \in \mathbb{N}}$ that converges to some limit L . Now we show that $L \in [a, b]$. If not, then either $L > b$ or $L < a$. If $L > b$, then $L - b > 0$, and so $\exists K \in \mathbb{N}$ such that for all $k > K$, $|x_{n_k} - L| < L - b$. But then for all $k > K$, we obtain

$$L - x_{n_k} \leq |L - x_{n_k}| = |x_{n_k} - L| < L - b,$$

that is, $x_{n_k} > b$, which contradicts (1.14). Similarly, if $L < a$, then $a - L > 0$, and so $\exists K \in \mathbb{N}$ such that for all $k > K$, $|x_{n_k} - L| < a - L$. But then for all $k > K$, we obtain

$$x_{n_k} - L \leq |x_{n_k} - L| < a - L,$$

that is, $x_{n_k} < a$, which contradicts (1.14). Thus $L \in [a, b]$. But by Theorem 1.2.2 it follows that $(f(x_{n_k}))_{k \in \mathbb{N}}$ cannot be convergent, since it is not bounded (indeed, $|f(x_{n_k})| > n_k!$). From Theorem 1.3.2, we see that this contradicts the continuity of f . So S must be bounded.

STEP 2. Let $M := \sup\{f(x) \mid x \in [a, b]\}$. We prove that there exists a $c \in [a, b]$ such that $f(c) = M$. For each $n \in \mathbb{N}$, $M - \frac{1}{n} < M$, and so $M - \frac{1}{n}$ cannot be an upper bound for $\{f(x) \mid x \in [a, b]\}$. So $\exists x_n \in [a, b]$ such that

$$M - \frac{1}{n} < f(x_n) \leq M. \quad (1.15)$$

By the Bolzano-Weierstrass theorem, $(x_n)_{n \in \mathbb{N}}$ has a convergent subsequence $(x_{n_k})_{k \in \mathbb{N}}$ with limit $c \in [a, b]$. Since f is continuous $(f(x_{n_k}))_{k \in \mathbb{N}}$ is convergent with limit $f(c)$. From (1.15), it follows

from the Sandwich theorem that $f(c) = M$. This proves the existence of $c \in [a, b]$ such that (1.12) holds.

STEP 3. Let $m := \inf\{f(x) \mid x \in [a, b]\}$. The proof that there exists a $d \in [a, b]$ such that $f(d) = m$ is similar to STEP 2 above, and is left as an exercise to the reader. ■

Examples.

1. The function $f : [0, 1] \rightarrow \mathbb{R}$ given by $f(x) = 1$ for all $x \in [0, 1]$ is continuous and for any $c, d \in [0, 1]$ we have

$$f(c) = 1 = \sup\{f(x) \mid x \in [0, 1]\} = \inf\{f(x) \mid x \in [0, 1]\} = 1 = f(d).$$

2. The function $f : (0, 1) \rightarrow \mathbb{R}$ given by $f(x) = x$ is continuous on $(0, 1)$. If

$$S = \{x \mid x \in (0, 1)\},$$

then $\sup S = 1$ and $\inf S = 0$, but these values are not attained. Thus the statement of Theorem 1.3.4 does not hold if $[a, b]$ is replaced by (a, b) . ◇

Exercises.

1. Complete the details in STEP 3 of the proof of Theorem 1.3.4.
2. Give an example of a function $f : [0, 1] \rightarrow \mathbb{R}$ that is not continuous on $[0, 1]$, but f satisfies the conclusion of Theorem 1.3.4.
3. A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is *periodic* if there exists a $T > 0$ such that for all $x \in \mathbb{R}$, $f(x + T) = f(x)$. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and periodic, then prove that f is bounded, that is, the set $S = \{f(x) \mid x \in \mathbb{R}\}$ is bounded.
4. Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous on $[a, b]$, and define f_* as follows:

$$f_*(x) = \begin{cases} f(a) & \text{if } x = a, \\ \max\{f(y) \mid y \in [a, x]\} & \text{if } x \in (a, b]. \end{cases}$$

- (a) Show that f_* is a well-defined function.
- (b) (*) Prove that f_* is continuous on $[a, b]$.
- (c) If $f : [-1, 1] \rightarrow \mathbb{R}$ is given by $f(x) = x - x^2$, then find f_* .

1.3.4 Intermediate value theorem

We now prove one of the most fundamental (and obvious!) theorems on continuous functions: a continuous function cannot “hop over” intermediate values. For instance, if the height of a mountain is 1976 meters above sea level, then given any number between 0 and 1976, say 399, there must exist a point on the mountain that is exactly 399 meters above sea level. See Figure 1.13.

Theorem 1.3.5 (Intermediate value theorem). *If $f : [a, b] \rightarrow \mathbb{R}$ is continuous and y is such that $f(a) \leq y \leq f(b)$, then there exists a $c \in [a, b]$ such that $f(c) = y$.*

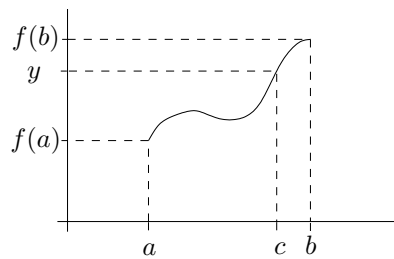


Figure 1.13: Intermediate value theorem.

Note that in the above statement of the theorem, the order $f(a) \leq y \leq f(b)$ can be reversed without changing the conclusion. Indeed, if $y \in \mathbb{R}$ is such that $f(b) \leq y \leq f(a)$, then we have

$$(-f)(a) \leq -y \leq (-f)(b),$$

and applying the above Theorem 1.3.5 to the continuous function $-f$ (Theorem 1.3.3.1 with $\alpha = -1!$), we get the existence of $c \in [a, b]$ such that $(-f)(c) = -y$, that is, $f(c) = y$.

Proof Suppose that $f(a) \leq y \leq f(b)$, and define

$$S = \{x \in [a, b] \mid y \leq f(x)\}.$$

S is not empty (since $b \in S$) and S is bounded below by a (since $S \subset [a, b]$). Thus S has an infimum, say c .

We claim that $f(c) = y$. Suppose that our claim is false. Then either $f(c) < y$ or $f(c) > y$.

1° Suppose that $f(c) < y$. Thus $\epsilon := y - f(c) > 0$, and since f is continuous at c , $\exists \delta > 0$ such that for all $x \in [a, b]$ such that $|x - c| < \delta$, $|f(x) - f(c)| < y - f(c)$. Since c is the infimum of S , $\exists x \in S$ such that $c \leq x < c + \delta$, that is, $x \in [a, b]$, $0 \leq x - c = |x - c| < \delta$ and $y \leq f(x)$. Thus we have

$$y - f(c) \leq f(x) - f(c) \leq |f(x) - f(c)| < y - f(c),$$

a contradiction.

2° Suppose that $f(c) > y$. Thus $\epsilon := f(c) - y > 0$, and since f is continuous at c , $\exists \delta > 0$ and moreover satisfying⁸ $\delta \leq c - a$, such that for all $x \in [a, b]$ such that $|x - c| < \delta$, $|f(x) - f(c)| < y - f(c)$. Let $x := c - \frac{\delta}{2}$. Then $x \in [a, c] \subset [a, b]$: indeed,

$$a \leq \frac{c+a}{2} \leq c - \frac{\delta}{2} \leq c \leq b.$$

Furthermore,

$$|x - c| = c - x = c - \left(c - \frac{\delta}{2}\right) = \frac{\delta}{2} < \delta.$$

Hence $f(c) - f(x) \leq |f(c) - f(x)| = |f(x) - f(c)| < f(c) - y$, and so $f(x) \geq y$. Consequently, $x \in S$ and since $c = \inf S$, we have $c \leq x$. Thus we obtain

$$c \leq x = c - \frac{\delta}{2} < c,$$

a contradiction.

⁸If not, then replace δ by $\frac{\delta}{N}$ with $N \in \mathbb{N}$ large enough to guarantee $\frac{\delta}{c-a} < N$.

So $f(c) = y$, and this completes the proof. ■

Examples.

1. Every odd polynomial with real coefficients has at least one real root.

Suppose that p is a polynomial with degree $2m + 1$, $m \in \mathbb{N} \cup \{0\}$. Let

$$p(x) = a_0 + a_1x + \cdots + a_{2m}x^{2m} + a_{2m+1}x^{2m+1},$$

where $a_{2m+1} \neq 0$.

In order to show that p has a real root, we will choose a large enough $N \in \mathbb{N}$ such that $p(N)$ and $p(-N)$ have opposite signs, and then restrict our attention to an interval $[-N, N]$. Then appealing to the intermediate value theorem, we can conclude that p must vanish at some point in this interval, that is, for some real $c \in [-N, N]$, $p(c) = 0$. The proof is long, and so we have divided it into a sequence of steps.

STEP 1. For large positive n , $p(n)$ has the same sign as a_{2m+1} .

Since

$$\lim_{n \rightarrow \infty} \left(\frac{a_0}{n^{2m+1}} + \frac{a_1}{n^{2m}} + \cdots + \frac{a_{2m}}{n} \right) = 0,$$

it follows that there exists $N_1 \in \mathbb{N}$ such that for all $n > N_1$,

$$\left| \frac{p(n)}{n^{2m+1}} - a_{2m+1} \right| < \frac{|a_{2m+1}|}{2}.$$

Now we show that for $n > N_1$, $p(n)$ has the same sign as a_{2m+1} . Since $a_{2m+1} \neq 0$, either $a_{2m+1} > 0$ or $a_{2m+1} < 0$. If $a_{2m+1} > 0$, then for all $n > N_1$

$$a_{2m+1} - \frac{p(n)}{n^{2m+1}} \leq \left| a_{2m+1} - \frac{p(n)}{n^{2m+1}} \right| = \left| \frac{p(n)}{n^{2m+1}} - a_{2m+1} \right| < \frac{|a_{2m+1}|}{2} = \frac{a_{2m+1}}{2},$$

and so

$$0 < \frac{a_{2m+1}}{2} < \frac{p(n)}{n^{2m+1}}.$$

Thus $p(n)$ is also positive for all $n > N_1$. Similarly, if $a_{2m+1} < 0$, then for all $n > N_1$

$$\frac{p(n)}{n^{2m+1}} - a_{2m+1} \leq \left| \frac{p(n)}{n^{2m+1}} - a_{2m+1} \right| < \frac{|a_{2m+1}|}{2} = -\frac{a_{2m+1}}{2},$$

and so

$$\frac{p(n)}{n^{2m+1}} < \frac{a_{2m+1}}{2} < 0.$$

Thus $p(n)$ is also negative for all $n > N_1$.

So it follows that for $n > N_1$, $p(n)$ has the same sign as a_{2m+1} .

STEP 2. For large positive n , $p(-n)$ has the same sign as $-a_{2m+1}$.

Similarly, using the fact that

$$\lim_{n \rightarrow \infty} \left(\frac{a_0}{(-n)^{2m+1}} + \frac{a_1}{(-n)^{2m}} + \cdots + \frac{a_{2m}}{(-n)} \right) = 0,$$

we now show that there exists $N_2 \in \mathbb{N}$ such that for all $n > N_2$, $p(-n)$ has the same sign as $-a_{2m+1}$. Let $N_2 \in \mathbb{N}$ be such that for all $n > N_2$,

$$\left| \frac{p(-n)}{n^{2m+1}} + a_{2m+1} \right| = \left| -\frac{p(-n)}{n^{2m+1}} - a_{2m+1} \right| < \frac{|a_{2m+1}|}{2}.$$

If $a_{2m+1} > 0$, then for all $n > N_2$

$$\frac{p(-n)}{n^{2m+1}} + a_{2m+1} \leq \left| \frac{p(-n)}{n^{2m+1}} + a_{2m+1} \right| < \frac{|a_{2m+1}|}{2} = \frac{a_{2m+1}}{2},$$

and so

$$\frac{p(-n)}{n^{2m+1}} < -\frac{a_{2m+1}}{2} < 0.$$

Thus $p(-n)$ is negative for all $n > N_2$. Similarly, if $a_{2m+1} < 0$, then for all $n > N_2$

$$-\frac{p(-n)}{n^{2m+1}} - a_{2m+1} \leq \left| \frac{p(-n)}{n^{2m+1}} + a_{2m+1} \right| < \frac{|a_{2m+1}|}{2} = -\frac{a_{2m+1}}{2},$$

and so

$$\frac{p(-n)}{n^{2m+1}} > -\frac{a_{2m+1}}{2} > 0.$$

Thus $p(-n)$ is positive for all $n > N_2$.

So it follows that for $n > N_2$, $p(-n)$ has the same sign as $-a_{2m+1}$.

STEP 3. Application of the intermediate value theorem and the conclusion.

Hence if $N := \max\{N_1, N_2\} + 1$, then $p(N)$ has the same sign as a_{2m+1} , while $p(-N)$ has the same sign as $-a_{2m+1}$. Thus the continuous function p must vanish at some point in the interval $[-N, N]$.

The polynomial $p(x) = x^{2005} - x^{1976} + \frac{1}{399}$ has a real root in $[-1, 1]$: indeed $p(1) = \frac{1}{399} > 0$ and $p(-1) = -2 + \frac{1}{399} < 0$ and so $\exists c \in [-1, 1]$ such that $p(c) = 0$.

- At any given time, there exists a pair of diametrically opposite points on the equator which have the same temperature.

Let $T(\Theta)$ denote the surface temperature at the point at longitude Θ . See Figure 1.14. (Note that $\Theta(0) = \Theta(2\pi)$.) Assuming that T is a continuous function of Θ , it follows that

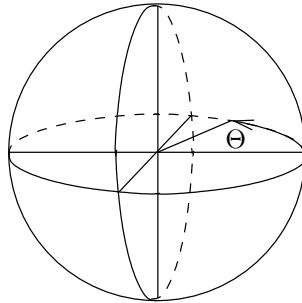


Figure 1.14: The point on the equator with longitude Θ .

the function $f : [0, \pi] \rightarrow \mathbb{R}$ defined by $f(\Theta) = T(\Theta) - T(\Theta + \pi)$ is continuous as well. If $f(0) = 0$, then it follows that the temperatures at 0 and 180° longitude are the same. If $f(0) \neq 0$, then since $f(0)$ and $f(\pi) = -f(0)$ have opposite signs, by the intermediate value theorem, it follows that f must vanish at some point, and so the claim follows. \diamond

Exercises.

- Suppose that $f : [0, 1] \rightarrow \mathbb{R}$ is a continuous function such that for all $x \in [0, 1]$, $0 \leq f(x) \leq 1$. Prove that there exists at least one $c \in [0, 1]$ such that $f(c) = c$.

HINT: Consider the continuous function $g(x) = f(x) - x$, and use the intermediate value theorem.

2. At 8:00 a.m. on Saturday, a hiker begins walking up the side of a mountain to his weekend campsite. On Sunday morning at 8:00 a.m., he walks back down the mountain along the same trail. It takes him one hour to walk up, but only half an hour to walk down. At some point on his way down, he realizes that he was at the same spot at exactly the same time on Saturday. Prove that he is right.

HINT: Let $u(t)$ and $d(t)$ be the position functions for the walks up and down, and apply the intermediate value theorem to $f(t) = u(t) - d(t)$.

3. Show that the polynomial function $p(x) = 2x^3 - 5x^2 - 10x + 5$ has a real root in the interval $[-1, 2]$.
4. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be continuous. If $S := \{f(x) \mid x \in \mathbb{R}\}$ is neither bounded above nor bounded below, prove that $S = \mathbb{R}$.

HINT: If $y \in \mathbb{R}$, then since S is neither bounded above nor bounded below, there exist $x_0, x_1 \in \mathbb{R}$ such that $f(x_0) < y < f(x_1)$.

5. (a) (*) Show that given any continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$, there exists a $x_0 \in [0, 1]$ and a $m \in \mathbb{Z}$ such that $f(x_0) = mx_0$. In other words, the graph of f intersects some line $y = mx$ at some point x_0 in $[0, 1]$.

HINT: If $f(0) = 0$, take $x_0 = 0$ and any $m \in \mathbb{Z}$. If $f(0) > 0$, then choose $N \in \mathbb{N}$ satisfying $N > f(1)$, and apply the intermediate value theorem to the continuous function $g(x) = f(x) - Nx$ on the interval $[0, 1]$. If $f(0) < 0$, then first choose a $N \in \mathbb{N}$ such that $N > -f(1)$, and consider the function $g(x) = f(x) + Nx$, and proceed in a similar manner.

- (b) (*) Prove that there does not exist a continuous function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that assumes rational values at irrational numbers, and irrational values at rational numbers, that is,

$$f(\mathbb{Q}) \subset \mathbb{R} \setminus \mathbb{Q} \text{ and } f(\mathbb{R} \setminus \mathbb{Q}) \subset \mathbb{Q}.$$

HINT: Note that for every $m \in \mathbb{Z}$, there does not exist a $x_0 \in \mathbb{R}$ such that $f(x_0) = mx_0$.

Chapter 2

Algebra

In this part of the course, we study two important algebraic objects: groups and vector spaces. We begin with a discussion about groups.

2.1 Groups

In this section, we study one of the most basic algebraic objects, namely a group. A group is a set on which a law of composition is defined, such that certain properties hold. The precise definition is given in the next subsection.

2.1.1 Definition of a Group

By a law of composition on a set S , we mean a rule for combining pairs a, b of S to get another element, say c , of S . We denote the set of all ordered pairs of elements from a set S by $S \times S$, that is, $S \times S = \{(a, b) \mid a, b \in S\}$.

Definition. A *law of composition* on a set S is a function $f : S \times S \rightarrow S$.

The functional notation $c = f(a, b)$ is not very convenient for what is going to follow, and so instead, the element obtained by applying the law of composition to a pair (a, b) is usually denoted using a notation resembling that used for addition or multiplication:

$$c = a * b, \quad \text{or} \quad ab, \quad \text{or} \quad a \circ b, \quad \text{or} \quad a + b \quad \text{and so on,}$$

with a fixed choice being made for the particular law in question.

Examples.

1. The addition of integers is a law of composition on \mathbb{Z} . Indeed, the sum of two integers is yet another integer, and addition is the function from the set $\mathbb{Z} \times \mathbb{Z}$ to the set \mathbb{Z} that assigns $a + b$ to the pair (a, b) , denoted by $(a, b) \mapsto a + b$.
2. The multiplication of real numbers is a law of composition on \mathbb{R} .

3. If a, b are rational numbers, then let $a * b = a + b - ab$. The function from $\mathbb{Q} \times \mathbb{Q}$ to \mathbb{Q} given by $(a, b) \mapsto a * b$ is a law of composition on \mathbb{Q} .
4. If a, b are real numbers, then define $a * b = \sqrt{a^2 + b^2}$. Then $(a, b) \mapsto a * b$ is not a law of composition on \mathbb{Q} , since $1 \in \mathbb{Q}$, but $1 * 1 = \sqrt{2} \notin \mathbb{Q}$. However, $(a, b) \mapsto a * b$ is a law of composition on \mathbb{R} .
5. Let $n \in \mathbb{N}$, and let S denote the set of all matrices of size $n \times n$ with real entries. Then matrix multiplication is a law of composition on S .
6. Let $n \in \mathbb{N}$, and let $GL(n, \mathbb{R})$ denote the set of all invertible matrices of size $n \times n$ with real entries. Then matrix multiplication is a law of composition on $GL(n, \mathbb{R})$. Indeed, if $A, B \in GL(n, \mathbb{R})$, then the matrix AB is again a matrix of size $n \times n$ with real entries, and moreover, since A and B are invertible, it follows that AB is also invertible.
7. Let a, b be real numbers such that $a < b$. Let $C[a, b]$ denote the set of all continuous functions on the interval $[a, b]$. Let addition of functions be defined as follows: if f, g belong to $C[a, b]$, then

$$(f + g)(x) = f(x) + g(x), \quad x \in [a, b].$$

Then addition of functions is a law of composition on $C[a, b]$, since the sum of continuous functions is again continuous; see Theorem 1.3.3.

8. If $a, b \in \mathbb{N}$, then let

$$a * b = \frac{a}{b}.$$

$*$ is not a law of composition on \mathbb{N} , since $1 * 2 = \frac{1}{2} \notin \mathbb{N}$. ◇

On a set there may be several different laws of compositions that can be defined. Some laws of compositions are nicer than others, that is, they possess some desirable properties. A group is a set G together with a law of composition on G that has three such desirable properties, and we give the definition below.

Definition. A *group* is a set G together with a law of composition $(a, b) \mapsto a * b : G \times G \rightarrow G$, which has the following properties:

- G1. (*Associativity.*) For all $a, b, c \in G$, $(a * b) * c = a * (b * c)$.
- G2. (*Identity.*) There exists an element¹ $e \in G$ such that for all $a \in G$, $a * e = a = e * a$.
- G3. (*Inverses.*) For every $a \in G$, there exists an element² $a^{-1} \in G$ such that $a * a^{-1} = e = a^{-1} * a$.

G1, G2, G3 are called group axioms. Sometimes we use the notation $(G, *)$ for the group.

Remarks. Note that hidden in the definition of a group, is the following axiom G0:

G0. For all $a, b \in G$, $a * b \in G$. (That is, $*$ is actually a *law of composition* on G .)

Hence when checking that a certain set G is a group with respect to a certain operation $*$ that combines pairs of elements from G , we have to check, first of all, that for every $a, b \in G$, $a * b$ belongs to G .

¹Such an element e is called an *identity element* of the group G .

²depending on a , and such an element a^{-1} is called an *inverse* of the element a in the group G .

Examples.

1. \mathbb{Z} with addition is a group. Indeed, the addition of integers is a law of composition on \mathbb{Z} that satisfies the group axioms:

G1. For all $a, b, c \in \mathbb{Z}$, $(a + b) + c = a + (b + c)$.

G2. 0 serves as an identity element: for all $a \in \mathbb{Z}$, $a + 0 = a = 0 + a$.

G3. If $a \in \mathbb{Z}$, then $-a \in \mathbb{Z}$ and $a + (-a) = 0 = -a + a$.

2. \mathbb{R} with multiplication is not a group. Indeed, although the multiplication of real numbers is a law of composition on \mathbb{R} that satisfies G1 and G2, but the group axiom G3 does not hold:

G1. For all $a, b, c \in \mathbb{R}$, $(ab)c = a(bc)$.

G2. (If e is an identity element, then we must have $ae = a = ea$ for all $a \in \mathbb{R}$, and in particular, with $a = 1$, we should have $1e = 1$, and so, $e = 1$. And so if e is an identity element, then it must be equal to 1!) 1 serves as an identity element: for all $a \in \mathbb{R}$, $a1 = a = 1a$.

G3. Does not hold, since $0 \in \mathbb{R}$, but for all $a^{-1} \in \mathbb{R}$, $0a^{-1} = a^{-1}0 = 0 \neq 1 = e$. So there is no inverse of the element $0 \in \mathbb{R}$.

However, the set of nonzero real numbers, or the set of positive real numbers are both groups with multiplication, since in addition to G1 and G2, now in each case G3 also holds.

3. Let $n \in \mathbb{N}$, and let $GL(n, \mathbb{R})$ denote the set of all invertible matrices of size $n \times n$ with real entries. Then $GL(n, \mathbb{R})$ is group with matrix multiplication. Indeed, matrix multiplication is a law of composition on $GL(n, \mathbb{R})$ that satisfies the group axioms:

G1. For all $A, B, C \in GL(n, \mathbb{R})$, $(AB)C = A(BC)$, since matrix multiplication is associative.

G2. The identity matrix

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

serves as an identity element. I_n is an invertible matrix of size $n \times n$ with real entries, and so it belongs to $GL(n, \mathbb{R})$, and moreover, for all $A \in GL(n, \mathbb{R})$, $AI_n = A = I_nA$.

G3. If $A \in GL(n, \mathbb{R})$, then A is an invertible matrix, and so there exists a matrix A^{-1} such that $AA^{-1} = I_n = A^{-1}A$. The matrix $A^{-1} \in GL(n, \mathbb{R})$ and serves as an inverse of A .

This group is called the *general linear group*.

4. Let a, b be real numbers such that $a < b$. Then $C[a, b]$ is a group with addition of functions. Indeed, the addition of functions is a law of composition on $C[a, b]$ that satisfies the group axioms:

G1. For all $f, g, h \in C[a, b]$, and any $x \in [a, b]$ we have

$$\begin{aligned} (f + (g + h))(x) &= f(x) + (g + h)(x) \\ &= f(x) + (g(x) + h(x)) \\ &= (f(x) + g(x)) + h(x) \text{ (since addition is associative in } \mathbb{R}\text{!)} \\ &= (f + g)(x) + h(x) \\ &= ((f + g) + h)(x). \end{aligned}$$

Hence $f + (g + h) = (f + g) + h$.

G2. The constant function $\mathbf{0}$ defined by $\mathbf{0}(x) = 0$ for all $x \in [a, b]$, serves as an identity element. $\mathbf{0}$ is a continuous function on $[a, b]$ and so $\mathbf{0} \in C[a, b]$. Moreover, for all $f \in C[a, b]$, we have for all $x \in [a, b]$:

$$\begin{aligned} (f + \mathbf{0})(x) &= f(x) + \mathbf{0}(x) \\ &= f(x) + 0 \\ &= f(x) \text{ (since 0 is an identity element for addition in } \mathbb{R}!) \\ &= 0 + f(x) \\ &= \mathbf{0}(x) + f(x) \\ &= (\mathbf{0} + f)(x). \end{aligned}$$

Hence $f + \mathbf{0} = f = \mathbf{0} + f$.

G3. If $f \in C[a, b]$, then define $-f$ by $(-f)(x) = -f(x)$, for $x \in [a, b]$. Given $f \in C[a, b]$, we have for all $x \in [a, b]$:

$$\begin{aligned} (f + (-f))(x) &= f(x) + (-f)(x) \\ &= f(x) + (-f(x)) \\ &= 0 = \mathbf{0}(x) = 0 \\ &= -f(x) + f(x) \\ &= (-f)(x) + f(x) \\ &= (-f + f)(x). \end{aligned}$$

Hence $f + (-f) = \mathbf{0} = -f + f$. ◇

We note that the groups in the above Examples 1 and 4, also satisfy

$$\boxed{\text{G4. (Commutativity.) For all } a, b \in G, a * b = b * a}$$

while the group in Example 3 (with $n = 2$) does not satisfy G4:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \neq \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

This gives rise to the following natural definition.

Definition. A group $(G, *)$ is said to be *abelian*³ if for all $a, b \in G$, $a * b = b * a$.

Examples.

1. The set of integers \mathbb{Z} with addition is an abelian group.
2. The set of positive real numbers with multiplication is an abelian group.
3. Let $n, m \in \mathbb{N}$. The set matrices $\mathbb{R}^{n \times m}$ of size $n \times m$ with entries in \mathbb{R} with matrix addition is an abelian group.
4. Let $n \in \mathbb{N}$. The set $GL(n, \mathbb{R})$ with matrix multiplication is a group, but it is not an abelian group if $n > 1$. ◇

³after the Norwegian mathematician Abel

We now prove a few elementary theorems concerning groups.

Theorem 2.1.1 *There is a unique identity element in a group.*

Proof Let e and e' be identity elements in $(G, *)$. Since $e \in G$ and e' is an identity, we obtain

$$e = e * e'.$$

Moreover, since $e' \in G$ and e is an identity, we also have

$$e * e' = e'.$$

Consequently, $e = e'$. ■

Theorem 2.1.2 *Let $(G, *)$ be a group and let $a \in G$. Then a has a unique inverse.*

Proof Let the group have the identity e . If a_1 and a_2 are inverses of a , then we have

$$\begin{aligned} a_1 &= a_1 * e \text{ (since } a_1 \in G \text{ and } e \text{ is the identity)} \\ &= a_1 * (a * a_2) \text{ (since } a_2 \text{ is an inverse of } a) \\ &= (a_1 * a) * a_2 \text{ (associativity)} \\ &= e * a_2 \text{ (since } a_1 \text{ is an inverse of } a) \\ &= a_2 \text{ (since } a_2 \in G \text{ and } e \text{ is the identity)}. \end{aligned}$$

■

Example. If A, B are matrices of size $n \times n$ with real entries and $I_n - AB$ is invertible, then $I_n - BA$ is also invertible, and $(I_n - BA)^{-1} = I_n + B(I_n - AB)^{-1}A$. Indeed, it is easy to check that

$$(I_n - BA)(I_n + B(I_n - AB)^{-1}A) = I_n = (I_n + B(I_n - AB)^{-1}A)(I_n - BA),$$

and so it follows that $I_n - BA$ is invertible. By the uniqueness of the inverse of the element $I_n - BA$ in the group $GL(n, \mathbb{R})$, it follows that $(I_n - BA)^{-1} = I_n + B(I_n - AB)^{-1}A$. ◇

Definitions. A group $(G, *)$ is said to be *finite*, if the set G has finite cardinality. The *order* of a finite group $(G, *)$ is the cardinality of G . A group is said to be *infinite* if it is not finite.

Examples.

1. The set $\{-1, 1\}$ with multiplication is a finite group of order 2.
2. The set \mathbb{Z} with addition is an infinite group. ◇

A finite group can be completely described by writing its *group table*. This is a table that displays the law of composition as follows: the elements of the group are listed in the first row and the first column, and then given $a, b \in G$, the element $a * b$ is entered in the row corresponding to a and the column corresponding to b , as shown below.

*	...	b	...
⋮			
a		$a * b$	
⋮			

We clarify this further by considering a simple example.

Example. The finite group $\{-1, 1\}$ with multiplication can be described by the group table given below.

\cdot	1	-1
1	1	-1
-1	-1	1

The table completely describes the law of composition: $1 \cdot 1 = 1$, $1 \cdot (-1) = -1$, $-1 \cdot 1 = -1$ and $-1 \cdot (-1) = 1$. \diamond

Exercises.

- Show that the set \mathbb{Z}_6 of integers modulo 6, with addition modulo 6 is a group. Write down its group table.
 - Is \mathbb{Z}_6 with multiplication modulo 6 also a group? If, instead, we consider the set \mathbb{Z}_6^* of nonzero integers modulo 6, then is \mathbb{Z}_6^* a group with multiplication modulo 6?
 - (*) Let m be an integer such that $m \geq 2$, and let \mathbb{Z}_m^* denote the set of nonzero integers modulo m . Prove that \mathbb{Z}_m^* is a group with multiplication modulo m iff m is a prime number.
HINT: If m is a prime number, then any $r \in \{1, \dots, m-1\}$ is coprime to m , and so there exist integers s and t such that $sm + tr = 1$.
- Given $n \in \mathbb{N}$, let S_n be the set of all bijections from the set $\{1, 2, 3, \dots, n\}$ onto itself. Prove that the set S_n with the composition of functions is a group. This is called the *symmetric group* S_n .
HINT: Composition of bijections is again a bijection, and composition of functions on a set is also associative. The identity element is simply the identity map $\iota : \{1, 2, 3, \dots, n\} \rightarrow \{1, 2, 3, \dots, n\}$ given by $\iota(m) = m$ for all $m \in \{1, 2, 3, \dots, n\}$, while the inverse of an element $f \in S_n$ is simply the inverse of that bijection.
 - What is the order of S_n ?
 - If $n = 3$, give examples of bijections f and g on $\{1, 2, 3\}$ such that $f \circ g \neq g \circ f$.
 - (*) Show that S_n is abelian iff $n \leq 2$.
- Consider the set

$$S = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mid a, b \in \mathbb{R} \text{ and } a^2 + b^2 \neq 0 \right\}.$$

Show that S with the operation of matrix multiplication forms a group.

- Let $(G, *)$ be a group.
 - Show that if $a, b, c \in G$ are such that $a * b = a * c$, then $b = c$.
 - Show that if $a, b \in G$, then the equation $a * x = b$ has a unique solution.
 - Show that if $a, b \in G$, then $(a * b)^{-1} = b^{-1} * a^{-1}$.

2.1.2 Subgroups

Definition. A subset H of a group G is called a *subgroup* of G , if it has the following properties:

- H1. (*Closure*) If $a, b \in H$, then $a * b \in H$.
- H2. (*Identity*) $e \in H$.
- H3. (*Inverses*) If $a \in H$, then $a^{-1} \in H$.

These conditions ensure that H is itself a group which is contained in the group G , and this is explained as follows:

- H1. This condition tells us that the law of composition $*$ on the group G can be used to define a law of composition on H , namely, the function from $H \times H$ to H given by $(a, b) \mapsto a * b$, which is called the *induced law of composition*. Since $*$ is associative, it follows that the induced law of composition is associative as well: for all $a, b, c \in H$, $a * (b * c) = (a * b) * c$.
- H2. This shows that H , with the induced law of composition, has an identity element. Indeed the identity element from G (which also belongs to H) serves as the identity element also in H : for every $a \in H$, $a * e = a = e * a$.
- H3. Finally this shows that every element in H possesses an inverse element in H . Of course, since G is a group, we already knew that a possesses an inverse element $a^{-1} \in G$. But now H3 says that this inverse element is in H . Thus for all $a \in H$, $\exists a^{-1} \in H$ such that $a * a^{-1} = e = a^{-1} * a$.

Thus the conditions H1, H2, H3 imply that the subset H , with the induced law of composition, is a group. Thus, a subgroup is itself a group which sits in a larger group.

Examples.

