Propositions of solutions for *Analysis I* by Terence Tao

Frédéric Santos

February 22, 2022

Contents

Remark. The numbering of the Exercises follows the fourth edition of *Analysis I*.

1. Introduction

No exercises in this chapter.

2. Starting at the beginning: the natural numbers

Exercise 2.2.1. — *Prove that the addition is associative, i.e. that for any natural numbers a, b, c, we have $(a+b)+c = a+(b+c)$.*

Let's use induction on c while keeping a and b fixed.

- Base case for $c = 0$: let's prove that $(a + b) + 0 = a + (b + 0)$. The left hand side is equal to $(a + b)$ according to Lemma 2.2.3. For the right hand side, if we apply the same lemma to the $(b + 0)$ part, we get $a + (b + 0) = a + b$. Both sides are equal to $a + b$, and the base case is thus done.

- Now let's suppose inductively that $(a + b) + c = a + (b + c)$: we have to prove that $(a + b) + c{+\!+} = a + (b + c{+\!+})$. Using Lemma 2.2.3 on the right hand side leads to $a + (b + c){+\!+}$. Now consider the left hand side. Using still the same lemma, we get $(a + b) + c{+\!+} = ((a + b) + c){+\!+}$. By the inductive hypothesis, this is also equal to $(a + (b + c)){+\!+}$. And, using the lemma 2.2.3 again, this also leads to $a + b + c{+\!+}$. Therefore, both sides are equal to $a + b + c{+\!+}$, and we have closed the induction.

Exercise 2.2.2. — *Let a be a positive number. Prove that there exists exactly one natural number b such that $b{+\!+} = a$.*

Let's use induction on a.

- Base case for $a = 1$: we know that $b = 0$ matches this property, since $0{+\!+} = 1$ by Definition 2.1.3. Furthermore, there is only one solution. Suppose that is another natural number b such that $b{+\!+} = 1$. Then, we would have $b{+\!+} = 0{+\!+}$, which would imply $b = 0$ by Axiom 2.4. The base case is demonstrated.

- Let's suppose inductively that there is exactly one natural number b such that $b{+\!+} = a$. We have to prove that there is exactly one natural number b' such that $b'{+\!+} = a{+\!+}$. By the induction hypothesis, and taking $b' = b{+\!+}$, we have $b'{+\!+} = (b{+\!+}){+\!+} = a{+\!+}$. So there exists a solution, with $b' = b{+\!+} = a$. Uniqueness is given by Axiom 2.4.: if $b'{+\!+} = a{+\!+}$, then we necessarily have $b' = a$.

Exercise 2.2.3. — *Let a, b, c be natural numbers. Prove the following properties of order for natural numbers:*

(a) Reflexivity: $a \geqslant a$. This is true since $a = 0 + a$ by Definition 2.2.1. By commutativity of addition, we can also write $a = a + 0$. So there is indeed a natural number n (with $n = 0$) such that $a = a + n$, i.e. $a \geqslant a$.

(b) Transitivity: if $a \geqslant b$ and $b \geqslant c$, then $a \geqslant c$. From the part $a \geqslant b$, there exists a natural number n such that $a = b+n$ according to Definition 2.2.11. A similar consideration for the part $b \geqslant c$ leads to $b = c+m$, m being a natural number. Combining together those two equalities, we can write $a = b+n = (c+m)+n = c+(m+n)$ by associativity (see Exercise 2.2.1). Then, $n+m$ being a natural number[1], the transitivity is demonstrated.

[1] This is a trivial induction from the definition of addition.

(c) Anti-symmetry: if $a \geqslant b$ and $b \geqslant a$, then $a = b$. From the part $a \geqslant b$, there exists a natural number n such that $a = b + n$. Similarly, there exists a natural number m such that $b = a + m$. Combining those two equalities leads to $a = b + n = (a + m) + n = a + (m + n)$. By cancellation law (Proposition 2.2.6), we can conclude that $0 = m + n$. According to Corollary 2.2.9, this leads to $m = n = 0$. Therefore, both m and n are null, meaning that $a = b + 0 = b$.

(d) Preservation of order: $a \geqslant b$ iff $a+c \geqslant b+c$. First, let's prove that $a+c \geqslant b+c \implies a \geqslant b$. If $a+c \geqslant b+c$, there exists a natural number n such that $a+c = b+c+n$. By cancellation law (Proposition 2.2.6)[2], we conclude that $a = b+n$, i.e. $a \geqslant b$, thus demonstrating the first implication. Conversely, let's suppose that $a \geqslant b$. There exists a natural number m such that $a = b + m$. Therefore, $a + c = b + m + c$ for any natural number c. Still by associativity and commutativity, we can rewrite this as $a + c = (b + c) + m$, i.e. $a + c \geqslant b + c$.

(e) $a < b$ iff $a{+}{+} \leqslant b$. First, let's prove that $a{+}{+} \leqslant b \implies a < b$. By definition of ordering, there exists a natural number n such that $b = (a{+}{+}) + n$. By definition of addition, we can re-write: $b = (a{+}{+} + n){+}{+}$. Then, by commutativity and yet again by definition of addition, $b = (n + a{+}{+}){+}{+} = (n{+}{+}) + (a{+}{+})$. Thus, there exists a natural number $n{+}{+}$ such that $b = n{+}{+} + a$, which means that $b \geqslant a$. But we still have to prove that $a \neq b$. Let's suppose that $a = b$: in this case, by cancellation law, we would have $n{+}{+} = 0$, which is impossible according to Axiom 2.3 (0 is not the successor of any natural number). Thus, $a \neq b$ et $b \geqslant a$: we have showed that $a < b$.

Conversely, let's prove that $a < b \implies a{+}{+} \leqslant b$. Starting from that strict inequality, there exists a *positive*[3] natural number n such that $b = a + n$. By Lemma 2.2.10, since n is positive, it has one unique antecessor m, so that n can be written $m{+}{+}$. Thus, $b = a + (m{+}{+}) = (a + m){+}{+} = (m + a){+}{+} = m + (a{+}{+}) = (a{+}{+}) + m$. And, m being a natural number, this corresponds to the statement $b \geqslant a$.

(f) $a < b$ iff $b = a + d$ for some positive number d. First, let's prove the first implication, $a < b \implies b = a + d$ with $d \neq 0$. Since $a < b$, we have in particular $a \leqslant b$, and there exists a natural number d such that $b = a + d$. For the sake of contradiction, let's suppose that $d = 0$. We would have $b = a$, which would contradict the condition $a \neq b$ of the strict inequality. Thus, d is a positive number, which demonstrates the left-to-right implication.

Conversely, let's suppose that $b = a + d$, with $d \neq 0$. This expression gives immediately $a \leqslant b$. But if $a = b$, by cancellation law, this would lead to $0 = d$, a contradiction with the fact that d is a positive number. Thus, $a \neq b$ and $a \leqslant b$, which demonstrates $a < b$.

Exercise 2.2.4. — *Demonstrate three lemmas used to prove the trichotomy of order for natural numbers.*

(a) Show that we have $0 \leqslant b$ for any natural number b. This is obvious since, by definition of addition, $0 + b = b$ for any natural number b. This is precisely the definition of $0 \leqslant b$.

[2] And also associativity and commutativity that we do not detail explicitly here.

[3] We make use here of the statement *(f)* demonstrated below. There is no circularity here, since proving *(f)* will not make use of *(e)*.

(b) Show that if $a > b$, then $a{+}{+} > b$. If $a > b$, then $a = b + d$, d being a positive natural number. Let's recall that $a{+}{+} = a + 1$. Thus, $a{+}{+} = a + 1 = b + d + 1 = b + (d + 1)$ by associativity of addition. Furthermore, $d+1$ is a positive natural number (by Proposition 2.2.8). Thus, $a{+}{+} > b$.

(c) Show that if $a = b$, then $a{+}{+} > b$. Once again, let's use the fact that $a{+}{+} = a + 1$. Thus, $a{+}{+} = a + 1 = b + 1$, and 1 is a positive natural number. This is the definition of $a{+}{+} > b$.

Exercise 2.2.5. — *Prove the strong principle of induction, formulated as follows: Let m_0 be a natural number, and let $P(m)$ be a property pertaining to an arbitrary natural number m. Suppose that for each $m \geqslant m_0$, we have the following implication: if $P(m')$ is true for all natural numbers $m_0 \leqslant m' < m$, then $P(m)$ is also true. (In particular, this means that $P(m_0)$ is true, since in this case the hypothesis is vacuous.) Then we can conclude that $P(m)$ is true for all natural numbers $m \geqslant m_0$.*

First let's introduce a small lemma (similar to Proposition 2.2.12*(e)*).

Lemma. *For any natural number a and b, $a < b{+}{+}$ iff $a \leqslant b$.*

Proof. If $a < b{+}{+}$, then $b{+}{+} = a + n$ for a given positive natural n. By Lemma 2.2.10, there exists one natural number m such as $n = m{+}{+}$. Thus $b{+}{+} = a + m{+}{+}$, which can be rewritten $b{+}{+} = (a + m){+}{+}$ by Lemma 2.2.3[4]. By Axiom 2.4., this is equivalent to $b = a + n$, which can also be written $a \leqslant b$.

Conversely, if $a \leqslant b$, there exists a natural number m such as $b = a + m$. Thus, $b{+}{+} = (a + m){+}{+} = a + (m{+}{+})$ by Definition of addition (2.2.1). And, $m{+}{+}$ being a positive number, this means that $b > a$ according to Proposition 2.2.12*(f)*. □

Now we can prove the main proposition. Let $Q(n)$ be the property "$P(m)$ is true for all m such that $m_0 \leqslant m < n$". Let's induct on n.

- (Although this is not necessary,) we could consider two types of base cases. If $n < m_0$, $Q(n)$ is the proposition "$P(m)$ is true for all m such that $m_0 \leqslant m < n$", but there is no such natural number m. Thus, $Q(n)$ is vacuously true. If $n = m_0$, $P(m_0)$ is true by hypothesis, thus $Q(m_0)$ is also true.

- Now let's suppose inductively that $Q(n)$ is true, and show that $Q(n{+}{+})$ is also true. If $Q(n)$ is true, $P(m)$ is true for all m such that $m_0 \leqslant m < n$. By hypothesis, this implies that $P(n)$ is true. Thus, $P(m)$ is true for any natural number m such that $m_0 \leqslant m \leqslant n$, i.e. such that $m_0 \leqslant m < n{+}{+}$ according to the lemma introduced above. This is precisely $Q(n{+}{+})$, and this closes the induction.

Thus, $Q(n)$ is true for all natural numbers n, which means in particular that $P(m)$ is true for any natural number $m \geqslant m_0$. This demonstrates the principle of strong induction.

Exercise 2.2.6. — *Let n be a natural number, and let $P(m)$ be a property pertaining to the natural numbers such that whenever $P(m{+}{+})$ is true, then $P(m)$ is true. Suppose that $P(n)$ is also true. Prove that $P(m)$ is true for all natural numbers $m \leqslant n$; this is known as the* principle of backwards induction.

[4] We could also rewrite $b + 1 = a + m + 1$ and then use the cancellation law.

Terence Tao suggests to use induction on n. So let $Q(n)$ be the following property: "if $P(n)$ is true, then $P(m)$ is true for all $m \leqslant n$. The goal is to prove $Q(n)$ for all natural numbers n.

- Base case $n = 0$: here, $Q(n)$ means that if $P(0)$ is true, then $P(m)$ is true for any $m \leqslant 0$. By Definition 2.2.11, if $m \leqslant 0$, there exists a natural number d such that $0 = m + d$. But, by Corollary 2.2.9, this implies that both $m = 0$ and $d = 0$. Thus, the only number m such that $m \leqslant 0$ is 0 itself. Therefore, $Q(0)$ is simply the tautology "if $P(0)$ is true, then $P(0)$ is true"— a statement that we can safely accept. The base case is the, demonstrated.

- Let's suppose inductively that $Q(n)$ is true: we must show that $Q(n{++})$ is also true. If $P(n{++})$ is true, then by definition of P, $P(n)$ is also true. Then, by induction hypothesis, $P(m)$ is true for all $m \leqslant n$. We have showed that $P(n{++})$ implies $P(m)$ for all $m \leqslant n{++}$[5], which is precisely $Q(n{++})$. This closes the induction.

Exercise 2.3.1. — *Show that multiplication is commutative, i.e., if n and m are natural numbers, show that $n \times m = m \times n$.*

We will use an induction of n while keeping m fixed. However, this is not a trivial result, and even the base case is not straightforward. We will first introduce some lemmas.

Lemma. *For any natural number n, $n \times 0 = 0$.*

Proof. Let's induct on n. For the base case $n = 0$, we know by Definition 2.3.1 of multiplication that $0 \times 0 = 0$, since $0 \times m = 0$ for any natural number m.

Now let's suppose that $n \times 0 = 0$. Thus, $n{++} \times 0 = (n \times 0) + 0$ by Definition 2.3.1. But by induction hypothesis, $n \times 0 = 0$, so that $n{++} \times 0 = 0 + 0 = 0$. This closes the induction. ☐

Lemma. *For all natural numbers m and n, we have $m \times n{++} = (m \times n) + m$.*

Proof. Let's induct on m. The base case $m = 0$ is easy to prove: $0 \times n{++} = 0$ by Definition 2.3.1 of multiplication, and $(0 \times n) + 0 = 0$.

Now suppose inductively that $m \times n{++} = (m \times n) + m$, and we must show that

$$m{++} \times n{++} = (m{++} \times n) + m{++} \tag{2.1}$$

We begin by the left hand side: by Definition 2.3.1, $m{++} \times n{++} = (m \times n{++}) + n{++}$. By induction hypothesis, this is equal to $(m \times n) + m + n{++}$.

Then, apply the definition of multiplication to the right hand side: $(m{++} \times n) + m{+}{+} = (m \times n) + n + m{++}$. The Lemma 2.2.3 and the commutativity of addition leads to $(m \times n) + n + m{++} = (m \times n) + (n + m){++} = (m \times n) + (m + n){++} = (m \times n) + m + n{++}$, which is equal to the left hand side.

Thus, both sides of equation (2.1) are equal, and we can close the induction. ☐

Now it is easier to prove the main result ($n \times m = m \times n$), by an induction on n.

[5] Actually, we use here yet another lemma, similar to the one introduced for the previous exercise. We use the fact that $m \leqslant n{++}$ is equivalent to $m = n{++}$ or $m \leqslant n$, which is easy to prove, but is not part of the "standard" results presented in the textbook.

- Base case $n = 0$: we already know by Definition 2.3.1 that $0 \times m = 0$. The first lemma introduced in this exercise also provides $m \times 0 = 0$. Thus, the base case is proved, since $0 \times m = m \times 0\ (= 0)$.

- Now we suppose inductively that $n \times m = m \times n$, and we must prove that:

$$n\text{++} \times m = m \times n\text{++} \tag{2.2}$$

  By Definition 2.3.1 of multiplication, the left hand side is equal to $(n \times m) + m$.

  Using the lemma introduced above, the right hand side is equal to $(m \times n) + m$. By induction hypothesis, this is also equal to $(n \times m) + m$, which closes the induction.

EXERCISE 2.3.2. — *Show that positive natural numbers have no zero divisors, i.e. that $nm = 0$ iff $n = 0$ or $m = 0$. In particular, if n and m are both positive, then nm is also positive.*

We will prove the second statement first. Suppose, for the sake of contradiction, that $nm = 0$ and that both n and m are positive numbers. Since they are positive, by Lemma 2.2.10, there exists two (unique) natural numbers a and b such that $n = a\text{++}$ and $m = b\text{++}$. Thus, the hypothesis $nm = 0$ can also be written $(a\text{++}) \times (b\text{++}) = 0$. But, by Definition 2.3.1 of multiplication, $(a\text{++}) \times (b\text{++}) = (a \times b\text{++}) + b\text{++}$. Thus, we should have $(a \times b\text{++}) + b\text{++} = 0$. By Corollary 2.2.9, this implies that both $(a \times b\text{++}) = 0$ and $b\text{++} = 0$, which is impossible since zero is the successor of no natural number (Axiom 2.3).

Thus, we have proved that if n and m are both positive, then nm is also positive. The main statement can now be proved more easily.

- The right-to-left implication is straightforward: if $n = 0$, then by Definition of multiplication, $n \times m = 0 \times m = 0$. Since multiplication is commutative, we have the same result if $m = 0$.

- The left-to-right implication is exactly the contrapositive of the statement we have just proved above.

EXERCISE 2.3.3. — *Show that multiplication is associative, i.e., for any natural numbers a, b, c, we have $(a \times b) \times c = a \times (b \times c)$.*

We will induct on c while keeping a and b fixed.

- Base case: for $c = 0$, we must prove that $(a \times b) \times 0 = a \times (b \times 0)$. The left hand side is equal to 0 by definition (and commutativity) of multiplication[6]. The right hand side is equal to $a0$, which is also 0. Both sides are null, and the base case is proved.

- Suppose inductively that $(a \times b) \times c = a \times (b \times c)$, and let's prove that $(a \times b) \times c\text{++} = a \times (b \times c\text{++})$. By definition (and commutativity) of multiplication, the left hand side is equal to $(a \times b) \times c + (a \times b)$. The right hand side is equal to $a \times (b \times c + b)$, and by distributive law (i.e., Proposition 2.3.4), this is also $a \times (b \times c) + a \times b$. But then, by inductive hypothesis, this can be rewritten $(a \times b) \times c + a \times b$, which is equal to the left hand side. The induction is closed.