1. The subset of even integers $\{2m \mid m \in \mathbb{Z}\}$ is a subgroup of the group of integers \mathbb{Z} with addition. Indeed, the sum of even numbers is even (and so H1 holds), 0 is even (and so H2 holds), and finally, given the even number $2m$, $-2m = 2(-m)$ is even as well (and so H3 holds).
2. The group of integers with addition $(\mathbb{Z}, +)$ is a subgroup of the group of rational numbers with addition $(\mathbb{Q}, +)$, which in turn is a subgroup of the group of real numbers with addition $(\mathbb{R}, +)$.
3. If G is a group with identity e , then $\{e\}$ and G are both subgroups of G .
4. The subset of *symmetric matrices* of size 2×2 , namely

$$\left\{ \left[\begin{array}{cc} a & b \\ b & d \end{array} \right] \mid a, b, d \in \mathbb{R} \right\}$$

with real entries is a subgroup of the set of all 2×2 matrices having real entries with matrix addition. Indeed, given any two symmetric matrices

$$\left[\begin{array}{cc} a & b \\ b & d \end{array} \right] \text{ and } \left[\begin{array}{cc} a' & b' \\ b' & d' \end{array} \right],$$

their sum

$$\begin{bmatrix} a & b \\ b & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ b' & d' \end{bmatrix} = \begin{bmatrix} a+a' & b+b' \\ b+b' & d+d' \end{bmatrix}$$

is also symmetric, and so H1 holds. Clearly the identity element

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

is symmetric, and so H2 also holds. Finally, the inverse (with respect to matrix addition) of any symmetric matrix

$$\begin{bmatrix} a & b \\ b & d \end{bmatrix},$$

is the element

$$\begin{bmatrix} -a & -b \\ -b & -d \end{bmatrix}$$

which is also symmetric, and so H3 holds.

5. The subset of *upper triangular* invertible matrices,

$$UT(2, \mathbb{R}) = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \mid a, b, d \in \mathbb{R} \text{ and } ad \neq 0 \right\}$$

is a subgroup of the group $GL(2, \mathbb{R})$ with matrix multiplication. Indeed, if

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}, \begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} \in UT(2, \mathbb{R}),$$

then

$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} a' & b' \\ 0 & d' \end{bmatrix} = \begin{bmatrix} aa' & ab' + bd' \\ 0 & dd' \end{bmatrix} \in UT(2, \mathbb{R}),$$

and so H1 holds. Clearly

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in UT(2, \mathbb{R}),$$

and so H2 also holds. Finally,

$$\text{if } \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \in UT(2, \mathbb{R}), \text{ then } \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{a} & -\frac{b}{ad} \\ 0 & \frac{1}{d} \end{bmatrix} \in UT(2, \mathbb{R}).$$

◇

Exercises.

- Determine if the following statements are TRUE or FALSE.
 - The nonnegative integers form a subgroup of \mathbb{Z} with addition.
 - The odd integers form a subgroup of \mathbb{Z} with addition.
 - If G is abelian and H is a subgroup of G , then H is abelian.
- Is there an infinite group with a finite subgroup?
- Consider the group of integers \mathbb{Z} with addition. Suppose that H is the subgroup of \mathbb{Z} comprising multiples of 4, and let K be the subgroup of \mathbb{Z} comprising multiples of 6. What is $H \cap K$?

- (b) If H and K are subgroups of a group G , then show that $H \cap K$ is also a subgroup of G .
4. Let $C[0, 1]$ denote the group comprising the set of continuous functions on the interval $[0, 1]$ with addition of functions defined in the usual way: if f, g belong to $C[0, 1]$, then for all $x \in [0, 1]$, $(f + g)(x) = f(x) + g(x)$.
- (a) Let $H_1 = \{f \in C[0, 1] \mid f(\frac{1}{2}) = 0\}$. Prove that H_1 is a subgroup of $C[0, 1]$.
- (b) Let H_2 denote the set of all polynomial functions, that is, the set of all functions $p : [0, 1] \rightarrow \mathbb{R}$ such that there exists an $n \in \mathbb{N} \cup \{0\}$ and real numbers $a_0, a_1, a_2, \dots, a_n$ such that $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, for all $x \in [0, 1]$. Show that H_2 is a subgroup of $C[0, 1]$.
5. Let G be a group. The *center of G* is the set $Z(G) = \{z \in G \mid \forall a \in G, z * a = a * z\}$.
- (a) Show that $Z(G)$ is not empty.
- (b) If G is abelian, then determine $Z(G)$.
- (c) Show that $Z(G)$ is a subgroup of G .
- (d) (*) If G is the group $GL(2, \mathbb{R})$ with matrix multiplication, then determine $Z(G)$.

HINT: Take $a = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

Let G be a group and let $a \in G$. We define

$$a^0 = e \text{ and } a^n = a^{n-1} * a \text{ for } n \in \mathbb{N}.$$

Moreover, if $n \in \mathbb{N}$, then we define

$$a^{-n} = (a^n)^{-1}.$$

It can be shown that the usual laws of exponents hold: for all $m, n \in \mathbb{Z}$,

$$a^m * a^n = a^{m+n} \text{ and } a^{mn} = (a^m)^n.$$

(Exercise!)

Definitions. Let G be a group and suppose that $a \in G$.

1. If there exists an $m \in \mathbb{N}$ such that $a^m = e$, then a is said to have *finite order*.
2. If a has finite order, then the *order of a* , denoted by $\text{ord}(a)$, is

$$\text{ord}(a) = \min\{m \in \mathbb{N} \mid a^m = e\}.$$

3. If a does not have finite order, then a is said to have *infinite order*.

Examples.

1. The element -1 has order 2 in the group of nonzero real numbers with multiplication.
2. The element 2 has infinite order in the group of integers with addition.

3. The element $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ is an element of order 3 in the group $GL(3, \mathbb{R})$. ◇

We now prove that given any element from a group, the set of its powers form a subgroup of the group.

Theorem 2.1.3 *Suppose that G is a group and $a \in G$. Let $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$. Then:*

1. $\langle a \rangle$ is a subgroup of G .
2. If a is an element with finite order m , then $\langle a \rangle = \{e, a, a^2, a^3, \dots, a^{m-1}\}$.

Proof

1. We prove that H1, H2, H3 hold.

H1. Given $m, n \in \mathbb{Z}$, clearly $m + n \in \mathbb{Z}$ and so $a^m * a^n = a^{m+n} \in \langle a \rangle$.

H2. $e = a^0 \in \langle a \rangle$.

H3. For any $m \in \mathbb{Z}$, $-m \in \mathbb{Z}$ and so $(a^m)^{-1} = a^{-1 \cdot m} = a^{-m} \in \langle a \rangle$.

So $\langle a \rangle$ is a subgroup of G .

2. Clearly $\{e, a, a^2, a^3, \dots, a^{m-1}\} \subset \langle a \rangle$. Conversely, if $n \in \mathbb{Z}$, then there exist integers q and r , such that $0 \leq r \leq m - 1$, and $n = q \cdot m + r$. So

$$a^n = a^{q \cdot m + r} = a^{q \cdot m} * a^r = (a^m)^q * a^r = (e)^q * a^r = e * a^r = a^r \in \{e, a, a^2, a^3, \dots, a^{m-1}\}.$$

■

The element $[1]$ in the group \mathbb{Z}_5 with addition modulo 5 has finite order, since

$$[1] \oplus [1] \oplus [1] \oplus [1] \oplus [1] = [0],$$

and the subgroup $\langle [1] \rangle = \{[0], [1], [2], [3], [4]\}$ is in fact the whole group. This example motivates the following definition.

Definition. A group G is said to be *cyclic* if there exists an element $a \in G$ such that $G = \langle a \rangle$. Such an element a is then called a *generator* of the group G .

Exercises.

1. If G is a finite group, then show that every element in the group has a finite order, which is at most equal to $|G|$.
HINT: If $a \in G$, then consider the set $S = \{e, a, a^2, a^3, \dots, a^{|G|}\}$, and use the pigeonhole principle.
2. If G is a group and $a \in G$, then show that for all $m, n \in \mathbb{Z}$, $a^m * a^n = a^{m+n}$ and $(a^m)^n = a^{mn}$.
3. Determine the order of $\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ in the group $GL(2, \mathbb{R})$.
4. Prove that in any group $(G, *)$, and for any a, b in G , the orders of $a * b$ and $b * a$ are the same.
5. Is the group of integers with addition cyclic? What is a generator of this group? Is it unique?

2.1.3 Homomorphisms and isomorphisms

Definition. Let $(G, *)$ and $(G', *')$ be groups. A *homomorphism* $\varphi : G \rightarrow G'$ is a function such that

$$\text{for all } a, b \in G, \quad \varphi(a * b) = \varphi(a) *' \varphi(b).$$

Examples.

1. Let G be \mathbb{R} with addition, and G' be the set of positive reals with multiplication. The exponential function from G to G' , given by $x \mapsto 2^x$, is a homomorphism. Indeed, we have $2^{x+y} = 2^x 2^y$ for all real x, y .
2. Let G be the group $GL(2, \mathbb{R})$ with matrix multiplication, and G' be the group of nonzero real numbers with multiplication. Then the determinant function $\det : G \rightarrow G'$ is a homomorphism, since for all 2×2 real matrices A, B we have $\det(AB) = \det(A) \det(B)$.
3. Let G be the group $C[0, 1]$ of continuous functions on the interval $[0, 1]$ with addition, and G' be the group \mathbb{R} with addition. Then the function $f \mapsto f(\frac{1}{2})$ is a homomorphism. Indeed, if f, g are continuous functions on the interval $[0, 1]$, then $(f + g)(\frac{1}{2}) = f(\frac{1}{2}) + g(\frac{1}{2})$, by the definition of $f + g$.
4. If G is a group, then the identity map $\iota : G \rightarrow G$ defined by $\iota(a) = a$ for all $a \in G$, and the trivial map $\zeta : G \rightarrow G$ defined by $\zeta(a) = e$ for all $a \in G$ are homomorphisms. \diamond

Thus a homomorphism is a function between two groups that respects the law of composition. In the next theorem we show that it preserves the identity element and the inverses of elements as well.

Theorem 2.1.4 *Let G be a group with identity e and G' be a group with identity e' . If $\varphi : G \rightarrow G'$ is a homomorphism, then:*

1. $\varphi(e) = e'$.
2. If $a \in G$, then $(\varphi(a))^{-1} = \varphi(a^{-1})$.

Proof We have

$$e' *' \varphi(e) = \varphi(e) = \varphi(e * e) = \varphi(e) *' \varphi(e),$$

and so canceling $\varphi(e)$ on both sides, we obtain $e' = \varphi(e)$. Next,

$$\varphi(a^{-1}) *' \varphi(a) = \varphi(a^{-1} * a) = \varphi(e) = e'$$

and similarly,

$$e' = \varphi(e) = \varphi(a * a^{-1}) = \varphi(a) *' \varphi(a^{-1}).$$

Thus $\varphi(a^{-1}) *' \varphi(a) = e' = \varphi(a) *' \varphi(a^{-1})$, and by the uniqueness of the inverse of $\varphi(a)$ in G' , we obtain $(\varphi(a))^{-1} = \varphi(a^{-1})$. \blacksquare

Every group homomorphism φ determines two important subgroups: its image and its kernel.

Definitions. Let G, G' be groups and let $\varphi : G \rightarrow G'$ be a group homomorphism.

1. The *kernel* of φ is the set $\ker(\varphi) = \{a \in G \mid \varphi(a) = e'\}$.
2. The *image* of φ is the set $\text{im}(\varphi) = \{a' \in G' \mid \exists a \in G \text{ such that } \varphi(a) = a'\}$.

Using Theorem 2.1.4, we now prove the following result.

Theorem 2.1.5 *Let G, G' be groups and let $\varphi : G \rightarrow G'$ be a group homomorphism. Then:*

1. $\ker(\varphi)$ is a subgroup of G .
2. $\text{im}(\varphi)$ is a subgroup of G' .

Proof It is easy to check that $\ker(\varphi)$ is a subgroup of G . Indeed if a, b belong to $\ker(\varphi)$, then $\varphi(a * b) = \varphi(a) *' \varphi(b) = e' *' e' = e'$, and so H1 holds. Moreover, as $\varphi(e) = e'$, $e \in \ker(\varphi)$ and so H2 holds too. Finally, H3 holds, since if $a \in \ker(\varphi)$, then $\varphi(a^{-1}) = (\varphi(a))^{-1} = (e')^{-1} = e'$, and so $a^{-1} \in \ker(\varphi)$. Hence $\ker(\varphi)$ is a subgroup of G .

We now also check that $\text{im}(\varphi)$ is a subgroup of G' . If a', b' belong to $\text{im}(\varphi)$, then there exist elements a, b in G such that $\varphi(a) = a'$ and $\varphi(b) = b'$. Consequently, $\varphi(a * b) = \varphi(a) *' \varphi(b) = a' *' b'$, and so there exists an element in G , namely $a * b$, such that $\varphi(a * b) = a' *' b'$, that is, $a' *' b' \in \text{im}(\varphi)$. Thus H1 holds. Since $\varphi(e) = e'$, it follows that $e' \in \text{im}(\varphi)$. Finally, if $a' \in \text{im}(\varphi)$, then there exists an $a \in G$ such that $\varphi(a) = a'$, and so $a'^{-1} = (\varphi(a))^{-1} = \varphi(a^{-1})$. Hence $a'^{-1} \in \text{im}(\varphi)$, and H3 holds. So $\text{im}(\varphi)$ is a subgroup of G' . ■

Examples.

1. Let G be \mathbb{R} with addition, and G' be the set of positive reals with multiplication. The exponential function from G to G' , given by $x \mapsto 2^x$, is a homomorphism with kernel the trivial subgroup comprising the element 0, and the image is the whole group of positive reals with multiplication, since it can be shown that given any $y > 0$, there exists a unique real number (called the *logarithm of y to the base 2*, denoted by $\log_2 y$) such that $y = 2^{\log_2 y}$.
2. Let G be the group $GL(2, \mathbb{R})$ with matrix multiplication, and G' be the group of nonzero real numbers with multiplication. Then the determinant function $\det : G \rightarrow G'$ is a homomorphism. Its kernel is the set of all invertible matrices with determinant equal to 1, and we denote this subgroup by $SL(2, \mathbb{R})$, and it is called the *special linear group*:

$$SL(2, \mathbb{R}) = \{A \in GL(2, \mathbb{R}) \mid \det(A) = 1\}.$$

The image of this homomorphism is the whole group of nonzero reals: indeed, given any real number a not equal to zero, we have that

$$A := \begin{bmatrix} 1 & 0 \\ 0 & a \end{bmatrix} \in GL(2, \mathbb{R}),$$

and $\det(A) = 1 \cdot a - 0 \cdot 0 = a$.

3. Let G be the group $C[0, 1]$ of continuous functions on the interval $[0, 1]$ with addition, and G' be the group \mathbb{R} with addition. Then the function $f \mapsto f\left(\frac{1}{2}\right)$ is a homomorphism, and its kernel is the set of all continuous functions on the interval $[0, 1]$ that have a root at $\frac{1}{2}$ (for instance the straight line $f(x) = x - \frac{1}{2}$ belongs to the kernel). The image is the set of all real numbers, since given any $a \in \mathbb{R}$, the constant function $f(x) = a$ for all $x \in [0, 1]$ is continuous, and $f\left(\frac{1}{2}\right) = a$. ◇

In Example 1 above, the homomorphism between the two groups was also bijective. We give a special name to such homomorphisms.

Definition. Let G, G' be groups. A homomorphism $\varphi : G \rightarrow G'$ is said to be an *isomorphism* if it is bijective.

Examples.

1. Let G be \mathbb{R} with addition, and G' be the set of positive reals with multiplication. The exponential function from G to G' , given by $x \mapsto 2^x$, is an isomorphism.
2. If G is a group, then the identity function $\iota : G \rightarrow G$ defined by $\iota(a) = a$ for all $a \in G$ is an isomorphism.
3. Let G be the subgroup of $GL(2, \mathbb{R})$ comprising all matrices of the form $\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}$, where $x \in \mathbb{R}$. Let G' be the group \mathbb{R} with addition. Then the function from G to G' , given by

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \mapsto x,$$

is an isomorphism. ◇

Isomorphisms are important because their existence between two groups means that the two groups are essentially the “same”, in the sense that as far as algebraic properties go, there is no real difference between them.

Exercises.

1. (a) Let G, G', G'' be groups and let $\varphi : G \rightarrow G'$ and $\psi : G' \rightarrow G''$ be group homomorphisms. Prove that the composition $\psi \circ \varphi : G \rightarrow G''$ defined by $(\psi \circ \varphi)(a) = \psi(\varphi(a))$, $a \in G$ is a group homomorphism.
 (b) Describe the kernel of $\psi \circ \varphi$.
2. A subgroup N of a group G is called a *normal subgroup* if for every $a \in N$ and every $b \in G$, then $b * a * b^{-1} \in N$. Prove that if G, G' are groups and $\varphi : G \rightarrow G'$ is a homomorphism, then $\ker(\varphi)$ is a normal subgroup of G .
3. If G is an abelian group, then show that the function from G to G , given by $a \mapsto a^{-1}$ is an isomorphism.
4. Let G, G' be groups and let $\varphi : G \rightarrow G'$ be an isomorphism. Prove that the inverse function $\varphi^{-1} : G' \rightarrow G$ is also an isomorphism.
5. Let G be a group and let a be an element of G .
 (a) Prove that the function from \mathbb{Z} to $\langle a \rangle$, given by $m \mapsto a^m$, is a homomorphism from the group of integers \mathbb{Z} with addition to the subgroup $\langle a \rangle$.
 (b) If a has infinite order, then prove that it is an isomorphism.

2.1.4 Cosets and Lagrange's theorem

Given a subgroup H of a group G , define the relation R on G by

$$aRb \text{ if } b = a * h \text{ for some } h \in H. \quad (2.1)$$

Then R is an equivalence relation:

- E1. (*Reflexivity*) For all $a \in G$, aRa since $e \in H$ and $a = a * e$.
- E2. (*Symmetry*) For all a, b in G , if aRb , then there exists a $h \in H$ such that $b = a * h$ and so $b * h^{-1} = a$. Since H is a subgroup, and $h \in H$, it follows that $h^{-1} \in H$. Thus bRa .
- E3. (*Transitivity*) For all a, b, c in G , if aRb and bRc , then there exist elements h_1, h_2 in H such that $b = a * h_1$ and $c = b * h_2$. Hence we obtain $c = b * h_2 = (a * h_1) * h_2 = a * (h_1 * h_2)$. Since $h_1, h_2 \in H$ and H is a subgroup, it follows that $h_1 * h_2 \in H$, and so aRc .

Definition. Let H be a subgroup of a group G , and let R be the equivalence relation given by (2.1). If $a \in G$, then the equivalence class of a , namely the set

$$\{b \in G \mid aRb\} = \{b \in G \mid \exists h \in H \text{ such that } b = a * h\} = \{a * h \mid h \in H\},$$

is called a *left coset of H* , and is denoted by $a * H$.

We know that the equivalence classes of an equivalence relation partition the set. (Recall that a *partition* of a set S , we mean a subdivision of the set S into nonoverlapping subsets:

$$S = \text{union of disjoint, nonempty subsets of } S.$$

For example, the sets $\{1, 3\}$, $\{2, 5\}$, $\{4\}$ form a partition of the $\{1, 2, 3, 4, 5\}$. The two sets, of even integers, and of odd integers, form a partition of the set \mathbb{Z} of all integers.)

Hence we obtain the following result:

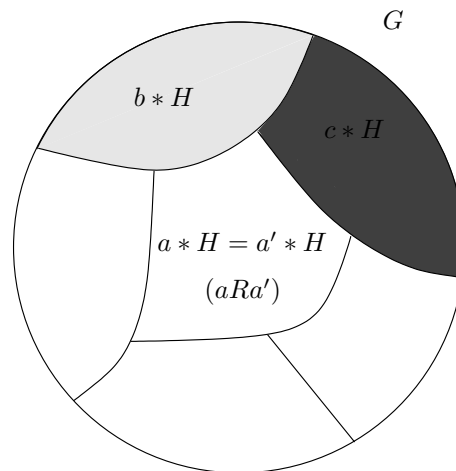


Figure 2.1: Distinct cosets partition the group.

Corollary 2.1.6 *Let H be a subgroup of a group G . Then the left cosets of H partition the group G .*

Remarks.

1. The notation $a * H$ denotes a certain subset of G . As with any equivalence relation, different notations may represent the same subset. In fact, we know that $a * H$ is the unique coset containing a , and so

$$a * H = b * H \text{ iff } aRb.$$

Corollary 2.1.6 says that if $a * H$ and $b * H$ have an element in common then they are equal.

2. One can also define the relation R' on G by $aR'b$ if $b = h * a$ for some $h \in H$. The associated equivalence classes are called *right cosets*.

Examples.

1. Consider the group G of the integers modulo 6, \mathbb{Z}_6 , with addition modulo 6, and let H be the subgroup $\langle [2] \rangle = \{[0], [2], [4]\}$. The left cosets, which we now denote by $a \oplus H$ are

$$\begin{aligned} [0] \oplus H &= [2] \oplus H = [4] \oplus H = \{[0], [2], [4]\}, \text{ and} \\ [1] \oplus H &= [3] \oplus H = 5 \oplus H = \{[1], [3], [5]\}. \end{aligned}$$

Note that the cosets $\{[0], [2], [4]\}$ and $\{[1], [3], [5]\}$ do form a partition of G :

$$G = \{[0], [1], [2], [3], [4], [5]\} = \{[0], [2], [4]\} \cup \{[1], [3], [5]\}, \text{ and } \{[0], [2], [4]\} \cap \{[1], [3], [5]\} = \emptyset.$$

2. Consider the group G of the integers \mathbb{Z} , with addition, and let H be the subgroup of even numbers $\{2m \mid m \in \mathbb{Z}\}$. The left cosets, which we now denote by $a + H$ are

$$\begin{aligned} \dots &= -2 + H = 0 + H = \{2m \mid m \in \mathbb{Z}\} = \{\dots, -4, -2, 0, 2, 4, \dots\} = 2 + H = \dots, \\ \dots &= -1 + H = 1 + H = \{2m + 1 \mid m \in \mathbb{Z}\} = \{\dots, -3, -1, 1, 3, \dots\} = 3 + H = \dots \end{aligned}$$

Note that the cosets $\{\dots, -4, -2, 0, 2, 4, \dots\}$ and $\{\dots, -3, -1, 1, 3, \dots\}$ do indeed partition the set of all integers. \diamond

Note that in Example 1 above, there are only two distinct cosets, and

$$|G| = 6 = 3 \cdot 2 = |H| \cdot (\text{number of cosets of } H).$$

In particular the order of H (namely 3) divides the order of G (namely 6). This is not a coincidence. We now prove an important result concerning the order of a group G and the number of cosets of a subgroup H , due to Lagrange.

Theorem 2.1.7 (Lagrange's theorem) *Let H be a subgroup of a finite group G . Then the order of H divides the order of G .*

Proof Note that there is a bijective function from the subgroup H to the coset $a * H$, given by $h \mapsto a * h$, for $h \in H$. Consequently, each coset $a * H$ has the same number of elements as H does.

Since G is the union of the cosets of H , and since these cosets do not overlap, we obtain the *counting formula*

$$|G| = |H| \cdot (\text{number of cosets of } H).$$

In particular, $|H|$ divides $|G|$. \blacksquare

Corollary 2.1.8 Let G be a finite group and let $a \in G$. Then the order⁴ of a divides the order of G . In particular $a^{|G|} = e$.

Proof Let a have order m . $\langle a \rangle$ is a subgroup of G , and from Theorem 2.1.3, it follows that $|\langle a \rangle| = m$ divides $|G|$, and so $|G| = m \cdot k$ for some $k \in \mathbb{N}$. Thus $a^{|G|} = a^{m \cdot k} = (a^m)^k = e^k = e$. ■

The following theorem characterizes all groups whose order is a prime number.

Corollary 2.1.9 If G is a group with prime order p , then G is cyclic, and $G = \langle a \rangle$ for every $a \in G \setminus \{e\}$.

Proof If $a \neq e$, then a has order > 1 , say m . Since m divides p , and p is prime, it follows that $m = p$. As G itself has order p , it now follows that $G = \langle a \rangle$, and so G is cyclic. ■

Exercises.

- Determine if the following statements are TRUE or FALSE.
 - If H is a subgroup of G and $a, b \in G$ are such that $a \neq b$, then $(a * H) \cap (b * H) = \emptyset$.
 - If H is a subgroup of G and $a \in G$ is such that $a * H$ has 4 elements, then H has 4 elements.
 - If H is a subgroup of a finite group G , then for any $a \in G$, the left coset $a * H$ has the same number of elements as the right coset $H * a$.

- (a) Verify that

$$G = \left\{ \left[\begin{array}{cc} x & y \\ 0 & 1 \end{array} \right] \mid x, y \in \mathbb{R} \text{ and } x > 0 \right\}$$

is a group with matrix multiplication.

- (b) Show that

$$H = \left\{ \left[\begin{array}{cc} x & 0 \\ 0 & 1 \end{array} \right] \mid x \in \mathbb{R} \text{ and } x > 0 \right\}$$

is a subgroup of G .

- (c) Any element of G can be represented by a point in the (x, y) -plane. Draw the partitions of the plane into left and into right cosets of H .
- Let H and K be subgroups of a group G of orders 3 and 5 respectively. Prove that $H \cap K = \{e\}$.
HINT: $H \cap K$ is a subgroup of H as well as K .
- Prove or disprove that the group S_4 has an element of order 16.
- (a) Let p be a prime, and let \mathbb{Z}_p^* denote the group of nonzero integers modulo p with multiplication modulo p . Show that if a is an integer such that a is not divisible by p , then $[a]^{p-1} = [1]$.
(b) (*) Prove *Fermat's little theorem*: for any integer a , $a^p \equiv a \pmod{p}$.
HINT: If $a \in \mathbb{Z}$ is not divisible by p , then $[a] \neq [0]$, and so by part (5a) above, $[a]^{p-1} = [1]$. Hence $p \mid (a^{p-1} - 1)$, and so $p \mid (a^p - a)$.
(c) (*) Show that 7 divides $2222^{5555} + 5555^{2222}$.
HINT: Note that $2222 = 7 \cdot 317 + 3$, so that in \mathbb{Z}_7 , $[2222^{5555}] = [3]^{5555}$. Now use the fact that $[3]^6 = [1]$ to conclude that $[2222^{5555}] = [3^5]$. Proceeding in a similar manner, show that $[5555^{2222}] = [3^2]$. Hence we obtain $[2222^{5555} + 5555^{2222}] = [3^5 + 3^2] = [3^2 \cdot 28] = [0]$.

⁴By Exercise 1 on page 48, it follows that every element in a finite group has a finite order.

2.2 Vector spaces

In this section we introduce a very important algebraic object, called a vector space. Roughly speaking it is a set of elements, called “vectors”. Any two vectors can be “added”, resulting in a new vector, and any vector can be multiplied by an element from \mathbb{R} so as to give a new vector. The precise definition is given in the next subsection.

2.2.1 Definition of a vector space

Definition. A *vector space* V is a set together with two functions, $+$: $V \times V \rightarrow V$, called *vector addition*, and \cdot : $\mathbb{R} \times V \rightarrow V$, called *scalar multiplication*, such that $(V, +)$ is an abelian group, and the following hold:

- V1. For all $v \in V$, $1 \cdot v = v$.
- V2. For all $\alpha, \beta \in \mathbb{R}$ and all $v \in V$, $\alpha \cdot (\beta \cdot v) = (\alpha\beta) \cdot v$.
- V3. For all $\alpha, \beta \in \mathbb{R}$ and all $v \in V$, $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.
- V4. For all $\alpha \in \mathbb{R}$ and all $v_1, v_2 \in V$, $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$.

V3 and V4 are called the *distributive laws*. The elements of a vector space are called *vectors*.

We observe that since $(V, +)$ is an abelian group, a vector space also has the following properties:

1. For all $v_1, v_2, v_3 \in V$, $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$.
2. There exists an element $\mathbf{0} \in V$ (called the⁵ *zero vector*) such that for all $v \in V$, $v + \mathbf{0} = v = \mathbf{0} + v$.
3. For every $v \in V$, there exists a unique⁶ element in V , denoted by $-v$, such that $v + (-v) = \mathbf{0} = -v + v$.
4. For all $v_1, v_2 \in V$, $v_1 + v_2 = v_2 + v_1$.

Examples.

1. Let $n, m \in \mathbb{N}$. The set $\mathbb{R}^{n \times m}$ of $n \times m$ matrices having real entries with matrix addition is an abelian group. Define scalar multiplication as follows: if $\alpha \in \mathbb{R}$ and

$$\text{if } \alpha \in \mathbb{R} \text{ and } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{n \times m}, \text{ then } \alpha \cdot A = \begin{bmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{bmatrix}. \quad (2.2)$$

Then $\alpha \cdot A \in \mathbb{R}^{n \times m}$, and moreover V1, V2, V3, V4 are satisfied:

⁵Since there is a unique identity element in a group, the zero vector is unique!

⁶In a group, every element has a unique inverse!

V1. If $A \in \mathbb{R}^{n \times m}$, then clearly

$$1 \cdot A = 1 \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} 1a_{11} & \dots & 1a_{1n} \\ \vdots & & \vdots \\ 1a_{m1} & \dots & 1a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = A.$$

V2. For all $\alpha, \beta \in \mathbb{R}$ and all $A \in \mathbb{R}^{n \times m}$,

$$\begin{aligned} \alpha \cdot (\beta \cdot A) &= \alpha \cdot \left(\beta \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \right) \\ &= \alpha \cdot \begin{bmatrix} \beta a_{11} & \dots & \beta a_{1n} \\ \vdots & & \vdots \\ \beta a_{m1} & \dots & \beta a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} \alpha(\beta a_{11}) & \dots & \alpha(\beta a_{1n}) \\ \vdots & & \vdots \\ \alpha(\beta a_{m1}) & \dots & \alpha(\beta a_{mn}) \end{bmatrix} \\ &= \begin{bmatrix} (\alpha\beta)a_{11} & \dots & (\alpha\beta)a_{1n} \\ \vdots & & \vdots \\ (\alpha\beta)a_{m1} & \dots & (\alpha\beta)a_{mn} \end{bmatrix} \\ &= (\alpha\beta) \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \\ &= (\alpha\beta) \cdot A. \end{aligned}$$

V3. For all $\alpha, \beta \in \mathbb{R}$ and all $A \in \mathbb{R}^{n \times m}$,

$$\begin{aligned} (\alpha + \beta) \cdot A &= (\alpha + \beta) \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} (\alpha + \beta)a_{11} & \dots & (\alpha + \beta)a_{1n} \\ \vdots & & \vdots \\ (\alpha + \beta)a_{m1} & \dots & (\alpha + \beta)a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} \alpha a_{11} + \beta a_{11} & \dots & \alpha a_{1n} + \beta a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} + \beta a_{m1} & \dots & \alpha a_{mn} + \beta a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} \alpha a_{11} & \dots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \dots & \alpha a_{mn} \end{bmatrix} + \begin{bmatrix} \beta a_{11} & \dots & \beta a_{1n} \\ \vdots & & \vdots \\ \beta a_{m1} & \dots & \beta a_{mn} \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} + \beta \cdot \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \\ &= \alpha \cdot A + \beta \cdot A. \end{aligned}$$

V4. For all $\alpha \in \mathbb{R}$ and all $A, B \in \mathbb{R}^{n \times m}$,

$$\begin{aligned}
\alpha \cdot (A + B) &= \alpha \cdot \left(\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} \right) \\
&= \alpha \cdot \begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix} \\
&= \begin{bmatrix} \alpha(a_{11} + b_{11}) & \cdots & \alpha(a_{1n} + b_{1n}) \\ \vdots & & \vdots \\ \alpha(a_{m1} + b_{m1}) & \cdots & \alpha(a_{mn} + b_{mn}) \end{bmatrix} \\
&= \begin{bmatrix} \alpha a_{11} + \alpha b_{11} & \cdots & \alpha a_{1n} + \alpha b_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} + \alpha b_{m1} & \cdots & \alpha a_{mn} + \alpha b_{mn} \end{bmatrix} \\
&= \begin{bmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{bmatrix} + \begin{bmatrix} \alpha b_{11} & \cdots & \alpha b_{1n} \\ \vdots & & \vdots \\ \alpha b_{m1} & \cdots & \alpha b_{mn} \end{bmatrix} \\
&= \alpha \cdot \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} + \beta \cdot \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} \\
&= \alpha \cdot A + \beta \cdot B.
\end{aligned}$$

Hence $\mathbb{R}^{m \times n}$ is a vector space with matrix addition and with scalar multiplication defined by (2.2). If $n = 1$, then we denote the vector space of column vectors $\mathbb{R}^{m \times 1}$ by \mathbb{R}^m .

2. Let a, b be real numbers with $a < b$. The set $C[a, b]$ of continuous functions on the interval $[a, b]$ with addition of functions is an abelian group. Let scalar multiplication be defined as follows:

$$\text{if } \alpha \in \mathbb{R} \text{ and } f \in C[a, b], \text{ then } (\alpha \cdot f)(x) = \alpha f(x), \quad x \in [a, b]. \quad (2.3)$$

Then $\alpha \cdot f \in C[a, b]$, and moreover V1, V2, V3, V4 are satisfied:

V1. Let $f \in C[a, b]$. For all $x \in [a, b]$, we have

$$(1 \cdot f)(x) = 1f(x) = f(x),$$

and so $1 \cdot f = f$.

V2. Let $\alpha, \beta \in \mathbb{R}$ and $f \in C[a, b]$. For all $x \in [a, b]$, we have

$$\begin{aligned}
(\alpha \cdot (\beta \cdot f))(x) &= \alpha(\beta \cdot f)(x) \\
&= \alpha(\beta f(x)) \\
&= (\alpha\beta)f(x) \\
&= ((\alpha\beta) \cdot f)(x),
\end{aligned}$$

and so $(\alpha \cdot (\beta \cdot f)) = (\alpha\beta) \cdot f$.

V3. Let $\alpha, \beta \in \mathbb{R}$ and $f \in C[a, b]$. For all $x \in [a, b]$, we have

$$\begin{aligned}
((\alpha + \beta) \cdot f)(x) &= (\alpha + \beta)f(x) \\
&= \alpha f(x) + \beta f(x) \\
&= (\alpha \cdot f)(x) + (\beta \cdot f)(x) \\
&= (\alpha \cdot f + \beta \cdot f)(x),
\end{aligned}$$

and so $(\alpha + \beta) \cdot f = \alpha \cdot f + \beta \cdot f$.

V4. Let $\alpha \in \mathbb{R}$ and $f, g \in C[a, b]$. For all $x \in [a, b]$, we have

$$\begin{aligned} (\alpha \cdot (f + g))(x) &= \alpha(f + g)(x) \\ &= \alpha(f(x) + g(x)) \\ &= \alpha f(x) + \alpha g(x) \\ &= (\alpha \cdot f)(x) + (\alpha \cdot g)(x) \\ &= (\alpha \cdot f + \alpha \cdot g)(x), \end{aligned}$$

and so $\alpha \cdot (f + g) = \alpha \cdot f + \alpha \cdot g$.

Hence $C[a, b]$ with addition and scalar multiplication is a vector space. \diamond

We now prove a few elementary properties of vector spaces.

Theorem 2.2.1 *Let V be a vector space. Then the following hold:*

1. For all $v \in V$, $0 \cdot v = \mathbf{0}$.
2. For all $\alpha \in \mathbb{R}$, $\alpha \cdot \mathbf{0} = \mathbf{0}$.
3. If $v \in V$, then $(-1) \cdot v = -v$.

Proof Please go through the entire proof carefully, noting which symbols refer to the number $0 \in \mathbb{R}$, and which refer to the zero vector $\mathbf{0} \in V$.

1. To see this, we use the distributive law to write

$$0 \cdot v + 0 \cdot v = (0 + 0) \cdot v = 0 \cdot v.$$

Now add $-(0 \cdot v)$ on both sides, to obtain $0 \cdot v = \mathbf{0}$.

2. Similarly, $\alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0} = \alpha \cdot (\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0}$, and hence $\alpha \cdot \mathbf{0} = \mathbf{0}$.

3. Finally, we have

$$\begin{aligned} v + (-1) \cdot v &= 1 \cdot v + (-1) \cdot v \text{ (since } 1 \cdot v = v) \\ &= (1 + (-1)) \cdot v \text{ (distributive law)} \\ &= 0 \cdot v \text{ (since } 1 + (-1) = 0) \\ &= \mathbf{0} \text{ (since } 0 \cdot v = \mathbf{0}), \end{aligned}$$

and so $v + (-1) \cdot v = \mathbf{0} = (-1) \cdot v + v$. Hence $(-1) \cdot v$ is an inverse of v . But the inverse of an element from a group is unique, and so $(-1) \cdot v = -v$. \blacksquare

Exercises.

1. Let $n \in \mathbb{N}$.
 - (a) Is the set of invertible $n \times n$ matrices having real entries with matrix addition and with scalar multiplication defined by (2.2) a vector space?

- (b) Is the set of invertible $n \times n$ matrices having real entries with matrix multiplication and with scalar multiplication defined by (2.2) a vector space?
2. Let V be a vector space. Prove that if $\alpha \in \mathbb{R}$ and $v \in V$ are such that $\alpha \cdot v = \mathbf{0}$, then either $\alpha = 0$ or $v = \mathbf{0}$.

HINT: If $\alpha \neq 0$, then $\alpha^{-1} \in \mathbb{R}$. Premultiply both sides of $\alpha \cdot v = \mathbf{0}$ by α^{-1} .

3. (*) Consider the set \mathbb{R}^∞ of all sequences with addition defined as follows:

$$\text{if } (a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty, \text{ then } (a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} = (a_n + b_n)_{n \in \mathbb{N}}, \quad (2.4)$$

and scalar multiplication defined as follows:

$$\text{if } \alpha \in \mathbb{R} \text{ and } (a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty, \text{ then } \alpha \cdot (a_n)_{n \in \mathbb{N}} = (\alpha a_n)_{n \in \mathbb{N}}. \quad (2.5)$$

Prove that \mathbb{R}^∞ is a vector space with the above addition and scalar multiplication.

2.2.2 Subspaces and linear combinations

Definition. Let V be a vector space. A subset U is called a *subspace* of V if

- S1. $\mathbf{0} \in U$.
- S2. If $v_1, v_2 \in U$, then $v_1 + v_2 \in U$.
- S3. If $v \in U$ and $\alpha \in \mathbb{R}$, then $\alpha \cdot v \in U$.

Subspaces of a vector space are just like subgroups of a group, that is, a subspace of a vector space is itself a vector space with the same addition and scalar multiplication as with V (this is easy to check). So a subspace is really a smaller vector space sitting inside a larger vector space.

Examples.

1. If V is a vector space, then the subset U comprising only the zero vector, namely $U = \{\mathbf{0}\}$, is a subspace of V .

Also, the entire vector space, that is $U = V$, is a subspace of V .

If a subspace U of V is neither $\{\mathbf{0}\}$ nor V , then it is called a *proper subspace* of V .

2. Consider the vector space $\mathbb{R}^{2 \times 2}$ with matrix addition and scalar multiplication defined by (2.2). Then the set of upper triangular matrices

$$U_1 = \left\{ \left[\begin{array}{cc} a & b \\ 0 & d \end{array} \right] \mid a, b, d \in \mathbb{R} \right\}$$

is a subspace of $\mathbb{R}^{2 \times 2}$.

Also, the set of *symmetric matrices*

$$U_2 = \left\{ \left[\begin{array}{cc} a & b \\ b & d \end{array} \right] \mid a, b, d \in \mathbb{R} \right\}$$

is a subspace of $\mathbb{R}^{2 \times 2}$.

3. Let $a, b \in \mathbb{R}$ and $a < b$. Consider the set of all polynomial functions

$$P[a, b] = \left\{ p : [a, b] \rightarrow \mathbb{R} \mid \begin{array}{l} \exists n \in \mathbb{N} \cup \{0\} \text{ and } a_0, a_1, a_2, \dots, a_n \in \mathbb{R} \text{ such that} \\ p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \text{ for all } x \in [a, b] \end{array} \right\}$$

Then $U = P[a, b]$ is a subspace of the vector space $C[0, 1]$ with addition of functions and scalar multiplication defined by (2.3). \diamond

Definitions. Let V be a vector space.

1. If v_1, \dots, v_n are vectors in V and $\alpha_1, \dots, \alpha_n$ belong to \mathbb{R} , then the vector $\alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$ is called a *linear combination of the vectors* v_1, \dots, v_n .
2. Let S be a nonempty subset of a vector space V . The *span of* S , denoted by $\text{span}(S)$, is defined as the set of all possible linear combinations⁷ of vectors from S :

$$\text{span}(S) = \{ \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n \mid n \in \mathbb{N}, v_1, \dots, v_n \in S, \alpha_1, \dots, \alpha_n \in \mathbb{R} \}.$$

Examples.

1. Let $m \in \mathbb{N}$. Any vector in the vector space \mathbb{R}^m is a linear combination of the vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, v_m = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Hence $\text{span}(\{v_1, \dots, v_m\}) = \mathbb{R}^m$.

2. Let $a, b \in \mathbb{R}$ with $a < b$. Any polynomial p on the interval $[a, b]$ is a linear combination of the functions from the set $S = \{1, x, x^2, \dots\}$. Hence $\text{span}(S) = P[a, b]$ in the vector space $C[a, b]$. \diamond

The span of a set of vectors turns out to be a special subspace of the vector space.

Theorem 2.2.2 *Let V be a vector space and S be a nonempty subset of V . Then $\text{span}(S)$ is the smallest subspace of V that contains S .*

Proof We first show that $\text{span}(S)$ is a subspace of V .

Let $v \in S$. Then $\mathbf{0} = 0 \cdot v \in \text{span}(S)$. If $u, v \in \text{span}(S)$, then we know that

$$u = \alpha_1 \cdot u_1 + \dots + \alpha_n \cdot u_n \text{ and } v = \beta_1 \cdot v_1 + \dots + \beta_m \cdot v_m$$

for some vectors $u_1, \dots, u_n, v_1, \dots, v_m \in S$ and scalars $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m \in \mathbb{R}$. Consequently,

$$u + v = \alpha_1 \cdot u_1 + \dots + \alpha_n \cdot u_n + \beta_1 \cdot v_1 + \dots + \beta_m \cdot v_m \in \text{span}(S).$$

⁷Note that although S might be infinite, a linear combination, by definition, is always a linear combination of a *finite* set of vectors from S .

Finally, if $v \in \text{span}(S)$, then we know that $v = \alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n$ for some $v_1, \dots, v_n \in S$ and scalars $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, and so for any $\beta \in \mathbb{R}$, we have

$$\beta \cdot v = \beta \cdot (\alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n) = \beta \cdot (\alpha_1 \cdot v_1) + \cdots + \beta \cdot (\alpha_n \cdot v_n) = (\beta \alpha_1) \cdot v_1 + \cdots + (\beta \alpha_n) \cdot v_n \in \text{span}(S).$$

So $\text{span}(S)$ satisfies S1,S2,S3, and so it is a subspace of V . Moreover, if $v \in S$, then $v = 1 \cdot v \in \text{span}(S)$, and so $S \subset \text{span}(S)$. Thus $\text{span}(S)$ contains S .

If U is another subspace that contains the vectors from S , then from S2 and S3 it follows that it certainly contains all linear combinations of vectors from S belong to U , and so $\text{span}(S) \subset U$.

Hence $\text{span}(S)$ is the smallest subspace of V containing S . ■

Definitions. Let V be a vector space, and suppose that v_1, \dots, v_n are vectors that belong to V .

1. The vectors v_1, \dots, v_n are called *linearly independent* if the following condition holds:

$$\text{if } \exists \alpha_1, \dots, \alpha_n \in \mathbb{R} \text{ such that } \alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n = \mathbf{0}, \text{ then } \alpha_1 = \cdots = \alpha_n = 0.$$

An arbitrary subset S of vectors from V is said to be *linearly independent* if every nonempty finite set of vectors from S is an independent set of vectors.

2. The vectors v_1, \dots, v_n are called *linearly dependent* if they are not linearly independent, that is,

$$\exists \alpha_1, \dots, \alpha_n \in \mathbb{R}, \text{ not all zeros, such that } \alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n = \mathbf{0}.$$

An arbitrary subset S of vectors from V is said to be a *linearly dependent* if there exists a nonempty finite set of dependent vectors from S .

Examples.

1. Let V be a vector space. Then any finite set of vectors from V containing the zero vector is linearly dependent. Indeed if $v_1, \dots, v_n \in V$ and $v_k = \mathbf{0}$, then

$$0 \cdot v_1 + \cdots + 0 \cdot v_{k-1} + 1 \cdot v_k + 0 \cdot v_{k+1} + \cdots + 0 \cdot v_n = \mathbf{0}.$$

2. The vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, v_m = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

are linearly independent in \mathbb{R}^m . Indeed if $\alpha_1 \cdot v_1 + \cdots + \alpha_m \cdot v_m = \mathbf{0}$, then

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} = \alpha_1 \cdot \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \alpha_m \cdot \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

and so $\alpha_1 = \cdots = \alpha_m = 0$.

3. Let $a, b \in \mathbb{R}$ with $a < b$. The functions $1, x$ on the interval $[a, b]$ are linearly independent. Indeed, if for all $x \in [a, b]$, $\alpha \cdot 1 + \beta \cdot x = \mathbf{0}(x)$, then in particular, we have

$$\alpha + \beta a = 0 \text{ and } \alpha + \beta b = 0,$$

and since $a \neq b$, it follows that $\alpha = \beta = 0$. ◇

Exercises.

1. Determine if the following statements are TRUE or FALSE:

- (a) The union of two subspaces of a vector space V is a subspace of V .
- (b) The intersection of two subspaces of a vector space V is a subspace of V .
- (c) $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \text{span} \left(\left\{ \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} \right)$ in the vector space \mathbb{R}^3 .
- (d) The vectors $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$ are linearly independent in the vector space \mathbb{R}^3 .
- (e) If v_1, v_2, v_3, v_4 are linearly independent, then v_1, v_2, v_3 are linearly independent.
- (f) If v_1, v_2, v_3, v_4 are linearly dependent, then v_1, v_2, v_3 are linearly dependent.

2. (a) Prove that if t_1 and t_2 are distinct real numbers in \mathbb{R} , then

$$\text{span} \left(\left\{ \begin{bmatrix} 1 \\ t_1 \end{bmatrix}, \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \right\} \right) = \mathbb{R}^2$$

in the vector space \mathbb{R}^2 .

(b) (*) Prove that \mathbb{R}^2 is not the union of a finite number of proper subspaces.

HINT: Consider the infinite subset $S = \left\{ \begin{bmatrix} 1 \\ t \end{bmatrix} \mid t \in \mathbb{R} \right\}$.

3. (*) Let \mathbb{R}^∞ be the vector space of all sequences, with addition and scalar multiplication defined by (2.4) and (2.5), respectively. We define the following subsets of \mathbb{R}^∞ :

- (a) ℓ^∞ is the set of all bounded sequences.
- (b) c is the set of all convergent sequences.
- (c) c_0 is the set of all convergent sequences with limit 0.
- (d) $c_{00} = \{(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty \mid \exists N \in \mathbb{N} \text{ such that } \forall n > N, a_n = 0\}$, is the set of all sequences that are *eventually zero*.

Prove that $c_{00} \subset c_0 \subset c \subset \ell^\infty \subset \mathbb{R}^\infty$, and that each is a subspace of the next one.

4. Consider the vector space $C[0, 1]$ with addition of functions and scalar multiplication defined by (2.3). Let $S(y_1, y_2) = \{f \in C[0, 1] \mid f(0) = y_1 \text{ and } f(1) = y_2\}$. Show that $S(y_1, y_2)$ is a subspace of $C[0, 1]$ iff $y_1 = 0 = y_2$.

2.2.3 Basis of a vector space

Definition. Let V be a vector space. Then a set of vectors B is said to be a *basis* of V if

B1. $\text{span}(B) = V$, and

B2. B is linearly independent.

Example. The vectors

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, v_m = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

form a basis of \mathbb{R}^m . ◇

Theorem 2.2.3 *Let V be a vector space and B be a basis of V such that B has n elements. Then any linearly independent set S of n vectors is also a basis of V .*

Proof We simply prove that $\text{span}(S) = V$. This is proved as follows: we claim that for each $k \in \{0, 1, \dots, n\}$, there is a set S_k with n elements such that it has $n - k$ elements from B and the other k elements belong to S , and such that $\text{span}(S_k) = V$. We prove this claim by induction on k .

For $k = 0$, we simply define $S_0 = B$. Then S_0 is a set with n elements such that it has $n - 0 = n$ elements from B and the other 0 elements belong to S , and since B is a basis, $V = \text{span}(B) = \text{span}(S_0)$.

Suppose that the claim is true for some k . Thus S_k is a set with n elements such that it has $n - k$ elements v_1, \dots, v_{n-k} from B and the other k elements, say u_1, \dots, u_k , belong to S , and such that $\text{span}(S_k) = V$. Now suppose that u_r is an element from S such that it does not belong to $\{u_1, \dots, u_k\}$. Since $\text{span}(S_k) = V$, there exist numbers $\alpha_1, \dots, \alpha_{n-k}, \beta_1, \dots, \beta_k \in \mathbb{R}$ such that

$$u_r = \alpha_1 \cdot v_1 + \dots + \alpha_{n-k} \cdot v_{n-k} + \beta_1 \cdot u_1 + \dots + \beta_k \cdot u_k.$$

Then we have the following two cases:

1° If $\alpha_1 = \dots = \alpha_{n-k} = 0$, then we get that $1 \cdot u_r - \beta_1 \cdot u_1 - \dots - \beta_k \cdot u_k = \mathbf{0}$, which contradicts the linear independence of S . So this case is not possible.

2° Hence there must exist an $\alpha_j \neq 0$. So the spans of the sets

$$\begin{aligned} S_k &= \{v_1, \dots, v_{j-1}, v_j, v_{j+1}, \dots, v_{n-k}, u_1, \dots, u_k\}, \text{ and} \\ S_{k+1} &:= \{v_1, \dots, v_{j-1}, u_r, v_{j+1}, \dots, v_{n-k}, u_1, \dots, u_k\} \end{aligned}$$

are the same. (Why?) Moreover, S_{k+1} has n elements, $n - k - 1$ of these belong to B , and the other elements (namely, u_r, u_1, \dots, u_k) belong to S .

Hence by induction, for $k = n$, we obtain that S_n has n elements, such that 0 of these are in B , and the other $n - 0$ elements are in S , and such that $\text{span}(S_n) = V$. But S has n elements, and so $S_n = S$. Thus $\text{span}(S) = V$. ■

Given a vector space, there are of course many bases. However, the next result says that the cardinality of the basis is unique for any given vector space.

Corollary 2.2.4 *If a vector space V has a basis with n elements, then every basis of V has the same number of elements.*

Proof Suppose that B is a basis of V with n elements and suppose that B' is another basis of V .

Let B' have $> n$ elements. Then take any n distinct elements v_1, \dots, v_n from B' . From the previous theorem, it follows that these span V , and so if $v \in B' \setminus \{v_1, \dots, v_n\}$, it can be written as a linear combination of $\{v_1, \dots, v_n\}$, which contradicts the independence of B' .

If B' has $< n$ elements, then by interchanging the roles of B and B' and proceeding as above, we once again arrive at a contradiction. ■

The above result motivates the following natural definitions.

Definition. Let V be a vector space.

1. If there exists a basis B of V such that B has n elements, then n is called the *dimension of V* , and it is denoted by $\dim(V)$.
2. If a vector space has a basis with a finite number of elements, then it is called a *finite dimensional vector space*.
3. If a vector space is not finite dimensional, then it is called an *infinite dimensional vector space*.

Exercises.

1. Prove or disprove that

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

is a basis for \mathbb{R}^3 .

2. Prove that if B is a basis of a finite dimensional vector space V , then every element $v \in V$ can be written as a unique linear combination of the vectors from B .
3. (*) Show that $C[0, 1]$ is infinite dimensional.

HINT: One can prove this by contradiction. Let $C[0, 1]$ be a finite dimensional vector space with dimension d , say. First show that the set $B = \{x, x^2, \dots, x^d\}$ is linearly independent. Then by Theorem 2.2.3, B is a basis for $C[0, 1]$, and so the constant function 1 should be a linear combination of the functions from B . Derive a contradiction.

4. (a) (*) For $k \in \mathbb{N}$, let e_k denote the sequence with the k th term equal to 1, and all other terms equal to zero:

$$e_k = (a_n)_{n \in \mathbb{N}}, \text{ where } a_n = \begin{cases} 1 & \text{if } n = k, \\ 0 & \text{if } n \neq k, \end{cases}$$

and let $B = \{e_k \mid k \in \mathbb{N}\}$. Prove that B is a basis for the vector space c_{00} , comprising all sequences that are eventually zero.

- (b) (*) Is $B = \{e_k \mid k \in \mathbb{N}\}$ also a basis for the vector space \mathbb{R}^∞ ?

HINT: Consider the constant sequence $(1)_{n \in \mathbb{N}}$.

2.2.4 Linear transformations

Definition. Let U, V be two vector spaces. A function $T : U \rightarrow V$ is called a *linear transformation* if it satisfies the following two properties:

L1. For all $u_1, u_2 \in U$, $T(u_1 + u_2) = T(u_1) + T(u_2)$.

L2. For all $u \in U$, and all $\alpha \in \mathbb{R}$, $T(\alpha \cdot u) = \alpha \cdot T(u)$.

Thus just like group homomorphisms that are functions that respect group operations, linear transformations are functions that respect vector space operations.

Examples.

1. Let $m, n \in \mathbb{N}$. Let

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Then the function $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by

$$T_A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n a_{1k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{bmatrix} \quad \text{for all } \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n, \quad (2.6)$$

is a linear transformation from the vector space \mathbb{R}^n to the vector space \mathbb{R}^m . Indeed, we have

$$\begin{aligned} T_A \left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right) &= T_A \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n a_{1k}(x_k + y_k) \\ \vdots \\ \sum_{k=1}^n a_{mk}(x_k + y_k) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^n (a_{1k}x_k + a_{1k}y_k) \\ \vdots \\ \sum_{k=1}^n (a_{mk}x_k + a_{mk}y_k) \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n a_{1k}x_k + \sum_{k=1}^n a_{1k}y_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k + \sum_{k=1}^n a_{mk}y_k \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^n a_{1k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{bmatrix} + \begin{bmatrix} \sum_{k=1}^n a_{1k}y_k \\ \vdots \\ \sum_{k=1}^n a_{mk}y_k \end{bmatrix} = T_A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + T_A \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \end{aligned}$$

for all

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n,$$

and so L1 holds. Moreover,

$$\begin{aligned} T_A \left(\alpha \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right) &= T_A \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n a_{1k} \alpha x_k \\ \vdots \\ \sum_{k=1}^n a_{mk} \alpha x_k \end{bmatrix} = \begin{bmatrix} \alpha \sum_{k=1}^n a_{1k} x_k \\ \vdots \\ \alpha \sum_{k=1}^n a_{mk} x_k \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} \sum_{k=1}^n a_{1k} x_k \\ \vdots \\ \sum_{k=1}^n a_{mk} x_k \end{bmatrix} = \alpha \cdot T_A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \end{aligned}$$

for all $\alpha \in \mathbb{R}$ and all vectors

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n,$$

and so L2 holds as well. Hence T_A is a linear transformation.

2. The function $T : C[0, 1] \rightarrow \mathbb{R}$ given by

$$Tf = f\left(\frac{1}{2}\right) \text{ for all } f \in C[0, 1],$$

is a linear transformation from the vector space $C[0, 1]$ to the vector space \mathbb{R} . Indeed, we have

$$T(f + g) = (f + g)\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right) + g\left(\frac{1}{2}\right) = T(f) + T(g), \text{ for all } f, g \in C[0, 1],$$

and so L1 holds. Furthermore

$$T(\alpha \cdot f) = (\alpha \cdot f)\left(\frac{1}{2}\right) = \alpha f\left(\frac{1}{2}\right) = \alpha T(f), \text{ for all } \alpha \in \mathbb{R} \text{ and all } f \in C[0, 1],$$

and so L2 holds too. Thus T is a linear transformation. \diamond

Just as in the case of homomorphisms between groups, there are two important subsets associated with a linear transformation between vector spaces, whose definitions are given below.

Definitions. Let U, V be vector spaces and $T : U \rightarrow V$ a linear transformation.

1. The *kernel* of T is defined to be the set $\ker(T) = \{u \in U \mid T(u) = \mathbf{0}_V\}$, where $\mathbf{0}_V$ denotes the zero vector in V .
2. The *image* of T is defined to be the set $\text{im}(T) = \{v \in V \mid \exists u \in U \text{ such that } T(u) = v\}$.

Examples.

1. Let $A \in \mathbb{R}^{m \times n}$, and let $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the linear transformation defined by (2.6). The kernel of the linear transformation is the set of all vectors

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

such that the system of linear equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

is simultaneously satisfied.

The range of T_A is the set of all vectors

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^m$$

such that there exists a vector

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

such that

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= y_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= y_m. \end{aligned}$$

2. The function $T : C[0, 1] \rightarrow \mathbb{R}$ given by $Tf = f(\frac{1}{2})$ for all $f \in C[0, 1]$, is a linear transformation from the vector space $C[0, 1]$ to the vector space \mathbb{R} , with kernel equal to the set of continuous functions on the interval $[0, 1]$ that vanish at the point $\frac{1}{2}$. The range of T is the whole vector space \mathbb{R} . \diamond

Analogous to Theorem 2.1.5 for homomorphisms between groups, we now prove the following result.

Theorem 2.2.5 *Let U, V be vector spaces and $T : U \rightarrow V$ a linear transformation. Then:*

1. $\ker(T)$ is a subspace of U .
2. $\text{im}(T)$ is a subspace of V .

Proof In each of the cases, we check that S1, S2, S3 hold.

Let $\mathbf{0}_U, \mathbf{0}_V$ denote the zero vectors in the vector spaces U, V , respectively. It is easy to check that $\ker(T)$ is a subspace of U . Indeed, as $T(\mathbf{0}_U) = T(\mathbf{0}_U + \mathbf{0}_U) = T(\mathbf{0}_U) + T(\mathbf{0}_U)$, it follows that $T(\mathbf{0}_U) = \mathbf{0}_V$, and so $\mathbf{0}_U \in \ker(T)$. Thus S1 holds. If u_1, u_2 belong to $\ker(T)$, then $T(u_1 + u_2) = T(u_1) + T(u_2) = \mathbf{0}_V + \mathbf{0}_V = \mathbf{0}_V$, and so S2 holds as well. Finally, S3 holds, since if $\alpha \in \mathbb{R}$ and $u \in \ker(T)$, then $T(\alpha \cdot u) = \alpha \cdot T(u) = \alpha \cdot \mathbf{0}_V = \mathbf{0}_V$. Hence $\ker(T)$ is a subspace of U .

We now also check that $\text{im}(T)$ is a subspace of V . Since $T(\mathbf{0}_U) = \mathbf{0}_V$, it follows that $\mathbf{0}_V \in \text{im}(T)$, and so S1 holds. If v_1, v_2 belong to $\text{im}(T)$, then there exist elements u_1, u_2 in U such that $T(u_1) = v_1$ and $T(u_2) = v_2$. Consequently, $T(u_1 + u_2) = T(u_1) + T(u_2) = v_1 + v_2$, and so there exists an element in U , namely $u_1 + u_2$, such that $T(u_1 + u_2) = v_1 + v_2$, that is, $v_1 + v_2 \in \text{im}(T)$. Thus S2 holds. Finally, if $\alpha \in \mathbb{R}$ and $v \in \text{im}(T)$, then there exists a $u \in U$ such that $T(u) = v$, and so $\alpha \cdot v = \alpha \cdot T(u) = T(\alpha \cdot u)$. Hence $\alpha \cdot v \in \text{im}(T)$, and S3 holds. So $\text{im}(T)$ is a subspace of V . ■

Exercises.

- Find the kernel and image of the linear transformations $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, where $A \in \mathbb{R}^{2 \times 2}$ is given by:

$$(a) \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

$$(b) \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

$$(c) \quad A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Each vector $\begin{bmatrix} x \\ y \end{bmatrix}$ in \mathbb{R}^2 can be represented by a point in the (x, y) -plane. In each of the cases, draw a picture of the subspaces $\ker(T_A)$ and $\text{im}(T_A)$ in the plane.

- Consider the vector space \mathbb{R}^2 with matrix addition and the usual scalar multiplication defined by (2.2). Define the function $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ as follows:

$$\text{if } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2, \text{ then } T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{cases} \begin{bmatrix} \frac{x_1^2}{x_2} \\ x_2 \end{bmatrix} & \text{if } x_1 x_2 \neq 0, \\ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & \text{if } x_1 x_2 = 0. \end{cases}$$

Show that T satisfies L2, but not L1, and hence it is not a linear transformation.

- (*) Let c denote the vector space of all convergent sequences. Consider the function $T : c \rightarrow \mathbb{R}$ given by

$$T((a_n)_{n \in \mathbb{N}}) = \lim_{n \rightarrow \infty} a_n, \text{ for all sequences } (a_n)_{n \in \mathbb{N}} \in c.$$

- Prove that T is a linear transformation from the vector space c to the vector space \mathbb{R} .
- What is the kernel of T ?
- Show that $\text{im}(T) = \mathbb{R}$.

Solutions

Analysis

Solutions to the exercises on page 6

1. (a) $S = (0, 1]$.

An upper bound of S . 1 is an upper bound, since for all $x \in S = (0, 1]$, we have $x \leq 1$. In fact any real number $u \geq 1$ is an upper bound.

A lower bound of S . 0 is a lower bound, since for all $x \in S = (0, 1]$, we have $0 < x$. In fact any real number $l \leq 0$ is a lower bound.

Is S bounded? Yes. S is bounded above, since 1 is an upper bound of S . S is also bounded below, since 0 is a lower bound. Since S is bounded above as well as bounded below, it is bounded.

Supremum of S . $\sup S = 1$. Indeed, 1 is an upper bound, and moreover, if u is also an upper bound, then $1 \leq u$ (since $1 \in S$).

Infimum of S . $\inf S = 0$. First of all, 0 is a lower bound. Let l be a lower bound of S . We prove that $l \leq 0$. (We do this by supposing that $l > 0$, and arriving at a contradiction. The contradiction is obtained as follows: if $l > 0$, then we will see that the average of 0 and l , namely $\frac{l}{2}$, is an element in S that is less than the lower bound l , which is a contradiction to the definition of a lower bound!) If $l > 0$, then $0 < \frac{l}{2}$. Moreover, since $l \leq 1$ (l is a lower bound of S and $1 \in S$) it follows that $\frac{l}{2} \leq \frac{1}{2} \leq 1$. Thus $\frac{l}{2} \in S$. But since $l > 0$, it follows that $\frac{l}{2} < l$, a contradiction. Hence $l \leq 0$.

If $\sup S$ exists, then is $\sup S \in S$? Yes, $\sup S = 1 \in (0, 1] = S$.

If $\inf S$ exists, then is $\inf S \in S$? No, $\inf S = 0 \notin (0, 1] = S$.

Maximum of S . $\max S = 1$, since $\sup S = 1 \in S$.

Minimum of S . $\min S$ does not exist since $\inf S = 0 \notin S$.

(b) $S = [0, 1]$.

An upper bound of S . 1 is an upper bound, since for all $x \in S = [0, 1]$, we have $x \leq 1$.

A lower bound of S . 0 is a lower bound, since for all $x \in S = [0, 1]$, we have $0 \leq x$.

Is S bounded? Yes, since S is bounded above (1 is an upper bound) and it is bounded below (0 is a lower bound).

Supremum of S . $\sup S = 1$. Indeed, 1 is an upper bound, and moreover, if u is also an upper bound, then $1 \leq u$ (since $1 \in S$).

Infimum of S . $\inf S = 0$. Indeed, 0 is a lower bound, and moreover, if l is also a lower bound, then $l \leq 0$ (since $0 \in S$).

If $\sup S$ exists, then is $\sup S \in S$? Yes, $\sup S = 1 \in [0, 1] = S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = 0 \in [0, 1] = S$.

Maximum of S . $\max S = 1$, since $\sup S = 1 \in S$.

Minimum of S . $\min S = 0$, since $\inf S = 0 \in S$.

(c) $S = (0, 1)$.

An upper bound of S . 1 is an upper bound, since for all $x \in S = (0, 1)$, $x < 1$.

A lower bound of S . 0 is a lower bound, since for all $x \in S = (0, 1)$, $0 < x$.

Is S bounded? Yes, since S is bounded above (1 is an upper bound) and it is bounded below (0 is a lower bound).

Supremum of S . $\sup S = 1$.

First of all, 1 is a upper bound. Let u be an upper bound of S . We prove that $1 \leq u$. (We do this by supposing that $u < 1$ and arriving at a contradiction. The contradiction is obtained as follows: if $u < 1$, then we will see that the average of u and 1, namely $\frac{u+1}{2}$, is an element in S that is larger than the upper bound u , which is a contradiction to the definition of an upper bound!) Since u is an upper bound and since $\frac{1}{2} \in S$, it follows that $0 < u$ (since $0 < \frac{1}{2} \leq u$). So if $u < 1$, then $0 < u = \frac{u+u}{2} < \frac{u+1}{2} < \frac{1+1}{2} = 1$. Hence $\frac{u+1}{2} \in S$. But $u < \frac{u+1}{2}$ contradicts the fact that u is an upper bound of S .

Infimum of S . $\inf S = 0$.

First of all, 0 is a lower bound. Let l be a lower bound of S . We prove that $l \leq 0$. If $l > 0$, then $0 < \frac{l}{2}$. Moreover, since $\frac{1}{2} \in S$ and l is a lower bound of S , it follows that $l \leq \frac{1}{2}$. Thus we have $0 < \frac{l}{2} < l \leq \frac{1}{2} < 1$, and so $\frac{l}{2} \in S$. But $\frac{l}{2} < l$ contradicts the fact that l is a lower bound of S .

If $\sup S$ exists, then is $\sup S \in S$? No, $\sup S = 1 \notin (0, 1) = S$.

If $\inf S$ exists, then is $\inf S \in S$? No, $\inf S = 0 \notin (0, 1) = S$.

Maximum of S . $\max S$ does not exist, since $\sup S = 1 \notin S$.

Minimum of S . $\min S$ does not exist, since $\inf S = 0 \notin S$.

(d) $S = \{\frac{1}{n} \mid n \in \mathbb{Z} \setminus \{0\}\} = \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{-\frac{1}{n} \mid n \in \mathbb{N}\}$.

An upper bound of S . 1 is an upper bound, since for all $n \in \mathbb{N}$, $\frac{1}{n} \leq 1$ and $-\frac{1}{n} \leq 0 \leq 1$.

A lower bound of S . -1 is a lower bound, since for all $n \in \mathbb{N}$, $-1 \leq 0 \leq \frac{1}{n}$ and $-1 \leq -\frac{1}{n}$.

Is S bounded? Yes, since S is bounded above (1 is an upper bound) and it is bounded below (-1 is a lower bound).

Supremum of S . $\sup S = 1$. 1 is an upper bound. Moreover, if u is also an upper bound, then since $1 = \frac{1}{1} \in S$, $1 \leq u$.

Infimum of S . $\inf S = -1$. -1 is a lower bound. Moreover, if l is also a lower bound, then since $-1 = \frac{1}{-1} \in S$, it follows that $l \leq -1$.

If $\sup S$ exists, then is $\sup S \in S$? Yes, $\sup S = 1 \in S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = -1 \in S$.

Maximum of S . $\max S = 1$, since $\sup S = 1 \in S$.

Minimum of S . $\min S = -1$, since $\inf S = -1 \in S$.

(e) $S = \{-\frac{1}{n} \mid n \in \mathbb{N}\}$.

An upper bound of S . 0 is an upper bound, since for all $n \in \mathbb{N}$, $-\frac{1}{n} < 0$.

A lower bound of S . -1 is a lower bound, since for all $n \in \mathbb{N}$, $-1 \leq -\frac{1}{n}$.

Is S bounded? Yes, since S is bounded above (0 is an upper bound) and it is bounded below (-1 is a lower bound).

Supremum of S . $\sup S = 0$. 0 is an upper bound. Moreover, if u is an upper bound and if $u < 0$, then let $N \in \mathbb{N}$ be such that $\frac{1}{-u} < N$ (Archimedean principle with $y = \frac{1}{-u}$ and $x = 1$!), and so we obtain $u < -\frac{1}{N} \in S$, a contradiction to the fact that u is an upper bound of S . Thus if u is an upper bound, then $0 \leq u$.