[6] Actually, we use the second lemma introduced for the resolution of Exercise 2.3.1.

Exercise 2.3.4. — *Prove the identity $(a+b)^2 = a^2 + 2ab + b^2$ for all natural numbers a, b.*

By distribution law (i.e., Proposition 2.3.4) and commutativity of multiplication, we have:

$$\begin{aligned}(a+b)^2 &= (a+b)(a+b) = (a+b)a + (a+b)b \\ &= a \times a + b \times a + a \times b + b \times b \\ &= a^2 + a \times b + a \times b + b^2 \\ &= a^2 + 2ab + b^2\end{aligned}$$

(For the last step, we recall that, by Definition 2.3.1, $2 \times m = m + m$ for any natural number m.)

Exercise 2.3.5. — *Euclidean algorithm. Let n be a natural number, and let q be a positive number. Prove that there exists natural numbers m, r such that $0 \leqslant r < q$ and $n = mq + r$.*

We will induct on n while remaining q fixed.

- Base case: if $n = 0$, there exists an obvious solution, namely $m = 0$ and $r = 0$.
- Suppose inductively that there exists m, r such that $n = mq + r$ with $0 \leqslant r < q$, and let's prove that there exists m', r' such that $n + 1 = m'q + r'$, with $0 \leqslant r' < q$.

  By the induction hypothesis, we have $n+1 = mq+r+1$. Since $r < q$, we have $r+1 \leqslant q$ (this is Proposition 2.2.12). Thus, we have two cases here:

  1. If $r + 1 < q$, then $n + 1 = mq + (r + 1)$, with $0 \leqslant r + 1 < q$, so that choosing $m' = m$ and $r' = r + 1$ is convenient.
  2. If $r + 1 = q$, then $n + 1 = mq + q = (m + 1)q$ according to the distributive law (Proposition 2.3.4). Thus, choosing $m' = m + 1$ and $r' = 0$ is convenient.

This closes the induction.

3. Set theory

EXERCISE 3.1.2. — *Using only Definition 3.1.4, Axiom 3.1, Axiom 3.2, and Axiom 3.3, prove that the sets $\varnothing$, $\{\varnothing\}$, $\{\{\varnothing\}\}$, and $\{\varnothing, \{\varnothing\}\}$ are all distinct (i.e., no two of them are equal to each other).*

As a general reminder, we recall that sets are objects (Axiom 3.1) and the empty set $\varnothing$ is such that no object is an element of $\varnothing$, thus $\varnothing \notin \varnothing$.

1. First let's show that $\varnothing$ is different from all other sets. $\varnothing$ is an element of $\{\varnothing\}$ and $\{\varnothing, \{\varnothing\}\}$, and $\{\varnothing\}$ is an element of $\{\{\varnothing\}\}$. But none of those two objects are elements of $\varnothing$ (by Axiom 3.2), thus $\varnothing$ is different from all three other sets.

2. Then let's show that $\{\varnothing\} \neq \{\{\varnothing\}\}$. By Axiom 3.3, the singleton $\{\varnothing\}$ is such that $x \in \{\varnothing\} \iff x = \varnothing$. Similarly, the singleton $\{\{\varnothing\}\}$ is such that $x \in \{\{\varnothing\}\} \iff x = \{\varnothing\}$. But we already know that $\varnothing \neq \{\varnothing\}$ so there exists an object, $\varnothing$, which is a element of $\{\varnothing\}$ but not an element of $\{\{\varnothing\}\}$. Those sets are not equal.

3. Now let's show that $\{\varnothing\} \neq \{\varnothing, \{\varnothing\}\}$. By Axiom 3.3, the pair $\{\varnothing, \{\varnothing\}\}$ is such that x is an element of this set iff $x = \varnothing$ or $x = \{\varnothing\}$. Thus, $\{\varnothing\}$ is an element of $\{\varnothing, \{\varnothing\}\}$, but is not an element of $\{\varnothing\}$ (if it was, we should have $\varnothing = \{\varnothing\}$, which would be a contradiction with the first point of this proof). Those two sets are thus different.

4. Finally, we also have $\{\{\varnothing\}\} \neq \{\varnothing, \{\varnothing\}\}$. Indeed, we have $\varnothing \in \{\varnothing, \{\varnothing\}\}$ by Axiom 3.3. However, $\varnothing \in \{\{\varnothing\}\} \iff \varnothing = \{\varnothing\}$ by definition of a singleton, and we know this latest statement is false by the first point of this proof. Those two sets are also different.

EXERCISE 3.1.3. — *Prove the remaining claims in Lemma 3.1.13.*

Those claims are the following:

1. $\{a, b\} = \{a\} \cup \{b\}$. By Axiom 3.3, the pair $\{a, b\}$ is such that $x \in \{a, b\} \iff x = a$ or $x = b$. Let's consider three cases:

   - if $x = a$, $x \in \{a\}$ by Axiom 3.3, thus $x \in \{a\} \cup \{b\}$ by Axiom 3.4
   - if $x = b$, $x \in \{b\}$ by Axiom 3.3, thus $x \in \{a\} \cup \{b\}$ by Axiom 3.4
   - if $x \neq a$ and $x \neq b$, $x \notin \{a\}$ and $x \notin \{b\}$ by Axiom 3.3, so that $x \notin \{a\} \cup \{b\}$

   Thus, $\{a, b\}$ and $\{a\} \cup \{b\}$ have the same elements, and are equal.

2. $A \cup B = B \cup A$ for all sets A and B. Indeed, $x \in A \cup B \iff x \in A$ or $x \in B$. If $x \in A$, then $x \in B \cup A$ by Axiom 3.4. A similar argument holds if $x \in B$. Thus, in both cases, $x \in B \cup A$. We can show in a similar fashion that any element of $B \cup A$ is in $A \cup B$.

3. $A \cup \varnothing = \varnothing \cup A = A$. Since we've just showed that union is commutative, proving $A \cup \varnothing = A$ is sufficient. If $x \in A$, then $x \in A \cup \varnothing$. The converse is also true: if $x \in A \cup \varnothing$, then $x \in A$ or $x \in \varnothing$. But there is no element in $\varnothing$, so that we have necessarily $x \in A$. Thus, $A \cup \varnothing$ and A have the same elements: they are equal.

Exercise 3.1.4. — *Prove the remaining claims from Proposition 3.1.18.*

Let A, B, C be sets. Those claims are the following:

1. If $A \subseteq B$ and $B \subseteq A$, then $B = A$. According to Definition 3.1.4, two sets A and B are equal iff every element of A is an element of B, and vice versa. This is precisely the present claim.

2. If $A \subsetneq B$ and $B \subsetneq C$, then $A \subsetneq C$. Let x be an element of A. Since $A \subsetneq B$, x is also an element of B. And since $B \subsetneq C$, x is also an element of C. This holds for any x in A, and thus it demonstrates that $A \subset C$. Furthermore, since $A \subsetneq B$, there exists an element $y \in B$ which is not an element of A. As $B \subsetneq C$, y is also an element of C. Thus we have y, an element of C which is not in A. Combined to the previous result $A \subset C$, this demonstrates $A \subsetneq C$.

Exercise 3.1.5. — *Let A, B be sets. Show that the three statements $A \subseteq B$, $A \cup B = B$ and $A \cap B = A$ are logically equivalent (i.e., any one of them implies the other two).*

1. First, we prove that $A \subseteq B \implies A \cup B = B$. The first inclusion $B \subseteq A \cup B$ is trivial, since any element of a set B is always either in A or B. For the converse inclusion, let x be an element of $A \cup B$, and let's prove that $x \in B$. By Axiom 3.4, we have $x \in A$ or $x \in B$. If $x \in B$, the result holds. If $x \in A$, then we also have $x \in B$ since $A \subseteq B$. Thus, any element of $A \cup B$ is an element of B, which demonstrates the equality $A \cup B = B$.

2. Then, we prove that $A \cup B = B \implies A \cap B = A$. The first inclusion is trivial: if $x \in A \cap B$, then we always have $x \in A$. Now let's prove the converse inclusion: let x be an element of A; we must show that $x \in A \cap B$. If $x \in A$, then $x \in A \cup B$. But, by hypothesis, $A \cup B = B$, thus $x \in B$. So, $x \in A$ and $x \in B$, i.e. $x \in A \cap B$. This demonstrates the implication.

3. Finally, we prove that $A \cap B = A \implies A \subseteq B$. Let $x \in A$. Since $A \cap B = A$, we have $x \in A \cap B$. It follows that $x \in B$. We have proved that any element $x \in A$ is also an element of B, i.e. $A \subseteq B$.

Exercise 3.1.8. — *Let A, B be sets. Prove the absorption laws $A \cap (A \cup B) = A$ and $A \cup (A \cap B) = A$.*

1. The first inclusion $A \cap (A \cup B) \subseteq A$ is trivial: if $x \in A \cap (A \cup B)$ then in particular $x \in A$ by Definition 3.1.23 of an intersection[7]. Thus, we have $A \cap (A \cup B) \subseteq A$.

   For the converse inclusion, let x be an element of A. Then by definition $x \in A$, and we have also $x \in A \cup B$ since $x \in A$. Thus, $x \in A \cap (A \cup B)$, which proves the converse inclusion.

   Consequently, $A = A \cap (A \cup B)$.

2. First we show that $A \cup (A \cap B) \subseteq A$. Let $x \in A \cup (A \cap B)$. By Definition of an union, we have either $x \in A$, or $x \in A \cap B$. In both cases[8], we have $x \in A$, so that the inclusion is proved.

[7] This intersection is not empty since A and $A \cup B$ are not disjoint.

[8] If A and B are disjoint, then the first case $x \in A$ necessarily holds, since $x \in A \cup B$ is impossible.

Conversely, let $x \in A$. Then in particular, we have $x \in A \cup (A \cap B)$ by Definition of an union, because $x \in A$. Thus, $x \in A \cup (A \cap B)$.

We have proved that $A \cup (A \cap B) = A$.

Exercise 3.1.9. — *Let A, B, X be sets such that $A \cup B = X$ and $A \cap B = \varnothing$. Show that $A = X \backslash B$ and $B = X \backslash A$.*

The two sets A and B play a symmetrical role here, so that proving one of these two assertions is sufficient. For instance, we prove that $A = X \backslash B$.

- Let x be an element of A. Since $x \in A$, we also have $x \in A \cup B$ by definition of an union. But $A \cup B = X$, and then $x \in X$. On the other hand, we cannot have $x \in B$, because $x \in A$ and the sets A, B are disjoint. Thus, $x \in X$ and $x \notin B$, which means that $x \in X \backslash B$. We have proved that $A \subseteq X \backslash B$.
- Conversely, let x be an element of $X \backslash B$. By definition, this means that $x \in X$, i.e. $x \in A \cup B$, and $x \notin B$. Since $x \in A \cup B$, we have either $x \in A$ or $x \in B$, but we know that the latter is impossible. Thus, we have necessarily $x \in A$. We have proved that $X \backslash B \subseteq A$.
- We can conclude that $X \backslash B = A$.

Exercise 3.1.11. — *Prove that the axiom of replacement (Axiom 3.6) implies the axiom of specification (Axiom 3.5).*

Let's recall the axiom of replacement. Let A be a set. For every $x \in A$, and for every (abstract) object y, let $P(x, y)$ be a statement pertaining to both x and y, such that for any $x \in A$ there is at most one y for which $P(x, y)$ is true. Then there exists a set $\{y : P(x, y) \text{ is true for some } x \in A\}$, such that for any object z,

$$z \in \{y : P(x, y) \text{ is true for some } x \in A\} \iff P(x, z) \text{ is true for some } x \in A$$

Now, let A be a set, x an element of A, and y an object. We accept the axiom of replacement, and show that it implies the axiom of specification.

Let $Q(x, y)$ be the property "$x = y$ and $P(x)$". According to the axiom of replacement, there exists a set $\{y : Q(x, y) \text{ is true for some } x \in A\}$ such that:

$$\begin{aligned}
&z \in \{y : Q(x, y) \text{ is true for some } x \in A\} \\
\iff &Q(x, z) \text{ is true for some } x \in A \\
\iff &x = z \text{ and } P(x) \text{ is true for some } x \in A \\
\iff &x = z \text{ and } P(z) \text{ is true for some } x \in A \text{ (by axiom of substitution)} \\
\iff &z \in A \text{ and } P(z) \text{ is true}
\end{aligned}$$

Thus, we have proved the existence of a set (the set $\{y : Q(x, y) \text{ is true for some } x \in A\}$) satisfying the axiom of specification: z belongs to this set iff $z \in A$ and $P(z)$ is true.

Exercise 3.3.1. — *Show that the definition of equality in Definition 3.3.7 is reflexive, symmetric and transitive. Also verify the substitution property: if $f_1, f_2 : X \to Y$ and $g_1, g_2 : Y \to Z$ are functions such that $f_1 f_2$ and $g_1 = g_2$, then $g_1 \circ f_1 = g_2 \circ f_2$.*

1. Definition 3.3.7 says that two functions f and g are equal if they have same domain X and range Y, and if, for all $x \in X$, $f(x) = g(x)$. This definition of equality is obviously reflexive, symmetric and transitive if we assume that the objects in the domain X and the range Y verify themselves the axioms of equality.

2. Since $f_1 = f_2$, they have same domain X and same range Y. This is also the case for g_1 and g_2, with domain Y and range Z. Thus, $g_1 \circ f_1$ has domain X and range Z, and so has $g_2 \circ f_2$. Furthermore, we have, for all $x \in X$:

$$\begin{aligned} g_2 \circ f_2(x) &= g_2 \circ f_1(x) \text{ (since } f_1 = f_2) \\ &= g_1 \circ f_1(x) \text{ (since } g_1 = g_2) \end{aligned}$$

which closes the demonstration.

Exercise 3.3.2. — *Let $f : X \to Y$ and $g : Y \to Z$ be functions. Show that if f and g are both injective, then so is $g \circ f$. Similarly, show that if f and g are both surjective, then so is $g \circ f$.*

First let's note that $g \circ f : X \to Z$.

1. Suppose that f and g are both injective, and let $x, x' \in X$. We have successively :

$$\begin{array}{rcl} g \circ f(x) & = & g \circ f(x') \\ g(f(x)) & = & g(f(x')) \\ f(x) & = & f(x') \text{ because } g \text{ is injective} \\ x & = & x' \text{ because } f \text{ is injective} \end{array}$$

We have showed that $g \circ f(x) = g \circ f(x') \to x = x'$ for all $x, x' \in X$, i.e. that $g \circ f$ is injective.

2. Suppose that f and g are both surjective, and let be $z \in Z$. Since g is surjective, there exists $y \in Y$ such that $z = g(y)$. And since f is surjective, there exists $x \in X$ such that $y = f(x)$. Thus, combining those two results, there exists $x \in X$ such that $z = g(f(x))$. This means precisely that $g \circ f$ is surjective.

Exercise 3.3.3. — *When is the empty function injective? surjective? bijective?*

Let f be the empty function, i.e. $f : \varnothing \to Y$ for a certain range Y.

1. f is injective iff $x \neq x' \Rightarrow f(x) \neq f(x')$. This can be considered as vacuously true since there are no such x and x' in $\varnothing$. f can be considered as always injective, for any range Y.

2. f is surjective iff for any $y \in Y$, there exists $x \in \varnothing$ such that $y = f(x)$. We can clearly see that this assertion is false if $Y \neq \varnothing$, since any $y \in Y$ will have no antecedent in $\varnothing$. Conversely, if $Y = \varnothing$, the assertion is vacuously true, since there is no element in Y. Thus, f is surjective iff $Y = \varnothing$.

3. Since f is always injective, and is surjective iff $Y = \varnothing$, it is clear that f is bijective iff $Y = \varnothing$.

Exercise 3.3.4. — *Let $f : X \to Y$, $\tilde{f} : X \to Y$, $g : Y \to Z$, $\tilde{g} : Y \to Z$ be functions. Show that if $g \circ f = g \circ \tilde{f}$ and g is injective, then $f = \tilde{f}$. Is this statement true if g is not injective? Also, show that if $g \circ f = \tilde{g} \circ f$ and f is surjective, then $g = \tilde{g}$. Is this statement true if f is not surjective?*

This exercise introduces some cancellation laws for composition.

1. First, note that f and $\tilde{f}$ have same domain and range, which is the first condition for two functions to be equal (by Definition 3.3.7). Then, suppose that $g \circ f = g \circ \tilde{f}$ and g is injective. For the sake of contradiction, suppose that there exists $x \in X$ such that $f(x) \neq \tilde{f}(x)$. Since g is injective, we would thus have $g(f(x)) \neq g(\tilde{f}(x))$, which would be a contradiction to the hypothesis $g \circ f = g \circ \tilde{f}$. Thus, there is no x such that $f(x) = \tilde{f}(x)$, or in other words, $f = \tilde{f}$.

   This property is false if g is not injective. As a counterexample, one can think of $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = x$, $\tilde{f} : \mathbb{R} \to \mathbb{R}$ with $\tilde{f}(x) = -x$, and $g : \mathbb{R} \to \mathbb{R}_+$ with $g(x) = |x|$.