Infimum of S . $\inf S = -1$. -1 is a lower bound, and if l is another lower bound, then since $-1 = -\frac{1}{1} \in S$, it follows that $l \leq -1$.

If $\sup S$ exists, then is $\sup S \in S$? No, since $\sup S = 0 \notin S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = -1 \in S$.

Maximum of S . $\max S$ does not exist since $\sup S = 0 \notin S$.

Minimum of S . $\min S = -1$ since $\inf S = -1 \in S$.

(f) $S = \left\{ \frac{n}{n+1} \mid n \in \mathbb{N} \right\}$.

An upper bound of S . 1 is an upper bound, since for all $n \in \mathbb{N}$, $\frac{n}{n+1} < 1$.

A lower bound of S . $\frac{1}{2}$ is a lower bound because for all $n \in \mathbb{N}$, $\frac{1}{2} \leq \frac{n}{n+1}$ (since $n+1 \leq 2n$, that is, $1 \leq n$).

Is S bounded? Yes, since S is bounded above (1 is an upper bound) and it is bounded below ($\frac{1}{2}$ is a lower bound).

Supremum of S . $\sup S = 1$. 1 is an upper bound of S . If $u < 1$ is an upper bound, then let $N \in \mathbb{N}$ be such that $\frac{u}{u-1} < N$ (Archimedean property). Then $u < \frac{N}{N+1}$, contradicting the fact that u is an upper bound.

Infimum of S . $\inf S = \frac{1}{2}$. $\frac{1}{2}$ is a lower bound, and if l is a lower bound, then since $\frac{1}{2} = \frac{1}{1+1} \in S$, it follows that $l \leq \frac{1}{2}$.

If $\sup S$ exists, then is $\sup S \in S$? No, since $\sup S = 1 \notin S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, since $\inf S = \frac{1}{2} \in S$.

Maximum of S . $\max S$ does not exist since $\sup S = 1 \notin S$.

Minimum of S . $\min S$ exists since $\inf S = \frac{1}{2} \in S$.

(g) $S = \{x \in \mathbb{R} \mid x^2 \leq 2\}$.

An upper bound of S . $\sqrt{2}$ is an upper bound. ($x > \sqrt{2}$ implies $x^2 > 2$, and so $x \notin S$. In other words, if $x \in S$, then $x \leq \sqrt{2}$, that is, $\sqrt{2}$ is an upper bound of S .)

A lower bound of S . $-\sqrt{2}$ is a lower bound. ($x < -\sqrt{2}$ implies that $-x > \sqrt{2} > 1 > 0$. So using the fact that if $a > b$ and $c > 0$, then $ac > bc$, we get $x^2 = (-x)(-x) > (-x)\sqrt{2} > \sqrt{2}\sqrt{2} = 2$, and so, $x \notin S$. In other words, if $x \in S$, then $x \geq -\sqrt{2}$, and so $-\sqrt{2}$ is a lower bound.)

Is S bounded? Yes, since S is bounded above ($\sqrt{2}$ is an upper bound) and it is bounded below ($-\sqrt{2}$ is a lower bound).

Supremum of S . $\sup S = \sqrt{2}$. First of all, $\sqrt{2}$ is an upper bound of S . Let u be an upper bound such that $u < \sqrt{2}$. Then $\frac{u+\sqrt{2}}{2} \in S$. (As $0 \in S$ and u is an upper bound of S , $0 \leq u$. As $u < \sqrt{2}$, $u < \frac{u+\sqrt{2}}{2} < \sqrt{2}$. Hence we have $0 < \frac{u+\sqrt{2}}{2} < \sqrt{2}$, and so $\left(\frac{u+\sqrt{2}}{2}\right)^2 = \left(\frac{u+\sqrt{2}}{2}\right)\left(\frac{u+\sqrt{2}}{2}\right) < \left(\frac{u+\sqrt{2}}{2}\right)\sqrt{2} < \sqrt{2}\sqrt{2} = 2$. Thus $\frac{u+\sqrt{2}}{2} \in S$.) But

$u < \frac{u+\sqrt{2}}{2}$ contradicts the fact that u is an upper bound. So if u is an upper bound of S , then $\sqrt{2} \leq u$.

Infimum of S . $\inf S = -\sqrt{2}$. $-\sqrt{2}$ is a lower bound. Let l be a lower bound such that $-\sqrt{2} < l$. Then we have $-l < \sqrt{2}$, and since $\sup S = \sqrt{2}$, it follows that $-l$ cannot be an upper bound of S . So there exists an $x \in S$ such that $-l < x$. But since $x \in S$, we have $(-x)^2 = x^2 \leq 2$. So it follows that $-x \in S$. As l is a lower bound, $l \leq -x$, that is, $x \leq -l$. From $-l < x$ and $x \leq -l$, we arrive at the contradiction that $-l < -l$. Consequently, if l is a lower bound, then $l \leq -\sqrt{2}$.

If $\sup S$ exists, then is $\sup S \in S$? Yes, $\sup S = \sqrt{2} \in S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = -\sqrt{2} \in S$.

Maximum of S . $\max S = \sqrt{2}$.

Minimum of S . $\min S = -\sqrt{2}$.

(h) $S = \{0, 2, 5, 2005\}$.

An upper bound of S . 2005 is an upper bound, since $0 \leq 2005$, $2 \leq 2005$, $5 \leq 2005$ and $2005 \leq 2005$.

A lower bound of S . 0 is a lower bound, since $0 \leq 0$, $0 \leq 2$, $0 \leq 5$ and $0 \leq 2005$.

Is S bounded? Yes, since S is bounded above (2005 is an upper bound) and it is bounded below (0 is a lower bound).

Supremum of S . $\sup S = 2005$. 2005 is an upper bound, and if u is also an upper bound, then since $2005 \in S$, it follows that $2005 \leq u$.

Infimum of S . $\inf S = 0$. 0 is a lower bound, and if l is also a lower bound, then since $0 \in S$, it follows that $l \leq 0$.

If $\sup S$ exists, then is $\sup S \in S$? Yes, $\sup S = 2005 \in \{0, 2, 5, 2005\} = S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = 0 \in \{0, 2, 5, 2005\} = S$.

Maximum of S . $\max S = 2005$ since $\sup S = 2005 \in S$.

Minimum of S . $\min S = 0$ since $\inf S = 0 \in S$.

(i) $S = \left\{(-1)^n \left(1 + \frac{1}{n}\right) \mid n \in \mathbb{N}\right\}$.
(This set has the elements $-\frac{2}{1}, \frac{3}{2}, -\frac{4}{3}, \frac{5}{4}, \dots$)

An upper bound of S . $\frac{3}{2}$ is an upper bound.

If $n \in \mathbb{N}$ and n is even, then $(-1)^n \left(1 + \frac{1}{n}\right) = 1 + \frac{1}{n} \leq 1 + \frac{1}{2} = \frac{3}{2}$.

If $n \in \mathbb{N}$ and n is odd, then $(-1)^n \left(1 + \frac{1}{n}\right) = -1 - \frac{1}{n} < 0 < \frac{3}{2}$.

A lower bound of S . -2 is a lower bound.

If $n \in \mathbb{N}$ and n is even, then $(-1)^n \left(1 + \frac{1}{n}\right) = 1 + \frac{1}{n} > 0 > -2$.

If $n \in \mathbb{N}$ and n is odd, then $(-1)^n \left(1 + \frac{1}{n}\right) = -1 - \frac{1}{n} \geq -1 - 1 = -2$.

Is S bounded? Yes, since S is bounded above ($\frac{3}{2}$ is an upper bound) and it is bounded below (-2 is a lower bound).

Supremum of S . $\sup S = \frac{3}{2}$. $\frac{3}{2}$ is an upper bound, and if u is also an upper bound, then since $\frac{3}{2} = (-1)^2 \left(1 + \frac{1}{2}\right) \in S$, it follows that $\frac{3}{2} \leq u$.

Infimum of S . $\inf S = -2$. -2 is a lower bound, and if l is also a lower bound, then since $-2 = (-1)^1 \left(1 + \frac{1}{1}\right) \in S$, it follows that $l \leq -2$.

If $\sup S$ exists, then is $\sup S \in S$? Yes, $\sup S = \frac{3}{2} \in S$.

Maximum of S . $\max S = \frac{3}{2}$.

Minimum of S . $\min S = -2$.

(j) $S = \{x^2 \mid x \in \mathbb{R}\}$.

An upper bound of S . There does not exist an upper bound of S . Let $u \in \mathbb{R}$ be an upper bound of S , then $u + \frac{1}{2} \in \mathbb{R}$, and so $u^2 + u + \frac{1}{4} = \left(u + \frac{1}{2}\right)^2 \in S$. But since $u^2 + \frac{1}{4} > 0$, it follows that $u < u^2 + u + \frac{1}{4} = \left(u + \frac{1}{2}\right)^2 \in S$, contradicting the fact that u is an upper bound of S . Thus S does not have an upper bound.

A lower bound of S . 0 is a lower bound of S , since for all $x \in \mathbb{R}$, $0 \leq x^2$.

Is S bounded? No, since the set is not bounded above.

Supremum of S . $\sup S$ does not exist, since the set does not have an upper bound, and so it cannot have a least upper bound. (Recall that every least upper bound is an upper bound).

Infimum of S . $\inf S = 0$. 0 is a lower bound, and if l is also a lower bound, then since $0 = 0^2 \in S$, it follows that $l \leq 0$.

If $\sup S$ exists, then is $\sup S \in S$? $\sup S$ does not exist.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = 0 \in S$.

Maximum of S . $\max S$ does not exist, since $\sup S$ does not exist.

Minimum of S . $\min S = 0$, since $\inf S = 0 \in S$.

$$(k) S = \left\{ \frac{x^2}{1+x^2} \mid x \in \mathbb{R} \right\}.$$

An upper bound of S . 1 is an upper bound of S , since for all $x \in \mathbb{R}$, $\frac{x^2}{1+x^2} < 1$.

A lower bound of S . 0 is a lower bound of S , since for all $x \in \mathbb{R}$, $0 \leq \frac{x^2}{1+x^2}$.

Is S bounded? Yes, since S is bounded above (1 is an upper bound) and it is bounded below (0 is a lower bound).

Supremum of S . $\sup S = 1$. 1 is an upper bound. Let u be an upper bound and let $u < 1$. Since $0 = \frac{0^2}{1+0^2} \in S$, $0 \leq u$. Let $N \in \mathbb{N}$ be such that $\sqrt{\frac{u}{1-u}} < N$. Then we have $u < \frac{N^2}{1+N^2} \in S$, a contradiction to the fact that u is an upper bound of S . Hence if u is an upper bound of S , then $1 \leq u$.

Infimum of S . $\inf S = 0$. 0 is a lower bound, and if l is also a lower bound, then since $0 = \frac{0^2}{1+0^2} \in S$, it follows that $l \leq 0$.

If $\sup S$ exists, then is $\sup S \in S$? No, $\sup S = 1 \notin S$.

If $\inf S$ exists, then is $\inf S \in S$? Yes, $\inf S = 0 \in S$.

Maximum of S . $\max S$ does not exist, since $\sup S \notin S$.

Minimum of S . $\min S = 0$, since $\inf S = 0 \in S$.

2. (a) FALSE (if $S = \{1\}$, then $u = 3$ is an upper bound of S , and although $u' = 2 < 3 = u$, $u' (= 2)$ is an upper bound of $\{1\} = S$).
- (b) TRUE (if $\epsilon > 0$, then $u_* - \epsilon < u_*$, and so $u_* - \epsilon$ cannot be an upper bound of S).
- (c) FALSE (\mathbb{N} has no maximum).
- (d) FALSE (\mathbb{N} has no supremum).
- (e) FALSE ($[0, 1)$ is bounded, but it does not have a maximum).
- (f) FALSE (\emptyset is bounded, but it does not have a supremum).
- (g) TRUE (least upper bound property of \mathbb{R}).
- (h) TRUE (the supremum itself is an upper bound).
- (i) TRUE (definition of maximum).
- (j) FALSE (the set $[0, 1)$ has supremum 1, but $1 \notin [0, 1)$).
- (k) FALSE ($-\mathbb{N}$ is bounded above since 0 is an upper bound, but $|-\mathbb{N}| = \mathbb{N}$ is not bounded).
- (l) TRUE ($l \leq x \leq u$ implies $x \leq u$ and $-x \leq -l$, and so we have

$$x \leq u \leq \max\{-l, u\} \text{ and } -x \leq -l \leq \max\{-l, u\}.$$

Thus $|x| \leq \max\{-l, u\}$ and so $\max\{-l, u\}$ is an upper bound of $|S|$. Moreover, for every $y \in |S|$, we have $y = |x|$ for some $x \in S$, and so $y = |x| \geq 0$. Thus 0 is a lower bound of $|S|$.)

- (m) FALSE (if $S = \{0, 1\}$, then $\inf S = 0 < \frac{1}{2} < 1 = \sup S$, but $\frac{1}{2} \notin S$).

3. Since S is bounded, in particular, it is bounded above, and furthermore, since it is nonempty, $\sup S$ exists, by the least upper bound property of \mathbb{R} .

Since S is bounded, in particular it is bounded below, and furthermore, since it is nonempty, $\inf S$ exists, by the greatest lower bound property of \mathbb{R} .

Let $x \in S$. Since $\inf S$ is a lower bound of S ,

$$\inf S \leq x. \quad (2.7)$$

Moreover, since $\sup S$ is an upper bound of S ,

$$x \leq \sup S. \quad (2.8)$$

From (2.7) and (2.8), we obtain $\inf S \leq \sup S$.

Let $\inf S = \sup S$. If $x \in S$, then we have

$$\inf S \leq x \leq \sup S, \quad (2.9)$$

and so $\inf S = x (= \sup S)$ (for if $\inf S < x$, then from (2.9), $\inf S < \sup S$, a contradiction). Thus S is a singleton set. Conversely, if $S = \{x\}$, then clearly x is an upper bound. If $u < x$ is an upper bound, then $x \leq u < x$ gives $x < x$, a contradiction. So $\sup S = x$. Clearly x is also a lower bound. If $l > x$ is also a lower bound, then $x > l \geq x$ gives $x > x$, a contradiction. So $\inf S = x = \sup S$.

4. Since $\sup B$ is an upper bound of B , we have $x \leq \sup B$ for all $x \in B$. Since $A \subset B$, in particular we obtain $x \leq \sup B$ for all $x \in A$. So $\sup B$ is an upper bound of A , and so by the definition of the least upper bound of A , we obtain $\sup A \leq \sup B$.

5. Since A is bounded above, $\exists M_A \in \mathbb{R}$ such that $\forall x \in A, x \leq M_A$.

Since B is bounded above, $\exists M_B \in \mathbb{R}$ such that $\forall y \in B, y \leq M_B$.

Consequently if $x \in A$ and $y \in B$, $x + y \leq M_A + M_B$.

So $\forall z \in A + B, z \leq M_A + M_B$, and so $M_A + M_B$ is an upper bound for $A + B$. So $A + B$ is bounded above.

Clearly $A + B$ is nonempty. Indeed, A is nonempty implies that $\exists x \in A$; B is nonempty implies that $\exists y \in B$; thus $x + y \in A + B$, and so $A + B$ is not empty.

Since $A + B$ is bounded above and it is not empty, by the least upper bound property of \mathbb{R} , it follows that $\sup(A + B)$ exists.

Since $\sup A$ is an upper bound of A , $\forall x \in A, x \leq \sup A$.

Since $\sup B$ is an upper bound of B , $\forall y \in B, y \leq \sup B$.

So for all $x \in A$ and $y \in B$, $x + y \leq \sup A + \sup B$.

Since every $z \in A + B$ is such that $z = x + y$ with $x \in A$ and $y \in B$, it follows that for all $z \in A + B$, $z \leq \sup A + \sup B$. So $\sup A + \sup B$ is an upper bound of $A + B$, and consequently, by the definition of the least upper bound of $A + B$,

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

6. Let l be a lower bound of S : $\forall x \in S, l \leq x$. So $\forall x \in S, -x \leq -l$, in other words,

$$\forall y \in -S, y \leq -l.$$

Thus $-S$ is bounded above because $-l$ is an upper bound of $-S$.

Since S is nonempty, it follows that $\exists x \in S$, and so we obtain that $-x \in -S$. Hence $-S$ is nonempty.

As $-S$ is nonempty and bounded above, it follows that $\sup(-S)$ exists, by the least upper bound property of \mathbb{R} .

Since $\sup(-S)$ is an upper bound of $-S$, we have:

$$\forall y \in -S, y \leq \sup(-S),$$

that is,

$$\forall x \in S, -x \leq \sup(-S),$$

that is,

$$\forall x \in S, -\sup(-S) \leq x.$$

So $-\sup(-S)$ is a lower bound of S .

Next we prove that $-\sup(-S)$ is the greatest lower bound of S . Suppose that l' is a lower bound of S such that $-\sup(-S) < l'$. Then we have

$$\forall x \in S, -\sup(-S) < l' \leq x,$$

that is,

$$\forall x \in S, -x \leq -l' < \sup(-S),$$

that is,

$$\forall y \in -S, y \leq -l' < \sup(-S).$$

So $-l'$ is an upper bound of $-S$, and $-l' < \sup(-S)$, which contradicts the fact that $\sup(-S)$ is the least upper bound of $-S$. Hence $l' \leq -\sup(-S)$.

Consequently, $\inf S$ exists and $\inf S = -\sup(-S)$.

7. (S is nonempty and bounded below (0 is a lower bound), and so by the greatest lower bound property of \mathbb{R} , $\inf S$ exists.)

IF: Let $\inf S > 0$. Since $\inf S$ is a lower bound, it follows that

$$\forall x \in S, \inf S \leq x,$$

that is,

$$\forall x \in S, \frac{1}{x} \leq \frac{1}{\inf S},$$

that is,

$$\forall y \in S^{-1}, y \leq \frac{1}{\inf S}.$$

So $\frac{1}{\inf S}$ is an upper bound of S^{-1} . Thus S^{-1} is bounded above.

ONLY IF: Suppose S^{-1} is bounded above (with an upper bound u , say). Then

$$\forall y \in S^{-1} y \leq u,$$

that is,

$$\forall x \in S, \frac{1}{x} \leq u. \tag{2.10}$$

Since S is not empty, $\exists x_* \in S$ and so $\frac{1}{x_*} \leq u$. But $x_* \in S$ implies that $x_* > 0$, and so $\frac{1}{x_*} > 0$. Consequently $u > 0$. Hence from (2.10), we have

$$\forall x \in S, \frac{1}{u} \leq x.$$

Thus $\frac{1}{u}$ is a lower bound of S , and so

$$\frac{1}{u} \leq \inf S.$$

But $u > 0$ implies $\frac{1}{u} > 0$, and consequently $\inf S (\geq \frac{1}{u}) > 0$.

If $\inf S > 0$, then as in the IF part, we see that $\frac{1}{\inf S}$ is an upper bound of S^{-1} and so we obtain

$$\sup S^{-1} \leq \frac{1}{\inf S}. \quad (2.11)$$

Furthermore, since $u := \sup S^{-1}$ is an upper bound of S^{-1} , as in the ONLY IF part, we see that $\frac{1}{u} = \frac{1}{\sup S^{-1}}$ is a lower bound of S , and so

$$\frac{1}{\sup S^{-1}} \leq \inf S,$$

that is,

$$\frac{1}{\inf S} \leq \sup S^{-1}. \quad (2.12)$$

From (2.11) and (2.12), it follows that

$$\sup S^{-1} = \frac{1}{\inf S}.$$

8. (a) If $x \in \bigcap_{n \in \mathbb{N}} \left(0, \frac{1}{n}\right)$, then

$$\text{for all } n \in \mathbb{N}, 0 < x < \frac{1}{n}. \quad (2.13)$$

Let $N \in \mathbb{N}$ be such that $\frac{1}{x} < N$ (Archimedean property). Thus $\frac{1}{N} < x$, which contradicts (2.13). So

$$\neg \left[\exists x \in \bigcap_{n \in \mathbb{N}} \left(0, \frac{1}{n}\right) \right], \text{ that is, } \bigcap_{n \in \mathbb{N}} \left(0, \frac{1}{n}\right) = \emptyset.$$

(b) Clearly $\{0\} \subset \left[0, \frac{1}{n}\right]$ for all $n \in \mathbb{N}$ and so

$$\{0\} \subset \bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right]. \quad (2.14)$$

Let $x \in \bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right]$. Then $x \in [0, 1]$ and so $x \geq 0$. If $x > 0$, then let $N \in \mathbb{N}$ be such that $\frac{1}{x} < N$ (Archimedean property), that is, $\frac{1}{N} < x$. So $x \notin \left[0, \frac{1}{N}\right]$, and hence $x \notin \bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right]$. Consequently, if $x \in \bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right]$, then $x = 0$, that is,

$$\bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right] \subset \{0\}. \quad (2.15)$$

From (2.14) and (2.15), $\bigcap_{n \in \mathbb{N}} \left[0, \frac{1}{n}\right] = \{0\}$.

(c) Let $n \in \mathbb{N}$. If $x \in \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$, then $0 < \frac{1}{n} \leq x \leq 1 - \frac{1}{n} < 1$, and so $x \in (0, 1)$. Hence

$$\bigcup_{n \in \mathbb{N}} \left[\frac{1}{n}, 1 - \frac{1}{n}\right] \subset (0, 1). \quad (2.16)$$

If $x \in (0, 1)$, then $0 < x < 1$. Let $N_1 \in \mathbb{N}$ be such that $\frac{1}{x} < N_1$ (Archimedean property), that is, $\frac{1}{N_1} < x$. Let $N_2 \in \mathbb{N}$ be such that $\frac{1}{1-x} < N_2$ (Archimedean property), that is, $x < 1 - \frac{1}{N_2}$. Thus with $N := \max\{N_1, N_2\}$, we have

$$\frac{1}{N} \leq \frac{1}{N_1} < x < 1 - \frac{1}{N_2} \leq 1 - \frac{1}{N},$$

that is, $x \in \left[\frac{1}{N}, 1 - \frac{1}{N}\right] \subset \bigcup_{n \in \mathbb{N}} \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$. So we have

$$(0, 1) \subset \bigcup_{n \in \mathbb{N}} \left[\frac{1}{n}, 1 - \frac{1}{n}\right]. \quad (2.17)$$

From (2.16) and (2.17), we obtain $(0, 1) = \bigcup_{n \in \mathbb{N}} \left[\frac{1}{n}, 1 - \frac{1}{n}\right]$.

(d) If $x \in [0, 1]$, then for any $n \in \mathbb{N}$,

$$-\frac{1}{n} < 0 \leq x \leq 1 < 1 + \frac{1}{n},$$

and so $x \in \left(-\frac{1}{n}, 1 + \frac{1}{n}\right)$. Hence

$$[0, 1] \subset \bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right). \quad (2.18)$$

Let $x \in \bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right)$. Then

$$-\frac{1}{n} \leq x \leq 1 + \frac{1}{n} \text{ for all } n \in \mathbb{N}. \quad (2.19)$$

We prove that this implies that $0 \leq x \leq 1$. For if $x < 0$, then let $N_1 \in \mathbb{N}$ be such that $-\frac{1}{x} < N_1$, that is, $x < -\frac{1}{N_1}$, a contradiction to (2.19). Similarly, if $x > 1$, then let $N_2 \in \mathbb{N}$ be such that $\frac{1}{x-1} < N_2$, that is, $x > 1 + \frac{1}{N_2}$, a contradiction to (2.18). Hence we see that neither $x < 0$ nor $x > 1$ are possible, and hence $x \in [0, 1]$. Thus

$$\bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right) \subset [0, 1]. \quad (2.20)$$

(2.18) and (2.20) imply $\bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, 1 + \frac{1}{n}\right) = [0, 1]$.

9. If $x \in (x_0 - \delta, x_0 + \delta)$, then $x_0 - \delta < x < x_0 + \delta$. Adding $-x_0$, we obtain $-\delta < x - x_0 < \delta$. So $x - x_0 < \delta$, and $-(x - x_0) < \delta$. Thus $|x - x_0| < \delta$, and $x \in \{x \in \mathbb{R} \mid |x - x_0| < \delta\}$. Consequently $(x_0 - \delta, x_0 + \delta) \subset \{x \in \mathbb{R} \mid |x - x_0| < \delta\}$.

If $x \in \{x \in \mathbb{R} \mid |x - x_0| < \delta\}$, then $|x - x_0| < \delta$. Since for any real number r , $|r| \geq r$ and $|r| \geq -r$, we obtain that $x - x_0 \leq |x - x_0| < \delta$ and $-(x - x_0) \leq |x - x_0| < \delta$. Hence $-\delta < x - x_0 < \delta$. Adding x_0 , this yields $x_0 - \delta < x < x_0 + \delta$, that is, $x \in (x_0 - \delta, x_0 + \delta)$. So $\{x \in \mathbb{R} \mid |x - x_0| < \delta\} \subset (x_0 - \delta, x_0 + \delta)$.

Thus $(x_0 - \delta, x_0 + \delta) = \{x \in \mathbb{R} \mid |x - x_0| < \delta\}$.

10. From the inequality (1.7) on page 6, we have that $|x| = |x - y + y| \leq |x - y| + |y|$, that is,

$$|x| - |y| \leq |x - y| \tag{2.21}$$

for all $x, y \in \mathbb{R}$. Interchanging x and y in (2.21), we obtain

$$|y| - |x| \leq |y - x| = |-(x - y)| = |-1||x - y| = 1 \cdot |x - y| = |x - y|,$$

and so,

$$-(|x| - |y|) \leq |x - y| \tag{2.22}$$

for all $x, y \in \mathbb{R}$. From (2.21) and (2.22), we obtain $||x| - |y|| \leq |x - y|$ for all $x, y \in \mathbb{R}$.

11. If S is bounded, then it is bounded above and it is bounded below. Thus S has an upper bound, say u , and a lower bound, say l . So for all $x \in S$, $l \leq x \leq u$, that is, $x \leq u$ and $-x \leq -l$, and so we have

$$x \leq u \leq \max\{-l, u\} \text{ and } -x \leq -l \leq \max\{-l, u\}.$$

Thus $|x| \leq \max\{-l, u\} =: M$.

Conversely, if there exists a M such that for all $x \in S$, $|x| \leq M$, we have $-M \leq x \leq M$. So $-M$ is a lower bound of S and M is an upper bound of S . Thus S is bounded.

Solutions to the exercises on page 13

1. (a) We prove that $(1)_{n \in \mathbb{N}}$ is convergent sequence with limit 1. Given $\epsilon > 0$, let $N \in \mathbb{N}$, say $N = 1$. Then for all $n > N = 1$, we have

$$|a_n - L| = |1 - 1| = |0| < \epsilon.$$

- (b) Yes. For instance, the constant sequence $(1)_{n \in \mathbb{N}}$ converges to 1.
 (c) Suppose that the terms of the convergent sequence $(a_n)_{n \in \mathbb{N}}$ (with limit, say, L) lie in the finite set $\{v_1, \dots, v_m\}$. If $L \notin \{v_1, \dots, v_m\}$, then with

$$\epsilon := \min\{|v_1 - L|, \dots, |v_m - L|\} > 0,$$

let $N \in \mathbb{N}$ be such that for all $n > N$, $|a_n - L| < \epsilon$. In particular, with $n = N+1 > N$, we have $|a_{N+1} - L| < \epsilon$. But $a_{N+1} \in \{v_1, \dots, v_m\}$. Let $a_{N+1} = v_k$ for some $k \in \{1, \dots, m\}$. Then we have

$$|v_k - L| = |a_{N+1} - L| < \epsilon = \min\{|v_1 - L|, \dots, |v_m - L|\} \leq |v_k - L|,$$

a contradiction. So $L \in \{v_1, \dots, v_m\}$, that is, L must be one of the terms. Thus we have shown that

$$\boxed{\begin{array}{c} \text{terms of the sequence} \\ \text{take finitely many values} \end{array}} \implies \boxed{\begin{array}{c} L \text{ must be one} \\ \text{of the terms} \end{array}},$$

that is,

$$\boxed{\begin{array}{c} L \text{ is not equal to} \\ \text{any of the terms} \end{array}} \implies \boxed{\begin{array}{c} \text{terms of the sequence cannot} \\ \text{consist of finitely many values} \end{array}}.$$

- (d) Suppose that $((-1)^n)_{n \in \mathbb{N}}$ is a convergent sequence with limit L . Then from part 1c above, it follows that $L = 1$ or $L = -1$: indeed the terms of the sequence take finitely many values, namely 1 and -1 , and so L must be one of these terms. So we have the following two cases:

1^o If the limit is 1, then given $\epsilon = 1 > 0$, let $N \in \mathbb{N}$ be such that for all $n > N$, $|(-1)^n - 1| < \epsilon = 1$. Let n be any odd number $> N$. Then we have $2 = |-2| = |-1 - 1| = |(-1)^n - 1| < \epsilon = 1$, a contradiction.

2^o If the limit is -1 , then given $\epsilon = 1 > 0$, let $N \in \mathbb{N}$ be such that for all $n > N$, $|(-1)^n - (-1)| < \epsilon = 1$. Let n be any even number $> N$. Then we have $2 = |2| = |1 + 1| = |(-1)^n - (-1)| < \epsilon = 1$, a contradiction.

So $((-1)^n)_{n \in \mathbb{N}}$ is divergent.

2. If $\lim_{n \rightarrow \infty} \frac{1}{n} = 1$, then given any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ such that $n > N$,

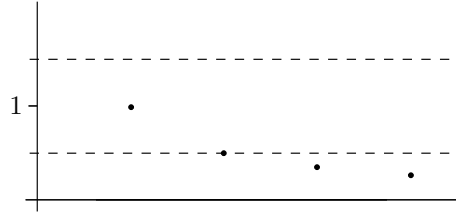
$$\left| \frac{1}{n} - 1 \right| < \epsilon.$$

Let $\epsilon = \frac{1}{2}$. (This choice is motivated by Figure 2.2, from which we see that, for instance $\epsilon = 1$ won't give us a contradiction, but $\epsilon = \frac{1}{2}$ would.) Then there exists a $N_* \in \mathbb{N}$ such that for all $n > N_*$,

$$\left| \frac{1}{n} - 1 \right| < \epsilon = \frac{1}{2},$$

that is,

$$1 - \frac{1}{n} = \left| \frac{1}{n} - 1 \right| < \frac{1}{2}.$$

Figure 2.2: $(\frac{1}{n})_{n \in \mathbb{N}}$.

Consequently, for all $n > N_*$,

$$1 - \frac{1}{n} < \frac{1}{2},$$

that is, $n < 2$. In particular, since $n = N_* + 1 > N_*$, we obtain $N_* + 1 < 2$, that is, $N_* < 1$. But there does not exist any natural number N_* that is less than 1. Hence we arrive at a contradiction, and so $\lim_{n \rightarrow \infty} \frac{1}{n} \neq 1$.

3. (a) We have seen that the sequence $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$ is divergent. Given any $\epsilon > 0$, let $N \in \mathbb{N}$ be such that $\frac{1}{\epsilon} < N$. Then for all even $n > N$ (there are obviously infinitely many such n), we have

$$|a_n - L| = \left| (-1)^n \left(1 + \frac{1}{n} \right) - 1 \right| = \left| 1 + \frac{1}{n} - 1 \right| = \frac{1}{n} < \frac{1}{N} < \epsilon.$$

- (b) Again, for the divergent sequence $((-1)^n (1 + \frac{1}{n}))_{n \in \mathbb{N}}$, with $\epsilon = 4 > 0$, for all $N \in \mathbb{N}$ and all $n > N$, we have

$$|a_n - L| \leq |a_n| + |L| \leq 2 + 1 = 3 < 4 = \epsilon.$$

4. S is nonempty and bounded above, and so by the least upper bound property of \mathbb{R} , it follows that $\sup S$ exists.

Given $n \in \mathbb{N}$, we have $\frac{1}{n} > 0$, and so $\sup S - \frac{1}{n} < \sup S$. Thus $S - \frac{1}{n}$ is not an upper bound of S . Hence there must exist an element in S , which we denote by a_n , such that $\neg[a_n \leq \sup S - \frac{1}{n}]$, that is, $a_n > \sup S - \frac{1}{n}$. In this way we construct the sequence $(a_n)_{n \in \mathbb{N}}$. As $\sup S$ is an upper bound of S and so we also have $a_n \leq \sup S$ for all $n \in \mathbb{N}$. Consequently,

$$\forall n \in \mathbb{N}, \quad \sup S - \frac{1}{n} < a_n \leq \sup S < \sup S + \frac{1}{n},$$

that is,

$$\forall n \in \mathbb{N}, \quad -\frac{1}{n} < a_n - \sup S < \frac{1}{n}, \quad \text{that is, } |a_n - \sup S| < \frac{1}{n}.$$

Given $\epsilon > 0$, let $N \in \mathbb{N}$ be such that $N > \frac{1}{\epsilon}$. Then for all $n > N$, we have

$$|a_n - \sup S| < \frac{1}{n} < \frac{1}{N} < \epsilon.$$

Hence $(a_n)_{n \in \mathbb{N}}$ is convergent with limit equal to $\sup S$.

5. Suppose $L < 0$. Then $\epsilon := -\frac{L}{2} > 0$, and so there exists a $N \in \mathbb{N}$ such that for all $n > N$, $|a_n - L| < \epsilon = -\frac{L}{2}$. Hence with $n = N + 1 (> N)$, we obtain

$$a_{N+1} - L \leq |a_{N+1} - L| < -\frac{L}{2},$$

that is, $a_{N+1} < \frac{L}{2} < 0$, a contradiction.

6. Let $(a_n)_{n \in \mathbb{N}}$ be a convergent sequence with limit L . Given $\epsilon > 0$, there exists a $N \in \mathbb{N}$ such that

$$\text{for all } n > N, \quad |a_n - L| < \frac{\epsilon}{2}. \quad (2.23)$$

Hence if $n, m > N$, then

$$\begin{aligned} |a_n - a_m| &= |a_n - L + L - a_m| \\ &\leq |a_n - L| + |L - a_m| \quad (\text{triangle inequality}) \\ &= |a_n - L| + |a_m - L| \quad (\text{using } |r| = |-r| \text{ for all real } r) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \quad (\text{using (2.23)}). \end{aligned}$$

Consequently $(a_n)_{n \in \mathbb{N}}$ is a Cauchy sequence.