2. As previously, first note that g and $\tilde{g}$ have same domain and range. Let be $y, y' \in Y$. Since f is surjective, there exist $x, x' \in X$ such that $y = f(x)$ and $y' = f(x')$ respectively. Since $g \circ f = g \circ \tilde{f}$, we have $g(f(x)) = g(f(x'))$, i.e. $g(y) = g(y')$. We have showed that, for any $y, y' \in Y$, we have $g(y) = g(y')$, which means that $g = \tilde{g}$.

   This statement is false if f is not surjective. For instance, let f be a constant function, e.g. $f : \mathbb{R} \to \mathbb{R}$ with $f(x) = 1$ for all x. Let $g, \tilde{g} : \mathbb{R} \to \mathbb{R}$ with $g(x) = 0$ and $\tilde{g}(x) = -x + 1$. We have $g(1) = \tilde{g}(1)$, i.e. $g(f(x)) = \tilde{g}(x)$ for all $x \in X$, but we obviously do not have $g = \tilde{g}$.

Exercise 3.3.5. — *Let $f : X \to Y$ and $g : Y \to Z$ be functions. Show that if $g \circ f$ is injective, then f must be injective. Is it true that g must also be injective? Show that if $g \circ f$ is surjective, then g must be surjective. Is it true that f must be surjective?*

1. If $g \circ f$ is injective, then for any given objects $x, x' \in X$, we have $g(f(x)) = g(f(x')) \implies x = x'$. For the sake of contradiction, suppose that f is not injective. In this case, there exist two elements $a, a' \in X$ such that $a \neq a'$ and $f(a) = f(a')$. We would thus have $g(f(a)) = g(f(a'))$ (axiom of substitution) and $a \neq a'$, which is incompatible with the hypothesis that $g \circ f$ is injective.

   Thus, $g \circ f$ injective implies that f is injective.

   However, g does not need to be injective. For instance, let's consider $X = \{1, 2\}$ and $Y = Z = \{1, 2, 3\}$. Let's define the function f as the mapping $f(1) = 1$, $f(2) = 2$. Let's define the function g as the mapping $g(1) = 1$, $g(2) = 2$, $g(3) = 2$. Here, f is injective, so is $g \circ f$, but g is not injective.

2. If $g \circ f$ is surjective, then for all $z \in Z$, there exists $x \in X$ such that $z = g(f(x))$. For the sake of contradiction, suppose that g is not surjective: then, there exists $z \in Z$ such that for all $y \in Y$, $z \neq g(y)$. In particular, for all $x \in X$, since $f(x) \in Y$, we would have $g(f(x)) \neq z$, which would be a contradiction with $g \circ f$ surjective.

   However, f does not need to be surjective. For instance, let's consider $X = Y = \{1, 2\}$ and $Z = \{1\}$. Let f be the mapping $f(1) = f(2) = 1$, and g be the mapping $g(1) = g(2) = 1$. Here, $g \circ f$ is surjective, but f is not.

Exercise 3.3.6. — *Let $f : X \to Y$ be a bijective function, and let $f^{-1} : Y \to X$ be its inverse. Verify the cancellation laws $f^{-1}(f(x)) = x$ for all $x \in X$ and $f(f^{-1}(y)) = y$ for all $y \in Y$. Conclude that f^{-1} is also invertible and has f as its inverse.*

Recall that, by definition, for all $y \in Y$, $f^{-1}(y)$ is the only element $x \in X$ such that $f(x) = y$.

1. Let a be an element of X, we thus have $f(a) \in Y$. Let's apply the definition to the element $y = f(a) \in Y$: by definition, $f^{-1}(f(a))$ is the only element $x \in X$ such that $f(x) = f(a)$. Since f is bijective, this implies $x = a$. We thus have proved that $f^{-1}(f(a)) = a$.

2. The proof for $f(f^{-1}(y)) = y$ is similar.

3. To prove that f^{-1} is also invertible, we need to prove that f^{-1} is bijective, i.e. injective and surjective.

   For any given $y \in Y$, since f is bijective, there exists exactly one $x \in X$ such that $y = f(x)$. Similarly, for any given $y' \in Y$, there exists exactly one $x' \in X$ such that $y' = f(x')$. In other words, $f^{-1}(y) = x$ and $f^{-1}(y') = x'$. Suppose that $f^{-1}(y) = f^{-1}(y')$. This can be written $x = x'$, which necessarily implies $f(x) = f(x')$ since f is a function (and by axiom of substitution). And this can also be written $y = y'$. We thus have proved that for any $y, y' \in Y$, $f^{-1}(y) = f^{-1}(y') \implies y = y'$. Thus, f^{-1} is injective.

   Furthermore, for any given $x \in X$, let's denote $y = f(x)$. Since f is bijective, this means that $f^{-1}(y) = x$. Thus, any $x \in X$ has a predecessor $y \in Y$ for f^{-1}, i.e. f^{-1} is surjective.

Exercise 3.3.7. — *Let $f : X \to Y$ and $g : Y \to Z$ be functions. Show that if f and g are bijective, then so is $g \circ f$, and we have $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

The first point is an immediate consequence of Exercise 3.3.2. We just have to show that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Let be any given element $z \in Z$. Since g is bijective, there exists one single element $y \in Y$ such that $z = g(y)$, i.e. $y = g^{-1}(z)$. And since f is also bijective, there exists exactly one single element $x \in X$ such that $y = f(x)$, i.e. $x = f^{-1}(y) = f^{-1}(g^{-1}(z))$.

Thus, for every $z \in Z$, there exists exactly one $x \in X$ such that $g \circ f(x) = z$, and this element is $f^{-1}(g^{-1}(z))$. This means exactly that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Exercise 3.4.1. — *Let $f : X \to Y$ be a bijective function, and let $f^{-1} : Y \to X$ be its inverse. Let V be any subset of Y. Prove that the forward image of V under f^{-1} is the same as the inverse image of V under f; thus the fact that both sets are denoted as f^{-1} will not lead to any inconsistency.*

Since "$f^{-1}(V)$" may refer to two different things here, let's first introduce some notations to avoid any confusion :

- Let F be the *forward image* of V under f^{-1}, i.e. $F = \{f^{-1}(y) \mid y \in V\}$.
- Let I be the *inverse image* of V under f, i.e. $I = \{x \in X \mid f(x) \in V\}$.

In this exercise we must show that $F = I$, so as to ensure that the two definitions of f^{-1} are equivalent. So, we will prove that $F \subseteq I$ and $I \subseteq F$.

1. Let be $x \in F$. Thus, there exists $y \in V$ such that $x = f^{-1}(y)$. By definition, this is equivalent to $f(x) = y$. But since $y \in V$, we can say that $f(x) \in V$. Thus, we have both $x \in X$ (beacuse $F \subseteq X$) and $f(x) \in V$, which means that $x \in I$.

2. Conversely, let be $x \in I$. By definition, this means that $x \in X$ and that $f(x) \in V$, i.e. there exists a certain element $y \in V$ such that $y = f(x) \in V$. This latter statement is equivalent to $x = f^{-1}(y)$. Thus, we have $x \in X$ and $x = f^{-1}(y)$ for a certain $y \in V$, which means that $x \in F$.

Exercise 3.4.2. — *Let $f : X \to Y$ be a function, let S be a subset of X and let U be a subset of Y. What, in general, can one say about $f^{-1}(f(S))$ and S? What about $f(f^{-1}(U))$ and U?*

This exercise gives a first taste of Exercise 3.4.5 below.

1. First consider $f^{-1}(f(S))$ and S.

   - Do we have $f^{-1}(f(S)) \subset S$? Generally, no. As an counterexample, let's consider $f(x) = x^2$ with $X = Y = \mathbb{R}$ and $S = \{0, 2\}$. We have $f^{-1}(f(S)) = f^{-1}(\{0, 4\}) = \{-2, 0, 2\}$. In this set, we have an element, -2, which is not an element of S.
   - Do we have $S \subset f^{-1}(f(S))$? Yes. Let be $x \in S$. Then, by definition, $f(x) \in f(S)$. So, $x \in X$ and is such that $f(x) \in f(S)$: this is precisely the definition of $x \in f^{-1}(f(S))$.
   - Conclusion: generally speaking, S and $f^{-1}(f(S))$ are not equal, but $S \subset f^{-1}(f(S))$.

2. Now consider $f(f^{-1}(U))$ and U.

   - Do we have $U \subset f(f^{-1}(U))$? Generally, no. As a counterexample, let's consider $f(x) = \sqrt{x}$ with $X = \mathbb{R}_+$, $Y = \mathbb{R}$ and $U = [-1, 1]$. We have $f(f^{-1}(U)) = f([0, 1]) = [0, 1]$, which is clearly not a subset of U.
   - Do we have $f(f^{-1}(U)) \subset U$? Yes. Let be $y \in f(f^{-1}(U))$. By definition, there exists $x \in f^{-1}(U)$ such that $y = f(x)$. But if $x \in f^{-1}(U)$, we have $f(x) \in U$. And since $y = f(x)$, this means that $y \in U$.
   - Conclusion: generally speaking, $U \neq f(f^{-1}(U))$, but $f(f^{-1}(U)) \subset U$.

Exercise 3.4.3. — *Let A, B be two subsets of X, and let be $f : X \to Y$. Show that $f(A \cap B) \subseteq f(A) \cap f(B)$, that $f(A) \backslash f(B) \subseteq f(A \backslash B)$, and $f(A \cup B) = f(A) \cup f(B)$. Is it true that, for the first two statements, the $\subseteq$ relation can be improved to $=$?*

Let's prove the three statements successively:

1. If $y \in f(A \cap B)$, then there exists $x \in A \cap B$ such that $f(x) = y$. Since $x \in A \cap B$, we have both $x \in A$ and $x \in B$, which implies $y = f(x) \in f(A)$ and $y = f(x) \in B$ respectively. Thus, $y \in f(A) \cap f(B)$, and we have proved that $f(A \cap B) \subseteq f(A) \cap f(B)$.

   However, the converse inclusion is false in general. For instance, let's consider the two sets $A = \{1, 2\}$, $B = \{2, 3\}$ and the (non injective) function f defined as the mapping $f(1) = 1$, $f(2) = 2$, $f(3) = 1$. We have $f(A) = \{1, 2\}$, $f(B) = \{1, 2\}$, thus $f(A) \cap f(B) = \{1, 2\}$. This is not a subset of $f(A \cap B) = f(\{2\}) = \{2\}$.

2. If $y \in f(A)\backslash f(B)$, then there exists $x_0 \in A$ such that $y = f(x_0)$, but we have $f(b) \neq y$ for all $b \in B$. Suppose that $x_0 \in B$: in this case, $f(x_0) \neq y$, a contradiction. Thus, $y = f(x_0)$ with $x_0 \in A\backslash B$, which proves that $f(A)\backslash f(B) \subseteq f(A\backslash B)$.

   However, the converse inclusion is false in general. For instance, let's consider the two sets $A = \{1, 2, 3\}$, $B = \{3\}$ and the function f defined as the mapping $f(1) = 1$, $f(2) = 2$, $f(3) = 1$. We have $f(A\backslash B) = \{1, 2\}$ but $f(A)\backslash f(B) = \{2\}$.

3. If $y \in f(A \cup B)$, then there exists $x \in A \cup B$ such that $y = f(x)$. If $x \in A$, then $f(x) \in f(A)$, which implies $x \in f(A) \cup f(B)$. There is an identical result if $x \in B$. Thus, $f(A \cup B) \subseteq f(A) \cup f(B)$.

   Conversely, if $y \in f(A) \cup f(B)$, then we have either $y \in f(A)$ or $y \in f(B)$ (or both). In the first case, there exists $x \in A$ such that $y = f(x)$. But since $x \in A$, we also have $x \in A \cup B$, so that $y \in f(A \cup B)$. The same result holds if $y \in B$. Thus, in both cases, $y \in f(A \cup B)$.

Exercise 3.4.4. — *Let be $f : X \to Y$ a function, and let A, B be subsets of Y. Show that $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$, that $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$, and that $f^{-1}(A\backslash B) = f^{-1}(A)\backslash f^{-1}(B)$.*

We prove only the first statement here; since only very small adjustments are required in its proof to prove the last two ones.

- Let be $x \in f^{-1}(A \cup B)$. By definition, $f(x) \in A \cup B$, so that we have either $f(x) \in A$ or $f(x) \in B$.

  If $f(x) \in A$, then $x \in f^{-1}(A)$ by definition. This implies that $x \in f^{-1}(A) \cup f^{-1}(B)$.

  The same conclusion holds if $f(x) \in B$. Thus, we have demonstrated that $f^{-1}(A \cup B) \subseteq f^{-1}(A) \cup f^{-1}(B)$.

- For the conserve inclusion, let be $x \in f^{-1}(A) \cup f^{-1}(B)$. We have either $x \in \inf f(A)$ or $x \in f^{-1}(B)$.

  If $x \in f^{-1}(A)$, then $f(x) \in A$, and since $A \subset A \cup B$, we have $f(x) \in A \cup B$. This implies $x \in f^{-1}(A \cup B)$.

  The same conclusion holds if $x \in f^{-1}(B)$. Thus, $f^{-1}(A) \cup f^{-1}(B) \subseteq f^{-1}(A \cup B)$.

- This proves the equality $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$.

Exercise 3.4.5. — *Let $f : X \to Y$ be a function. Show that $f^{-1}(f(S)) = S$ for every $S \subseteq Y$ iff f is surjective. Show that $f(f^{-1}(S)) = S$ for every $S \subseteq X$ iff f is injective.*

This exercise is a continuation of Exercise 3.4.2. Let's recall its results, that will reduce the amount of things to be proven here:

- we always have $f^{-1}(f(S)) \subseteq S$ for every $S \subseteq Y$, thus we just have to prove that f is surjective iff $S \subseteq f^{-1}(f(S))$ for every $S \subseteq Y$.

- we always have $S \subseteq f(f^{-1}(S))$ for every $S \subseteq X$, thus we just have to prove that f is injective iff then $f(f^{-1}(S)) \subseteq S$ for every $S \subseteq X$.

Let's prove those two statements.

1. Let f be surjective: let's show that $S \subseteq f(f^{-1}(S))$ for all $S \subseteq Y$. Let S be a subset of Y, and $y \in S$[9]. Since f is surjective, there exists $x \in X$ such that $y = f(x)$. Recall that $y \in S$: this means that$f(x) \in S$, i.e. $x \in f^{-1}(S)$. Thus, $y = f(x) \in f(f^{-1}(S))$. We have proved that, if f is surjective, $y \in S \to y \in f(f^{-1}(S))$, i.e. $S \subseteq f(f^{-1}(S))$.

   Conversely, suppose that $S \subseteq f(f^{-1}(S))$ for every $S \subseteq Y$ and let's show that f is surjective. Let's choose $S = Y$: by hypothesis, we have $Y \subseteq f(f^{-1}(Y))$. Then, let be $y \in Y$. There exists $x \in f^{-1}(Y) \subseteq X$ such that $y = f(x)$. This means precisely that f is surjective.

   The first equivalence is proved.

2. Let f be injective, and $S \subseteq X$. Let be $x \in f^{-1}(f(S))$. Thus, by definition, $f(x) \in f(S)$. This means that there exists $x' \in f(S)$ such that $f(x) = f(x')$. And since f is injective, $x = x' \in S$. Thus, if f is injective, $f^{-1}(f(S)) \subseteq S$ for every $S \subseteq X$.

   Conversely, suppose that $f^{-1}(f(S)) \subseteq S$ for every $S \subseteq X$. In particular, this is true for any singleton $S = \{x_0\}$, with $x_0 \in S$. In such a case, we obtain $f^{-1}(f(\{x_0\})) = \{x_0\}$. For any element $x \in X$, if $x \neq x_0$, we have $x \notin \{x_0\}$ by definition of a singleton, thus $x \notin f^{-1}(f(\{x_0\}))$, and thus $f(x) \neq f(x_0)$. This means that f is injective.

   The second equivalence is proved.

Exercise 3.4.6. — *Prove lemma 3.4.9. (Hint: start with the set $\{0,1\}^X$ and apply the replacement axiom, replacing each function f with the object $f^{-1}(\{1\})$.)*

First, let's recall the main propositions involved in this exercise:

- **Replacement axiom**. Let A be a set. For any object $x \in A$, and any object y, suppose we have a statement $P(x, y)$ pertaining to x and y, such that for each $x \in A$ there is at most one y for which $P(x, y)$ is true. Then there exists a set $\{y \mid P(x, y) \text{ is true for some } x \in A\}$, such that for any object z,

$$z \in \{y \mid P(x, y) \text{ is true for some } x \in A\} \iff P(x, z) \text{ is true for some } x \in A$$

- **Power set axiom**. Let X and Y be sets. There there exists a set, denoted Y^X, which consists of all the function from X to Y:

$$f \in Y^X \iff (f \text{ is a function from } X \text{ to } Y\}$$

- **Lemma 3.4.9**. Let X be a set. Then the set $\{Y \mid Y \text{ is a subset of X}\}$ is a set.

The aim is to prove this lemma using the two axioms recalled here.

1. Let X be a set, and $Y = \{0, 1\}$. Per the power set axiom, $\{0,1\}^X$ is a set, and it contains all the functions $f : X \to \{0, 1\}$.

2. Let A be a subset of X. One can define the function $f_A : X \to \{0, 1\}$, such that for all $x \in X$, $f(x) = 1$ if $x \in A$, and $f(x) = 0$ otherwise. We can even say more:

[9] If S is empty, the statement is vacuously true, so that we can suppose $S \neq \varnothing$.

- If A is a subset of X, then there exists an element $f \in \{0,1\}^X$ such that $A = f^{-1}(\{1\})$: this is precisely f_A as defined above.
- Conversely, if $f \in \{0,1\}^X$, then $A = f^{-1}(\{1\})$ is by definition a subset of X.

Thus, the two statements "A is a subset of X" and "there exists $f \in \{0,1\}^X$ such that $A = f^{-1}(\{1\})$" are equivalent.

3. Finally, let be $A \subset X$ and $f \in \{0,1\}^X$. Let's define $P(f, A)$ the statement "$A = f^{-1}(\{1\})$". For each f, there is at most one A (in fact, *exactly* one A) such that $P(f, A)$ is true. Thus, per the axiom of replacement, there exists a set:

$$\mathcal{P} = \{A \mid A = f^{-1}(\{1\}) \text{ for some } f \in \{0,1\}^X\}$$

And, thanks to the equivalence demonstrated in 2.:

$$\mathcal{P} = \{A \mid A \text{ is a subset of } X\}$$

is a well-defined set, which we wanted to prove.

Exercise 3.4.7. — *Let X, Y be sets. Define a partial function from X to Y to be any function $f : X' \to Y'$ with $X' \subseteq X$ and $Y' \subseteq Y$. Show that the collection of all partial functions from X to Y is itself a set.*

- Let be $X' \subseteq X$ and $Y' \subseteq Y$. If both X' and Y' are fixed, then per the power set axiom (3.10), there exists a set $Y'^{X'}$ which consists of all the functions from X' to Y'.

- By lemma 3.4.9, there exist a set 2^X which consists of all the subsets of X, and a set 2^Y which consists of all the subsets of Y.

- Now we *fix* an element X' of 2^X. Let be Y' an element of the set 2^Y, A a set, and P the property "$P(Y', A)$: $A = Y'^{X'}$". Per the replacement axiom, there exists exactly one (and thus, at most one) set:

$$\begin{aligned}\{A \,|\, P(Y', A) \text{ is true for some } Y' \in 2^Y\} &= \{A \,|\, A = Y'^{X'} \text{ for some } Y' \in 2^Y\} \\ &= \{Y'^{X'} \,|\, Y' \in 2^Y\}\end{aligned}$$

- Each element of this set is itself a set. Thus we can apply the axiom of union (3.11): there exists a set $\bigcup\{Y'^{X'} \,|\, Y' \in 2^Y\}$ whose elements are those objects which are elements of elements of $\{Y'^{X'} \,|\, Y' \in 2^Y\}$, i.e.:

$$\bigcup\{Y'^{X'} \,|\, Y' \in 2^Y\} = \{f | f : X' \to Y' \text{ for some } Y' \in 2^Y\}$$

This set is obtained for one given *fixed* subset $X' \subseteq X$, so let's denote this set:

$$S_{X'} = \{f | f : X' \to Y' \text{ for some } Y' \in 2^Y\}$$

- Now we apply once again the union set (3.11), especially in its second formulation. If we denote $I = 2^X$, then for each element $X' \in I$ we do have one set $S_{X'}$, which is defined above. Thus, there exists a set $\bigcup_{X' \in 2^X} S_{X'} := \bigcup\{S_{X'} \,|\, X' \in 2^X\}$. And, for every function f, we have $f \in \bigcup\{S_{X'} \,|\, X' \in 2^X\}$ iff there exists $X' \in 2^X$ such that $f \in S_{X'}$, i.e. if there exists $X' \subset X$ and $Y' \subset Y$ such that $f : X' \to Y'$.

- Consequently, we have proved that there exists a set which consists of the collection of all partial functions from X to Y.

Exercise 3.4.8. — *Prove that Axiom 3.4 can be deduced from Axiom 3.1, Axiom 3.3 and Axiom 3.11.*

Let's recall very briefly the four axioms involved here :

- Axiom 3.4 (to be proved) says that if A and B are sets, then there exists a union set $A \cup B$ such that $x \in A \cup B$ iff $x \in A$ or $x \in B$.
- Axiom 3.1 essentially says that sets are objects.
- Axiom 3.3 says that singletons are pair sets do exist.
- Axiom 3.11: let A be a set, whose all elements are themselves sets. Then there exists a set $\bigcup A$ whose elements are those objects which are elements of elements of A, i.e., $x \in \bigcup A$ iff $x \in S$ for some $S \in A$.

Here is a sketch of proof for Axiom 3.4. Let A and B be sets. According to Axiom 3.1, A and B are themselves objects: they can be elements of other sets. Consequently, according to Axiom 3.3, it makes sense to talk about the singleton sets $\{A\}$ and $\{B\}$, and the set $\{A, B\}$.

Now we consider this latter set, which we denote $\mathcal{A} = \{A, B\}$. According to axiom 3.11, there exists a set $\bigcup \mathcal{A}$ whose elements those objects which are the elements of $\mathcal{A}$, i.e., $x \in \mathcal{A}$ iff there exists $S \in \mathcal{A}$ such that $x \in S$. But $\mathcal{A}$ is a pair set with only two elements, so that S must necessarily be equal to A or B.

This leads to the following conclusion: if A and B are sets, then there exists a set $\mathcal{A}$ such that $x \in \mathcal{A}$ iff $x \in A$ or $x \in B$. This is precisely the Axiom 3.4.

Exercise 3.5.1. — *Suppose we define the ordered pair (x, y) for any objects x and y by the formula $(x, y) := \{\{x\}, \{x, y\}\}$. Show that this definition obeys the property (3.5), and also whenever X and Y are sets, the cartesian product $X \times Y$ is also a set.*

Recall that property (3.5) says that $(x, y) = (x', y')$ iff $x = x'$ and $y = y'$. The proof below is heavily inspired by the sketch given by Paul Halmos in his famous book, *Naive Set Theory.* (The proof below is just immensely more verbose.)

1. First, we go back to Remark 3.1.9 by Terence Tao (page 37). In this remark, Tao says that for any object a, the pair set $\{a, a\}$ is in fact the singleton $\{a\}$. Tao asks *why?* to the reader. This is easy to prove using the Definition 3.1.4 (equality of sets): both sets have the same elements, thus they are equal. This fact is a crucial point for the current proof.

   Indeed, first note that for any object x, the ordered pair (x, x) will be (by definition) equal to $\{\{x\}, \{x, x\}\}$. Applying twice the Remark 3.1.9 made by Terence Tao, we can conclude that $(x, x) = \{\{x\}\}$ for any object x.

   Conversely, if any ordered pair (x, y) is a singleton, this means that $\{\{x\}, \{x, y\}\}$ is a singleton. This implies that both elements of this pair set are equal, i.e. $\{x\} = \{x, y\}$. Thus, (by Definition 3.1.4,) $y \in \{x\}$, i.e. $x = y$.

   This gives a handy conclusion, that we can write as a lemma:

Lemma. *An ordered pair (x, y) is a singleton if and only if $x = y$ (and in this case, this singleton is $\{\{x\}\} = \{\{y\}\}$).*

We can now prove more easily that property (3.5) is met.

2. Let's prove that the property (3.5) is satisfied.

   - First, let be two ordered pairs $(a, b) = \{\{a\}, \{a, b\}\}$ and $(x, y) = \{\{x\}, \{x, y\}\}$. If $a = x$ and $y = b$, then we obviously have $\{\{a\}, \{a, b\}\} = \{\{x\}, \{x, y\}\}$.
   - For the reciprocal, suppose that $(a, b) = (x, y)$, and let's show that $a = x$ and $b = y$. We will consider two distinct cases.

     (a) First consider the case where $a = b$ (note that this also covers the case $x = y$, since they play symmetrical roles). Thus $(a, b) = \{\{a\}\}$. Since $(a, b) = (x, y)$, we have $\{\{x\}, \{x, y\}\} = \{\{a\}\}$.
     This means that $\{x\} \in \{\{a\}\}$, i.e. $a = x$.
     But we also have $\{x, y\} \in \{\{a\}\}$, i.e. $\{x, y\} = \{a\}$. This means in particular that $\{x, y\}$ is a singleton, which is only possible if $x = y$ according to the lemma introduced above.
     Thus, $a = b$ by hypothesis, and $a = x$, and $x = y$. This finally means that all four elements are equal: $a = b = x = y$.
     We can insist: if we have either $a = b$ or $x = y$, then all four elements are equal, and property (3.5) is met.

     (b) The other case is $a \neq b$ (which also implies $x \neq y$, otherwise all four elements would be equal). Since $(a, b) = (x, y)$, we have $\{a\} \in \{\{x\}, \{x, y\}\}$, so that we have either $\{a\} = \{x\}$ or $\{a\} = \{x, y\}$. The latter case can be excluded: $\{a\} = \{x, y\}$ would mean that $\{x, y\}$ is a singleton, thus $x = y$, a contradiction with our hypothesis. Thus, the only possibility is $\{a\} = \{x\}$, i.e. $a = x$.
     We also have $\{a, b\} \in \{\{x\}, \{x, y\}\}$, and for the same reason, the only possibility is $\{a, b\} = \{x, y\}$. But we have showed that $a = x$, so that $\{a, b\} = \{a, y\}$. The conclusion is $y = b$.
   - Conclusion: in both cases, (a, b) implies both $a = x$ and $y = b$, which is our initial goal. Property (3.5) is met.

3. Finally, if we adopt this definition, $X \times Y$ is a set. Indeed, for every $x \in X$ and $y \in Y$, both x and y are elements of $X \cup Y$. Thus, the singleton $\{x\}$ and the pair set $\{x, y\}$ are elements of the power set of $X \cup Y$ (which is indeed a set, by Lemma 3.4.9: see Exercise 3.4.6). In other words, if $\mathcal{P}(A)$ denotes the power set of a set A, we have $\{x\} \in \mathcal{P}(X \cup Y)$ and $\{x, y\} \in \mathcal{P}(X \cup Y)$.

   Thus, for every objects $x \in X$ and $y \in Y$, $\{\{x\}, \{x, y\}\} \subset \mathcal{P}(X \cup Y)$. This latter statement is equivalent to $\{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(X \cup Y))$, which is also a well-defined set by a (recursive) application of Lemma 3.4.9.

   Then, for any element $S \in \mathcal{P}(\mathcal{P}(X \cup Y))$, let $P(S)$ be the property "There exists $x \in X$ and $y \in Y$ such that $S = \{\{x\}, \{x, y\}\}$". By the axiom of specification (Axiom 3.5), there exists a set $\{S \in \mathcal{P}(\mathcal{P}(X \cup Y)) \mid P(S) \text{ is true}\}$: this set is precisely the cartesian product $X \times Y$ we were looking for.

EXERCISE 3.6.2. — *Show that a set X has cardinality 0 if and only if X is the empty set.*

First suppose that $X = \varnothing$ and let's show that $\#X = 0$. By Definition 3.6.5, we must show that there exists a bijection between X and the set $\{i \in \mathbb{N} : 1 \leqslant i \leqslant 0\}$. It turns out that $X = \varnothing$ by hypothesis, and that $\{i \in \mathbb{N} : 1 \leqslant i \leqslant 0\}$ is also equal to $\varnothing$. And, by Exercise 3.3.3, we know that the empty function $f : \varnothing \to \varnothing$ is bijective. Thus, we have indeed $\#X = 0$.

Now suppose that we have a set X such that $\#X = 0$, and let's show that $X = \varnothing$. If $\#X = 0$, we know that there exists a bijection from $\{i \in \mathbb{N} : 1 \leqslant i \leqslant 0\}$ to X, i.e. from $\varnothing$ to X. Still by Exercise 3.3.3, a function $f : \varnothing \to X$ can be bijective if and only if $X = \varnothing$, so we are done.

EXERCISE 3.6.3. — *Let n be a natural number, and let $f : \{i \in \mathbb{N} : 1 \leqslant i \leqslant n\} \to \mathbb{N}$ be a function. Show that there exists a natural number M such that $f(i) \leqslant M$ for all $1 \leqslant i \leqslant n$. Thus finite subsets of the natural numbers are bounded.*

Let's induct on n.

- Let's take $n = 1$ for the base case. f is thus a function from the singleton $\{1\}$ to $\mathbb{N}$. If we choose the natural number $M := f(1)$, we have indeed $f(i) \leqslant M$ for all $i \in \{1\}$.

- Now suppose inductively that the proposition is true for a natural number n, and let's show that it is still true for $n + 1$.

  Let be $f : \{i \in \mathbb{N} : 1 \leqslant i \leqslant n + 1\} \to \mathbb{N}$. By the induction hypothesis, we already know that there exists an M such that $f(i) \leqslant M$ for $1 \leqslant i \leqslant n$. Now let be $M' := M+f(n+1)$. Since $f(n+1)$ is a natural number, we have $M \leqslant M'$, so that $f(i) \leqslant M'$ for all $1 \leqslant i \leqslant n$ by transitivity.

  Also, we have $f(n + 1) \leqslant f(n + 1) + M = M'$ because M is a natural number.

  Thus, we have found a natural number M' such that $f(i) \leqslant M'$ for all $i \in [\![1, n + 1]\!]$, as required. This closes the proof.

EXERCISE 3.6.7. — *Let A and B be sets. Let us say that A has* lesser or equal cardinality *to B if there exists an injection $f : A \to B$ from A to B. Show that if A and B are finite sets, then A has lesser or equal cardinality to B if and only if $\#(A) \leqslant \#(B)$.*

Since A is finite, there exists a natural number n such that $\#A = n$; and thus a bijection $g : [\![1, n]\!] \to A$.

Similarly, there exists $m \in \mathbb{N}$ such that $\#B = m$; and thus a bijection $h : [\![1, m]\!] \to B$.

- First suppose that $n \leqslant m$, and let's show that there exists a bijection $f : A \to B$. Let's define a function $f : A \to B$, such that for all $i \in [\![1, n]\!]$, we have $f(g(i)) = h(i)$. We claim that f is injective. Indeed, let's suppose that we have $f(x) = f(x')$ for $x, x' \in A$. Since g is bijective, there exists $i, j \in [\![1, n]\!]$ such that $f(g(i)) = f(g(j))$. By definition of f, this means that $h(i) = h(j)$. Since h is also bijective, it means that $i = j$. And since $i = i$, we have $g(i) = g(j)$, and thus $x = x'$.

  We have showed that $f(x) = f(x')$ for all $x, x' \in A$, i.e., that f is injective, as required.

- Now suppose that there exists an injection $f : A \to B$, and let's show that $\#A \leqslant \#B$. Obviously, f is a bijection from A to $f(A)$. It means that $f(A)$ and A have the same cardinality, n.

  But $f(A) \subseteq B$, so that by Proposition 3.6.14(c), we have $\#B \geqslant \#(f(A)) = n$, which closes the proof.

4. Integers and rationals

Exercise 4.1.1. — *Verify that the definition of equality on the integers is both reflexive and symmetric.*

Recall the Definition 4.1.1 of equality on integers: two integers $a — b$ and $c — d$ are equal iff $a + d = c + b$. This defines a binary relation on $\mathbb{Z}$, denoted "=". Let's show that this relation is reflexive and symmetric.

- Reflexivity: let a and b be natural numbers, so that $a — b$ is an integer. We know that, on natural numbers, equality is reflexive, i.e. $a + b = a + b$. This equality means precisely that $a — b = a — b$.
- Symmetry: let a, b, c, d be natural numbers. If $a — b = c — d$, do we also have $c — d = a — b$?

$$\begin{array}{rrcl} & a — b & = & c — d \\ \Longleftrightarrow & a + d & = & c + b \\ \Longleftrightarrow & c + b & = & a + d \text{ (equality is symmetric on naturals)} \\ \Longleftrightarrow & c — d & = & a — b \end{array}$$

Exercise 4.1.2. — *Show that the definition of negation on the integers is well-defined in the sense that if $(a — b) = (a' — b')$, then $-(a — b) = -(a' — b')$ (so equal integers have equal negations).*

Since $a — b = a' — b'$, we have $a + b' = a' + b$.
Also, by Definition 4.1.4 of negation, we have:

$$-(a — b) = b — a$$
$$-(a' — b') = b' — a'$$

Then, we have successively:

$$\begin{aligned} b + a' &= a' + b \text{ (addition is commutative on naturals, Prop. 2.2.4)} \\ &= a + b' \text{ (initial hypothesis)} \\ &= b' + a \text{ (by commutativity on naturals once again)} \end{aligned}$$

and this equality $b + a' = b' + a$ precisely means that $b — a = b' — a'$, i.e. that $-(a — b) = -(a' — b')$.

Exercise 4.1.3. — *Show that $(-1) \times a = -a$ for every integer a.*

Since a is an integer, there exist two natural numbers n and m such that $a = m — n$.
On the one hand, by Definition 4.1.4, $-a = n — m$.
On the other hand, using once again Definition 4.1.4 and Definition 4.1.2,

$$\begin{aligned} (-1) \times a &= (0 — 1) \times (m — n) \\ &= (0 \times m + 1 \times n) — (0 \times n + 1 \times m) \\ &= n — m \end{aligned}$$

Thus, we have indeed $(-1) \times a = -a$.