Solutions to the exercises on page 17

1. We prove that the sequence is monotone and bounded, and hence it must be convergent.

We prove that $|a_n| \leq 1$ for all $n \in \mathbb{N}$. We prove this using induction. We have $|a_1| = |1| = 1$. If $k \in \mathbb{N}$ is such that $|a_k| \leq 1$, then

$$|a_{k+1}| = \left| \frac{2k+3}{3k+3} a_k \right| = \left| \frac{2k+3}{3k+3} \right| |a_k| = \left(\frac{2k+3}{3k+3} \right) |a_k| \leq 1 \cdot 1 = 1,$$

and so the claim follows from induction. So the sequence is bounded.

Since $n \geq 1$, it follows that $2n+1 \leq 3n$, and so $\frac{2n+1}{3n} \leq 1$ for all $n \in \mathbb{N}$. Furthermore, note that for all $n \in \mathbb{N}$, $a_n \geq 0$ (induction!). Hence for all $n \geq 2$, $a_n = \frac{2n+1}{3n} a_{n-1} \leq 1 \cdot a_{n-1} = a_{n-1}$. So $(a_n)_{n \in \mathbb{N}}$ is decreasing.

As the sequence is bounded and monotone, it is convergent.

2. Let $M > 0$ be such that for all $n \in \mathbb{N}$, $|b_n| \leq M$. Given $\epsilon > 0$, let $N \in \mathbb{N}$ be such that $\frac{M}{\epsilon} < N$, that is, $\frac{1}{N} < \frac{\epsilon}{M}$. Then for all $n > N$,

$$\left| \frac{b_n}{n} - 0 \right| = \frac{|b_n|}{n} \leq \frac{M}{n} < \frac{M}{N} < M \cdot \frac{\epsilon}{M} = \epsilon.$$

Hence $\left(\frac{b_n}{n}\right)_{n \in \mathbb{N}}$ is convergent with limit 0.

3. (a) Given $\epsilon > 0$, $\exists N_1 \in \mathbb{N}$ such that $\forall n > N_1$, $|a_n - L| < \frac{\epsilon}{2}$. Since $(a_n)_{n \in \mathbb{N}}$ is convergent, it is bounded: $\exists M > 0$ such that $\forall n \in \mathbb{N}$, $|a_n| \leq M$. Choose⁸ $N \in \mathbb{N}$ such that

$$\max \left\{ N_1, \frac{N_1(M + |L|)}{\frac{\epsilon}{2}} \right\} < N,$$

and so,

$$N > N_1 \text{ and } \frac{N_1(M + |L|)}{N} < \frac{\epsilon}{2}.$$

⁸This is arrived at by working backwards; we wish to make $\left| \frac{a_1 + \dots + a_n}{n} - L \right|$ less than ϵ for all $n > N$, so we manipulate this (as shown in the chain of inequalities that follow) to see if can indeed achieve this by choosing the N large enough.

Then for $n > N$, we have:

$$\begin{aligned}
& \left| \frac{a_1 + \cdots + a_{N_1} + a_{N_1+1} + \cdots + a_n}{n} - L \right| \\
&= \left| \frac{a_1 + \cdots + a_{N_1} + a_{N_1+1} + \cdots + a_n - nL}{n} \right| \\
&= \frac{|a_1 + \cdots + a_{N_1} + a_{N_1+1} + \cdots + a_n - nL|}{n} \\
&\leq \frac{|a_1 - L| + \cdots + |a_{N_1} - L| + |a_{N_1+1} - L| + \cdots + |a_n - L|}{n} \\
&\leq \frac{(|a_1| + |L| + \cdots + |a_{N_1}| + |L|) + \frac{\epsilon}{2} + \cdots + \frac{\epsilon}{2}}{n} \\
&\leq \frac{N_1(M + |L|) + (n - N_1)\frac{\epsilon}{2}}{n} \\
&\leq \frac{N_1(M + |L|)}{n} + \left(1 - \frac{N_1}{n}\right) \cdot \frac{\epsilon}{2} \\
&< \frac{N_1(M + |L|)}{N} + 1 \cdot \frac{\epsilon}{2} \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\
&= \epsilon.
\end{aligned}$$

So $\left(\frac{a_1 + \cdots + a_n}{n}\right)_{n \in \mathbb{N}}$ is a convergent sequence with limit L .

(b) If $a_n = (-1)^n$, $n \in \mathbb{N}$, then $(a_n)_{n \in \mathbb{N}}$ is divergent, but the sequence with n th term

$$\frac{a_1 + \cdots + a_n}{n} = \frac{(-1)^1 + (-1)^2 + \cdots + (-1)^n}{n} = \begin{cases} 0 & \text{if } n \text{ is even} \\ -\frac{1}{n} & \text{if } n \text{ is odd} \end{cases},$$

is convergent with limit equal to 0. Indeed, given $\epsilon > 0$, let $N \in \mathbb{N}$ be such that $\frac{1}{\epsilon} < N$. Then for $n > N$, we have

$$|a_n - 0| = |a_n| = \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{1}{n} & \text{if } n \text{ is odd} \end{cases} \leq \frac{1}{n} < \frac{1}{N} < \epsilon.$$

So $\left(\frac{a_1 + \cdots + a_n}{n}\right)_{n \in \mathbb{N}}$ is convergent with limit 0.

4. Since $(a_n)_{n \in \mathbb{N}}$ is bounded, it follows that there exists a M such that for all $n \in \mathbb{N}$, $|a_n| \leq M$, that is, $-M \leq a_n \leq M$. If $k \in \mathbb{N}$, then in particular, for all $n \geq k$, $-M \leq a_n \leq M$, and so the set $\{a_n \mid n \geq k\}$ is bounded. By the least upper bound property of \mathbb{R} , it then follows that $\inf\{a_n \mid n \geq k\}$ and $\sup\{a_n \mid n \geq k\}$ exist, that is, l_k and u_k are well-defined. Furthermore, for each k ,

$$-M \leq \inf\{a_n \mid n \geq k\} \leq \sup\{a_n \mid n \geq k\} \leq M,$$

and so we see that the sequences $(l_k)_{k \in \mathbb{N}}$ and $(u_k)_{k \in \mathbb{N}}$ are bounded.

Clearly $\{a_n \mid n \geq k+1\} \subset \{a_n \mid n \geq k\}$, and by Exercise 4 on page 7, we then obtain that

$$u_{k+1} = \sup\{a_n \mid n \geq k+1\} \leq \sup\{a_n \mid n \geq k\} = u_k,$$

and so $(u_k)_{k \in \mathbb{N}}$ is a decreasing sequence.

Similarly, $\{-a_n \mid n \geq k+1\} \subset \{-a_n \mid n \geq k\}$, and so we have

$$\sup\{-a_n \mid n \geq k+1\} \leq \sup\{-a_n \mid n \geq k\}.$$

Using Exercise 6 on page 7, we obtain

$$\inf\{a_n \mid n \geq k+1\} = -\sup\{-a_n \mid n \geq k+1\} \geq -\sup\{-a_n \mid n \geq k\} = \inf\{a_n \mid n \geq k\},$$

that is, $l_{k+1} \geq l_k$. Consequently, $(l_k)_{k \in \mathbb{N}}$ is an increasing sequence.

As the sequences $(u_k)_{k \in \mathbb{N}}$, $(l_k)_{k \in \mathbb{N}}$ are both bounded and monotone, by Theorem 1.2.3, it follows that they are convergent.

5. Let $(a_n)_{n \in \mathbb{N}}$ be a Cauchy sequence. Then given $\epsilon := 1 > 0$, there exists a $N \in \mathbb{N}$ such that for all $n, m > N$, $|a_n - a_m| < \epsilon = 1$. In particular, with $m := N + 1 (> N)$, $|a_n - a_{N+1}| < 1$ for all $n > N$, that is,

$$|a_n| = |a_{N+1} + a_n - a_{N+1}| \leq |a_{N+1}| + |a_n - a_{N+1}| < |a_{N+1}| + 1 \quad \text{for all } n > N.$$

Defining $M = \max\{|a_1|, \dots, |a_N|, |a_{N+1}| + 1\}$, we see that $|a_n| \leq M$ for all $n \in \mathbb{N}$, and so $(a_n)_{n \in \mathbb{N}}$ is bounded.

Solutions to the exercises on page 20

1. First we prove the

CLAIM: $(a_n)_{n \in \mathbb{N}}$ is convergent with limit L , then $(a_{n+1})_{n \in \mathbb{N}}$ is also convergent with limit L .

Proof Let $\epsilon > 0$. Then $\exists N \in \mathbb{N}$ such that for all $n > N$, $|a_n - L| < \epsilon$. Thus for all $n > N$, we have $n + 1 > N + 1 > N$, and so $|a_{n+1} - L| < \epsilon$. Hence $(a_{n+1})_{n \in \mathbb{N}}$ is convergent with limit L .

Alternately, observe that $(a_{n+1})_{n \in \mathbb{N}}$ is a subsequence of $(a_n)_{n \in \mathbb{N}}$, and use Theorem 1.2.6. ■

We now apply this result to our sequence $(a_n)_{n \in \mathbb{N}}$, which satisfies:

$$\begin{aligned} a_{n+1} &= \frac{2(n+1)+1}{3(n+1)} a_n \\ &= \frac{2 + \frac{3}{n}}{3 + \frac{3}{n}} a_n, \end{aligned}$$

for all $n \in \mathbb{N}$. Since the sequence $(\frac{1}{n})_{n \in \mathbb{N}}$ is convergent with limit 0, by the theorem on algebra of limits, it follows that the sequence $(\frac{2 + \frac{3}{n}}{3 + \frac{3}{n}})_{n \in \mathbb{N}}$ is convergent with limit $\frac{2+3 \cdot 0}{3+3 \cdot 0} = \frac{2}{3}$. Again applying the theorem on algebra of limits, we obtain

$$\begin{aligned} L &= \lim_{n \rightarrow \infty} a_{n+1} \\ &= \lim_{n \rightarrow \infty} \left(\frac{2 + \frac{3}{n}}{3 + \frac{3}{n}} a_n \right) \\ &= \lim_{n \rightarrow \infty} \left(\frac{2 + \frac{3}{n}}{3 + \frac{3}{n}} \right) \lim_{n \rightarrow \infty} a_n \\ &= \frac{2}{3} L. \end{aligned}$$

Hence $\frac{1}{3}L = 0$, that is, $L = 0$. So $\lim_{n \rightarrow \infty} a_n = 0$.

2. We have

$$c_n = \frac{a_n b_n + 5n}{a_n^2 + n} = \frac{a_n \cdot \frac{b_n}{n} + 5}{a_n \cdot a_n \cdot \frac{1}{n} + 1} \text{ for all } n \in \mathbb{N}.$$

N. The sequence $(a_n \cdot \frac{b_n}{n} + 5)_{n \in \mathbb{N}}$ is convergent.

The sequence $(a_n)_{n \in \mathbb{N}}$ is convergent with limit, say L . Since $(b_n)_{n \in \mathbb{N}}$ is bounded, the sequence $(\frac{b_n}{n})_{n \in \mathbb{N}}$ is convergent with limit 0 (see Exercise 2 on page 17). Hence $(a_n \cdot \frac{b_n}{n})_{n \in \mathbb{N}}$ is convergent with limit $L \cdot 0 = 0$. The sequence $(5)_{n \in \mathbb{N}}$ is convergent with limit 5. So the sequence $(a_n \cdot \frac{b_n}{n} + 5)_{n \in \mathbb{N}}$ is convergent with limit $0 + 5 = 5$.

D. The sequence $(\frac{a_n^2}{n} + 1)_{n \in \mathbb{N}}$ has nonzero terms for all $n \in \mathbb{N}$ and it is convergent with the nonzero limit 1.

We have $\frac{a_n^2}{n} + 1 \geq 1$ for all $n \in \mathbb{N}$, and so $\frac{a_n^2}{n} + 1 \neq 0$ for all $n \in \mathbb{N}$.

Since the sequence $(a_n)_{n \in \mathbb{N}}$ is convergent with limit L , it follows that the sequence $(a_n^2)_{n \in \mathbb{N}}$ is convergent with limit L^2 . Since $(\frac{1}{n})_{n \in \mathbb{N}}$ is convergent with limit 0, it follows that $(a_n^2 \cdot \frac{1}{n})_{n \in \mathbb{N}}$ is convergent with limit $L^2 \cdot 0 = 0$. Finally, as $(1)_{n \in \mathbb{N}}$ is convergent with limit 1, it follows that the sequence $(\frac{a_n^2}{n} + 1)_{n \in \mathbb{N}}$ is convergent with limit $0 + 1 = 1 (\neq 0)$.

From the parts N and D above and the theorem on the algebra of limits, it follows that $(c_n)_{n \in \mathbb{N}}$ is convergent with limit $\frac{5}{1} = 5$.

3. (a) We begin by showing that $L \geq 0$. If $L < 0$, then $\epsilon := -\frac{L}{2} > 0$. Let $N \in \mathbb{N}$ be such that for all $n > N$, $|a_n - L| < \epsilon = -\frac{L}{2}$. Then we have

$$a_n - L \leq |a_n - L| < -\frac{L}{2},$$

and so $a_n < \frac{L}{2} < 0$ for all $n > N$, a contradiction to the fact that

$$a_n \geq 0 \text{ for all } n \in \mathbb{N}.$$

So $L \geq 0$.

Now we show that $(\sqrt{a_n})_{n \in \mathbb{N}}$ is convergent with limit \sqrt{L} . Let $\epsilon > 0$. We consider the only two possible cases, namely $L = 0$ and $L > 0$:

1° If $L = 0$, then let $N \in \mathbb{N}$ be such that for all $n > N$,

$$|a_n - L| = |a_n - 0| = |a_n| = a_n < \epsilon^2.$$

Then for $n > N$, we have $\sqrt{a_n} < \epsilon$, that is,

$$|\sqrt{a_n} - \sqrt{L}| = |\sqrt{a_n} - \sqrt{0}| = |\sqrt{a_n}| = \sqrt{a_n} < \epsilon.$$

So $(\sqrt{a_n})_{n \in \mathbb{N}}$ is convergent with limit \sqrt{L} .

2° If $L > 0$, then let $N \in \mathbb{N}$ be such that for $n > N$, $|a_n - L| < \epsilon\sqrt{L}$. Then for all $n > N$, we obtain

$$\begin{aligned} \epsilon &> |a_n - L| \\ &= |(\sqrt{a_n} - \sqrt{L})(\sqrt{a_n} + \sqrt{L})| \\ &= |\sqrt{a_n} - \sqrt{L}| |\sqrt{a_n} + \sqrt{L}| \\ &= |\sqrt{a_n} - \sqrt{L}| (\sqrt{a_n} + \sqrt{L}) \end{aligned}$$

and so

$$|\sqrt{a_n} - \sqrt{L}| < \frac{\epsilon\sqrt{L}}{\sqrt{a_n} + \sqrt{L}} \leq \frac{\epsilon\sqrt{L}}{\sqrt{L}} = \epsilon.$$

Hence $(\sqrt{a_n})_{n \in \mathbb{N}}$ is convergent with limit \sqrt{L} .

- (b) For all $n \in \mathbb{N}$, we have

$$\begin{aligned} \sqrt{n^2 + n} - n &= (\sqrt{n^2 + n} - n) \cdot \frac{\sqrt{n^2 + n} + n}{\sqrt{n^2 + n} + n} \\ &= \frac{n^2 + n - n^2}{\sqrt{n^2 + n} + n} \\ &= \frac{n}{\sqrt{n^2 + n} + n} \\ &= \frac{n(1)}{n(\frac{1}{n}\sqrt{n^2 + n} + 1)} \\ &= \frac{1}{\sqrt{\frac{n^2 + n}{n^2}} + 1} \\ &= \frac{1}{\sqrt{1 + \frac{1}{n}} + 1}. \end{aligned}$$

As $(1)_{n \in \mathbb{N}}$ is convergent with limit 1 and $(\frac{1}{n})_{n \in \mathbb{N}}$ is convergent with limit 0, it follows that $(1 + \frac{1}{n})_{n \in \mathbb{N}}$ is convergent with limit 1. Furthermore $1 + \frac{1}{n} \geq 0$ for all $n \in \mathbb{N}$, and so by the previous part, it follows that $(\sqrt{1 + \frac{1}{n}})_{n \in \mathbb{N}}$ is convergent with limit $\sqrt{1} = 1$. Hence $(\sqrt{1 + \frac{1}{n}} + 1)_{n \in \mathbb{N}}$ is convergent with limit $1 + 1 = 2 (\neq 0)$. Also $\sqrt{1 + \frac{1}{n}} + 1 > 1 > 0$. Thus by the theorem on algebra of limits, we conclude that the sequence $(\frac{1}{\sqrt{1 + \frac{1}{n}} + 1})_{n \in \mathbb{N}}$ is convergent with limit $\frac{1}{2}$, that is, $(\sqrt{n^2 + n} - n)_{n \in \mathbb{N}}$ is convergent with limit $\frac{1}{2}$.

4. Consider the sequence $(b_n - a_n)_{n \in \mathbb{N}}$. As $a_n \leq b_n$, it follows that $b_n - a_n \geq 0$ for all $n \in \mathbb{N}$. From the theorem on algebra of limits, it follows that the sequence $(b_n - a_n)_{n \in \mathbb{N}}$ is convergent (being the sum of the convergent sequence $(b_n)_{n \in \mathbb{N}}$ and the convergent sequence $(-a_n)_{n \in \mathbb{N}}$). Moreover, its limit is $\lim_{n \rightarrow \infty} b_n - \lim_{n \rightarrow \infty} a_n$. From Exercise 5 on page 13, we obtain $\lim_{n \rightarrow \infty} b_n - \lim_{n \rightarrow \infty} a_n \geq 0$, that is $\lim_{n \rightarrow \infty} b_n \geq \lim_{n \rightarrow \infty} a_n$.

Solutions to the exercises on page 22

1. For all $n \in \mathbb{N}$, we have

$$0 \leq \frac{n!}{n^n} = \frac{1}{n} \cdot \frac{2}{n} \cdots \frac{n-1}{n} \cdot \frac{n}{n} \leq \frac{1}{n} \cdot 1 \cdots 1 \cdot 1 = \frac{1}{n}.$$

Since $(0)_{n \in \mathbb{N}}$ and $(\frac{1}{n})_{n \in \mathbb{N}}$ are both convergent with the same limit 0, from the Sandwich theorem, it follows that $(\frac{n!}{n^n})_{n \in \mathbb{N}}$ is convergent with the limit 0.

2. Let $k \in \mathbb{N}$. For all $n \in \mathbb{N}$, we have

$$0 \leq \frac{1^k + 2^k + 3^k + \cdots + n^k}{n^{k+2}} \leq \frac{n^k + n^k + n^k + \cdots + n^k}{n^{k+2}} \leq \frac{n \cdot n^k}{n^{k+2}} = \frac{1}{n}.$$

Thus

$$0 \leq \frac{1^k + 2^k + 3^k + \cdots + n^k}{n^{k+2}} \leq \frac{1}{n}$$

for all $n \in \mathbb{N}$. Since the sequences $(0)_{n \in \mathbb{N}}$ and $(\frac{1}{n})_{n \in \mathbb{N}}$ are both convergent with limit 0, from the Sandwich theorem, we obtain that

$$\lim_{n \rightarrow \infty} \frac{1^k + 2^k + 3^k + \cdots + n^k}{n^{k+2}} = 0.$$

3. (a) We prove the claim using induction. Let $x \geq -1$. Clearly $(1+x)^1 = 1+x = 1+1 \cdot x$. If for some $k \in \mathbb{N}$, $(1+x)^k \geq 1+kx$, then we have

$$\begin{aligned} (1+x)^{k+1} &= (1+x)^k(1+x) \\ &\geq (1+kx)(1+x) \text{ (by the induction hypothesis and since } 1+x \geq 0) \\ &= 1+kx+x+x^2 \\ &= 1+(k+1)x+x^2 \\ &\geq 1+(k+1)x. \end{aligned}$$

Hence by induction, the result follows.

(b) For all $n \in \mathbb{N}$,

$$n^{\frac{1}{n}} \geq 1 \quad (2.24)$$

(for if $n^{\frac{1}{n}} < 1$, then $n = (n^{\frac{1}{n}})^n < 1^n = 1$, a contradiction!). Clearly for all $n \in \mathbb{N}$,

$$n^{\frac{1}{n}} = (\sqrt{n^2})^{\frac{1}{n}} = \sqrt{n^{\frac{2}{n}}} < (1 + \sqrt{n})^{\frac{2}{n}}. \quad (2.25)$$

Finally,

$$\left(1 + \frac{1}{\sqrt{n}}\right)^n \geq 1 + n \cdot \frac{1}{\sqrt{n}} = 1 + \sqrt{n},$$

and so

$$\left(1 + \frac{1}{\sqrt{n}}\right)^2 = \left(\left(1 + \frac{1}{\sqrt{n}}\right)^n\right)^{\frac{2}{n}} \geq (1 + \sqrt{n})^{\frac{2}{n}}. \quad (2.26)$$

Combining (2.24), (2.25), and (2.26), we obtain

$$1 \leq n^{\frac{1}{n}} < (1 + \sqrt{n})^{\frac{2}{n}} \leq \left(1 + \frac{1}{\sqrt{n}}\right)^2 \quad (2.27)$$

for all $n \in \mathbb{N}$.

(c) Since $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$, using Exercise 3a on page 20, it follows that $\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} = 0$. Hence

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{\sqrt{n}}\right)^2 = (1 + 0)^2 = 1 = \lim_{n \rightarrow \infty} 1.$$

Using (2.27), it follows from the Sandwich theorem that $(n^{\frac{1}{n}})_{n \in \mathbb{N}}$ is convergent and $\lim_{n \rightarrow \infty} n^{\frac{1}{n}} = 1$.

4. Consider the sequence $(a_n - a)_{n \in \mathbb{N}}$. As $a_n \in (a, b)$ for all $n \in \mathbb{N}$, we have $a_n - a \geq 0$. From the theorem on algebra of limits, it follows that the sequence $(a_n - a)_{n \in \mathbb{N}}$ is convergent (being the sum of the convergent sequence $(a_n)_{n \in \mathbb{N}}$ and the convergent sequence $(-a)_{n \in \mathbb{N}}$). Moreover, its limit is $L - a$. From Exercise 5 on page 13, we obtain $L - a \geq 0$, that is $a \leq L$.

Next consider the sequence $(b - a_n)_{n \in \mathbb{N}}$. As $a_n \in (a, b)$ for all $n \in \mathbb{N}$, we have $b - a_n \geq 0$. From the theorem on algebra of limits, it follows that the sequence $(b - a_n)_{n \in \mathbb{N}}$ is convergent (being the sum of the convergent sequence $(-a_n)_{n \in \mathbb{N}}$ and the convergent sequence $(b)_{n \in \mathbb{N}}$). Moreover, its limit is $-L + b$. From Exercise 5 on page 13, we obtain $-L + b \geq 0$, that is $L \leq b$.

Consequently $a \leq L \leq b$, that is, $L \in [a, b]$.

Consider the sequence $(\frac{1}{n})_{n \in \mathbb{N}}$ contained in $(0, 1)$. It is convergent with limit $0 \notin (0, 1)$.

5. For all $n \in \mathbb{N}$, we have $-\frac{1}{n} < b_n - a_n < \frac{1}{n}$, and so by adding a_n , we have $-\frac{1}{n} + a_n < b_n < \frac{1}{n} + a_n$. By the theorem on algebra of limits, we know that

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(-\frac{1}{n} + a_n\right) &= \lim_{n \rightarrow \infty} -\frac{1}{n} + \lim_{n \rightarrow \infty} a_n = 0 + \lim_{n \rightarrow \infty} a_n \\ &= \lim_{n \rightarrow \infty} a_n \end{aligned}$$

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} + a_n\right) = \lim_{n \rightarrow \infty} \frac{1}{n} + \lim_{n \rightarrow \infty} a_n = 0 + \lim_{n \rightarrow \infty} a_n.$$

So by the sandwich theorem, it follows that $(b_n)_{n \in \mathbb{N}}$ is convergent with the limit $\lim_{n \rightarrow \infty} a_n$.

Solution to the exercise on page 24

Observe that

the terms 2, 8, 2, 8 appear adjacently
 and so the terms 1, 6, 1, 6 appear adjacently
 and so the terms 6, 6, 6 appear adjacently
 and so the terms 3, 6, 3, 6 appear adjacently
 and so the terms 1, 8, 1, 8 appear adjacently

and so the terms 8, 8, 8 appear adjacently
 and so the terms 6, 4, 6, 4 appear adjacently
 and so the terms 2, 4, 2, 4 appear adjacently
 and so the terms 8, 8, 8 appear adjacently

⋮

Hence we get the loop

$$\dots, 8, 8, 8, \dots \rightarrow \dots, 6, 4, 6, 4, \dots \rightarrow \dots, 2, 4, 2, 4, \dots \rightarrow \dots, 8, 8, 8, \dots \rightarrow,$$

which contains 6, and so 6 appears infinite number of times. Thus we can choose indices

$$n_1 < n_2 < n_3 < \dots$$

such that for all $k \in \mathbb{N}$, $a_{n_k} = 6$. So $(6)_{k \in \mathbb{N}}$ is a subsequence of the given sequence.

Solution to the exercise on page 26

Let $(a_n)_{n \in \mathbb{N}}$ be a Cauchy sequence. From Exercise 5 on page 17, it follows that $(a_n)_{n \in \mathbb{N}}$ is bounded. By the Bolzano-Weierstrass theorem, it follows that $(a_n)_{n \in \mathbb{N}}$ has a convergent subsequence, say $(a_{n_k})_{k \in \mathbb{N}}$ with limit L . Using the fact that $(a_n)_{n \in \mathbb{N}}$ is Cauchy, we now prove that $(a_n)_{n \in \mathbb{N}}$ is itself convergent with limit L .

Let $\epsilon > 0$. Since $(a_n)_{n \in \mathbb{N}}$ is Cauchy, there exists a $N \in \mathbb{N}$ such that for all $n, m > N$, $|a_n - a_m| < \frac{\epsilon}{2}$.

As $(a_{n_k})_{k \in \mathbb{N}}$ is convergent with limit L , there exists a $K \in \mathbb{N}$ such that for all $n_K > N$ and $|a_{n_K} - L| < \frac{\epsilon}{2}$. Then for all $n > N$, we have

$$\begin{aligned} |a_n - L| &= |a_n - a_{n_K} + a_{n_K} - L| \\ &\leq |a_n - a_{n_K}| + |a_{n_K} - L| \quad (\text{triangle inequality}) \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Consequently, $(a_n)_{n \in \mathbb{N}}$ is convergent with limit L .

Solutions to the exercises on page 29

1. (a) Let $\epsilon > 0$. If $\delta := \sqrt{\epsilon}$, then we note that $\delta > 0$. Moreover, if $x \in \mathbb{R}$ and $|x - 0| = |x| < \delta = \sqrt{\epsilon}$, then we obtain

$$|x^2 - 0^2| = |x^2| = |x| \cdot |x| < \delta \cdot \delta = \sqrt{\epsilon} \cdot \sqrt{\epsilon} = \epsilon.$$

So f is continuous at 0.

- (b) Let $c \in \mathbb{R}$ and suppose that $c \neq 0$. Let $\epsilon > 0$. If $x \in \mathbb{R}$, then

$$|x^2 - c^2| = |(x - c)(x + c)| = |x - c| \cdot |x + c|.$$

If $x \in \mathbb{R}$ is such that $|x - c| < \delta$, then

$$\begin{aligned} x < c + \delta &\leq |c + \delta| \leq |c| + |\delta| = |c| + \delta, \text{ and} \\ -x < \delta - c &\leq |\delta - c| \leq |\delta| + |-c| = \delta + |c|. \end{aligned}$$

Thus if $x \in \mathbb{R}$ satisfies $|x - c| < \delta$, then $|x| < \delta + |c|$, and so

$$|x + c| \leq |x| + |c| < \delta + |c| + |c| = \delta + 2|c|.$$

So if $|x - c| < \delta$, we have $|x^2 - c^2| = |x - c| \cdot |x + c| < \delta \cdot (\delta + 2|c|)$. Thus in order to make $|x^2 - c^2|$ less than ϵ , we choose δ such that $\delta(\delta + 2|c|) < \epsilon$: indeed, let

$$\delta := \min \left\{ \frac{\epsilon}{2|c| + 1}, 1 \right\}.$$

Since $\epsilon > 0$, it follows that δ is positive. Furthermore, if $x \in \mathbb{R}$ satisfies $|x - c| < \delta$, then we obtain

$$|x^2 - c^2| < \delta(\delta + 2|c|) \leq \frac{\epsilon}{2|c| + 1} (1 + 2|c|) = \epsilon.$$

2. (a) Let $c' \in \mathbb{R}$ and $\epsilon > 0$. Since f is continuous at c , there exists a $\delta > 0$ such that for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$. Then for all $x \in \mathbb{R}$ satisfying $|x - c'| < \delta$, we have⁹

$$|f(x) - f(c')| = |f(x) - f(c') + f(c) - f(c)| = |f(x - c' + c) - f(c)| < \epsilon,$$

since $|(x - c' + c) - c| = |x - c'| < \delta$. So f is continuous at c' . Since the choice of $c' \in \mathbb{R}$ was arbitrary, it follows that f is continuous on \mathbb{R} .

- (b) Let $\alpha \in \mathbb{R}$, and let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $f(x) = \alpha x$, for all $x \in \mathbb{R}$. Then

$$f(x + y) = \alpha(x + y) = \alpha x + \alpha y = f(x) + f(y).$$

3. We observe that as $|f(0)| \leq M|0| = M \cdot 0 = 0$, it follows that $|f(0)| = 0$, that is, $f(0) = 0$. Given $\epsilon > 0$, we define $\delta = \frac{\epsilon}{M}$. Then for all $x \in \mathbb{R}$ satisfying $|x| = |x - 0| < \delta$, we have

$$|f(x) - f(0)| = |f(x) - 0| = |f(x)| \leq M|x| = M|x - 0| < M\delta = M \frac{\epsilon}{M} = \epsilon.$$

Hence f is continuous at 0.

4. Let $c \in \mathbb{R}$. Suppose that f is continuous at c . Consider $\epsilon = \frac{1}{2} > 0$. Then $\exists \delta > 0$ such that for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon = \frac{1}{2}$. We have the following two cases:

⁹Here we use the fact that f is 'additive'. First of all, $f(0) = f(0 + 0) = f(0) + f(0)$, and so $f(0) = 0$. Hence it follows that $-f(c') = f(-c')$, since $0 = f(0) = f(c' - c') = f(c') + f(-c')$. Finally we have $f(x) - f(c') + f(c) = f(x) + f(-c') + f(c) = f(x - c') + f(c) = f(x - c' + c)$.

- 1° $c \in \mathbb{Q}$. Then there exists $x \in \mathbb{R} \setminus \mathbb{Q}$ such that $|x - c| < \delta$.
 But $|f(x) - f(c)| = |1 - 0| = |1| = 1 > \frac{1}{2}$, a contradiction.
- 2° $c \in \mathbb{R} \setminus \mathbb{Q}$. Then there exists $x \in \mathbb{Q}$ such that $|x - c| < \delta$.
 But $|f(x) - f(c)| = |0 - 1| = |-1| = 1 > \frac{1}{2}$, a contradiction.

Hence f is not continuous at c .

5. Since $\frac{f(c)}{2} > 0$, there exists a $\delta > 0$ such that for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta$,

$$|f(x) - f(c)| < \frac{f(c)}{2}.$$

Thus for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta$ (that is, $c - \delta < x < c + \delta$), we have

$$f(c) - f(x) \leq |f(c) - f(x)| = |f(x) - f(c)| < \frac{f(c)}{2},$$

and so $f(x) > \frac{f(c)}{2} > 0$.

Solutions to the exercises on page 31

- Let $(x_n)_{n \in \mathbb{N}}$ be a convergent sequence with limit c . Then $(x_n^2)_{n \in \mathbb{N}}$ is also convergent with limit c^2 . Thus $(f(x_n))_{n \in \mathbb{N}}$ is convergent with limit $f(c)$. So by Theorem 1.3.2, f is continuous.
- If $n \in \mathbb{N}$, then there exists $q_n \in \mathbb{Q}$ such that $|q_n - c| < \frac{1}{n}$.

CLAIM: $(q_n)_{n \in \mathbb{N}}$ is convergent with limit c .

Proof Let $\epsilon > 0$. By the Archimedean principle, there exists an $N \in \mathbb{N}$ such that $\frac{1}{\epsilon} < N$. Thus for $n > N$,

$$|q_n - c| < \frac{1}{n} < \frac{1}{N} < \epsilon.$$

This proves the claim. ■

Since f is continuous at c , by Theorem 1.3.2, we have $f(c) = \lim_{n \rightarrow \infty} f(q_n) = \lim_{n \rightarrow \infty} 0 = 0$.

- If $x_1 \neq x_2$, then the sequence $x_1, x_2, x_1, x_2, \dots$ is divergent. (Indeed, the subsequence x_1, x_1, x_1, \dots converges to x_1 , while the subsequence x_2, x_2, x_2, \dots converges to x_2 , and so by Theorem 1.2.6, it follows that the sequence $x_1, x_2, x_1, x_2, \dots$ is divergent.)

Thus the sequence $f(x_1), f(x_2), f(x_1), f(x_2), \dots$ is divergent. Consequently $f(x_1) \neq f(x_2)$: for otherwise if $f(x_1) = f(x_2)$, then the sequence

$$\begin{array}{ccccccc} f(x_1), & f(x_2), & f(x_1), & f(x_2) & \dots & & \\ \parallel & \parallel & \parallel & \parallel & & & \\ f(x_1), & f(x_1), & f(x_1), & f(x_1) & \dots & & \end{array}$$

is a constant sequence, and so it is convergent with limit $f(x_1)$ ($= f(x_2)$).