EXERCISE 4.1.4. — *Prove the remaining identities in Proposition 4.1.6.*

Let $x = a \text{---} b$, $y = c \text{---} d$ and $z = e \text{---} f$ be three integers. Those identities are the following:

1. $x + y = y + x$, i.e., we must prove that addition is commutative on the integers. We have:

$$\begin{aligned} x + y &= (a \text{---} b) + (c \text{---} d) \\ &= (a + c) \text{---} (b + d) \text{ (by Definition 4.1.2)} \\ &= (c + a) \text{---} (d + b) \text{ (addition is commutative on naturals)} \\ &= (c \text{---} d) + (a \text{---} b) \text{ (by Definition 4.1.2 again)} \\ &= y + x \end{aligned}$$

2. $(x + y) + z = x + (y + z)$, i.e. prove that addition is associative on the integers. This is a very similar proof, and this is a direct consequence of associativity of addition on the naturals.

3. $x + 0 = 0 + x = x$. We have already showed that addition is commutative on the integers, so we just have to prove that $x + 0 = x$.

$$\begin{aligned} x + 0 &= (a \text{---} b) + (0 \text{---} 0) \\ &= (a + 0) \text{---} (b + 0) \\ &= a \text{---} b = x. \end{aligned}$$

4. $x+(-x) = (-x)+x = 0$. Once again, thanks to the previous result about commutativity, we just have to prove that $x + (-x) = 0$.

$$\begin{aligned} x + (-x) &= (a \text{---} b) + (b \text{---} a) \text{ (by Definition 4.1.4)} \\ &= (a + b) \text{---} (b + a) \text{ (by Definition 4.1.2)} \\ &= (a + b) \text{---} (a + b) \text{ (addition is commutative on naturals)} \\ &= 0 \text{ (because } m \text{---} m = 0 \text{---} 0 \text{ for all integer } m\text{)} \end{aligned}$$

5. $xy = yx$, i.e. multiplication is commutative on the integers.

$$\begin{aligned} xy &= (a \text{---} b) \times (c \text{---} d) \\ &= (ac + bd) \text{---} (ad + bc) \text{ (by Definition 4.1.2)} \\ &= (ca + db) \text{---} (da + cb) \text{ (multiplication is commutative on the naturals)} \\ &= yx \text{ (by Definition 4.1.2)} \end{aligned}$$

6. $(xy)z = x(yz)$, i.e. multiplication is associative on the integers. This is actually the only identity proved in the main text by Terence Tao.

7. $x1 = 1x = x$. The equality between the first two terms is a direct consequence of commutativity of multiplication on the integers. We just have to prove that $x1 = x$. And indeed, $x1 = (a \text{——} b) \times (1 \text{——} 0) = (a1 + b0) \text{——} (b1 + a0) = a \text{——} b = x$.

8. $x(y + z) = xy + xz$, i.e. show distributivity on the integers. On the left side, we have:

$$\begin{aligned} x(y + z) &= (a \text{——} b)\,((c \text{——} d) + (e \text{——} f)) \\ &= (a \text{——} b)\,((c + e) \text{——} (d + f)) \\ &= (a(c + e) + b(d + f)) \text{——} (a(d + f) + b(c + e)) \\ &= ((ac + ae + bd + bf)) \text{——} ((ad + af + bc + be)) \end{aligned}$$

   and then on the left side:

$$\begin{aligned} xy + xz &= (a \text{——} b)(c \text{——} d) + (a \text{——} b)(e \text{——} f) \\ &= ((ac + bd) \text{——} (ad + bc)) + ((ae + bf) \text{——} (af + be)) \\ &= ((ac + ae + bd + bf)) \text{——} ((ad + af + bc + be)) \end{aligned}$$

9. $(y + z)x = yx + zx$. This latter identity is a direct consequence of commutativity of multiplication on the integers, and distributivity on the integers, both being already proved earlier in this exercise.

EXERCISE 4.1.5. — *Prove Proposition 4.1.8, i.e.: let x and y be integers such that $xy = 0$, then either $x = 0$ or $y = 0$ (or both).*

We will use here Lemma 4.1.5 (trichotomy of integers, which says that any integer is either zero, or equal to a positive natural number, or the negation of a positive natural number), and Lemma 2.3.3 (which provides an equivalent of Proposition 4.1.8 for natural numbers only). We will prove the proposition for (all) three possible cases: $x = 0$, x is a positive natural number, $-x$ is a positive natural number.

y will be considered as a fixed integer, $y = c \text{——} d$ with c, d natural numbers.

1. First let's take the case $x = 0$. There is nothing to prove here, the proposition is obviously true.

2. Then let's take the case where x is a positive natural number (and, consequently, is not equal to zero). In this case, as an integer, x can be written $n \text{——} 0$ with n a positive natural number.

   We have $xy = (n \text{——} 0) \times (c \text{——} d) = (nc + 0d) \text{——} (nd + 0c) = nc \text{——} nd$.

   Thus, $xy = 0$ iff $nc \text{——} nd = 0 \text{——} 0$, and by Definition 4.1.1, this means that $nc = nd$. But since all three n, c, d are natural numbers, we can use the cancellation law for natural numbers and conclude that $c = d$.

   This means that $y = c \text{——} c = 0 \text{——} 0 = 0$. Thus, in this case, if $xy = 0$ with x non-zero, we have showed that y is necessarily equal to 0.

3. Finally, let's take the case where x is the opposite of a positive natural number n, i.e. $x = 0 \text{——} n$. A very similar proof also leads to $c = d$, and to $y = 0$.

Exercise 4.1.6. — *Prove Corollary 4.1.9, i.e. if a, b, c are integers such that $ac = bc$ and c is non-zero, then $a = b$.*

If $ac = bc$, then $ac + (-bc) = bc + (-bc) = 0$. Thus, $ac - bc = 0$.

Let's use the property of distributivity (Proposition 4.1.6): we obtain $(a - b)c = 0$. According to Proposition 4.1.8 (see also the previous exercise), this implies either $c = 0$ or $a - b = 0$. The first option ($c = 0$) must be discarded since it does not fit the initial hypothesis. The only possibility is thus $a - b = 0$, and adding b to both sides finally leads to $a = b$.

Exercise 4.1.7. — *Prove Lemma 4.1.11.*

The statements to prove are the following:

1. Show that $a > b$ if and only if $a - b$ is a positive natural number.

   First suppose that $a > b$. This means (Definition 4.1.10) that there exists a natural number n such that $a = b + n$, and $a \neq b$. Then, we add to both sides the opposite of b, and we get $a + (-b) = b + n + (-b)$, i.e. $a - b = n$. In this latest equality, n cannot be zero, otherwise we would have $a = b$, which is excluded. The first implication is proved.

   Then suppose that $a - b = n$ with n a positive natural number. Adding b to both sides leads to $a = b + n$, i.e. $a \geqslant b$. We cannot have $a = b$, because this would be a contradiction with the fact that $n \neq 0$. Thus, $a > b$.

2. Show that addition preserves order, i.e. if $a > b$, then $a + c > b + c$.

   Suppose that $a > b$. According to the previous point, this means that $a - b = n$, with n a positive natural number. Then, we get successively:

$$\begin{array}{rcll} a & = & b+n & \text{(by adding } b \text{ to both sides)} \\ a+c & = & b+c+n & \text{(by adding } c \text{ to both sides)} \\ a+c-b-c & = & n & \text{(by adding } (-b)+(-c) \text{ to both sides)} \\ a+c-(b+c) & = & n & \text{(by using the distributive law and Exercise 4.1.3)} \end{array}$$

   Using again the previous point, since $(a + c) - (b + c)$ is equal to a positive natural number, we can conclude that $a + c > b + c$.

3. Show that positive multiplication preserves order, i.e. if $a > b$ and c is positive, then $ac > bc$.

   Since $a > b$, according to the first point of this exercise, we have $a - b = n$, with n a positive natural number. According to the distributive law, $(a - b)c = ac - bc$. But we also have $(a - b)c = nc$, and nc is a positive natural number (as product of two positive numbers, see Lemma 2.3.3). Thus, $ac - bc$ is equal to a positive number, which means that $ac > bc$.

4. Show that negation reverses order, i.e. if $a > b$, then $-a < -b$.

   Here, we will need a small lemma, which says that for any natural number n, we have $n = -(-n)$. There are several ways to show this result: either by proving that $(-1) \times (-1) = 1$ and using Exercise 4.1.3, or simply by noting that $n + (n) = 0$ for all n, which means that n is the opposite of $-n$ (i.e., $n = -(-n)$).

Now this point is easy to prove. $a > b$ means that $a - b$ is a positive number, as shown earlier in this exercise. We want to prove that $-a < -b$, and proving this assertion requires to show that $-b - (-a)$ is a positive number. But $-b - (-a) = -b + a = a - b$, which is a positive natural number. Thus we are done.

5. Show that order is transitive, i.e. if $a > b$ and $b > c$, then $a > c$.

   Still using the first point of this exercise, we have $a - b = n$ and $b - c = m$, with n, m two positive natural numbers. We know that $n + m$ is positive as the sum of two positive numbers, thus $n + m = a - b + b - c = a - c$ is positive. This means that $a > c$.

6. Show order trichotomy, i.e.: exactly one of the statements $a > b$, $a < b$, or $a = b$ is true.

   - If $a = b$, then obviously (exactly) one of those statements is true.
   - Now consider the case $a \neq b$, and let's show that we have either $a > b$ or $a < b$ (and cannot have both).

     Let's consider the integer $a - b$. By trichotomy of integers (Lemma 4.1.5), we know that we have either $a - b = 0$ (which is excluded here), or $a - b = n$ with n positive, or $a - b = -n$ with n positive.

     If $a - b = n$, then $a > b$ according to the first point of this exercise. If $a - b = -n$, then $-n = -(a - b) = b - a$, thus $b > a$.

     Finally, we just have to show that we cannot have both $a > b$ and $b > a$ at the same time. If $a > b$, then the integer $a - b$ is positive. If $b > a$, then $b - a$ is positive, i.e. $-(b - a) = a - b$ is the opposite of a positive natural. Thus, $a - b$ is both positive and the opposite of a positive number, which is excluded by the trichotomy of integers.

Exercise 4.1.8. — *Show that the principle of induction (Axiom 2.5) does not apply directly to the integers. More precisely, give an example of a property $P(n)$ pertaining to an integer n such that $P(0)$ is true, and that $P(n)$ implies $P(n{+}{+})$ for all integers n, but that $P(n)$ is not true for all integers n.*

According to Lemma 4.1.5, an integer is either equal to 0, or equal to a positive natural number, or equal to the negation of a positive natural number.

Let's define $P(n)$ as the property "The integer n is a natural number, i.e. is either equal to 0 or equal to a positive natural number". Obviously, $P(0)$ is true. Furthermore, if n is a natural number, then $n{+}{+}$ is also a natural number (Axiom 2.2), so that if $P(n)$ is true, then $P(n{+}{+})$ is true. Thus, $P(n)$ matches the required condtions.

However, $P(-1)$ is obviously false.

Exercise 4.2.1. — *Show that the definition of equality for the rational numbers is reflexive, symmetric, and transitive.*

This exercise ressembles Exercise 4.1.1, and the same approach applies. Recall the Definition 4.2.1: two rational numbers $a /\!/ b$ and $c /\!/ d$ are equal iff $ad = bc$. This defines a binary relation on $\mathbb{Q}$, denoted "=". Let's show that this relation is reflexive, symmetric and transitive.

Hereafter, a, b, c, d, e, f are integers (and b, d, f are non-zero).

- Reflexivity: here we must prove that $a \,//\, b = a \,//\, b$. This is the case iff $ab = ba$, with is true because of (commutativity of multiplication on $\mathbb{Z}$ and) reflexivity of equality on $\mathbb{Z}$.
- Symmetry: here we must prove that if $a \,//\, b = c \,//\, d$, we also have $c \,//\, d = a \,//\, b$. We have successively:

$$\begin{array}{rrcl} & a \,//\, b & = & c \,//\, d \\ \iff & ad & = & bc \\ \iff & da & = & cb \text{ ($\times$ is commutative on $\mathbb{Z}$)} \\ \iff & cb & = & da \text{ ($=$ is symmetric on $\mathbb{Z}$)} \\ \iff & c \,//\, d & = & a \,//\, b \end{array}$$

- Transitivity: here we must prove that if $a \,//\, b = c \,//\, d$ and $c \,//\, d = e \,//\, f$, then $a \,//\, b = e \,//\, f$. I.e., we must prove that if $ad = bc$ and $cf = de$, then $af = be$.

  Let's multiply by e both sides of the equality $ad = bc$: we get $ade = bce$. Since $de = cf$, we also get $acf = bce$.

  In this latest equality, using the cancellation law (Corollary 4.1.9) for c would lead to $af = be$, which would close the proof. However, unlike b, d or f, the integer c may be equal to 0, and in this case we cannot use the cancellation law. There are thus two different cases:

  - If $c \neq 0$, we simply use the cancellation law: since $acf = bce$, then $af = be$, which means that $a \,//\, b = e \,//\, f$.
  - If $c = 0$, then $bc = 0$. But since $ad = bc$, we also have $ad = 0$, and we know that $d \neq 0$. According to Proposition 4.1.8, this leads to $a = 0$. A similar reasoning leads to $e = 0$. Thus, $a = c = e = 0$, and $0 = af = be$, which means $a \,//\, b = e \,//\, f$.

Exercise 4.2.2. — *Prove the remaining components of Lemma 4.2.3.*

Let $a \,//\, b = a' \,//\, b'$ be two rationals; let $c \,//\, d = c' \,//\, d'$ be two rationals. The remaining claims are the following:

- Prove that $-(a' \,//\, b') = -(a \,//\, b)$. This equality holds iff $(-a') \,//\, b' = (-a) \,//\, b$, i.e. iff $(-a')b = b'(-a)$. We have successively:

$$\begin{aligned} (-a')b &= (-1)a'b \text{ (see Exercise 4.1.3)} \\ &= (-1)ab' \text{ (because $a \,//\, b = a' \,//\, b'$)} \\ &= (-a)b' \text{ (using Exercise 4.1.3 one again)} \end{aligned}$$

  Thus we are done.

- Prove that $(a \,//\, b) \times (c \,//\, d) = (a' \,//\, b') \times (c \,//\, d)$. To prove this equality, we must show that $(ac) \,//\, (bd) = (a'c) \,//\, (b'd)$. By definition of equality between rationals, this holds iff $acb'd = bda'c$. Since $ab' = ba'$, the claim follows (using commutativity of multiplication on integers[10]).

[10] This kind of precision about very basic properties of naturals and integers will not be mentioned anymore.

- Prove that $(a \,//\, b) \times (c \,//\, d) = (a \,//\, b) \times (c' \,//\, d')$. To prove this equality, we must show that $(ac) \,//\, (bd) = (ac') \,//\, (bd')$. By definition of equality between rationals, this holds iff $acbd' = bdac'$. Since $cd' = dc'$, the claim follows.

Exercise 4.2.3. — *Prove the remaining components of Proposition 4.2.4.*

Let $x = a \,//\, b$, $y = c \,//\, d$, and $z = e \,//\, f$ be rational numbers, with a, c, e integers, and b, d, f non-zero integers. The remaining claims are the following:

1. $x + y = y + x$, i.e. addition is commutative for the rationals.

   On the one hand, we have $x + y = a \,//\, b + c \,//\, d = (ad + bc) \,//\, bd$.

   On the other hand, $y + x = c \,//\, d + a \,//\, b = (cb + da) \,//\, db = (ad + bc) \,//\, bd$ using the commutativity of addition and multiplication on the integers. Thus, the two expressions are equal.

2. $(x + y) + z = x + (y + z)$: already proved in the book.

3. $x + 0 = 0 + x = x$. By the first point of this exercise, we already know that $x + 0 = 0 + x$, so we have just to show that $x + 0 = x$. We have $x + 0 = a \,//\, b + 0 \,//\, 1 = (a1 + b0) \,//\, (b1) = a \,//\, b = x$, which is the required result.

4. $x + (-x) = (-x) + x = 0$. Once again, the part $x + (-x) = (-x) + x$ comes from the first point of this exercise, so we just have to prove that $x + (-x) = 0$. We have $x + (-x) = a \,//\, b + -(a \,//\, b) = a \,//\, b + (-a) \,//\, b = (ab - ba) \,//\, b^2 = 0 \,//\, b^2$. But we know that $0 \,//\, m = 0 \,//\, 1 = 0$ for all non-zero integer m, since $0 \times 1 = m \times 0$. Thus, $x + (-x) = 0 \,//\, b^2 = 0$, as required.

5. $xy = yx$, i.e. multiplication is commutative on the rationals. Indeed, $xy = (a \,//\, b) \times (c \,//\, d) = (ac) \,//\, (bd)$ by definition. On the other hand, $yx = (c \,//\, d) \times (a \,//\, b) = (ca) \,//\, (db) = (ac) \,//\, (bd)$ by commutativity of multiplication on the integers. Thus, $xy = yx$.

6. $(xy)z = x(yz)$, i.e. multiplication is associative on the rationals. We have $(xy)z = (ace) \,//\, (bdf) = x(yz)$ by associativity of multiplication on the integers.

7. $x1 = 1x = x$. Once again, we already know that $x1 = 1x$, thanks to the fifth point of this exercise. So we just have to show that $x1 = x$. We have $x1 = (a \,//\, b) \times (1 \,//\, 1) = (a1) \,//\, (b1) = a \,//\, b = x$.

8. $x(y + z) = xy + xz$, i.e. distributivity of multplication for the rationals. On the one hand, we have:

$$\begin{aligned} x(y+z) &= (a \,//\, b)(c \,//\, d + e \,//\, f) \\ &= (a \,//\, b)((cf + de) \,//\, (df)) \\ &= (acf + ade) \,//\, (bdf) \end{aligned}$$

On the other hand[11]:

$$\begin{aligned} xy + xz &= (a \,/\!/\, b)(c \,/\!/\, d) + (a \,/\!/\, b)(e \,/\!/\, f) \\ &= (ac) \,/\!/\, (bd) + (ae) \,/\!/\, (bf) \\ &= (acbf + bdae) \,/\!/\, (b^2 df) \\ &= (acf + ade) \,/\!/\, (bdf) \end{aligned}$$

Thus we have indeed $x(y+z) = xy + xz$.

9. $(y+z)x = yx + zx$. This can be deduced immediately from commutativity of multiplication and the eighth point of this exercise.

10. For all $x \neq 0$, $xx^{-1} = x^{-1}x = 1$. Once again, the part $xx^{-1} = x^{-1}x$ comes from the fifth point of this exercise, so that we just have have to show that $xx^{-1} = 1$.

$$\begin{aligned} xx^{-1} &= (a \,/\!/\, b) \times (b \,/\!/\, a) \\ &= (ab) \,/\!/\, (ba) \\ &= 1 \,/\!/\, 1 = 1 \end{aligned}$$

Exercise 4.2.4. — *Prove Lemma 4.2.7. (trichotomy of rationals), i.e., if x is a rational number, then exactly one of the following three statements is true: (a) x is equal to 0, (b) x is a positive rational number, or (c) x is a negative rational number.*

Following the hint given by Terence Tao, we'll first prove that *at least* one of those statements is true, and then that *at most* one of them is true. Let be $x = a \,/\!/\, b$, where a is an integer and b a non-zero integer.

1. Let's prove that at least one of those statements is true.

   First, an obvious case: if $a = 0$, then $x = a \,/\!/\, b = 0$, thus one of the statements is true. Now consider the case where $a \neq 0$. By the trichotomy of integers, a can be either positive or negative. Similarly, b can also be either positive or negative (it cannot be null, by definition). Thus, there are four main cases:

   - $a > 0$ and $b > 0$. Here, by Definition 4.2.6, $x = a \,/\!/\, b$ is positive.
   - $a > 0$ and $b < 0$. Here, $b = -m$, with m a positive natural number. Thus, $x = a \,/\!/\, (-m)$. But $a \,/\!/\, (-m) = (-a) \,/\!/\, m$ (this is easy to verify: $am = (-a)(-m)$). This means that $x = (-a) \,/\!/\, m$, with both a and m positive, i.e. x is negative.
   - $a < 0$ and $b > 0$. Here, $x = a \,/\!/\, b$ is negative by Definition 4.2.6.
   - $a < 0$ and $b < 0$. Here, we can say that $a = -n$ and $b = -m$, with n, m positive natural numbers. Thus, $x = (-n) \,/\!/\, (-m) = n \,/\!/\, m$ (once again, this latest equality is easy to verify). Thus, x is positive.

   Conclusion: if all four cases, at least one of the three properties is true.

2. Now prove that at most one of those statements is true.

[11] We use implicitly here the fact that $(nm) \,/\!/\, n = m \,/\!/\, 1$ for all integers n, m with $m \neq 0$, which is straightforward to prove.

- Suppose, for the sake of contradiction, that we have both $x = 0$ and x positive. On the one hand, "$x = a // b = 0$" implies that $a = 0$ (see Terence Tao's remark, page 83). On the other hand, "x is positive" implies that $a > 0$. So we would have both $a = 0$ and a positive, which is not compatible with the trichotomy of integers.
- A similar argument holds if we suppose both $x = 0$ and x negative.
- Now suppose that we have both x positive and x negative, i.e. $x = c // d = (-n) // m$, with c, d, n, m positive natural numbers. Thus, we should have $cm = (-n)d$. On the one hand, cm is positive, as the multiplication of two positive natural numbers. On the other hand, $(-n)d = -(nd)$ is negative. The equality $cm = -nd$ is thus impossible.

Conclusion: all three statements are mutually exclusive.

Exercise 4.2.5. — *Prove Proposition 4.2.9.*

Let x, y, z be rational numbers. This proposition includes the following statements:

1. Prove that exactly one of the three statements $x = y$, $x < y$, or $x > y$ is true.

   This statement is very close to Lemma 4.2.7, proved in the previous exercise. Let's consider the rational number $x - y$. According to the trichotomy of rationals, this number can be either zero, positive or negative (exactly one of these statements is true).

   If $x - y = 0$, then $x = y$. If $x - y$ is positive, then $x > y$. And if $x - y$ is negative, then $x < y$. Thus, the order trichotomy is a direct consequence of the ordering of rationals.

2. Prove that one has $x < y$ if and only if $y > x$.

   Since $x < y$, the rational number $x - y$ is negative, and can be written $(-a) // b$ for positive integers a, b. And since $x - y = -a // b$, we have $a // b = y - x$, i.e. $y - x$ is positive, i.e. $y > x$.

3. Prove that if $x < y$ and $y < z$, then $x < z$.

   Since $x < y$, $x - y$ is negative. Similarly, $y - z$ is negative. This means that $x - y$ can be written $(-a) // b$, and $y - z$ can be written $(-c) // d$, with a, b, c, d positive integers.

   On the one hand, their sum is $x - z$. On the other hand, their sum is $((-ad) + (-cb)) // (bd)$. This latest expression is a negative rational, thus we have $x < z$.

4. Prove that if $x < y$, then $x + z < y + z$.

   Suppose that $x < y$. Thus, $x - y$ is negative. But we have, for any rational z, $x - y = (x + z) - (y + z)$, and thus this latter expression is also negative. This means that $x + z < y + z$.

5. Prove that if $x < y$ and z is positive, then $xz < yz$.

   Since $x < y$, the rational number $x - y$ is negative. Furthermore, we know (by Proposition 4.2.4) that $xz - yz = (x - y)z$. In the expression $(x - y)z$, z is supposed to be positive and $x - y$ is negative, thus their product is negative[12]. This means that $xz < yz$.

EXERCISE 4.2.6. — *Show that if x, y, z are rational numbers such that $x < y$ and z is negative, then $xz > yz$.*

If $x < y$, then $x - y$ is negative. Thus, $(x - y)z$ is the product of two negative rationals: it is a positive rational[13].

But $(x-y)z = xz - yz$ by Proposition 4.2.4. And since we have showed that this number is positive, we have $xz > yz$.

Note: in particular, this exercise says that if $x > y$, then $-x < -y$ (with $z = -1$).

EXERCISE 4.3.1. — *Prove Proposition 4.3.3.*

Let x, y, z be rational numbers. The statements to prove are the following:

(a) Show that $|x| \geqslant 0$ for all x, and that $|x| = 0$ iff $x = 0$.

There are three cases:

- if $x = 0$, then $|x| := 0$, thus we have in particular $|x| \geqslant 0$.
- if $x > 0$, then $|x| := x$, thus $|x| > 0$. And in particular, this means that $|x| \geqslant 0$.
- if $x < 0$, then $|x| := -x$, thus $|x| > 0$. And in particular, $|x| \geqslant 0$.

We can note that the only case where $|x| = 0$ is when $x = 0$. Thus, by trichotomy of rationals, $|x| = 0$ iff $x = 0$.

(b) Show that $|x + y| \leqslant |x| + |y|$.

- If $x = 0$ or $y = 0$, this is immediate.
- If $x > 0$ and $y > 0$, $x + y$ is positive, thus $|x + y| = x + y = |x| + |y|$.
- If $x < 0$ and $y < 0$, $x + y$ is negative, thus $|x + y| = -(x + y) = -x - y$. On the other hand, $|x| + |y| = -x - y$.
- Finally, the case where x and y are of opposite signs. Say that x is positive and y negative, but they are exchangeable. On the one hand, $|x| + |y| = x - y > 0$. On the other hand, $|x + y|$ can be either equal to $x + y$ if $x + y > 0$, i.e. if $x > -y$; or equal to $-x - y$ if $x + y < 0$, i.e. if $x < -y$.

  In the first case, since $-y < 0 < y$ by hypothesis, we have $|x + y| = x + y < x - y = |x| + |y|$.

  In the second case, since $-x < 0 < x$ by hypothesis, we have $|x + y| = -x - y < x - y = |x| + |y|$.

Conclusion: in all cases, we have indeed $|x + y| \leqslant |x| + |y|$.

---

[12] This is never explicitly mentioned in the book. However, using Exercise 4.1.3, we know that for every integer a, we have $-a = (-1) \times a$. So let's consider the product $n(-m)$ where n, m are positive integers: this product is $(-1)nm = -(nm)$, and thus is negative.

[13] Similarly, if x and y are negative, then $-x$ and $-y$ are positive, and their product $(-x)(-y) = (-1)(-1)xy = xy$ is positive by definition. This can also be deduced from Proposition 4.2.9(e), by choosing $x = 0$.

(c) Show that $-y \leqslant x \leqslant y$ iff $y \geqslant |x|$. (Thus, in particular, $-|x| \leqslant x \leqslant |x|$.)

- First suppose that $y \geqslant |x|$. Note that, whatever could be the value of x, we have necessarily $y \geqslant 0$ according to the first point of this exercise. Now we can split into three cases.

  If $x = 0$ then $y \geqslant 0$ and the claim is immediate.

  If $x > 0$, then $|x| = x$, and the part $y \geqslant x$ is immediate. Furthermore, the other part $-y \leqslant x$ is also immediate since $-y$ is negative and x is positive.

  If $x < 0$, then $|x| = -x$, thus we have $y \geqslant -x$, i.e. $-y \leqslant x$ according to Exercise 4.2.6. Additionally, the part $x \leqslant y$ is immediate since x is negative and y is positive.
- Conversely, suppose that $-y \leqslant x \leqslant y$. If $x \geqslant 0$, then $|x| = x$, thus the rightmost inequality gives $x = |x| \leqslant y$. In the other case, if $x < 0$, then $|x| = -x$. The leftmost inequality $-y \leqslant x$ leads (according to Exercise 4.2.6) to $y \geqslant -x$, i.e. $y \geqslant |x|$.

(d) Show that $|xy| = |x| \times |y|$. (In particular, $|-x| = |x|$.)

Once again, we can split into several cases, as in the second point of this exercise.

- If $x = 0$ or $y = 0$, both sides of the equality are zero (cf. the first point of this exercise), thus the claim is immediate.
- If $x > 0$ and $y > 0$, the product xy is also positive. Thus, $|xy| = xy$, and $|x| \times |y| = xy$, and the claim follows.
- If $x < 0$ and $y < 0$, then the product xy is positive, and $|xy| = xy$. On the other hand, $|x| \times |y| = (-x)(-y) = xy$, and the claim follows.
- If x and y are of opposite signs (say x positive and y negative, but they are exchangeable), then xy is negative, and $|xy| = -xy$. On the other hand, $|x| \times |y| = -xy$, thus the claim follows.

(e) Show that $d(x, y) \leqslant 0$ for all x, y, and that $d(x, y) = 0$ iff $x = y$.

We have $d(x, y) = |x - y| \geqslant 0$ according to the first point of this exercise. Furthermore, still according to the first point, $d(x, y) = |x - y| = 0$ iff $x - y = 0$, i.e. $x = y$.

(f) Show that $d(x, y) = d(y, x)$.

We have $d(x, y) = |x - y|$ and $d(y, x) = |y - x|$ by definition. But $|y - x| = |-(x - y)| = |x - y|$ according to the fourth point of this exercise.

(g) Show that $d(x, z) = d(x, y) + d(y, z)$.

We have $d(x, z) = |x - z| = |(x - y) + (y - z)| \leqslant |x - y| + |y - z|$ according to the second point of this exercise. The claim follows.

Exercise 4.3.2. — *Prove the remaining claims in Proposition 4.3.7.*

Let x, y, z, w be rational numbers.

(a) If $x = y$, then $d(x, y) = 0$ according to Proposition 4.3.3(e). Thus, $d(x, y) \leqslant \varepsilon$ for any positive number ε.

Conversely, suppose that $d(x, y) \leqslant \varepsilon$ for any $\varepsilon > 0$, and let's prove that $x = y$. Suppose, for the the sake of contradiction, that $x \neq y$; and let be $\varepsilon = |x - y|/2$. Since $x \neq y$, we have $|x - y| > 0$, thus ε is a positive number. Furthermore, $d(x, y) = \varepsilon + \varepsilon$, thus $d(x, y) > \varepsilon$, which is a contradiction.

(b) This is a direct consequence from Proposition 4.3.3(f). Indeed, since $d(x, y) = d(y, x)$, we obviously have $d(y, x) \leqslant \varepsilon$ when $d(x, y) \leqslant \varepsilon$.

(c) Suppose that $d(x, y) \leqslant \varepsilon$ and $d(y, z) \leqslant \delta$. Thus, by triangle inequality, $d(x, z) \leqslant d(x, y) + d(y, z) \leqslant \varepsilon + \delta$.

(d) Suppose that $d(x, y) \leqslant \varepsilon$ and $d(z, w) \leqslant \delta$. Thus,

$$\begin{aligned} d(x + z, y + w) &= |(x + z) - (y + w)| \\ &= |(x - y) + (z - w)| \\ &\leqslant |x - y| + |z - w| \\ &\leqslant \varepsilon + \delta \end{aligned}$$

which means that $x + z$ and $y + w$ are $(\varepsilon + \delta)$-close.

Similarly, $d(x-z, y-w) = |(x - y) + (w - z)|$, and using just the symmetry of distance, we can conclude that $x - z$ and $y - w$ are $(\varepsilon + \delta)$-close according to the previous result.

(e) This is clear: we have $d(x, y) \leqslant \varepsilon < \varepsilon'$.

(f) Since $d(x, y) \leqslant \varepsilon$, we have $-\varepsilon \leqslant y - x \leqslant \varepsilon$. Similarly, we have $-\varepsilon \leqslant z - x \leqslant \varepsilon$.

y and z are exchangeable here, so we can suppose that $y \leqslant w \leqslant z$. From this inequality, we can get $y - x \leqslant w - x \leqslant z - x$. Extending this with the former inequalities, we have:

$$-\varepsilon \leqslant y - x \leqslant w - x \leqslant z - x \leqslant \varepsilon$$

and in particular $-\varepsilon \leqslant w - x \leqslant \varepsilon$, which means $d(w, x) \leqslant \varepsilon$.

(g) We have $d(x, y) = |x - y| \leqslant \varepsilon$. Since z is positive, we have $|z| > 0$, thus $|x - y|\,|z| \leqslant \varepsilon\,|z|$. But according to Proposition 4.3.3(d), $|x - y|\,z = |(x - y)z| = |xz - yz|$. Thus, $|xz - yz| \leqslant \varepsilon\,|z|$, i.e., xz and yz are $\varepsilon\,|z|$-close.

Exercise 4.3.3. — *Prove Proposition 4.3.10.*

Let x, y be rationals, and n, m be natural numbers. The claims to prove are the following (they are re-ordered and re-numbered here):

(a) Show that $x^n x^m = x^{n+m}$. We induct on n while keeping m fixed.

For the base case $n = 0$, we have on the one hand $x^n x^m = x^0 x^m = 1 \cdot x^m = x^m$. On the other hand, $x^{n+m} = x^{0+m} = x^m$. Thus, both sides are equal, and the base case is done.

Now suppose that $x^n x^m = x^{n+m}$, and let's show that $x^{n+1}x^m = x^{(n+1)m}$. We have:

$$\begin{aligned}
x^{n+1}x^m &= (x^n x)x^m && \text{(by Definition 4.3.9)}\\
&= x^n x^m x && \text{(by associativity and commutativity of multiplication)}\\
&= x^{n+m}x && \text{(by induction hypothesis)}\\
&= x^{n+m+1} && \text{(by Definition 4.3.9 once again)}
\end{aligned}$$

This closes the induction.

(b) Show that $(xy)^n = x^n y^n$. Let's induct on n. The base case $n = 0$ is obvious, since both sides are equal to 1. Now suppose inductively that $(xy)^n = x^n y^n$. Thus we have:

$$\begin{aligned}
(xy)^{n+1} &= (xy)^n(xy) && \text{(by Definition 4.3.9)}\\
&= x^n y^n xy && \text{(by inductive hypothesis)}\\
&= x^n x y^n y && \text{(by commutativity of multiplication)}\\
&= x^{n+1}y^{n+1} && \text{(by Definition 4.3.9 once again)}
\end{aligned}$$

(c) Show that $(x^n)^m = x^{nm}$. We induct on n while keeping m fixed.

For the base case $n = 0$, we have $(x^n)^m = 1^m = 1$, since $1^m = 1$ for all natural number m[14]. On the other hand, $x^{nm} = x^{0m} = 1$. Thus, both sides are equal, and the base case is done.

Now suppose inductively that $(x^n)^m = x^{nm}$. Then we have:

$$\begin{aligned}
(x^{n+1})^m &= (x^n x)^m && \text{(by Definition 4.3.9)}\\
&= (x^n)^m x^m && \text{(proved in 2. from this exercise)}\\
&= x^{nm}x^m && \text{(by inductive hypothesis)}\\
&= x^{nm+m} && \text{(proved in 1. from this exercise)}\\
&= x^{(n+1)m}
\end{aligned}$$

This closes the induction.