So we have shown that if $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$, that is, the function f is one-to-one.

Solution to the exercise on page 32

We apply Theorem 1.3.3 several times in order to prove this.

Since the function $x \mapsto x$ is continuous on \mathbb{R} , it follows that the function $x \mapsto x^2$ is continuous on \mathbb{R} as well. Moreover the function $x \mapsto 1$ is continuous on \mathbb{R} , and so we obtain that the function $x \mapsto 1 + x^2$ is continuous on \mathbb{R} . As $1 + x^2 \geq 1 > 0$ for all real x , we conclude that the function $x \mapsto \frac{1}{1+x^2}$ is continuous on \mathbb{R} . Hence the function $x \mapsto x^2 \cdot \frac{1}{1+x^2} = \frac{x^2}{1+x^2}$ is continuous on \mathbb{R} , that is, f is continuous on \mathbb{R} .

Solutions to the exercises on page 34

1. Let $m := \inf\{f(x) \mid x \in [a, b]\}$. We prove that there exists a $d \in [a, b]$ such that $f(d) = m$. For each $n \in \mathbb{N}$, $m < m + \frac{1}{n}$, and so $m + \frac{1}{n}$ cannot be a lower bound for $\{f(x) \mid x \in [a, b]\}$. So there exists an $x_n \in [a, b]$ such that

$$m \leq f(x_n) < m + \frac{1}{n}. \quad (2.28)$$

By the Bolzano-Weierstrass theorem, $(x_n)_{n \in \mathbb{N}}$ has a convergent subsequence $(x_{n_k})_{k \in \mathbb{N}}$ with limit $d \in [a, b]$. Since f is continuous at d , it follows that $(f(x_{n_k}))_{k \in \mathbb{N}}$ is convergent with limit $f(d)$. From (2.28), using the Sandwich theorem, we conclude that $f(d) = m$.

2. Let $f : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 0 & \text{for all } 0 \leq x < \frac{1}{2}, \\ 1 & \text{for all } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Then f is not continuous at $\frac{1}{2}$: indeed the sequence $\left(\frac{1}{2} - \frac{1}{n+1}\right)_{n \in \mathbb{N}}$ is contained in $[0, 1]$,

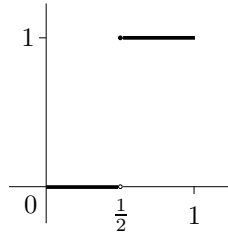


Figure 2.3: A discontinuous function on $[0, 1]$ that attains its extreme values.

and it is convergent with limit $\frac{1}{2} - 0 = \frac{1}{2}$. However,

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{2} - \frac{1}{n+1}\right) = \lim_{n \rightarrow \infty} 0 = 0 \neq 1 = f\left(\frac{1}{2}\right).$$

But $\{f(x) \mid x \in [0, 1]\} = \{0, 1\}$, and so

$$\begin{aligned} \sup\{f(x) \mid x \in [0, 1]\} &= \sup\{0, 1\} = 1 = f(1), \text{ and} \\ \inf\{f(x) \mid x \in [0, 1]\} &= \inf\{0, 1\} = 0 = f(0). \end{aligned}$$

3. Consider the function $g : [0, T] \rightarrow \mathbb{R}$, given by

$$g(x) = f(x) \text{ for all } x \in [0, T].$$

Then g is continuous on $[0, T]$.

(Indeed, the continuity of g on $[0, T]$ is a trivial consequence of the continuity of f on \mathbb{R} : If $c \in [0, T]$, and $\epsilon > 0$, then there exists a positive δ such that for all $x \in \mathbb{R}$ satisfying $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$. Hence for all $x \in [0, T]$ satisfying $|x - c| < \delta$, $|g(x) - g(c)| = |f(x) - f(c)| < \epsilon$. So g is continuous at c . But the choice of c was arbitrary, and so g is continuous on $[0, T]$.)

Applying the extreme value theorem to g , we conclude that there exist $c, d \in [0, T]$ such that

$$\begin{aligned} g(c) &= \max\{g(x) \mid x \in [0, T]\} \text{ and} \\ g(d) &= \min\{g(x) \mid x \in [0, T]\}. \end{aligned}$$

So for all $x \in [0, T]$, $g(d) \leq g(x) \leq g(c)$, that is, $f(d) \leq f(x) \leq f(c)$. So far we have proved

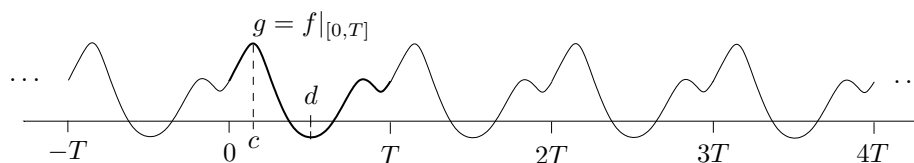


Figure 2.4: A continuous periodic function is bounded.

the fact that f is bounded on $[0, T]$. We now prove that f is bounded on \mathbb{R} by using the periodicity of f . See Figure 2.4.

Now if x is any real number, there exists a $n \in \mathbb{Z}$ such that $x = nT + r$, where $r \in \mathbb{R}$ is such that $r \in [0, T)$. (Indeed, we have

$$\frac{x}{T} = \left\lfloor \frac{x}{T} \right\rfloor + \Theta$$

where $\Theta \in [0, 1)$. Consequently, we obtain $x = nT + r$, where $n := \left\lfloor \frac{x}{T} \right\rfloor \in \mathbb{Z}$ and $r := T \cdot \Theta \in [0, T)$. Thus $f(x) = f(nT + r) = f(r)$. As $f(d) \leq f(r) \leq f(c)$, it follows that $f(d) \leq f(x) \leq f(c)$. Since the choice of $x \in \mathbb{R}$ was arbitrary, it follows that $f(d) \leq f(x) \leq f(c)$ for all $x \in \mathbb{R}$. So $f(c)$ and $f(d)$ are upper and lower bounds, respectively, of the set $\{f(x) \mid x \in \mathbb{R}\}$, and so it is bounded.

4. (a) If $x \in (a, b]$, then let $f|_{[x,b]}$ denote the restriction of f to $[x, b]$, defined by $f|_{[x,b]}(y) = f(y)$ for all $y \in [x, b]$. We note that $f|_{[x,b]}$ is a continuous function. Applying the intermediate value theorem to $f|_{[x,b]}$, we see that

$$\max\{f|_{[x,b]}(y) \mid y \in [a, x]\} = \max\{f(y) \mid y \in [a, x]\}$$

exists, and so f_* is well-defined.

- (b) First we observe that f_* is an increasing function, that is, if $x < y$ then, $f_*(x) \leq f_*(y)$. Furthermore, since $\sup(A \cup B) = \max\{\sup A, \sup B\}$ (why?), it follows that if $x < y$, then

$$f_*(y) \leq \max\{f_*(x), \max\{f(z) \mid z \in [x, y]\}\}.$$

Let $c \in [a, b]$, and let $\epsilon > 0$. Choose $\delta > 0$ such that for all $x \in [a, b]$ such that $|x - c| < \delta$, $|f(x) - f(c)| < \epsilon$. Let $x \in [a, b]$ be such that $|x - c| < \delta$. Then we have the following two cases:

1^o Let $c < x < c + \delta$. Then we have

$$\begin{aligned} f_*(x) &= \max\{f_*(c), \max\{f(z) \mid z \in [c, x]\}\} \\ &\leq \max\{f_*(c), f(c) + \epsilon\} \\ &\leq \max\{f_*(c), f_*(c) + \epsilon\} \\ &= f_*(c) + \epsilon, \end{aligned}$$

and so $|f_*(x) - f_*(c)| = f_*(x) - f_*(c) \leq \epsilon$.

2^o Let $c - \delta < x \leq c$. If $x = c$, then $|f_*(x) - f_*(c)| = 0 < \epsilon$ trivially. If $c - \delta < x < c$, then for any $z \in (c - \delta, c)$, we have

$$|f(z) - f(x)| = |f(z) - f(c) + f(c) - f(x)| \leq |f(z) - f(c)| + |f(c) - f(x)| = 2\epsilon,$$

and so $f(z) \leq f(x) + 2\epsilon$. Thus

$$\begin{aligned} f_*(c) &= \max\{f_*(x), \max\{f(z) \mid z \in [x, c]\}\} \\ &\leq \max\{f_*(x), f(x) + 2\epsilon\} \\ &\leq \max\{f_*(x), f_*(x) + 2\epsilon\} \\ &= f_*(x) + 2\epsilon, \end{aligned}$$

and so $|f_*(x) - f_*(c)| = f_*(c) - f_*(x) \leq 2\epsilon$.

Hence for all $x \in [a, b]$ satisfying $|x - c| < \delta$, we have $|f_*(x) - f_*(c)| < 2\epsilon$. Consequently f_* is continuous at c . Since the choice of c was arbitrary, it follows that f_* is continuous on $[a, b]$.

(c) f_* is given by

$$f_*(x) = \begin{cases} x - x^2 & \text{if } 0 \leq x \leq \frac{1}{2}, \\ \frac{1}{4} & \text{if } \frac{1}{2} < x \leq 1. \end{cases}$$

See Figure 2.5.

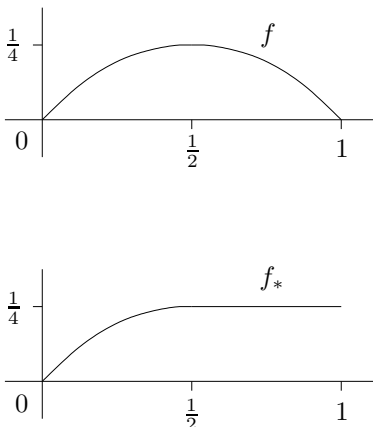


Figure 2.5: f and f_* .

Solutions to the exercises on page 37

1. Since the function $f : [0, 1] \rightarrow \mathbb{R}$ and the function $x \mapsto x$ from $[0, 1]$ to \mathbb{R} are both continuous on the interval $[0, 1]$, it follows that also the function $g : [0, 1] \rightarrow \mathbb{R}$ defined by

$$g(x) = f(x) - x, \text{ for all } x \in [0, 1]$$

is continuous on $[0, 1]$. Since $0 \leq f(x) \leq 1$ for all $x \in [0, 1]$, we have

$$g(0) = f(0) - 0 = f(0) \geq 0, \text{ and}$$

$$g(1) = f(1) - 1 \leq 0.$$

So by the intermediate value theorem, there exists a $c \in [0, 1]$ such that $g(c) = 0$, that is, $f(c) = c$.

2. Let the weekend campsite be at altitude H . Let $u : [0, 1] \rightarrow \mathbb{R}$ be the position function for the walk up, and $d : [0, \frac{1}{2}] \rightarrow \mathbb{R}$ be the position function for the walk down. (We assume that these are continuous functions.) Consider the function $f : [0, \frac{1}{2}] \rightarrow \mathbb{R}$ given by

$$f(t) = u(t) - d(t), \quad t \in \left[0, \frac{1}{2}\right].$$

Then f is also continuous, and moreover

$$f(0) = u(0) - d(0) = 0 - H = -H < 0, \text{ while}$$

$$f\left(\frac{1}{2}\right) = u\left(\frac{1}{2}\right) - d\left(\frac{1}{2}\right) = u\left(\frac{1}{2}\right) - 0 = u\left(\frac{1}{2}\right) > 0.$$

See Figure 2.6. Hence by the intermediate value theorem, it follows that there exists a $c \in [0, \frac{1}{2}]$ such that $f(c) = 0$, that is, $u(c) = d(c)$. So at time c past 8:00, the hiker was exactly at the same spot on Saturday and Sunday.

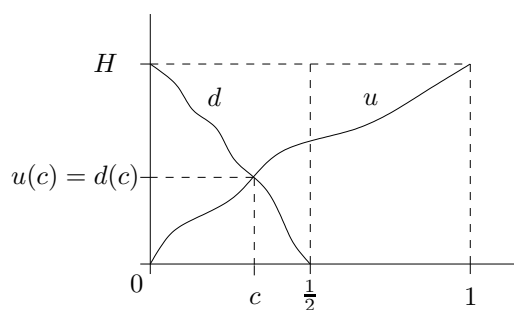


Figure 2.6: The walks up and down.

3. The polynomial function $p : [-1, 2] \rightarrow \mathbb{R}$ is continuous on the interval $[-1, 2]$. Moreover,

$$p(-1) = 2 \cdot (-1)^3 - 5 \cdot (-1)^2 - 10 \cdot (-1) + 5 = -2 - 5 + 10 + 5 = 8 > 0, \text{ and}$$

$$p(2) = 2 \cdot (2)^3 - 5 \cdot (2)^2 - 10 \cdot (2) + 5 = 16 - 20 - 20 + 5 = -19 < 0.$$

Since $p(-1) > 0 > p(2)$, from the intermediate value theorem applied to the continuous function p on the interval $[-1, 2]$, we conclude that there must exist a $c \in [-1, 2]$ such that $p(c) = 0$. So p has a real root in the interval $[-1, 2]$.

4. Obviously $S \subset \mathbb{R}$. We now show the reverse inclusion. Let $y \in \mathbb{R}$.

As S is not bounded above, y is not an upper bound of S , that is, there exists a $x_0 \in \mathbb{R}$ such that $f(x_0) < y$.

Similarly, since S is not bounded below, y is not a lower bound of S , and so there exists a $x_1 \in \mathbb{R}$ such that $f(x_1) > y$.

Now consider the restriction of f to the interval with endpoints x_0 and x_1 with the endpoints included in the interval. Applying the intermediate value theorem to this continuous function, it follows that there exists a real number c such that $f(c) = y$.

This shows that $S = \mathbb{R}$.

5. (a) The following three cases are possible:

- $\underline{1}^\circ$ $f(0) = 0$. Then let $x_0 = 0$ and let $m \in \mathbb{Z}$. Clearly $f(x_0) = f(0) = 0 = m0 = mx_0$.
- $\underline{2}^\circ$ $f(0) > 0$. Choose $N \in \mathbb{N}$ satisfying $N > f(1)$ (that such a N exists follows from the Archimedean property). Consider the function $g : [0, 1] \rightarrow \mathbb{R}$ defined by $g(x) = f(x) - Nx$, $x \in [0, 1]$. As the functions f and $x \mapsto Nx$ are continuous, so is g . Note that $g(0) = f(0) - N0 = f(0) > 0$, while $g(1) = f(1) - N < 0$. Applying the intermediate value theorem to g (with $y = 0$), it follows that there exists a $x_0 \in [0, 1]$ such that $g(x_0) = 0$, that is, $f(x_0) = Nx_0$.
- $\underline{3}^\circ$ $f(0) < 0$. Choose a $N \in \mathbb{N}$ such that $N > -f(1)$ (again the Archimedean property guarantees the existence of such a N), and consider the continuous function $g : [0, 1] \rightarrow \mathbb{R}$ defined by $g(x) = f(x) + Nx$. We observe that $g(0) = f(0) < 0$, and $g(1) = f(1) + N > 0$, and so by the intermediate value theorem, it follows that there exists a $x_0 \in [0, 1]$ such that $g(x_0) = 0$, that is, $f(x_0) = -Nx_0$.

This completes the proof.

- (b) Suppose that such a continuous function exists. From the part above, it follows that there exists a $x_0 \in \mathbb{R}$ and a $m \in \mathbb{Z}$ such that $f(x_0) = mx_0$. We have the following two possible cases:

- $\underline{1}^\circ$ $x_0 \in \mathbb{Q}$. But then $f(x_0)$ is irrational, while mx_0 is rational, a contradiction.
- $\underline{2}^\circ$ $x_0 \notin \mathbb{Q}$. But then $f(x_0)$ is rational, whereas mx_0 is irrational, a contradiction.

So f cannot be continuous.

Algebra

Solutions to the exercises on page 44

1. (a) Addition modulo 6 is a law of composition on \mathbb{Z}_6 , since if $[a], [b] \in \mathbb{Z}_6$, then $[a] \oplus [b] = [a + b] \in \mathbb{Z}_6$. Moreover, we verify below that the group axioms are satisfied.

G1. Addition modulo 6 is associative: if $[a], [b], [c] \in \mathbb{Z}_6$, then

$$\begin{aligned} ([a] \oplus [b]) \oplus [c] &= [a + b] \oplus [c] \\ &= [(a + b) + c] \\ &= [a + (b + c)] \\ &= [a] \oplus [b + c] \\ &= [a] \oplus ([b] \oplus [c]). \end{aligned}$$

G2. $[0]$ is the identity element: for all $[a] \in \mathbb{Z}_6$, we have

$$[0] \oplus [a] = [0 + a] = [a] = [a + 0] = [a] \oplus [0].$$

G3. If $[a] \in \mathbb{Z}_6$, then $[-a] \in \mathbb{Z}_6$, and

$$[a] \oplus [-a] = [a + (-a)] = [0] = [-a + a] = [-a] \oplus [a].$$

So \mathbb{Z}_6 with addition modulo 6 forms a group.

We give its group table below:

\oplus	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[1]	[2]	[3]	[4]	[5]
[1]	[1]	[2]	[3]	[4]	[5]	[0]
[2]	[2]	[3]	[4]	[5]	[0]	[1]
[3]	[3]	[4]	[5]	[0]	[1]	[2]
[4]	[4]	[5]	[0]	[1]	[2]	[3]
[5]	[5]	[0]	[1]	[2]	[3]	[4]

- (b) No; multiplication modulo 6 is a law of composition on \mathbb{Z}_6 that satisfies G1 and G2, but G3 is not true:

G1. Multiplication modulo 6 is associative: if $[a], [b], [c] \in \mathbb{Z}_6$, then

$$\begin{aligned} ([a] \otimes [b]) \otimes [c] &= [ab] \otimes [c] \\ &= [(ab)c] \\ &= [a(bc)] \\ &= [a] \otimes [bc] \\ &= [a] \otimes ([b] \otimes [c]). \end{aligned}$$

G2. (If $e = [k]$ is the identity element for some integer k , then for all $[a] \in \mathbb{Z}_6$, $[a] \otimes e = [a] = e \otimes [a]$. In particular, with $[a] = [1] \in \mathbb{Z}_6$, we obtain $e = [k] = [1k] = [1] \otimes [k] = [1] \otimes e = [1]$. So if there is an identity element e , then it must be $[1]$.) Indeed $[1]$ serves as an identity element, since for all $[a] \in \mathbb{Z}_6$,

$$[a] \otimes [1] = [a \cdot 1] = [a] = [1 \cdot a] = [a] \otimes [1].$$

G3 is not satisfied. For instance, $[0] \in \mathbb{Z}_6$, but for all $[a] \in \mathbb{Z}_6$, $[0] \otimes [a] = [a] \otimes [0] = [0] \neq [1]$. So there does not exist an inverse of the element $[0]$ in \mathbb{Z}_6 .

Hence \mathbb{Z}_6 with multiplication modulo 6 is not a group.

\mathbb{Z}_6^* with multiplication modulo 6 is also not a group, since multiplication modulo 6 is not a law of composition on the set \mathbb{Z}_6^* : indeed, $[2], [3] \in \mathbb{Z}_6^*$, but $[2] \otimes [3] = [2 \cdot 3] = [6] = [0] \notin \mathbb{Z}_6^*$.

(c) IF: Let m be a prime number.

First we show that multiplication modulo m is a law of composition on \mathbb{Z}_m^* . Let $[a], [b] \in \mathbb{Z}_m^*$. Then a and b are not divisible by m , and since m is prime, it follows that ab is also not divisible by m . Thus $[a] \otimes [b] = [ab] \neq [0]$, and so $[a] \otimes [b] \in \mathbb{Z}_m^*$.

Next we show that the group axioms are satisfied:

G1. If $[a], [b], [c] \in \mathbb{Z}_m^*$, then

$$\begin{aligned} ([a] \otimes [b]) \otimes [c] &= [ab] \otimes [c] \\ &= [(ab)c] \\ &= [a(bc)] \\ &= [a] \otimes [bc] \\ &= [a] \otimes ([b] \otimes [c]). \end{aligned}$$

G2. $[1] \neq [0]$ and so $[1] \in \mathbb{Z}_m^*$. Furthermore, For all $[a] \in \mathbb{Z}_m^*$, we have

$$[a] \otimes [1] = [a \cdot 1] = [a] = [1 \cdot a] = [1] \otimes [a].$$

G3. If $[a] \in \mathbb{Z}_m^*$, then m does not divide a (for otherwise $[a] = [0]$). As m is prime, it follows that $\gcd(a, m) = 1$. (Indeed, $\gcd(a, m)$ divides m , and since m is prime, the only factors of m are $1, m, -1, -m$. As $\gcd(a, m) \geq 0$, the only possible values are m and 1 . Also, $\gcd(a, m)$ divides a , and since a is not divisible by m , it follows that $\gcd(a, m) \neq m$. So the only possible value of $\gcd(a, m)$ is 1 .) So there exist $s, t \in \mathbb{Z}$ such that

$$as + mt = 1.$$

Consequently

$$\begin{aligned} [s] \otimes [a] = [sa] = [as] &= [a] \otimes [s] \\ &= [as] \\ &= [1 - mt] \\ &= [1] \oplus [m(-t)] \\ &= [1] \oplus [0] \\ &= [1]. \end{aligned}$$

Furthermore $[s] \neq [0]$, since otherwise $m|s$, and so $m|as + mt$, that is $m|1$, a contradiction. Hence $[s] \in \mathbb{Z}_m^*$, and so $[s]$ is the inverse of $[a]$.

Hence \mathbb{Z}_m^* with multiplication modulo m is a group.

ONLY IF: Suppose that $m = p \cdot q$, where $p, q \in \{2, 3, \dots, m-1\}$. Clearly, as m does not divide p or q , it follows that $[p] \neq [0]$ and $[q] \neq [0]$. So $[p], [q] \in \mathbb{Z}_m^*$. However,

$$[p] \otimes [q] = [p \cdot q] = [m] = [0] \notin \mathbb{Z}_m^*,$$

and so multiplication modulo m is not a law of composition on \mathbb{Z}_m^* . Consequently, \mathbb{Z}_m^* with multiplication modulo m is not a group.

2. (a) Composition of functions gives a law of composition on the set S_n of all bijections from $\{1, 2, 3, \dots, n\}$ onto itself. Indeed if f, g are bijections from $\{1, 2, 3, \dots, n\}$ onto itself, then $f \circ g$ is again a bijection from $\{1, 2, 3, \dots, n\}$ onto itself:

- i. $f \circ g$ is one-to-one. Let $k_1, k_2 \in \{1, 2, 3, \dots, n\}$. If $(f \circ g)(k_1) = (f \circ g)(k_2)$, then $f(g(k_1)) = f(g(k_2))$, and since f is one-to-one, we get $g(k_1) = g(k_2)$. Furthermore, since g is one-to-one, $k_1 = k_2$.
- ii. $f \circ g$ is onto. If $k \in \{1, 2, 3, \dots, n\}$, then since f is onto, there exists a $k' \in \{1, 2, 3, \dots, n\}$ such that $f(k') = k$. Moreover, since g is onto, there exists a $k'' \in \{1, 2, 3, \dots, n\}$ such that $g(k'') = k'$. Thus $(f \circ g)(k'') = f(g(k'')) = f(k') = k$. So $f \circ g$ is onto.

Next we show that the group axioms are satisfied:

G1. If $f, g, h \in S_n$, then for all $k \in \{1, 2, 3, \dots, n\}$,

$$((f \circ g) \circ h)(k) = (f \circ g)(h(k)) = f(g(h(k))) = f((g \circ h)(k)) = (f \circ (g \circ h))(k).$$

So $(f \circ g) \circ h = f \circ (g \circ h)$.

G2. The identity function $\iota : \{1, 2, 3, \dots, n\} \rightarrow \{1, 2, 3, \dots, n\}$ defined by $\iota(k) = k$ for all $k \in \{1, 2, 3, \dots, n\}$, serves as the identity element in S_n . Indeed, for all $f \in S_n$, we have

$$(f \circ \iota)(k) = f(\iota(k)) = f(k) = \iota(f(k)) = (\iota \circ f)(k), \quad \forall k \in \{1, 2, 3, \dots, n\}.$$

Thus $f \circ \iota = f = \iota \circ f$.

G3. If $f \in S_n$, then since f is a bijection, it follows that for all $k \in \{1, 2, 3, \dots, n\}$, there exists a unique element k' such that $f(k') = k$. Define $f^{-1} : \{1, 2, 3, \dots, n\} \rightarrow \{1, 2, 3, \dots, n\}$ as follows: if $k \in \{1, 2, 3, \dots, n\}$, then $f^{-1}(k) = k'$, where $k' \in \{1, 2, 3, \dots, n\}$ is such that $f(k') = k$. Then for all $k \in \{1, 2, 3, \dots, n\}$, we have

$$(f \circ f^{-1})(k) = f(f^{-1}(k)) = k = \iota(k) = k = f^{-1}(f(k)) = (f^{-1} \circ f)(k),$$

and so $f \circ f^{-1} = \iota = f^{-1} \circ f$.

Hence S_n with the composition of functions is a group.

- (b) The number of bijections from a set with n elements onto itself is $n!$, and so the order of S_n is $n!$.

Indeed, in order to specify the bijection f , one needs to specify $f(1), \dots, f(n)$. The number of ways of specifying $f(1)$ is n (as it can be any one of the numbers $1, \dots, n$). As f is one-to-one, $f(2)$ can be only be any one of the elements of the set $\{1, 2, 3, \dots, n\} \setminus \{f(1)\}$ (which has $n - 1$ elements). Proceeding in this way, the number of distinct bijections we get are $n \cdot (n - 1) \cdot \dots \cdot 2 \cdot 1 = n!$.

- (c) Let $f : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ be the function

$$\begin{aligned} f(1) &= 2 \\ f(2) &= 1 \\ f(3) &= 3, \end{aligned}$$

and $g : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ be the function

$$\begin{aligned} g(1) &= 1 \\ g(2) &= 3 \\ g(3) &= 2. \end{aligned}$$

Then

$$(f \circ g)(1) = f(g(1)) = f(1) = 2 \neq 3 = g(2) = g(f(1)) = (g \circ f)(1),$$

and so $f \circ g \neq g \circ f$.

- (d) IF: S_1 has the only element $\iota : \{1\} \rightarrow \{1\}$, and so $\iota \circ \iota = \iota = \iota \circ \iota$. Thus S_1 is abelian. S_2 has the two elements $\iota : \{1, 2\} \rightarrow \{1, 2\}$, and $r : \{1, 2\} \rightarrow \{1, 2\}$ given by

$$\begin{aligned} r(1) &= 2 \\ r(2) &= 1. \end{aligned}$$

Since $\iota \circ r = r \circ \iota$ we see that S_2 is abelian.

So if $n \leq 2$, then S_n is abelian.

ONLY IF: If $n > 2$, then consider the bijections $f : \{1, 2, 3, \dots, n\} \rightarrow \{1, 2, 3, \dots, n\}$ and $g : \{1, 2, 3, \dots, n\} \rightarrow \{1, 2, 3, \dots, n\}$ given by

$$\boxed{\begin{array}{l} f(1) = 2 \\ f(2) = 1 \\ f(n) = n \quad \forall n \in \{3, \dots, n\} \end{array}} \quad \text{and} \quad \boxed{\begin{array}{l} g(1) = 1 \\ g(2) = 3 \\ g(3) = 2 \\ g(n) = n \quad \forall n \in \{1, 2, 3, \dots, n\} \setminus \{1, 2, 3\} \end{array}}.$$

Then

$$(f \circ g)(1) = f(g(1)) = f(1) = 2 \neq 3 = g(2) = g(f(1)) = (g \circ f)(1),$$

and so $f \circ g \neq g \circ f$. Thus S_n is not abelian.

Hence S_n is abelian iff $n \geq 2$.

3. We note that if

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}, \begin{bmatrix} a' & b' \\ -b' & a' \end{bmatrix} \in S,$$

then

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} a' & b' \\ -b' & a' \end{bmatrix} = \begin{bmatrix} aa' - bb' & ab' + ba' \\ -(ab' + ba') & aa' - bb' \end{bmatrix},$$

and

$$\begin{aligned} (aa' - bb')^2 + (ab' + ba')^2 &= a^2a'^2 - 2aa'bb' + b^2b'^2 + a^2b'^2 - 2ab'ba' + b^2a'^2 \\ &= a^2(a'^2 + b'^2) + b^2(b'^2 + a'^2) \\ &= (a^2 + b^2)(a'^2 + b'^2) \\ &\neq 0. \end{aligned}$$

Hence matrix multiplication is a law of composition on S . Next we verify that the group axioms are satisfied:

G1. Matrix multiplication is associative, and so for all $A, B, C \in S$, $(AB)C = A(BC)$.

G2. Let I denote the identity matrix. Then we have

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -0 & 1 \end{bmatrix}$$

and moreover, $1^2 + 0^2 = 1 \neq 0$. So I belongs to S . Furthermore, for every $A \in S$, $AI = A = IA$.

G3. Let

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in S.$$

Then $a^2 + b^2 \neq 0$. If

$$B := \begin{bmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ \frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{bmatrix},$$

then

$$\left(\frac{a}{a^2+b^2}\right)^2 + \left(\frac{b}{a^2+b^2}\right)^2 = \frac{a^2+b^2}{(a^2+b^2)^2} = \frac{1}{a^2+b^2} \neq 0.$$

So $B \in S$, and furthermore, $AB = I = BA$.

So S is a group with matrix multiplication.

4. (a) Let $a, b, c \in G$, and suppose that $a * b = a * c$. Since $a \in G$, there exists $a^{-1} \in G$ such that $a * a^{-1} = e = a^{-1} * a$, where e denotes the identity element in G . Then we have

$$\begin{aligned} b &= e * b \quad (\text{since } e \text{ is the identity element}) \\ &= (a^{-1} * a) * b \quad (\text{since } a^{-1} \text{ is the inverse of } a) \\ &= a^{-1} * (a * b) \quad (\text{associativity}) \\ &= a^{-1} * (a * c) \quad (\text{since } a * b = a * c) \\ &= (a^{-1} * a) * c \quad (\text{associativity}) \\ &= e * c \quad (\text{since } a^{-1} \text{ is the inverse of } a) \\ &= c \quad (\text{since } e \text{ is the identity element}). \end{aligned}$$

- (b) Clearly

$$a * (a^{-1} * b) = (a * a^{-1}) * b = e * b = b,$$

and so $x = a^{-1} * b$ is a solution to $a * x = b$.

If there are two solutions, say x_1 and x_2 , then

$$a * x_1 = b = a * x_2,$$

and by part 4a above, it follows that $x_1 = x_2$. Hence $a * x = b$ has the unique solution $x = a^{-1} * b$.

- (c) Since $a, b \in G$, there exist elements $a^{-1}, b^{-1} \in G$ such that

$$a * a^{-1} = e = a^{-1} * a \text{ and } b * b^{-1} = e = b^{-1} * b.$$

We have

$$\begin{aligned} (a * b) * (b^{-1} * a^{-1}) &= a * (b * (b^{-1} * a^{-1})) \quad (\text{associativity}) \\ &= a * ((b * b^{-1}) * a^{-1}) \quad (\text{associativity}) \\ &= a * (e * a^{-1}) \quad (\text{since } b^{-1} \text{ is the inverse of } b) \\ &= a * a^{-1} \quad (\text{since } e \text{ is the identity element}) \\ &= e \quad (\text{since } a^{-1} \text{ is the inverse of } a), \end{aligned}$$

and

$$\begin{aligned} (b^{-1} * a^{-1}) * (a * b) &= ((b^{-1} * a^{-1}) * a) * b \quad (\text{associativity}) \\ &= (b^{-1} * (a^{-1} * a)) * b \quad (\text{associativity}) \\ &= (b^{-1} * e) * b \quad (\text{since } a^{-1} \text{ is the inverse of } a) \\ &= b^{-1} * b \quad (\text{since } e \text{ is the identity element}) \\ &= e \quad (\text{since } b^{-1} \text{ is the inverse of } b). \end{aligned}$$

Thus $b^{-1} * a^{-1}$ is an inverse of $a * b$. Since the inverse of $a * b$ is unique, it follows that $(a * b)^{-1} = b^{-1} * a^{-1}$.

Solutions to the exercises on page 46

1. (a) FALSE. Although H1 and H2 hold, H3 does not hold. Indeed, 1 is a nonnegative integer, but its inverse -1 is not a nonnegative integer.
 - (b) FALSE. H1 and H2 are not true. Indeed, the sum of the odd integers 1 and -1 is 0, which is not an odd integer. So H1 and H2 do not hold. But H3 is true, since if m is an odd integer, then so is $-m$.
 - (c) TRUE. For all $a, b \in H$, we have that a, b also belong to G , and as G is abelian, we know that $a * b = b * a$.
2. Yes. For instance, given any infinite group G , the set $\{e\}$ comprising the identity element e is a subgroup of the group G . Thus $\{0\}$ is a finite subgroup of the infinite group \mathbb{Z} of integers with addition. The subgroup $\{-1, 1\}$ of the set of nonzero real numbers with multiplication is another example.
 3. (a) We note that

$$H = \{4m \mid m \in \mathbb{Z}\} = \{\dots, -12, -8, -4, 0, 4, 8, 12, \dots\}, \text{ and}$$

$$K = \{6m \mid m \in \mathbb{Z}\} = \{\dots, -18, -12, -6, 0, 6, 12, 18, \dots\}.$$

CLAIM: $H \cap K = \{12m \mid m \in \mathbb{Z}\} = \{\dots, -12, 0, 12, \dots\}$.

Proof If $k \in H \cap K$, then $k \in H$ and $k \in K$. Hence $4 \mid k$ and $6 \mid k$. So there exist integers k_1 and k_2 such that $k = 4k_1$ and $k = 6k_2$. Thus $4k_1 = 6k_2$ and so $2k_1 = 3k_2$. Since integers possess a unique factorization into primes, it follows that $2 \mid k_2$. Consequently, $k_2 = 2k'_2$, and so $k = 6k_2 = 6(2k'_2) = 12k'_2$. Hence $k \in \{12m \mid m \in \mathbb{Z}\}$. So we have shown that $H \cap K \subset \{12m \mid m \in \mathbb{Z}\}$.