(d) Show that if $n > 0$, then $x^n = 0$ iff $x = 0$. For that, let's induct on n. Here the base case starts with $n = 1$ since we suppose $n > 0$. For $n = 1$, $x^1 = x$, thus we obviously have $x^1 = 0 \Leftrightarrow x = 0$ since both objects are equal.

Now suppose inductively that $x^n = 0$ iff $x = 0$. We must show that $x^{n+1} = 0$ iff $x = 0$. Here we'll need the following lemma:

Lemma. *Let x, y be rational numbers. Then, if $xy = 0$, we have either $x = 0$ or $y = 0$.*

Proof. Let's denote $x = a /\!/ b$ and $y = c /\!/ d$. By Definition 4.2.2, $xy = (ac) /\!/ (bd)$. Thus, since $xy = 0$, we have $ac = 0$ (see Tao's remark p. 83). And, by Proposition 4.1.8, we have either $a = 0$ or $c = 0$. In the first case, this means that $x = 0$; in the second case this means that $y = 0$. □

Now go back to the main proof. First, if $x = 0$, we have $x^{n+1} = x^n x = 0^n \times 0 = 0$. Conversely, if $x^{n+1} = 0$, then $x^n x = 0$. According to the previous lemma, this means that either $x^n = 0$ or $x = 0$. In the second case, we are done. In the first case, the induction hypothesis also allows to conclude that $x^n = 0$. This closes the induction.

[14] This can easily be proved by induction, which we'll not write formally here.

(e) Show that if $x \geqslant y \geqslant 0$, then $x^n \geqslant y^n \geqslant 0$. Let's induct on n.

For the base case $n = 0$, $x^0 = y^0 = 1$. Thus in particular we have indeed $x^0 \geqslant y^0 \geqslant 0$.

Now suppose inductively that $x^n \geqslant y^n \geqslant 0$, and show that $x^{n+1} \geqslant y^{n+1} \geqslant 0$. We start from $x^n \geqslant y^n \geqslant 0$ and multiply all terms by x (which preserves inequality since x is supposed to be positive): we get $x^{n+1} \geqslant xy^n \geqslant 0$. If we start from $x \geqslant y \geqslant 0$ and multiply all terms by y^n (which is also positive by induction hypothesis), we get $y^n x \geqslant y^{n+1} \geqslant 0$. Now combine all those inequalities:

$$x^{n+1} \geqslant xy^n \geqslant y^{n+1} \geqslant 0$$

This closes the induction.

(f) Show that $|x^n| = |x|^n$. Let's induct on n.

For the base case $n = 0$, we have $|x^n| = |x^0| = |1| = 1$; and $|x|^n = |x|^0 = 1$. Thus both sides are equal, and the base case is done.

Now suppose that $|x^n| = |x|^n$ and show that $|x^{n+1}| = |x|^{n+1}$. We have:

$$\begin{aligned} |x^{n+1}| &= |x^n x| && \text{(by Definition 4.3.9)} \\ &= |x^n| \cdot |x| && \text{(by Proposition 4.3.3d)} \\ &= |x|^n \cdot |x| && \text{(by inductive hypothesis)} \\ &= |x|^{n+1} \end{aligned}$$

This closes the induction.

Exercise 4.3.4. — *Prove Proposition 4.3.12.*

This is essentially the same exercise as 4.3.3, but dealing with integer exponents (instead of natural exponents). The claims to prove are the following (and once again, they are relabeled):

(a) Prove that $x^n x^m = x^{n+m}$. Let's distinguish three cases:

- If $n, m \geqslant 0$, then this is simply Proposition 4.3.10(a).
- If $n, m < 0$, then $n = -p$ and $m = -q$ with p, q positive natural numbers. Thus, $x^n x^m = (1/x^p) \cdot (1/x^q) = 1/(x^p x^q)$ by Definition 4.2.2. But since p, q are positive, this is also equal to $1/(x^{p+q})$ according to Proposition 4.3.10(a). This can also be written $x^{-(p+q)}$ by Definition 4.3.11, which is finally equal to x^{n+m}.
- If $n \geqslant 0$ and $m < 0$ (or inversely, since they are exchangeable), then $m = -q$ with q a positive natural number. Thus, $x^n x^m = x^n \times (1/x^q) = x^n/x^q$. We will (once again) split into two cases:
  - if $n \geqslant -m$, i.e. if $n - q \geqslant 0$, then we can note that $x^{n-q} \cdot x^q = x^n$ according to Proposition 4.3.10(a). Thus, let's multiply both sides of this equality by x^{-q} to get $x^{n-q} = x^n x^{-q}$; which can be rewritten $x^{n+m} = x^n x^m$ as required.
  - if $n < -m$, i.e. $q - n > 0$, then we can note that $x^{q-n} x^n = x^q$ according to Proposition 4.3.10(a). Also, since $n - q < 0$, according to Definition 4.3.11, we have $x^{n-q} = 1/x^{q-n}$. Let's multiply both sides by $1/x^n$, to get $x^{n-q}/x^n = 1/(x^{q-n}x^n) = 1/x^q = x^{-q}$. Finally, multiply both sides by x^n to get $x^{n+m} = x^n x^m$.

(b) Prove that $(x^n)^m = x^{nm}$.

(c) Show that $(xy)^n = x^n y^n$. If $n \geqslant 0$, this is simply Proposition 4.3.10(a). So let's consider the case $n < 0$. In this case, $n = -p$, with p a positive natural number. Thus we have successively:

$$\begin{array}{rcll}
(xy)^n & = & (xy)^{-p} & \\
& = & 1/(xy)^p & \text{(by Definition 4.3.11)} \\
& = & 1/(x^p y^p) & \text{(Proposition 4.3.10(a))} \\
& = & 1/(x^p) \times 1/(y^p) & \text{(Definition 4.2.2)} \\
& = & x^{-p} \times y^{-p} & \text{(by Definition 4.3.11)} \\
& = & x^n \times y^n &
\end{array}$$

(d) Show that if $x \geqslant y > 0$, then $x^n \geqslant y^n > 0$ if n is positive, and $0 < x^n \leqslant y^n$ if n is negative.

- If $n > 0$, according to Proposition 4.3.10(c), we already have $x^n \geqslant y^n \geqslant 0$, so that we just have to show that the rightmost inequality is strict, i.e. that $y^n > 0$. To show that, we only need to prove $y^n \neq 0$. For the sake of contradiction, let's suppose that $y^n = 0$. Our starting hypothesis was $x \geqslant y > 0$, thus we know that $y \neq 0$. According to Proposition 4.3.10(b), we can't have both $y \neq 0$ and $y^n = 0$, this is a contradiction. Thus, we indeed have $y^n \neq 0$, which shows the inequality $x^n \geqslant y^n > 0$ as required.
- If $n < 0$, this includes an important result, which is that taking the inverse reverses order. Indeed, let's begin by proving that if $x \geqslant y > 0$, then $1/x \leqslant 1/y$. Since both x and y are positive, their product xy is also positive, and $1/(xy)$ is also positive. Following Proposition 4.2.9(e), we can multiply both sides of $x \geqslant y$ by $1/(xy)$ to get $1/y \geqslant 1/x$. Then, we immediately get $(1/y)^p \geqslant (1/x)^p$ for any positive number p by Proposition 4.3.10(c), which can be rewritten $y^n \geqslant x^n$ with $n = -p$ negative. And since both numbers are positive (because x and y are positive), the claim follows.

(e) Prove that if $x, y > 0$ and $n \neq 0$, then $x^n = y^n \implies x = y$. Let's consider two cases: $n > 0$ and $n < 0$.

First, if $n > 0$, suppose for the sake of contradiction that we have both $x^n = y^n$ and $x \neq y$. According to the trichotomy of rationals (Lemma 4.2.7), this last claim means that we have either $x > y$ or $y > x$. Since x and y are exchangeable, we only prove the first case here, $x > y$. In this case, Proposition 4.3.10(c) leads to $x^n > y^n$, which is obviously not compatible with our initial hypothesis $x^n = y^n$. A similar contradiction follows in the case $y > x$. Thus, both $x > y$ and $y > x$ are impossible, and the only possibility is $x = y$.

Now, if $n < 0$, then $n = -p$, with p a positive natural number. Suppose that $x^n = y^n$, i.e. that $x^{-p} = y^{-p}$, or finally $1/x^p = 1/y^p$. From this last equality, by multiplying both sides by $x^p y^p$, we get $y^p = x^p$. We are thus back in the previous case, and obtain $x = y$.

(f) Prove that $|x^n| = |x|^n$. If $n \geqslant 0$, this is simply Proposition 4.3.10(d). So let's consider the case $n < 0$. We'll need a quick lemma:

Lemma. *For all rationals $x \neq 0$, we have $|1/x| = 1/|x|$.*

Proof. If $x > 0$, there is nothing to show. If $x < 0$, then $1/x$ is also negative[15]. Thus, $1/|x| = 1/(-x)$; and $|1/x| = -(1/x)$. And we have clearly $1/(-x) = -(1/x)$ because $1/(-x) + 1/x = 0$. □

In this case, $n = -p$, with p a positive natural number. We have successively:

$$\begin{array}{rcll} |x^n| & = & |x^{-p}| & \\ & = & |1/(x^p)| & \text{(by Definition 4.3.11)} \\ & = & |(1/x)^p| & \text{(Proposition 4.3.12(a))} \\ & = & |1/x|^p & \text{(Proposition 4.3.10(d))} \\ & = & (|1|/|x|)^p & \text{(lemma introduced just above)} \\ & = & 1/|x|^p & \text{(Proposition 4.3.12(a))} \\ & = & |x|^{-p} & \text{(Definition 4.3.11)} \\ & = & |x|^n & \end{array}$$

Exercise 4.3.5. — *Prove that $2^N \geqslant N$ for all positive integers N.*

Let's use induction on N. Since we only consider positive integers, we have here $N \geqslant 1$, and in particular, the base case starts at $N = 1$.

For the base case $N = 1$, the assertion is true, since we have indeed $2^1 \geqslant 1$.

Now suppose inductively that $2^N \geqslant N$, and show that $2^{N+1} \geqslant N + 1$. We have $2^{N+1} = 2^N \times 2 \geqslant N \times 2$ by induction hypothesis. But we know that $2N = N + N$ (recall Definition 2.3.1 for instance), thus we can rewrite this as $2^{N+1} \geqslant N + N$. And since $N \geqslant 1$, we finally get $2^{N+1} \geqslant N + 1$.

Exercise 4.4.1. — *Prove Proposition 4.4.1.*

We have to prove that, for any rational number x, there exists an integer n such that $n \leqslant x < n + 1$. Let's proceed through the following four steps:

- Suppose that $x \in \mathbb{Q}_+$. Thus, $x = a/b$, with a and b natural numbers. According to Proposition 2.3.9, there exists $n, r \in \mathbb{N}$ such that $a = bn + r$, with $0 \leqslant r < b$. By dividing all terms by b, this also means that $x = a/b = n + r/b$, with $0 \leqslant r/b < 1$.

  Since $0 \leqslant r/b < 1$, we have $n \leqslant n + r/b < n + 1$, i.e. $n \leqslant x < n + 1$, as required.

- Now suppose that $x \in \mathbb{Q}_-^\star$. Consequently, $-x \in \mathbb{Q}_+$, and we are back in the previous case: there exists a natural number n such that $n \leqslant -x < n + 1$, i.e. $-n - 1 < x \leqslant -n$. Now we have two possible cases:

  - if $x = -n$, then let be $m = -n$. Thus, $m - 1 < x \leqslant m$, and then $m \leqslant x < m + 1$, as required.
  - if $x \neq -n$, then let be $m = -n - 1$. Thus, $m < x \leqslant m + 1$, i.e. $m - 1 \leqslant x < m$. And by denoting $p = m - 1$, we have $p \leqslant x < p + 1$ as required.

[15]Formally, see Definition 4.2.6, and note that $a \,//\, (-b) = (-a) \,//\, b$ if a and b are positive integers.

- Let's prove that this integer n is unique. Suppose that we have two integers m, n such that:

$$n \leqslant x < n+1 \tag{4.1}$$

$$m \leqslant x < m+1 \tag{4.2}$$

  From (4.2), we also have $-m-1 < -x \leqslant -m$. And, by adding this inequality to (4.1), we get $n-m-1 < 0 < n-m+1$. The left-hand side says that $n < m+1$, i.e. that $n \leqslant m$ (recall Proposition 2.2.12 (e)). Similarly, the right-hand side says that $n > m-1$, i.e. that $n \geqslant m$. Thus, we have both $n \leqslant m$ and $n \geqslant m$, which means that $n = m$.

- Finally, this means in particular that there exists a natural number N such that $N > x$. Indeed, if x is negative, then $N = 0$ is suitable; and if x is positive, then N is directly given by $N = \lfloor x \rfloor + 1$.

Exercise 4.4.2. — *A sequence $a_0, a_1, a_2, \ldots$ of numbers (natural numbers, integers, rationals, or reals) is said to be in* infinite descent *if we have $a_n > a_{n+1}$ for all natural numbers n (i.e., $a_0 > a_1 > a_2 > \ldots$).*

1. *Prove the principle of infinite descent: that it is not possible to have a sequence of natural numbers which is in infinite descent.*
2. *Does the principle of infinite descent work if the sequence $a_1, a_2, a_3, \ldots$ is allowed to take integer values instead of natural number values? What about if it is allowed to take positive rational values instead of natural numbers? Explain.*

We follow the hints given by Terence Tao.

1. Assume for the sake of contradiction that we have a sequence of natural numbers (a_n) which is in infinite descent. Let k be a natural number, and P_k be the property "$a_n \geqslant k$ for all natural numbers n". Let's induct on k.

   For the base case, P_0 is true since a_n are natural numbers for all n, so that $a_n \geqslant 0$ for all n by definition.

   Now let's suppose inductively that P_k is true, i.e. that $a_0 > a_1 > a_2 > \ldots \geqslant k$. If we had $a_p = k$ for one given natural number p, then we would have $k = a_p > a_{p+1}$. But also, $a_{p+1} > a_{p+2} > \ldots > k$ by induction hypothesis. However, the inequality $k \geqslant a_{p+1} > k$ is a contradiction, so that $a_n \neq k$ for all n. Thus, P_{k+1} is also true: we have $a_n > k+1$ for all n.

   However, having $a_n > k$ for all natural numbers k, n is a contradiction. Indeed, for $k = a_0$ and $n = 1$, we have $a_1 > a_0$, which contradicts the fact that (a_n) is in infinite descent.

   Thus, there are no such sequence of natural numbers.

2. A general note: to prove that the infinite descent principle does not work for integers or rationals, it is enough to find *one* sequence of such numbers which is actually in infinite descent. Instead of a formal proof as in the previous case, a simple counterexample will do the trick.

- If the sequence $a_0 > a_1 > \dots$ can take integer values, lets define the sequence by $a_n = -n$. By definition, we have $a_n > a_{n+1}$ for all natural number n (since $-n > -n-1$, as a simple induction will show).
- If the sequence $a_0 > a_1 > \dots$ can take rational values, lets define the sequence by $a_n = 1/n$. Thus, we have $a_n > a_{n+1}$ for all natural number n, since $1/n > 1/(n+1)$. (This can be shown as follows: $1/n - 1/(n+1) = 1/(n(n+1)) > 0$.)

5. The real numbers

Exercise 5.1.1. — *Prove Lemma 5.1.15, i.e. that every Cauchy sequence is bounded.*

Let $(a_n)_{n=1}^{\infty}$ be a Cauchy sequence.

- By definition 5.1.8, for every rational $\varepsilon > 0$, there exists a natural number N such that if $j, k \geqslant N$, then $d(a_j, a_k) \leqslant \varepsilon$. In particular, let's rephrase this statement with the arbitrary value $\varepsilon = 1$ (valid, since 1 is a positive rational): there exists a natural number N such that if $j, k \geqslant N$, then $|a_j - a_k| \leqslant 1$.

  Since $N \geqslant N$, we can take in particular $k = N$ to get yet another particular formulation: if $j \geqslant N$, then $|a_j - a_N| \leqslant 1$.

  According to Proposition 4.3.3(b), we have $|x + y| \leqslant |x| + |y|$ for all rationals x, y. Let's consider $x = a_j - a_N$ and $y = a_N$: this leads to $|a_j| \leqslant |a_j + a_N| + |a_N|$, i.e. $|a_j| - |a_N| \leqslant |a_j - a_N|$.

  Thus, this means that $|a_j| - |a_N| \leqslant |a_j - a_N| \leqslant 1$ as soon as $j \geqslant N$, i.e. that $|a_j| \leqslant 1 + |a_N|$ for $j \geqslant N$. We have bounded part of the infinite sequence.

- The other part is simply the finite sequence $a_0, a_1, \cdots, a_{N-1}$. By Lemma 5.1.14, this finite sequence is necessarily bounded by a rational number M.

- Finally, let's consider the rational number $B = 1 + |a_N| + M$. Since we have both $B \geqslant M$ and $B \geqslant 1 + |a_N|$, both the infinite sequence $(a_n)_{n=N}^{\infty}$ and the finite sequence $a_0, \cdots, a_{N-1}$ are bounded by B. Thus, the whole Cauchy sequence $(a_n)_{n=1}^{\infty}$ is bounded by B.