Conversely, if $k \in \{12m \mid m \in \mathbb{Z}\}$, then $k = 12m$ for some $m \in \mathbb{Z}$. Thus $k = 4(3m) = 6(2m)$, and so k is a multiple of 4 and 6. Consequently k belongs to H as well as K . Hence $\{12m \mid m \in \mathbb{Z}\} \subset H \cap K$. ■

- (b) We check that H1, H2, H3 hold:
 - H1. If $a, b \in H \cap K$, then a, b belong to both H and K . As H is a subgroup and $a, b \in H$, it follows that $a * b \in H$. Similarly, as K is a subgroup and $a, b \in K$, it follows that $a * b \in K$ as well. Hence $a * b \in H \cap K$.
 - H2. As H is a subgroup, $e \in H$. Also, as K is a subgroup, it follows that $e \in K$ as well. Thus $e \in H \cap K$.
 - H3. If $a \in H \cap K$, then $a \in H$ and $a \in K$. Since H is a subgroup and $a \in H$, it follows that $a^{-1} \in H$. Also, since K is a subgroup and $a \in K$, it follows that $a^{-1} \in K$. Consequently $a^{-1} \in H \cap K$.

So $H \cap K$ is a subgroup of G .
4. (a) We check that H1, H2, H3 hold:
 - H1. If $f, g \in H_1$, then f and g are continuous on the interval $[0, 1]$, and moreover, $f(\frac{1}{2}) = 0$ and $g(\frac{1}{2}) = 0$. As f, g are continuous on the interval $[0, 1]$, it follows that $f + g$ is also continuous on $[0, 1]$. Furthermore,

$$(f + g)\left(\frac{1}{2}\right) = f\left(\frac{1}{2}\right) + g\left(\frac{1}{2}\right) = 0 + 0 = 0.$$

Consequently, $f + g \in H_1$.

H2. The constant function $\mathbf{0} : [0, 1] \rightarrow \mathbb{R}$, defined by $\mathbf{0}(x) = 0$ for all $x \in [0, 1]$, is the identity element in the group $C[0, 1]$ with addition of functions. Moreover,

$$\mathbf{0} \left(\frac{1}{2} \right) = 0,$$

and so $\mathbf{0} \in H_1$.

H3. If $f \in H_1$, then $f \left(\frac{1}{2} \right) = 0$. The inverse of f in $C[0, 1]$ is the function $-f : [0, 1] \rightarrow \mathbb{R}$, defined by $(-f)(x) = -f(x)$ for all $x \in [0, 1]$. As

$$(-f) \left(\frac{1}{2} \right) = -f \left(\frac{1}{2} \right) = -0 = 0,$$

it follows that $-f \in H_1$.

So H_1 is a subgroup of the group $C[0, 1]$ with addition of functions.

(b) We check that H1, H2, H3 hold:

H1. If $p, q \in H_2$, then p and q are continuous on the interval $[0, 1]$, and moreover, there exist $n, m \in \mathbb{N} \cup \{0\}$ and real numbers $a_0, a_1, a_2, \dots, a_n, b_0, b_1, \dots, b_m$, such that

$$\begin{aligned} p(x) &= a_0 + a_1x + a_2x^2 + \dots + a_nx^n, \text{ for all } x \in [0, 1], \text{ and} \\ q(x) &= b_0 + b_1x + b_2x^2 + \dots + b_mx^m, \text{ for all } x \in [0, 1]. \end{aligned}$$

If $n \leq m$, then for all $x \in [0, 1]$,

$$\begin{aligned} (p+q)(x) &= p(x) + q(x) \\ &= a_0 + a_1x + a_2x^2 + \dots + a_nx^n + b_0 + b_1x + b_2x^2 + \dots + b_mx^m \\ &= a_0 + b_0 + (a_1 + b_1)x + (a_2 + b_2)x^2 + \dots + (a_n + b_n)x^n + b_{n+1}x^{n+1} + \dots + b_mx^m, \end{aligned}$$

and so $p+q \in H_2$.

On the other hand, if $n > m$, then for all $x \in [0, 1]$,

$$\begin{aligned} (p+q)(x) &= p(x) + q(x) \\ &= a_0 + a_1x + a_2x^2 + \dots + a_nx^n + b_0 + b_1x + b_2x^2 + \dots + b_mx^m \\ &= a_0 + b_0 + (a_1 + b_1)x + (a_2 + b_2)x^2 + \dots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + \dots + a_nx^n, \end{aligned}$$

and so $p+q \in H_2$.

H2. The constant function $\mathbf{0} : [0, 1] \rightarrow \mathbb{R}$, defined by $\mathbf{0}(x) = 0$ for all $x \in [0, 1]$, is the identity element in the group $C[0, 1]$ with addition of functions. Clearly, $\mathbf{0} \in H_2$ (with $n = 0$ and $a_0 = 0$).

H3. If $p \in H_2$, then there exists an $n \in \mathbb{N} \cup \{0\}$ and real numbers $a_0, a_1, a_2, \dots, a_n$, such that $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$, for all $x \in [0, 1]$. The inverse of p in $C[0, 1]$ is the function $-p : [0, 1] \rightarrow \mathbb{R}$, defined by $(-p)(x) = -p(x)$ for all $x \in [0, 1]$. As

$$\begin{aligned} (-p)(x) &= -p(x) \\ &= -(a_0 + a_1x + a_2x^2 + \dots + a_nx^n) \\ &= -a_0 + (-a_1)x + (-a_2)x^2 + \dots + (-a_n)x^n, \end{aligned}$$

for all $x \in [0, 1]$, it follows that $-p \in H_2$.

So H_2 is a subgroup of the group $C[0, 1]$ with addition of functions.

5. Let e denote the identity in the group G .

(a) Since $e \in G$, and for every $a \in G$, $a * e = a = e * a$, it follows that $e \in Z(G)$. Thus $Z(G)$ is not empty.

(b) Clearly $Z(G) \subset G$.

Let $z \in G$. For every $a \in G$, $z * a = a * z$, since G is abelian. So $z \in Z(G)$. Hence $G \subset Z(G)$.

Thus $Z(G) = G$.

(c) We check that H1, H2, H3 hold:

H1. If $z_1, z_2 \in Z(G)$, then for every $a \in G$, we have

$$\begin{aligned} (z_1 * z_2) * a &= z_1 * (z_2 * a) && \text{(associativity)} \\ &= z_1 * (a * z_2) && \text{(since } z_2 \in Z(G)) \\ &= (z_1 * a) * z_2 && \text{(associativity)} \\ &= (a * z_1) * z_2 && \text{(since } z_1 \in Z(G)) \\ &= a * (z_1 * z_2) && \text{(associativity)}. \end{aligned}$$

Thus $z_1 * z_2 \in Z(G)$.

H2. $e \in Z(G)$ since for every $a \in G$, $e * a = a = a * e$.

H3. Let $z \in Z(G)$. Then for every $a \in G$, we have $a^{-1} \in G$, and so

$$z * a^{-1} = a^{-1} * z.$$

Consequently $(z * a^{-1})^{-1} = (a^{-1} * z)^{-1}$, that is, $(a^{-1})^{-1} * z^{-1} = z^{-1} * (a^{-1})^{-1}$. But $a * a^{-1} = e = a^{-1} * a$, and from the uniqueness of inverses, it follows that $(a^{-1})^{-1} = a$. Thus we obtain $a * z^{-1} = (a^{-1})^{-1} * z^{-1} = z^{-1} * (a^{-1})^{-1} = z^{-1} * a$. Hence $z^{-1} \in Z(G)$.

So $Z(G)$ is a subgroup of the group G .

(d) If $z = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in Z(GL(2, \mathbb{R}))$, then since

$$a = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in GL(2, \mathbb{R}),$$

it follows that

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \alpha & \alpha + \beta \\ \gamma & \gamma + \delta \end{bmatrix} = \begin{bmatrix} \alpha + \gamma & \beta + \delta \\ \gamma & \delta \end{bmatrix},$$

and so it follows that $\gamma = 0$ and $\alpha = \delta$. Also, since

$$b = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \in GL(2, \mathbb{R}),$$

it follows that

$$\begin{bmatrix} \alpha & \beta \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ 0 & \alpha \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \alpha + \beta & \beta \\ \alpha & \alpha \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \alpha & \beta + \alpha \end{bmatrix},$$

and so it follows that $\beta = 0$. Hence we have shown that

$$Z(GL(2, \mathbb{R})) \subset \left\{ \alpha \cdot I = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \mid \alpha \in \mathbb{R} \setminus \{0\} \right\}. \quad (2.29)$$

Let $\alpha \in \mathbb{R} \setminus \{0\}$. Then $\alpha \cdot I \in GL(2, \mathbb{R})$, and for all

$$\begin{bmatrix} p & q \\ r & s \end{bmatrix} \in GL(2, \mathbb{R}),$$

we have

$$\begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} \alpha p & \alpha q \\ \alpha r & \alpha s \end{bmatrix} = \begin{bmatrix} p\alpha & q\alpha \\ r\alpha & s\alpha \end{bmatrix} = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}.$$

Hence $\alpha \cdot I \in Z(GL(2, \mathbb{R}))$. Thus

$$\left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \mid \alpha \in \mathbb{R} \setminus \{0\} \right\} \subset Z(GL(2, \mathbb{R})). \quad (2.30)$$

From (2.29) and (2.30), we conclude that

$$Z(GL(2, \mathbb{R})) = \left\{ \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \mid \alpha \in \mathbb{R} \setminus \{0\} \right\}.$$

Solutions to the exercises on page 48

1. We know that $S = \{e, a, a^2, a^3, \dots, a^{|G|}\} \subset G$, and that S has $|G| + 1$ elements, while G has $|G|$ elements. So from the pigeonhole principle, it follows that there exist distinct $k_1, k_2 \in \{0, 1, 2, 3, \dots, |G|\}$ such that $a^{k_1} = a^{k_2}$. We may assume that $k_2 > k_1$ (otherwise we can interchange them), and so $a^{k_2 - k_1} = e$. As $k_2 - k_1 \in \mathbb{N}$, it follows that a has finite order. Furthermore, as $k_2 \leq |G|$ and $k_1 \geq 0$, we obtain $k_2 - k_1 \leq |G|$. Thus

$$\text{ord}(a) = \min\{m \in \mathbb{N} \mid a^m = e\} \leq k_2 - k_1 \leq |G|.$$

So a has order at most equal to $|G|$.

2. CLAIM: For all $m, n \in \mathbb{N}$, $a^m * a^n = a^{m+n}$.

Proof We prove this in several steps.

STEP 1. Let $m \geq 0$. We show that $a^m * a^n = a^{m+n}$ for all $n \geq 0$ by induction on n . Clearly

$$a^m * a^0 = a^m * e = a^m = a^{m+0},$$

and so the result is true for $n = 0$. If $a^m * a^k = a^{m+k}$ for some $k \geq 0$, then we have

$$\begin{aligned} a^m * a^{k+1} &= a^m * (a^k * a) && \text{(definition of } a^{k+1}\text{)} \\ &= (a^m * a^k) * a && \text{(associativity)} \\ &= a^{m+k} * a && \text{(induction hypothesis)} \\ &= a^{(m+k)+1} && \text{(definition of } a^{(m+k)+1}\text{)} \\ &= a^{m+(k+1)}. \end{aligned}$$

So by induction, we have $a^m * a^n = a^{m+n}$ for all $n \geq 0$. But the choice of $m \geq 0$ was arbitrary, and so we have shown that

$$\text{for all } m \geq 0, \text{ and all } n \geq 0, a^m * a^n = a^{m+n}. \quad (2.31)$$

STEP 2. Let $m < 0$ and $n \geq 0$. Then we have the following two cases:

1^o If $n + m > 0$, then from (2.31), we have $a^{-m} * a^{n+m} = a^{-m+(n+m)} = a^n$, and so $a^{-m} * a^{n+m} = a^n$. We observe that $a^m * a^{-m} = (a^{-m})^{-1} * a^{-m} = e$. So we obtain $a^{n+m} = e * a^{n+m} = (a^m * a^{-m}) * a^{n+m} = a^m * (a^{-m} * a^{n+m}) = a^m * a^n$.

2^o If $n + m \leq 0$, then from (2.31), we have $a^n * a^{-n-m} = a^{n+(-n-m)} = a^{-m}$, and so $a^n * a^{-n-m} = a^{-m}$. So by premultiplying by $a^m = (a^{-m})^{-1}$ and postmultiplying by $a^{n+m} = (a^{-n-m})^{-1}$, we obtain $a^m * a^n = a^{n+m} = a^{m+n}$.

Hence we have shown that for all $m < 0$, and all $n \geq 0$, $a^m * a^n = a^{m+n}$. Combining this result with the result from STEP 1, we obtain

$$\text{for all } m \in \mathbb{Z}, \text{ and all } n \geq 0, a^m * a^n = a^{m+n}. \quad (2.32)$$

STEP 3. Let $m \in \mathbb{Z}$ and $n < 0$. Then from STEP 2, we have $a^{m+n} * a^{-n} = a^{(m+n)-n} = a^m$. By postmultiplying both sides by $a^n = (a^{-n})^{-1}$, we then obtain $a^{m+n} = a^m * a^n$. Hence we have shown that for all $m \in \mathbb{Z}$ and all $n < 0$, $a^m * a^n = a^{m+n}$. Combining this result with the result from STEP 2, namely (2.32), we obtain that for all integers m and n , $a^m * a^n = a^{m+n}$. ■

Next we prove the following claim.

CLAIM: For all $m, n \in \mathbb{N}$, $(a^m)^n = a^{mn}$.

Proof First we show that for all $m \in \mathbb{Z}$ and all $n \geq 0$, $(a^m)^n = a^{mn}$, by induction on n . Let $m \in \mathbb{Z}$. We have $(a^m)^0 = e = a^0 = a^{m0}$, and so the statement holds with $n = 0$. If for some $k \geq 0$ there holds that $(a^m)^k = a^{mk}$, then we have

$$(a^m)^{k+1} = (a^m)^k * a^m = a^{mk} * a^m = a^{mk+m} = a^{m(k+1)}.$$

So by induction, we have that $(a^m)^n = a^{mn}$ for all $n \geq 0$. But the choice of $m \in \mathbb{Z}$ was arbitrary, and so it follows that for all $m \in \mathbb{Z}$ and all $n \geq 0$, $(a^m)^n = a^{mn}$.

If $n < 0$, then we have $(a^m)^n = ((a^m)^{-n})^{-1} = (a^{m(-n)})^{-1} = (a^{-mn})^{-1} = a^{mn}$, where the last equality follows from the following fact: for all $k \in \mathbb{Z}$, $(a^{-k})^{-1} = a^k$ (Indeed, for $k < 0$, we know that $a^k = (a^{-k})^{-1}$, by the definition of a^k for a negative k . If $k \geq 0$, then $-k \leq 0$, and so $a^{-k} = (a^{-(-k)})^{-1} = (a^k)^{-1}$. Hence by taking inverses, we obtain $(a^{-k})^{-1} = ((a^k)^{-1})^{-1} = a^k$.) ■

3. We have

$$\begin{aligned} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^1 &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \\ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^2 &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^1 \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \\ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^3 &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^2 \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \\ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^4 &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^3 \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \\ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^5 &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^4 \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \\ \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^6 &= \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^5 \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Thus

$$\text{ord} \left(\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \right) = \min \left\{ m \in \mathbb{N} \mid \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}^m = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} = 6.$$

4. If $a * b$ has finite order, say m , then $(a * b)^m = e$, that is,

$$\underbrace{(a * b) * \cdots * (a * b)}_{m \text{ times}} = e,$$

and so

$$a * \underbrace{(b * a) * \cdots * (b * a)}_{(m-1) \text{ times}} * b = e,$$

that is,

$$a * (b * a)^{m-1} * b = e. \quad (2.33)$$

Hence by premultiplying (2.33) by b , and postmultiplying by b^{-1} , we obtain

$$(b * a) * (b * a)^{m-1} = e,$$

that is, $(b * a)^m = e$. So we see that $b * a$ has finite order, and $\text{ord}(b * a) \leq m$. By interchanging the roles of a and b , we obtain that $m \leq \text{ord}(b * a)$, and so $\text{ord}(b * a) = m$.

So we have shown that

if $a * b$ has finite order m , then $b * a$ has finite order m .

By interchanging a and b , we obtain

$a * b$ has finite order m iff $b * a$ has finite order m .

Hence if $a * b$ has infinite order, then $b * a$ has infinite order as well.

So the orders of $a * b$ and $b * a$ are the same.

5. Yes, since $\langle 1 \rangle = \mathbb{Z}$.

1 is a generator of \mathbb{Z} with addition.

As $\langle -1 \rangle = \mathbb{Z}$, we see that -1 is also a generator of \mathbb{Z} with addition. So the cyclic group \mathbb{Z} with addition does not have a unique generator.

Solutions to the exercises on page 51

1. (a) For all a, b in G , we have

$$\begin{aligned} (\psi \circ \varphi)(a * b) &= \psi(\varphi(a * b)) && \text{(definition of } \psi \circ \varphi) \\ &= \psi(\varphi(a) *' \varphi(b)) && \text{(since } \varphi \text{ is a homomorphism)} \\ &= \psi(\varphi(a)) *'' \psi(\varphi(b)) && \text{(since } \psi \text{ is a homomorphism)} \\ &= (\psi \circ \varphi)(a) *'' (\psi \circ \varphi)(b) && \text{(definition of } \psi \circ \varphi), \end{aligned}$$

and so $\psi \circ \varphi : G \rightarrow G''$ is a homomorphism.

(b) We have

$$\begin{aligned} \ker(\psi \circ \varphi) &= \{a \in G \mid (\psi \circ \varphi)(a) = e''\} \\ &= \{a \in G \mid \psi(\varphi(a)) = e''\} \\ &= \{a \in G \mid \varphi(a) \in \ker(\psi)\} \\ &= \varphi^{-1}(\ker(\psi)). \end{aligned}$$

2. If $a \in \ker(\varphi)$, then $\varphi(a) = e'$. Thus for all $b \in G$,

$$\varphi(b * a * b^{-1}) = \varphi(b) *' \varphi(a) *' \varphi(b^{-1}) = \varphi(b) *' e' *' \varphi(b^{-1}) = \varphi(b) *' \varphi(b^{-1}) = \varphi(b) *' (\varphi(b))^{-1} = e',$$

and so $b * a * b^{-1} \in \ker(\varphi)$. Since the choice of $a \in \ker(\varphi)$ was arbitrary, it follows that

$$\forall a \in \ker(\varphi) \text{ and } \forall b \in G, \quad b * a * b^{-1} \in \ker(\varphi),$$

and so $\ker(\varphi)$ is a normal subgroup of G .

3. For all $a, b \in G$, we have

$$\begin{aligned} \varphi(a * b) &= (a * b)^{-1} && \text{(definition of } \varphi) \\ &= b^{-1} * a^{-1} && \text{(Exercise 4c from the exercises on page 44)} \\ &= a^{-1} * b^{-1} && \text{(since } G \text{ is abelian)} \\ &= \varphi(a) * \varphi(b) && \text{(definition of } \varphi), \end{aligned}$$

and so $\varphi : G \rightarrow G$ is a homomorphism.

Let $a, b \in G$ and $a \neq b$. Then $a^{-1} \neq b^{-1}$ (for otherwise, $a = (a^{-1})^{-1} = (b^{-1})^{-1} = b$). Hence $\varphi(a) = a^{-1} \neq b^{-1} = \varphi(b)$, and so $\varphi : G \rightarrow G$ is one-to-one. For all $a \in G$, we have $a^{-1} \in G$ and $\varphi(a^{-1}) = (a^{-1})^{-1} = a$, and so $\varphi : G \rightarrow G$ is onto. As $\varphi : G \rightarrow G$ is one-to-one and onto, it is a bijection.

Consequently $\varphi : G \rightarrow G$ is an isomorphism.

4. First we show that $\varphi^{-1} : G' \rightarrow G$ is a homomorphism.

Let $a', b' \in G'$. Since $\varphi : G \rightarrow G'$ is a bijection, it follows that there exist unique $a, b \in G$ such that $\varphi(a) = a'$ and $\varphi(b) = b'$. We have

$$\varphi(a * b) = \varphi(a) *' \varphi(b) = a' *' b',$$

and so $\varphi^{-1}(a' *' b') = a * b = \varphi^{-1}(a') * \varphi^{-1}(b')$. So φ^{-1} is a homomorphism.

If $a' \neq b'$, then clearly $\varphi^{-1}(a') \neq \varphi^{-1}(b')$ (for otherwise $a' = \varphi(\varphi^{-1}(a')) = \varphi(\varphi^{-1}(b')) = b'$). Thus $\varphi^{-1} : G' \rightarrow G$ is one-to-one. For all $a \in G$, we have $\varphi^{-1}(\varphi(a)) = a$, and so $\varphi^{-1} : G' \rightarrow G$ is onto. Hence $\varphi^{-1} : G' \rightarrow G$ is a bijection.

Consequently $\varphi^{-1} : G' \rightarrow G$ is an isomorphism.

5. Let us denote the homomorphism $m \mapsto a^m$ from \mathbb{Z} to $\langle a \rangle$ by φ .

(a) For all $m, n \in \mathbb{Z}$, we have $\varphi(m+n) = a^{m+n} = a^m * a^n = \varphi(m) * \varphi(n)$, and so $\varphi : \mathbb{Z} \rightarrow \langle a \rangle$ is a homomorphism.

(b) If $m \neq n$, then $\varphi(m) = a^m \neq a^n = \varphi(n)$ (for otherwise, if $m > n$, then $a^{m-n} = e$, and if $m < n$, then $a^{n-m} = e$; in either case we get a contradiction to the fact that a has infinite order). Hence φ is one-to-one.

Furthermore, if $b \in \langle a \rangle$, then $b = a^m$ for some $m \in \mathbb{Z}$. We have $\varphi(m) = a^m = b$, and so φ is onto.

So φ is a bijection and consequently, $\varphi : \mathbb{Z} \rightarrow \langle a \rangle$ is an isomorphism.

Solutions to the exercises on page 54

1. (a) FALSE.

For example, consider the group \mathbb{Z} of integers with addition, and let H be the subgroup of even integers. Then $0 \neq 2$, but $0 + H = 2 + H$ (both these cosets are equal to H).

- (b) TRUE.

Consider the function $f : H \rightarrow a * H$ given by $h \mapsto a * h$, for all $h \in H$. Then it can be seen that f is a bijection as follows. If $a \in G$ and $h_1, h_2 \in H$ are such that $a * h_1 = a * h_2$, then by premultiplication with a^{-1} , we obtain $h_1 = h_2$. Thus f is one-to-one. Furthermore, if $b \in a * H$, then there exists a $h \in H$ such that $b = a * h$. Consequently $f(h) = a * h = b$. Hence f is onto. So f is a bijection, and it follows that the cardinalities of H and $a * H$ are the same.

- (c) TRUE.

As in part 1b above, one can show that the function $g : H \rightarrow H * a$, given by $h \mapsto h * a$, for all $h \in H$, is a bijection. Thus $a * H$, H and $H * a$ all have the same cardinalities.

2. (a) Matrix multiplication is a law of composition on G :

$$\text{if } A = \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} x' & y' \\ 0 & 1 \end{bmatrix} \text{ belong to } G, \text{ then } x > 0 \text{ and } x' > 0.$$

We have

$$AB = \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x' & y' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} xx' & xy' + y \\ 0 & 1 \end{bmatrix} \text{ and } xx' > 0,$$

and so $AB \in G$. Moreover the group axioms are satisfied:

G1. As matrix multiplication is associative, it follows that in particular for elements A, B, C from G , there holds $(AB)C = A(BC)$.

G2. The identity matrix

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in G \quad (x := 1 > 0, y := 0),$$

and furthermore, for all $A \in G$, clearly we have $AI = A = IA$.

G3. If

$$\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \in G,$$

then $x > 0$. Thus $\frac{1}{x} > 0$, and

$$\begin{bmatrix} \frac{1}{x} & -\frac{y}{x} \\ 0 & 1 \end{bmatrix} \in G.$$

Moreover,

$$\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{x} & -\frac{y}{x} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{x} & -\frac{y}{x} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix}.$$

So G is a group with matrix multiplication.

- (b) Clearly $H \subset G$. We now check that H1, H2, H3 are satisfied:

H1. If

$$\begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} x' & 0 \\ 0 & 1 \end{bmatrix} \in H,$$

then $x > 0$ and $x' > 0$. We have

$$\begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x' & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} xx' & 0 \\ 0 & 1 \end{bmatrix}$$

and $xx' > 0$. So

$$\begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x' & 0 \\ 0 & 1 \end{bmatrix} \in H.$$

H2. Clearly the identity element from G , namely the identity matrix

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

also belongs to H .

H3. If

$$\begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix} \in H,$$

then $x > 0$. So $\frac{1}{x} > 0$, and the inverse of

$$\begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}$$

in the group G , namely the matrix

$$\begin{bmatrix} \frac{1}{x} & 0 \\ 0 & 1 \end{bmatrix},$$

also belongs to H .

Thus H is a subgroup of the group G .

(c) Let

$$a = \begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} \in G.$$

Then the corresponding left coset of H is

$$\begin{aligned} aH &= \left\{ \begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix} \mid x \in \mathbb{R} \text{ and } x > 0 \right\} \\ &= \left\{ \begin{bmatrix} xx_0 & y_0 \\ 0 & 1 \end{bmatrix} \mid x \in \mathbb{R} \text{ and } x > 0 \right\}. \end{aligned}$$

Thus the left coset aH is a straight line parallel to the x -axis, and it passes through the point (x_0, y_0) . See Figure 2.7. The right coset of H corresponding to the element

$$a = \begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} \in G$$

is

$$\begin{aligned} Ha &= \left\{ \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 & y_0 \\ 0 & 1 \end{bmatrix} \mid x \in \mathbb{R} \text{ and } x > 0 \right\} \\ &= \left\{ \begin{bmatrix} xx_0 & xy_0 \\ 0 & 1 \end{bmatrix} \mid x \in \mathbb{R} \text{ and } x > 0 \right\}. \end{aligned}$$

Thus the right coset Ha is a straight line passing through the origin and the point (x_0, y_0) . See Figure 2.8.

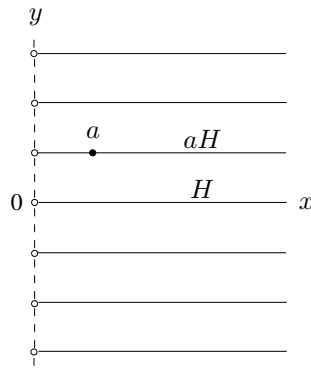


Figure 2.7: Partition of the group G (represented by the right half plane: $x > 0$, $y \in \mathbb{R}$) by left cosets aH , $a \in G$ (represented by lines parallel to the x -axis).

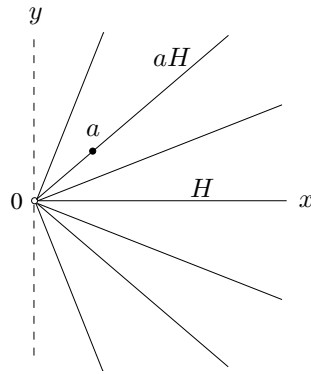


Figure 2.8: Partition of the group G by right cosets Ha , $a \in G$.

3. $H \cap K$ is a subset of the group H . Furthermore, the following hold:

- H1. If $a, b \in H \cap K$, then $a, b \in H$ and $a, b \in K$. As H and K are subgroups of G , it follows that $a * b \in H$ and $a * b \in K$. Hence $a * b \in H \cap K$.
- H2. The identity element in the group H is the identity element e from G . As K is a subgroup of G , it follows that e also belongs to K . Thus $e \in H \cap K$.
- H3. If $a \in H \cap K$, then $a \in H$ and $a \in K$. As H and K are subgroups, $a^{-1} \in H$ and $a^{-1} \in K$. Consequently $a^{-1} \in H \cap K$.

So $H \cap K$ is a subgroup of H .

By interchanging the roles of H and K , it follows that $H \cap K$ is also a subgroup of K .

Thus by Lagrange's theorem, we have that

$$|H \cap K| \text{ divides } |H| = 3, \text{ and } |H \cap K| \text{ divides } |K| = 5.$$

As $\gcd(3, 5) = 1$, it follows that $|H \cap K|$ divides 1. So $|H \cap K| = 1$, and so it comprises just one element. As $e \in H \cap K$, it follows that $H \cap K = \{e\}$.

- 4. The cardinality of the group S_4 is $4! = 24$. As 16 does not divide 24, it follows from Corollary 2.1.8 that S_4 cannot have an element of order 16.
- 5. (a) If p does not divide a , then $[a] \neq [0]$, and so $[a] \in \mathbb{Z}_p^*$. The set $\mathbb{Z}_p^* = \{[1], \dots, [p-1]\}$ has cardinality equal to $p-1$. As $[1]$ is the identity element in the group \mathbb{Z}_p^* with multiplication modulo p , from Corollary 2.1.8 it follows that $[a]^{p-1} = [1]$.

(b) If p divides a , then p also divides a^p , and hence it divides $a^p - a$ as well, that is, $a^p \equiv a \pmod{p}$.

If p does not divide a , then $[a] \neq [0]$, and so by part 5a above, $[a]^{p-1} = [1]$. Hence $[a^{p-1}] = [1]$, and so $[a^{p-1} - 1] = [0]$. Thus p divides $a^{p-1} - 1$, and consequently, it also divides $a(a^{p-1} - 1) = a^p - a$. So $a^p \equiv a \pmod{p}$.

(c) Note that $2222 = 317 \cdot 7 + 3$, so that in \mathbb{Z}_7 ,

$$[2222^{5555}] = [2222]^{5555} = [3]^{5555}.$$

As $[3] \neq [0]$ in \mathbb{Z}_7 , from part (5b) above, it follows that $[3]^{7-1} = [3]^6 = [1]$. Thus we obtain

$$[2222^{5555}] = [3]^{5555} = [3]^{925 \cdot 6 + 5} = ([3]^6)^{925} [3]^5 = [1]^{925} [3]^5 = [1][3]^5 = [3]^5.$$

Proceeding in a similar manner, we can show that $[5555^{2222}] = [3^2]$. Indeed, $5555 = 793 \cdot 7 + 4$, so that

$$[5555^{2222}] = [5555]^{2222} = [4]^{2222}.$$

Again, $[4] \neq [0]$ in \mathbb{Z}_7 , and so by part (5b), $[4]^{7-1} = [4]^6 = [1]$. Hence

$$[5555^{2222}] = [4]^{2222} = [4]^{370 \cdot 6 + 2} = ([4]^6)^{370} [4]^2 = [1][4]^2 = [4]^2 = [-3]^2 = [(-3)^2] = [3^2].$$

Finally,

$$[2222^{5555} + 5555^{2222}] = [3^5 + 3^2] = [3^2 \cdot 28] = [0],$$

which proves that $7 \mid 2222^{5555} + 5555^{2222}$.

Solutions to the exercises on page 58

1. (a) NO.

For instance the sum of two invertible matrices may not be invertible. The identity matrix I is invertible, and so is its additive inverse, $-I$. However, $I + (-I) = 0$, which is not invertible. So matrix addition, $(A, B) \mapsto A + B$ is not a law of composition on the set of invertible matrices, and so it does not form an abelian group. Consequently it is not a vector space.

Alternately, we can observe that the scalar multiplication of the real number 0 with an invertible matrix is the zero matrix, which is not invertible.

(b) NO.

The scalar multiplication of the real number 0 with an invertible matrix is the zero matrix, which is not invertible.

2. Let $\alpha \in \mathbb{R}$ and $v \in V$ be such that

$$\alpha \cdot v = \mathbf{0}. \quad (2.34)$$

Then if $\alpha \neq 0$, then we obtain

$$\begin{aligned} \mathbf{0} &= \alpha^{-1} \cdot \mathbf{0} \quad (\text{Theorem 2.2.1}) \\ &= \alpha^{-1} \cdot (\alpha \cdot v) \quad (\text{using (2.34)}) \\ &= (\alpha^{-1}\alpha) \cdot v \quad (\text{using V2}) \\ &= (1) \cdot v \quad (\text{since } \alpha^{-1}\alpha = 1) \\ &= v \quad (\text{using V1}). \end{aligned}$$

This proves the claim.

3. First we check that \mathbb{R}^∞ , with addition defined by (2.4), is an abelian group. Clearly (2.4) gives a law of composition on \mathbb{R}^∞ . Moreover, we have:

G1. For all $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}, (c_n)_{n \in \mathbb{N}}$ in \mathbb{R}^∞ , we have

$$\begin{aligned} ((a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}}) + (c_n)_{n \in \mathbb{N}} &= (a_n + b_n)_{n \in \mathbb{N}} + (c_n)_{n \in \mathbb{N}} \\ &= ((a_n + b_n) + c_n)_{n \in \mathbb{N}} \\ &= (a_n + (b_n + c_n))_{n \in \mathbb{N}} \\ &= (a_n)_{n \in \mathbb{N}} + (b_n + c_n)_{n \in \mathbb{N}} \\ &= (a_n)_{n \in \mathbb{N}} + ((b_n)_{n \in \mathbb{N}} + (c_n)_{n \in \mathbb{N}}). \end{aligned}$$

G2. The sequence $(0)_{n \in \mathbb{N}}$ serves as the identity element: for all $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$, we have

$$(a_n)_{n \in \mathbb{N}} + (0)_{n \in \mathbb{N}} = (a_n + 0)_{n \in \mathbb{N}} = (a_n)_{n \in \mathbb{N}} = (0 + a_n)_{n \in \mathbb{N}} = (0)_{n \in \mathbb{N}} + (a_n)_{n \in \mathbb{N}}.$$

G3. If $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$, then $(-a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$, and furthermore,

$$(a_n)_{n \in \mathbb{N}} + (-a_n)_{n \in \mathbb{N}} = (a_n + -a_n)_{n \in \mathbb{N}} = (0)_{n \in \mathbb{N}} = (-a_n + a_n)_{n \in \mathbb{N}} = (-a_n)_{n \in \mathbb{N}} + (a_n)_{n \in \mathbb{N}}.$$

G4. Finally, commutativity holds, since for all $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$, we have

$$(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} = (a_n + b_n)_{n \in \mathbb{N}} = (b_n + a_n)_{n \in \mathbb{N}} = (b_n)_{n \in \mathbb{N}} + (a_n)_{n \in \mathbb{N}}.$$

Next we check that V1, V2, V3, V4 are also satisfied:

V1. For all $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$, $1 \cdot (a_n)_{n \in \mathbb{N}} = (1a_n)_{n \in \mathbb{N}} = (a_n)_{n \in \mathbb{N}}$.

V2. For all $\alpha, \beta \in \mathbb{R}$ and all $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$,

$$\alpha \cdot (\beta \cdot (a_n)_{n \in \mathbb{N}}) = \alpha \cdot (\beta a_n)_{n \in \mathbb{N}} = (\alpha(\beta a_n))_{n \in \mathbb{N}} = ((\alpha\beta)a_n)_{n \in \mathbb{N}} = (\alpha\beta) \cdot (a_n)_{n \in \mathbb{N}}.$$

V3. For all $\alpha, \beta \in \mathbb{R}$ and all $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$,

$$\begin{aligned} (\alpha + \beta) \cdot (a_n)_{n \in \mathbb{N}} &= ((\alpha + \beta)a_n)_{n \in \mathbb{N}} = (\alpha a_n + \beta a_n)_{n \in \mathbb{N}} = (\alpha a_n)_{n \in \mathbb{N}} + (\beta a_n)_{n \in \mathbb{N}} \\ &= \alpha \cdot (a_n)_{n \in \mathbb{N}} + \beta \cdot (a_n)_{n \in \mathbb{N}}. \end{aligned}$$

V4. For all $\alpha \in \mathbb{R}$ and all $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty$,

$$\begin{aligned} \alpha \cdot ((a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}}) &= \alpha \cdot (a_n + b_n)_{n \in \mathbb{N}} = (\alpha(a_n + b_n))_{n \in \mathbb{N}} \\ &= (\alpha a_n + \alpha b_n)_{n \in \mathbb{N}} = (\alpha a_n)_{n \in \mathbb{N}} + (\alpha b_n)_{n \in \mathbb{N}} \\ &= \alpha \cdot (a_n)_{n \in \mathbb{N}} + \alpha \cdot (b_n)_{n \in \mathbb{N}}. \end{aligned}$$

So \mathbb{R}^∞ is a vector space with addition defined by (2.4) and scalar multiplication defined by (2.5).

Solutions to the exercises on page 62

1. (a) FALSE.

Consider the subspaces

$$U_1 = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \text{ and } U_2 = \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

of \mathbb{R}^2 . Then

$$v_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in U_1 \subset U_1 \cup U_2 \text{ and } v_2 := \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in U_2 \subset U_1 \cup U_2,$$

but

$$v_1 + v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin U_1 \cup U_2,$$

since

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \text{ and } \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin \text{span} \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Thus S2 is not satisfied, and so $U_1 \cup U_2$ is not a vector space.

(Each vector $\begin{bmatrix} x \\ y \end{bmatrix}$ in \mathbb{R}^2 can be represented by a point in the (x, y) -plane, and the above can be seen pictorially in Figure 2.9).

(b) TRUE.

Let U_1, U_2 be subspaces of the vector space V . We check that S1, S2, S3 hold for $U_1 \cap U_2$:

S1. As U_1, U_2 are subspaces, $\mathbf{0} \in U_1$ and $\mathbf{0} \in U_2$. Thus $\mathbf{0} \in U_1 \cap U_2$.

S2. If v_1, v_2 belong to $U_1 \cap U_2$, then v_1, v_2 belong to U_1 and v_1, v_2 belong to U_2 . As U_1 is a subspace, $v_1 + v_2 \in U_1$. Moreover, as U_2 is a subspace, $v_1 + v_2 \in U_2$. Hence $v_1 + v_2 \in U_1 \cap U_2$.

S3. Let $\alpha \in \mathbb{R}$ and $v \in U_1 \cap U_2$. Thus $v \in U_1$ and $v \in U_2$. As U_1 is a subspace, it follows that $\alpha \cdot v \in U_1$. Also, as U_2 is a subspace, it follows that $\alpha \cdot v \in U_2$. Consequently $\alpha \cdot v \in U_1 \cap U_2$.

So $U_1 \cap U_2$ is a subspaces of the vector space V .

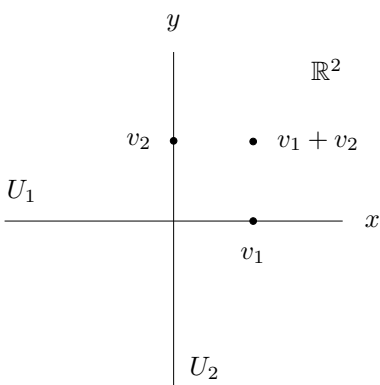


Figure 2.9: The union $U_1 \cup U_2$ of subspaces U_1 and U_2 is not a subspace.

(c) TRUE.

Indeed,

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} + (-1) \cdot \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

(d) FALSE.

We have

$$(3) \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-3) \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + (2) \cdot \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

(e) TRUE.

Suppose that there exist scalars such that $\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \alpha_3 \cdot v_3 = \mathbf{0}$. Then we have

$$\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \alpha_3 \cdot v_3 + 0 \cdot v_4 = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \alpha_3 \cdot v_3 + \mathbf{0} = \alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \alpha_3 \cdot v_3 = \mathbf{0},$$

and by the independence of v_1, v_2, v_3, v_4 , it follows that $\alpha_1 = \alpha_2 = \alpha_3 = 0$. So v_1, v_2, v_3 are linearly independent.

(f) FALSE.

For instance, in \mathbb{R}^3 , the vectors

$$v_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad v_4 := \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

are linearly dependent, since $1 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3 + (-1) \cdot v_4 = \mathbf{0}$, while the vectors v_1, v_2, v_3 are linearly independent.

2. (a) Clearly,

$$\text{span} \left(\left\{ \begin{bmatrix} 1 \\ t_1 \end{bmatrix}, \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \right\} \right) \subset \mathbb{R}^2.$$

(In order to prove the reverse inclusion, observe that if

$$\begin{bmatrix} x \\ y \end{bmatrix} = \alpha \cdot \begin{bmatrix} 1 \\ t_1 \end{bmatrix} + \beta \cdot \begin{bmatrix} 1 \\ t_2 \end{bmatrix},$$

then we obtain the linear equations

$$\begin{aligned} \alpha + \beta &= x, \\ \alpha t_1 + \beta t_2 &= y, \end{aligned}$$

which have the solution $\alpha = \frac{y-xt_2}{t_1-t_2}$ and $\beta = \frac{xt_1-y}{t_1-t_2}$.) We now complete the proof as follows: suppose that

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2.$$

As $t_1 - t_2 \neq 0$, we have

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{y-xt_2}{t_1-t_2} \cdot \begin{bmatrix} 1 \\ t_1 \end{bmatrix} + \frac{xt_1-y}{t_1-t_2} \cdot \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \in \text{span} \left(\left\{ \begin{bmatrix} 1 \\ t_1 \end{bmatrix}, \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \right\} \right).$$

So

$$\mathbb{R}^2 \subset \text{span} \left(\left\{ \begin{bmatrix} 1 \\ t_1 \end{bmatrix}, \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \right\} \right).$$

Hence the claim follows.

- (b) Suppose that U_1, \dots, U_n are subspaces of \mathbb{R}^2 , such that $U_1 \cup \dots \cup U_n = \mathbb{R}^2$. Since the infinite set

$$S := \left\{ \begin{bmatrix} 1 \\ t \end{bmatrix} \mid t \in \mathbb{R} \right\}$$

is contained in $\mathbb{R}^2 = U_1 \cup \dots \cup U_n$, it follows that infinitely many elements from S belong to one of the subspaces, say U_k for some $k \in \{1, \dots, n\}$. Thus there exist real numbers t_1, t_2 such that

$$\begin{bmatrix} 1 \\ t_1 \end{bmatrix}, \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \in U_k.$$

As U_k is a subspace of \mathbb{R}^2 , it follows that

$$\text{span} \left(\left\{ \begin{bmatrix} 1 \\ t_1 \end{bmatrix}, \begin{bmatrix} 1 \\ t_2 \end{bmatrix} \right\} \right) \subset U_k.$$

From the result of the part above, we obtain $\mathbb{R}^2 \subset U_k$, and so $U_k = \mathbb{R}^2$, that is U_k is not a proper subspace.

3. ℓ^∞ is a subspace of \mathbb{R}^∞ . We have:

- S1. The sequence $(0)_{n \in \mathbb{N}}$ is the zero vector in \mathbb{R}^∞ , and it belongs to ℓ^∞ , since it is bounded: indeed for all $n \in \mathbb{N}$, the n th term of the sequence $(0)_{n \in \mathbb{N}}$ is equal to 0, and $|0| = 0 \leq 1$.
- S2. If $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ belong to ℓ^∞ , then there exist $M_1, M_2 > 0$ such that for all $n \in \mathbb{N}$, $|a_n| \leq M_1$ and $|b_n| \leq M_2$. Hence for all $n \in \mathbb{N}$, $|a_n + b_n| \leq |a_n| + |b_n| \leq M_1 + M_2$, and so the sequence $(a_n + b_n)_{n \in \mathbb{N}}$ is bounded, that is, $(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}}$ belongs to ℓ^∞ .
- S3. Let $\alpha \in \mathbb{R}$ and $(a_n)_{n \in \mathbb{N}} \in \ell^\infty$. Then there exists an $M > 0$ such that for all $n \in \mathbb{N}$, $|a_n| \leq M$. Thus for all $n \in \mathbb{N}$, $|\alpha a_n| = |\alpha| |a_n| \leq (|\alpha| + 1)M$, and so the sequence $(\alpha a_n)_{n \in \mathbb{N}}$ is bounded, that is, $\alpha \cdot (a_n)_{n \in \mathbb{N}}$ belongs to ℓ^∞ .

So ℓ^∞ is a subspace of \mathbb{R}^∞ .

c is a subspace of ℓ^∞ . Indeed, there holds:

- S1. The sequence $(0)_{n \in \mathbb{N}}$ is the zero vector in ℓ^∞ , and it belongs to c , since it is convergent.
- S2. If $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ belong to c , then by Theorem 1.2.4, it follows that the sequence $(a_n + b_n)_{n \in \mathbb{N}}$ is also convergent, and so $(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}}$ belongs to c .
- S3. Let $\alpha \in \mathbb{R}$ and $(a_n)_{n \in \mathbb{N}} \in c$. Then by Theorem 1.2.4, the sequence $(\alpha a_n)_{n \in \mathbb{N}}$ is also convergent, that is, $\alpha \cdot (a_n)_{n \in \mathbb{N}}$ belongs to c .

Thus c is a subspace of ℓ^∞ .

c_0 is a subspace of c . We have:

- S1. The sequence $(0)_{n \in \mathbb{N}}$ is the zero vector in c , and it belongs to c_0 , since it is a convergent sequence with limit equal to 0.
- S2. If $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ belong to c_0 , then by Theorem 1.2.4, it follows that the sequence $(a_n + b_n)_{n \in \mathbb{N}}$ is also convergent and

$$\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n = 0 + 0 = 0,$$

and so $(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}}$ belongs to c_0 .

- S3. Let $\alpha \in \mathbb{R}$ and $(a_n)_{n \in \mathbb{N}} \in c_0$. Then by Theorem 1.2.4, the sequence $(\alpha a_n)_{n \in \mathbb{N}}$ is also convergent and

$$\lim_{n \rightarrow \infty} \alpha a_n = \alpha \lim_{n \rightarrow \infty} a_n = \alpha 0 = 0,$$

that is, $\alpha \cdot (a_n)_{n \in \mathbb{N}}$ belongs to c_0 .

Hence c_0 is a subspace of c .

c_{00} is a subspace of c_0 . Indeed, there holds:

- S1. The sequence $(0)_{n \in \mathbb{N}}$ is the zero vector in c_0 , and it belongs to c_{00} , since for all $n > 1$, $a_n = 0$.
- S2. If $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ belong to c_{00} , then there exist $N_1, N_2 \in \mathbb{N}$ such that

$$\forall n > N_1, a_n = 0 \quad \text{and} \quad \forall n > N_2, b_n = 0.$$

Thus with $N := \max\{N_1, N_2\} \in \mathbb{N}$, for all $n > N$, we obtain $a_n + b_n = 0 + 0 = 0$. Hence $(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} = (a_n + b_n)_{n \in \mathbb{N}}$ belongs to c_{00} .

- S3. Let $\alpha \in \mathbb{R}$ and $(a_n)_{n \in \mathbb{N}} \in c_{00}$. Then there exists an $N \in \mathbb{N}$ such that for all $n > N$, $a_n = 0$, and so $\alpha a_n = \alpha 0 = 0$. Thus $(\alpha a_n)_{n \in \mathbb{N}} = \alpha \cdot (a_n)_{n \in \mathbb{N}}$ belongs to c_{00} .

Hence c_{00} is a subspace of c_0 .

(The examples

$$\left(\frac{1}{n}\right)_{n \in \mathbb{N}} \in c_0, \quad (1)_{n \in \mathbb{N}} \in c, \quad ((-1)^n)_{n \in \mathbb{N}} \in \ell^\infty, \quad (n)_{n \in \mathbb{N}} \in \mathbb{R}^\infty,$$

show that each of the inclusions are strict.)

4. IF: Let $y_1 = 0 = y_2$. Then we have:

- S1. The zero function $\mathbf{0} \in S(0, 0)$, since $\mathbf{0}$ is continuous on $[0, 1]$ and moreover, $\mathbf{0}(0) = 0$ and $\mathbf{0}(1) = 0$.
- S2. If $f, g \in S(0, 0)$, then $f + g \in S(0, 0)$. Indeed, as f, g are continuous on $[0, 1]$, so is $f + g$, and also $(f + g)(0) = f(0) + g(0) = 0 + 0 = 0$ and $(f + g)(1) = f(1) + g(1) = 0 + 0 = 0$.
- S3. Let $f \in S(0, 0)$ and $\alpha \in \mathbb{R}$. Then $\alpha \cdot f$ is continuous on $[0, 1]$, and furthermore, $(\alpha \cdot f)(0) = \alpha f(0) = \alpha 0 = 0$ and $(\alpha \cdot f)(1) = \alpha f(1) = \alpha 0 = 0$.

Hence $S(0, 0)$ is a subspace of $C[0, 1]$.

ONLY IF: Suppose that $S(y_1, y_2)$ is a subspace of $C[0, 1]$. If $f \in S(y_1, y_2)$, then $2 \cdot f \in S(y_1, y_2)$. Thus $(2 \cdot f)(0) = y_1$ and $(2 \cdot f)(1) = y_2$, and so $2y_1 = y_1$ and $2y_2 = y_2$. Consequently $y_1 = 0 = y_2$.

Solutions to the exercises on page 64

1. We have:

B1. If

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3,$$

then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x-y) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (y-z) \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \text{span} \left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \right).$$

Thus

$$\text{span} \left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \right) = \mathbb{R}^3.$$

B2. If α, β, γ are scalars such that

$$\alpha \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \gamma \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

then we obtain

$$\begin{aligned} \alpha + \beta + \gamma &= 0, \\ \beta + \gamma &= 0, \\ \gamma &= 0, \end{aligned}$$

and so it follows that $\alpha = \beta = \gamma = 0$. So B is linearly independent.

Since B1, B2 hold it follows that B is a basis.

2. Suppose that $B = \{v_1, \dots, v_n\}$ is a basis of the vector space V , and let $v \in V$. As $v \in V = \text{span}(B)$, it follows that there exist scalars $\alpha_1, \dots, \alpha_n$ such that $v = \alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n$. So v is a linear combination of vectors from V . In order to prove uniqueness, suppose that there exist scalars $\alpha'_1, \dots, \alpha'_n$ such that

$$\alpha_1 \cdot v_1 + \dots + \alpha_n \cdot v_n = v = \alpha'_1 \cdot v_1 + \dots + \alpha'_n \cdot v_n.$$

So we obtain $(\alpha_1 - \alpha'_1) \cdot v_1 + \dots + (\alpha_n - \alpha'_n) \cdot v_n = \mathbf{0}$, and by the independence of v_1, \dots, v_n , it then follows that $\alpha_1 - \alpha'_1 = \dots = \alpha_n - \alpha'_n = 0$, that is $\alpha_1 = \alpha'_1, \dots, \alpha_n = \alpha'_n$.

3. We prove this by contradiction. Suppose that $C[0, 1]$ is a finite dimensional vector space with dimension d , say.

Consider the d functions f_1, \dots, f_d defined by

$$\left. \begin{aligned} f_1(x) &= x \\ f_2(x) &= x^2 \\ f_3(x) &= x^3 \\ &\vdots \\ f_d(x) &= x^d \end{aligned} \right\} \text{ for all } x \in [0, 1]. \quad (2.35)$$

The functions f_1, \dots, f_d are all polynomials and so they are all continuous on $[0, 1]$, that is, $f_1, \dots, f_d \in C[0, 1]$. Now we prove the following.

CLAIM: The functions f_1, \dots, f_d given by (2.35) are linearly independent in $C[0, 1]$.

Proof Suppose that there exist scalars $\alpha_1, \dots, \alpha_d$, not all zeros, such that

$$\alpha_1 \cdot f_1 + \dots + \alpha_d \cdot f_d = \mathbf{0}.$$

Then for all $x \in [0, 1]$, $(\alpha_1 \cdot f_1 + \dots + \alpha_d \cdot f_d)(x) = \mathbf{0}(x)$, that is,

$$\text{for all } x \in [0, 1], \quad \alpha_1 x + \dots + \alpha_d x^d = 0.$$

Let $k \in \{1, 2, \dots, d\}$ be the smallest number such that $\alpha_k \neq 0$. Then we obtain

$$\text{for all } x \in [0, 1], \quad \alpha_k x^k + \alpha_{k+1} x^{k+1} + \dots + \alpha_d x^d = 0,$$

and so for all $x \in [0, 1]$, $x^k(\alpha_k + \alpha_{k+1}x + \alpha_{k+2}x^2 + \dots + \alpha_d x^{d-k}) = 0$. For all $n \in \mathbb{N}$, $0 < \frac{1}{n} < 1$, and so we have

$$\text{for all } n \in \mathbb{N}, \quad \left(\frac{1}{n}\right)^k \left(\alpha_k + \alpha_{k+1} \left(\frac{1}{n}\right) + \alpha_{k+2} \left(\frac{1}{n}\right)^2 + \dots + \alpha_d \left(\frac{1}{n}\right)^{d-k} \right) = 0,$$

that is,

$$\text{for all } n \in \mathbb{N}, \quad \alpha_k + \alpha_{k+1} \left(\frac{1}{n}\right) + \alpha_{k+2} \left(\frac{1}{n}\right)^2 + \dots + \alpha_d \left(\frac{1}{n}\right)^{d-k} = 0.$$

Hence

$$\begin{aligned} 0 &= \lim_{n \rightarrow \infty} 0 \\ &= \lim_{n \rightarrow \infty} \left(\alpha_k + \alpha_{k+1} \left(\frac{1}{n}\right) + \alpha_{k+2} \left(\frac{1}{n}\right)^2 + \dots + \alpha_d \left(\frac{1}{n}\right)^{d-k} \right) \\ &= \alpha_k + 0 + \dots + 0 \\ &= \alpha_k, \end{aligned}$$

a contradiction. So f_1, \dots, f_d are linearly independent in $C[0, 1]$. ■

Consequently, by Theorem 2.2.3, it follows that $\{f_1, \dots, f_d\}$ is a basis of $C[0, 1]$, and in particular, they span the whole space $C[0, 1]$.

The constant function $\mathbf{1} : [0, 1] \rightarrow \mathbb{R}$, defined by $\mathbf{1}(x) = 1$ for all $x \in [0, 1]$ clearly belongs to $C[0, 1]$, and so, from the above, there must exist scalars β_1, \dots, β_d such that

$$\mathbf{1} = \beta_1 \cdot f_1 + \dots + \beta_d \cdot f_d.$$

Thus for all $x \in [0, 1]$, $\mathbf{1}(x) = (\beta_1 \cdot f_1 + \dots + \beta_d \cdot f_d)(x)$, that is,

$$\text{for all } x \in [0, 1], \quad 1 = \beta_1 x + \dots + \beta_d x^d.$$

In particular, with $x = 0$, we obtain $1 = 0$, a contradiction. So $C[0, 1]$ is not a finite dimensional vector space.

4. (a) We have

B1. Let $(a_n)_{n \in \mathbb{N}} \in c_{00}$. Then there exists an $N \in \mathbb{N}$ such that for all $n > N$, $a_n = 0$. Clearly

$$(a_n)_{n \in \mathbb{N}} = a_1 \cdot e_1 + \dots + a_n \cdot e_n \in \text{span}(\{e_k \mid k \in \mathbb{N}\}).$$

So $c_{00} \subset \text{span}(B)$.

On the other hand, suppose that $k_1, \dots, k_m \in \mathbb{N}$, and $\alpha_1, \dots, \alpha_m$ be scalars. If $N := \max\{k_1, \dots, k_m\}$, then for all $n > N$, the n th term of the sequence

$$\alpha_1 \cdot e_{k_1} + \dots + \alpha_m \cdot e_{k_m}$$

is $\alpha_1 0 + \dots + \alpha_m 0 = 0$. Thus $\alpha_1 \cdot e_{k_1} + \dots + \alpha_m \cdot e_{k_m} \in c_{00}$. Consequently $\text{span}(B) \subset c_{00}$.

So $\text{span}(B) = c_{00}$.

- B2. Suppose that $k_1 < \dots < k_m$ are natural numbers, and $\alpha_1, \dots, \alpha_m$ be scalars, not all zeros, such that

$$\alpha_1 \cdot e_{k_1} + \dots + \alpha_m \cdot e_{k_m} = (0)_{n \in \mathbb{N}}.$$

Let $l \in \{1, \dots, m\}$ be the largest number such that $\alpha_l \neq 0$. Then equating the k_l th term on both sides, we obtain that $\alpha_l = 0$, a contradiction. So B is linearly independent.

- (b) Indeed, if B were a basis for \mathbb{R}^∞ , then the sequence $(1)_{n \in \mathbb{N}}$ would be a linear combination of elements from B . But as seen above, *any* linear combination of elements from B is in c_{00} , that is, it is a sequence that is eventually zero. Clearly $(1)_{n \in \mathbb{N}} \notin c_{00}$, and so it is not a linear combination of elements from c_{00} .

Solutions to the exercises on page 68

1. (a) We have

$$\begin{aligned} \ker(T_A) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid \begin{bmatrix} x+y \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x = y = 0 \right\} \\ &= \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}. \end{aligned}$$

We depict this set in the (x, y) -plane in Figure 2.10.

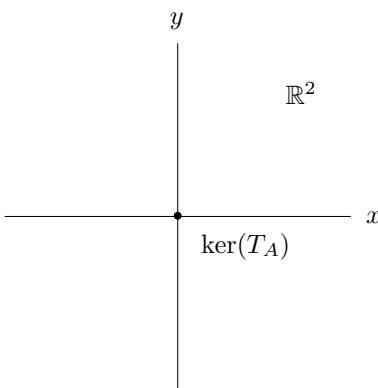
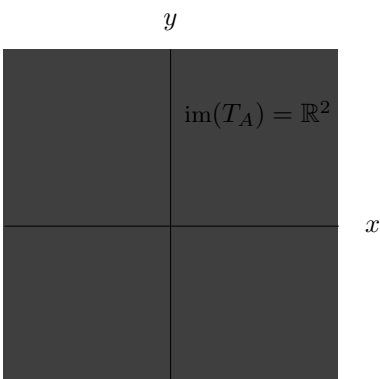


Figure 2.10: Kernel of T_A .

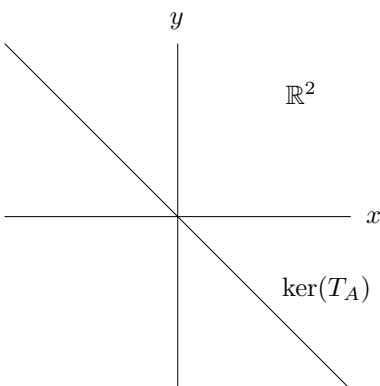
Clearly $\text{im}(T_A) \subset \mathbb{R}^2$. On the other hand, if $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$, then $\begin{bmatrix} x \\ y \end{bmatrix} = T_A \begin{bmatrix} x-y \\ y \end{bmatrix}$, and so $\mathbb{R}^2 \subset \text{im}(T_A)$. Hence $\text{im}(T_A) = \mathbb{R}^2$. We depict this set in the (x, y) -plane in Figure 2.11.

Figure 2.11: Image of T_A .

(b) We have

$$\begin{aligned} \ker(T_A) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid \begin{bmatrix} x+y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x+y=0 \right\}. \end{aligned}$$

We depict this set in the (x, y) -plane in Figure 2.12.

Figure 2.12: Kernel of T_A .

We have

$$\text{im}(T_A) \subset S := \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \in \mathbb{R}^2 \mid x \in \mathbb{R} \right\},$$

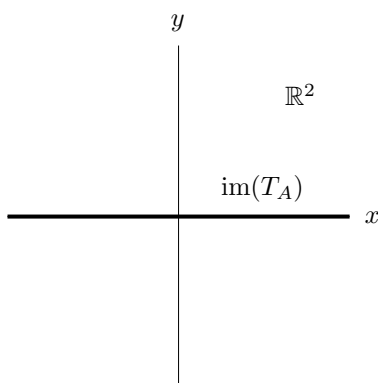
since

$$\text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2, \quad T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ 0 \end{bmatrix} \in S.$$

Conversely, since

$$\begin{bmatrix} x \\ 0 \end{bmatrix} = T_A \begin{bmatrix} x \\ 0 \end{bmatrix},$$

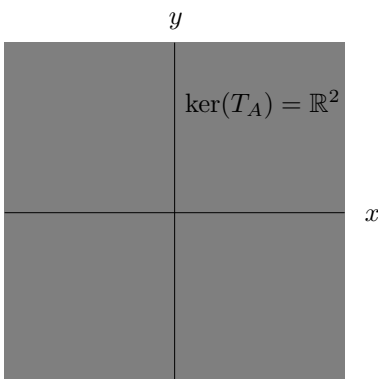
it follows that $S \subset \text{im}(T_A)$. Hence $\text{im}(T_A) = S$. We depict this set in the (x, y) -plane in Figure 2.13.

Figure 2.13: Image of T_A .

(c) We have

$$\begin{aligned} \ker(T_A) &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid T_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \mid x \in \mathbb{R} \text{ and } y \in \mathbb{R} \right\} \\ &= \mathbb{R}^2. \end{aligned}$$

We depict this set in the (x, y) -plane in Figure 2.14.

Figure 2.14: Kernel of T_A .

Clearly

$$\text{im}(T_A) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

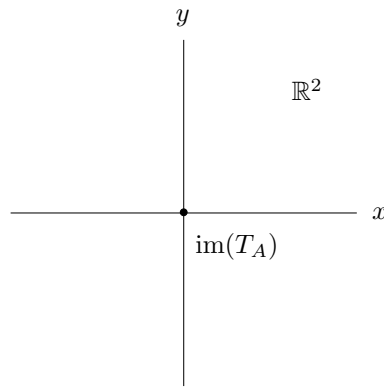
We depict this set in the (x, y) -plane in Figure 2.15.

2. L1 is not satisfied, since

$$T \left(\begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = T \left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} \frac{(-1)^2}{1} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

while

$$T \begin{bmatrix} -1 \\ 0 \end{bmatrix} + T \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Figure 2.15: Image of T_A .

If $\alpha \in \mathbb{R} \setminus \{0\}$ and $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, then we have

$$\begin{aligned}
 T\left(\alpha \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= T\begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix} \\
 &= \begin{cases} \begin{bmatrix} \frac{\alpha^2 x_1^2}{\alpha x_2} \\ \alpha x_2 \end{bmatrix} & \text{if } x_1 x_2 \neq 0 \\ \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \end{bmatrix} & \text{if } x_1 x_2 = 0 \end{cases} \\
 &= \begin{cases} \alpha \cdot \begin{bmatrix} \frac{x_1^2}{x_2} \\ x_2 \end{bmatrix} & \text{if } x_1 x_2 \neq 0 \\ \alpha \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & \text{if } x_1 x_2 = 0 \end{cases} \\
 &= \alpha \cdot \left(T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right).
 \end{aligned}$$

If $\alpha = 0$, and $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$, we have

$$T\left(\alpha \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = T\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \cdot \left(T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right).$$

Thus for all $\alpha \in \mathbb{R}$, and all $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$,

$$T\left(\alpha \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \alpha \cdot \left(T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right),$$

and so L2 holds.

3. (a) We verify that L1 and L2 hold:

L1. For all $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ in c , we have

$$\begin{aligned} T((a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}}) &= T((a_n + b_n)_{n \in \mathbb{N}}) \\ &= \lim_{n \rightarrow \infty} (a_n + b_n) \\ &= \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n \quad (\text{Theorem 1.2.4}) \\ &= T((a_n)_{n \in \mathbb{N}}) + T((b_n)_{n \in \mathbb{N}}). \end{aligned}$$

L2. For all $\alpha \in \mathbb{R}$, and all $(a_n)_{n \in \mathbb{N}} \in c$, we have

$$\begin{aligned} T(\alpha \cdot (a_n)_{n \in \mathbb{N}}) &= T((\alpha a_n)_{n \in \mathbb{N}}) \\ &= \lim_{n \rightarrow \infty} \alpha a_n \\ &= \alpha \lim_{n \rightarrow \infty} a_n \quad (\text{Theorem 1.2.4}) \\ &= \alpha \cdot T((a_n)_{n \in \mathbb{N}}). \end{aligned}$$

So T is a linear transformation from c to \mathbb{R} .

(b) We have

$$\ker(T) = \{(a_n)_{n \in \mathbb{N}} \in c \mid T((a_n)_{n \in \mathbb{N}}) = 0\} = \{(a_n)_{n \in \mathbb{N}} \in c \mid \lim_{n \rightarrow \infty} a_n = 0\},$$

and so $\ker(T)$ is the set comprising all convergent sequences with limit 0, that is, the subspace c_0 .

(c) Clearly, $\text{im}(T) \subset \mathbb{R}$. Moreover, if $L \in \mathbb{R}$, then the constant sequence L, L, L, \dots is convergent with limit L . Thus $T((L)_{n \in \mathbb{N}}) = L$, and so $L \in \text{im}(T)$. So $\mathbb{R} \subset \text{im}(T)$. Consequently, $\text{im}(T) = \mathbb{R}$.

Bibliography

- [1] N.L. Biggs. *Discrete Mathematics*. Clarendon Press, 2002.
(Chapter 20 on Groups.)
- [2] K.G. Binmore. *Mathematical Analysis: A Straightforward Approach*. Cambridge University Press, 1982.
- [3] V. Bryant. *Yet Another Introduction to Analysis*. Cambridge University Press, 1990.
- [4] P.R. Halmos. *Finite Dimensional Vector Spaces*. Springer, 1996.
- [5] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 3rd Edition, 1976.

Index

- k th power of a function, 31
- abelian group, 42
- absolute value, 5
- absolute value of a function, 31
- Archimedean property, 4
- associativity, 40
- basis of a vector space, 63
- Bolzano-Weierstrass theorem, 25
- bounded above, set, 1
- bounded below, set, 1
- bounded function, 33, 34
- bounded sequence, 14
- bounded set, 1
- Cauchy sequence, 13
- center of a group, 47
- commutativity, 42
- continuity of a function at a point, 27
- continuous function, 27
- convergent sequence, 10
- counting formula, 53
- cyclic, 48
- decreasing sequence, 15
- dimension of a vector space, 64
- distance, 5
- distributive laws, 55
- divergent sequence, 11
- eventually zero sequence, 62
- extreme value theorem, 33
- Fermat's little theorem, 54
- finite dimensional vector space, 64
- finite group, 43
- finite order, group element with, 47
- fractional part of a real number, 26
- general linear group, 41
- generator of a group, 48
- greatest integer part of a real number, 26
- greatest upper bound, 2
- group, 40
- group axioms, 40
- group table, 43
- identity element in a group, 40
- iff, 7
- image of a homomorphism, 50
- image of a linear transformation, 66
- increasing sequence, 15
- induced law of composition, 45
- infimum, 2
- infinite dimensional of a vector space, 64
- infinite group, 43
- infinite order, group element with, 47
- intermediate value theorem, 34
- interval, 5
- inverse of an element in a group, 40
- isomorphism, 51
- kernel of a homomorphism, 50
- kernel of a linear transformation, 66
- Lagrange's theorem, 53
- law of composition on a set, 39
- least upper bound, 2
- least upper bound property, 3
- linear combination of vectors, 60
- linear transformation, 65
- linearly dependent vectors, 61
- linearly independent vectors, 61
- linearly dependent set, 61
- linearly independent set, 61
- lower bound, 1
- maximum, 3
- minimum, 3
- monotone sequence, 15
- normal subgroup, 51
- order of a group, 43
- order of an element in a group, 47
- partition, 52
- periodic function, 34
- polynomial function, 32
- positive definiteness, 6
- product of functions, 31

proper subspace, 59

rational function, 32

right cosets, 53

Sandwich theorem, 21

scalar multiplication, 55

seaview property, 25

sequence, 9

span of a subset of a vector space, 60

special linear group, 50

subgroup, 45

subsequence, 23

subspace of a vector space, 59

sum of functions, 31

supremum, 2

symmetric group, 44

symmetric matrix, 45, 59

symmetry, 6

triangle inequality, 6

upper bound, 1

upper triangular matrix, 46, 59

vector, 55

vector addition, 55

vector space, 55

zero vector, 55