Exercise 5.2.1. — *Show that if $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are eventually ε-close, then $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence if and only if $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence.*

First note that $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are exchangeable here, so that showing only one direction ("if $(a_n)_{n=1}^{\infty}$ is Cauchy, then $(b_n)_{n=1}^{\infty}$ is Cauchy") will be enough.

Let be $\varepsilon > 0$ a positive rational. Since $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are eventually ε-close, there exists a natural number N_1 such that $n \geqslant N_1 \Longrightarrow |a_n - b_n| \leqslant \frac{\varepsilon}{3}$. Furthermore, since $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence, there exists a natural number N_2 such that $j, k \geqslant N_2 \Longrightarrow |a_j - a_k| \leqslant \frac{\varepsilon}{3}$.

Let be $N = \max(N_1, N_2)$. If $j, k \geqslant N$, then we have:

$$\begin{aligned}
|b_j - b_k| &= |b_j - a_j + a_j - a_k + a_k - b_k| \\
&\leqslant |b_j - a_j| + |a_j - a_k| + |a_k - b_k| \quad \text{(by triangle inequality)} \\
&\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&\leqslant \varepsilon
\end{aligned}$$

which means that $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence.

Exercise 5.2.2. — *Let $\varepsilon > 0$. Show that if $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are eventually ε-close, then $(a_n)_{n=1}^{\infty}$ is bounded if and only if $(b_n)_{n=1}^{\infty}$ is bounded.*

As in the previous exercise, $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are exchangeable here, so that showing only one direction will be enough.

- Since $(a_n)_{n=1}^{\infty}$ is bounded, there exists a rational number M_1 such that $|a_n| \leqslant M_1$ for all natural n.

- Since $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are eventually ε-close, there exists a positive natural number N such that $n \geqslant N \implies |a_n - b_n| \leqslant \varepsilon$.

- Let's decompose $(b_n)_{n=1}^{\infty}$ into a finite and an infinite part, and show that both parts are bounded.

  By Lemma 5.1.14, there exists a positive rational M_2 such that the finite sequence $b_0, \ldots, b_{N-1}$ is bounded by M_2.

  Furthermore, we know by triangle inequality that, if $n \geqslant N$, we have $|b_n| - |a_n| \leqslant |b_n - a_n| \leqslant \varepsilon$. Consequently, $|b_n| \leqslant |a_n| + \varepsilon \leqslant M_1 + \varepsilon$.

  Finally, let be $M = M_1 + M_2 + \varepsilon$: we have indeed $|b_n| \leqslant M$ for all natural n.

Exercise 5.3.1. — *Prove Proposition 5.3.3.*

The three laws of equality must be verified.

1. Reflexivity: let's prove that $x = x$. By Definition 5.3.1, we have $x = x$ if and only if $\mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} a_n$, i.e. if and only if $(a_n)_{n=1}^{\infty}$ and $(a_n)_{n=1}^{\infty}$ are equivalent Cauchy sequences. Let be $\varepsilon > 0$ a positive rational, and let be $N = 1$. For all $n \geqslant N$, we have $|a_n - a_n| = 0 \leqslant \varepsilon$, QED.

2. Symmetry: let's suppose that $x = y$, and let's show that $y = x$. Let $\varepsilon > 0$ be a positive rational. We have:

$$\begin{aligned} x = y &\implies (a_n)_{n=1}^{\infty} \text{ and } (b_n)_{n=1}^{\infty} \text{ are equivalent Cauchy sequences} \\ &\implies \exists N \geqslant 1 \text{ such that } |a_n - b_n| \leqslant \varepsilon \text{ for } n \geqslant N \\ &\implies \exists N \geqslant 1 \text{ such that } |b_n - a_n| \leqslant \varepsilon \text{ for } n \geqslant N \\ &\implies (b_n)_{n=1}^{\infty} \text{ and } (a_n)_{n=1}^{\infty} \text{ are equivalent Cauchy sequences} \\ &\implies y = x \end{aligned}$$

3. Transitivity: if $x = y$ and $y = z$. Let be $\varepsilon > 0$ a positive rational. Since $x = y$, there exists $N_1 \geqslant 1$ such that $|a_n - b_n| \leqslant \varepsilon/2$ for $n \geqslant N_1$. Since $y = z$, there exists $N_2 \geqslant 1$ such that $|b_n - c_n| \leqslant \varepsilon/2$ for $n \geqslant N_2$. Thus, if $n \geqslant \max(N_1, N_2)$, a_n and b_n are $\varepsilon - 2$-close, and b_n and c_n are $\varepsilon - 2$-close. Thus, according to Proposition 4.3.7(c), a_n and c_n are ε-close, which means that $(a_n)_{n=1}^{\infty}$ and $(c_n)_{n=1}^{\infty}$ are eventually ε-close for all ε, i.e. are equivalent Cauchy sequences. This closes the proof.

Exercise 5.3.2. — *Prove Proposition 5.3.10.*

To prove that xy is a real number, we must show that $(a_n b_n)_{n=1}^{\infty}$ is a Cauchy sequence. Let $\varepsilon > 0$ be a positive rational. We must show that there exists a natural number N such that $j, k \geqslant N \implies |a_j b_j - a_k b_k| \leqslant \varepsilon$.

We already know that:

- Since $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence, it is bounded by a rational number M_a. Furthermore, there exists a natural number N_a such that $j, k \geqslant N_a \implies |a_j - a_k| \leqslant \frac{\varepsilon}{2M_a}$.

- In a similar fashion, since $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence, it is bounded by a rational number M_b. Furthermore, there exists a natural number N_b such that $j, k \geqslant N_b \implies |b_j - b_k| \leqslant \frac{\varepsilon}{2M_b}$.

Now let's consider $N = \max(N_a, N_b)$. If $j, k \geqslant N$, we have:

$$\begin{aligned}
|a_j b_j - a_k b_k| &= |a_j b_j - a_j b_k + a_j b_k - a_k b_k| \\
&\leqslant |a_j| \cdot |b_j - b_k| + |b_k| \cdot |a_j - a_k| \\
&\leqslant M_a \frac{\varepsilon}{2M_a} + M_b \frac{\varepsilon}{2M_b} \\
&\leqslant \varepsilon
\end{aligned}$$

This proves that $(a_n b_n)_{n=1}^{\infty}$ is a Cauchy sequence, as required.

Exercise 5.3.3. — *Let a, b be rational numbers. Show that $a = b$ if and only if $LIM_{n\to\infty} a = LIM_{n\to\infty} b$ (i.e., the Cauchy sequences $a, a, a, a, \ldots$ and $b, b, b, b, \ldots$ are equivalent if and only if $a = b$).*

In what follows, we denote $(a_n)_{n=1}^{\infty}$ the constant sequence $a, a, a, \ldots$, and $(b_n)_{n=1}^{\infty}$ the constant sequence $b, b, b, \ldots$.

- If $a = b$, we have $|a - b| = 0$, i.e. $|a_n - b_n| = 0$ for all $n \geqslant 1$. Thus, $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are ε-close for all $\varepsilon > 0$: they are equivalent Cauchy sequences. This means that $a = b$.
- If $\mathrm{LIM}_{n\to\infty} a = \mathrm{LIM}_{n\to\infty} b$, let's suppose (for the sake of contradiction) that $a \neq b$, and let's denote $\varepsilon = \frac{|a-b|}{2}$. Since $a \neq b$, we have $\varepsilon > 0$., and we also have $|a - b| > \varepsilon$. In other words, for all $n \geqslant 1$, we have found an ε such that $|a_n - b_n| > \varepsilon$: this is a contradiction with the fact that $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent Cauchy sequences. Thus, $a = b$.

Exercise 5.3.4. — *Let $(a_n)_{n=1}^{\infty}$ be a sequence of rational numbers which is bounded. Let $(b_n)_{n=1}^{\infty}$ be another sequence of rational numbers which is equivalent to $(a_n)_{n=1}^{\infty}$. Show that $(b_n)_{n=1}^{\infty}$ is also bounded.*

This exercise is actually very close to Exercise 5.2.2, and the same proof could apply. But for short: in Exercise 5.2.2, we showed that if two sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are ε-close for a given positive rational ε, then $(a_n)_{n=1}^{\infty}$ is bounded if and only if $(b_n)_{n=1}^{\infty}$ is bounded. Here, since $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent, they are ε-close for any positive ε, and $(a_n)_{n=1}^{\infty}$ is bounded by hypothesis. Thus, $(b_n)_{n=1}^{\infty}$ is also bounded.

Exercise 5.3.5. — *Show that $LIM_{n\to\infty} 1/n = 0$.*

As a real number, by definition, 0 is the formal limit of the constant sequence $0, 0, 0, \ldots$. Thus, we must show that $\mathrm{LIM}_{n\to\infty} 1/n = \mathrm{LIM}_{n\to\infty} 0$. Still by definition, we must prove that the sequence $(1/n)_{n=1}^{\infty}$ is equivalent to the sequence $(0)_{n=1}^{\infty}$. This can be achieved by proving that for all $\varepsilon > 0$, there exists a natural $N \geqslant 1$ such that $n \geqslant N \implies |1/n - 0| \leqslant \varepsilon$, i.e. such that $n \geqslant 1/\varepsilon$.

By Proposition 4.4.1, there always exists a natural number N such that $N > 1/\varepsilon$. Then, if $n \geqslant N$, we have the desired proprty, which closes the proof.

Exercise 5.4.1. — *Prove Proposition 5.4.4.*

Let x be a real number.

1. First we show that at most one of the three statements above is true. To do this, we show that all those statements are "pairwise incompatible".

   - First suppose that we have both $x = 0$ and x positive. If $x = 0$, then x is the formal limit of the sequence $0, 0, \ldots$. If x is positive, then $x = \mathrm{LIM}_{n\to\infty} a_n$ with $a_n \geqslant c$ for all n and a certain rational $c > 0$. By Definition 5.3.1, this implies that both sequences $0, 0, \ldots$ and $(a_n)_{n=1}^{\infty}$ must be equivalent Cauchy sequences. I.e., for all $\varepsilon > 0$, there must exist $M \geqslant 1$ such that $|a_n - 0| \leqslant \varepsilon$ for all $n \geqslant M$. But taking $\varepsilon = c/2$ leads to an obvious contradiction: we cannot have both $a_n \geqslant c$ and $|a_n| \leqslant c/2$ for all $n \geqslant M$. Thus, x cannot be both zero and positive.
   - A similar argument shows that x cannot be both zero and negative.
   - Finally, we show that x cannot be both positive and negative. If x is positive, then $x = \mathrm{LIM}_{n\to\infty} a_n$, with $(a_n)_{n=1}^{\infty}$ positively bounded away from 0, i.e. $a_n \geqslant c$ for a certain rational $c > 0$. Similarly, if x is negative, then $x = \mathrm{LIM}_{n\to\infty} b_n$, with $(b_n)_{n=1}^{\infty}$ negatively bounded away from 0, i.e. $b_n \leqslant -d$, or $-b_n \geqslant d$, for a certain rational $d > 0$. But then, according to Definition 5.3.1, $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ should be equivalent Cauchy sequences, i.e. for any rational $\varepsilon > 0$, there should exist $M \geqslant 1$ such that $|a_n - b_n| \leqslant \varepsilon$ for $n \geqslant M$. But since $a_n - b_n$ is positive, this also can be written simply $a_n - b_n \leqslant \varepsilon$. But since we have both $a_n \geqslant c$ and $-b_n \geqslant d$, we know that $a_n - b_n \geqslant c + d$ for all natural numbers n, an obvious contradiction (take $\varepsilon = (c + d)/2$).
   - Thus, at most one of the three statements is true.

2. Next we show that at least one of the statements "x is zero", "x is positive", "x is negative" is true. Actually, if $x = 0$, we know that the statement "x is zero" is true, so we're done. We can thus suppose that $x \neq 0$, and we have just to show that x is either positive or negative.

   If $x \neq 0$, then by Lemma 5.3.14, $x = \mathrm{LIM}_{n\to\infty} a_n$ with $(a_n)_{n=1}^{\infty}$ bounded away from 0, i.e. there exists $c > 0$ such that $|a_n| \geqslant c$ for any natural n.

   It turns out that, for any Cauchy sequence $(a_n)_{n=1}^{\infty}$, if $(a_n)_{n=1}^{\infty}$ is bounded away from 0, it is either eventually positively bounded away from 0, or eventually negatively bounded away from 0.

   Indeed, suppose that we have at the same time $|a_n| \geqslant \varepsilon$ for all $n \geqslant N$, $a_k \geqslant \varepsilon$ for some $k \geqslant N$ and $a_j \leqslant -\varepsilon$ for some $j \geqslant N$. In such a case, we would have by triangular inequality:
   $$|a_k - a_j| \geqslant ||a_k| - |a_j|| \geqslant |\varepsilon + \varepsilon| \geqslant 2\varepsilon$$

   Thus, for all N, we could find two indexes $j, k \geqslant N$ such that $|a_k - a_j| > \varepsilon$, which contradicts that fact that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence. Consequently, the Cauchy sequence $(a_n)_{n=1}^{\infty}$ is either eventually positively bounded away from 0, or eventually negatively bounded away from 0.

   In the first case, $x = \mathrm{LIM}_{n\to\infty} a_n$ is a positive number[16], and in the second case it is a negative number.

[16] Actually, we also have to define an equivalent sequence $(b_n)_{n=1}^{\infty}$ defined by $a_n = b_n$ if $n \geqslant N$ and $b_n = 1$ otherwise to get what we want.

Exercise 5.4.2. — *Prove the remaining claims in Proposition 5.4.7.*

Let x, y, z be real numbers. Those claims are as follows:

(a) (Order trichotomy) Prove that exactly one of the three statements $x = y$, $x < y$, or $x > y$ is true.

This is actually a simple rephrasing of the trichotomy of real numbers. Indeed, let's consider the real number $x - y$. According to Proposition 5.4.4, this real number is either null (in this case, $x = y$), positive (in this case, $x > y$) or negative (in this case, $x < y$).

(b) (Order is anti-symmetric) Prove that one has $x < y$ if and only if $y > x$.

The following statements are equivalent:

$$\begin{aligned} x < y &\iff x - y \text{ is positive} \\ &\iff -(x - y) \text{ is negative (Prop. 5.4.4)} \\ &\iff y - x \text{ is negative} \\ &\iff y > x \end{aligned}$$

(c) (Order is transitive) Prove that if $x < y$ and $y < z$, then $x < z$.

Note that $x - z = (x - y) + (y - z)$. In this sum, both terms are negative, so that their sum is also negative (this can easily be deduced from Proposition 5.4.4). Thus, $x - z$ is negative, which closes the proof.

(d) (Addition preserves order) Prove that if $x < y$, then $x + z < y + z$.

Note that if $x < y$, then $x - y$ is negative. Since $x - y = (x + z) - (y + z)$, we get the statement required.

(e) is already proven in the main text.

Exercise 5.4.4. — *Show that for any positive real number $x > 0$ there exists a positive integer N such that $x > 1/N > 0$.*

Let's use the archimedean property of the reals (Proposition 5.4.13), with $\varepsilon = 1$. Since x is a positive real number, there exists a positive integer N such that $Nx > 1$. Multiplying this inequality by the (positive[17]) rational $1/N$ leads to $x > 1/N$. Since we have $1/N > 0$, the claim follows.

Exercise 5.4.5. — *Prove Proposition 5.4.14: Given any two real numbers $x < y$, we can find a rational number q such that $x < q < y$.*

Following the hint given by Terence Tao, we can make use of Exercise 5.4.4. Actually, we do have a positive real number here: since $x < y$, then $y - x$ is positive. Thus, according to Exercise 5.4.4, we can find a positive integer N such that $y - x > 1/N > 0$.

Now let's multiply all terms by N. Since N is positive, we get the following inequality: $Ny - Nx > 1 > 0$. Intuitively: since $Ny - Nx > 1$, we should be able to find explicitly an integer lying between them. Then, dividing by N will provide the desired inequality.

[17] According to Proposition 5.4.8.

Nx is a real number. According to Exercise 5.4.3, there exists an integer n such that $n+1 > Nx \geqslant n$. In particular, since $Nx \geqslant n$, we also have $Nx+1 \geqslant n+1$.

Thus, gathering all the inequalities we know:

$$Ny > Nx + 1 \geqslant n + 1 > Nx \tag{5.1}$$

Then let's divide by N: we finally get $y > \frac{n+1}{N} > x$, which is the required property with $q = \frac{n+1}{N}$.

Exercise 5.4.6. — *Let x, y be real numbers and let $\varepsilon > 0$ be a positive real. Show that $|x-y| < \varepsilon$ if and only if $y - \varepsilon < x < y + \varepsilon$; and that $|x-y| \leqslant \varepsilon$ iff $y - \varepsilon \leqslant x \leqslant y + \varepsilon$.*

We only give the proof for the strict version; the other one being totally similar.

- First suppose that $|x-y| < \varepsilon$, and let's consider two cases depending on the sign of $x - y$.
  - If $x - y \geqslant 0$, i.e. if $x \geqslant y$, then $|x-y| = x - y < \varepsilon$ by hypothesis. If we add y to both sides of this inequality, we get $x < y + \varepsilon$, which is the first part of the required result. Furthermore, we know that $y \leqslant x$. And since ε is positive, we have $y - \varepsilon < y \leqslant x$. Thus, by combining all those results, we finally get $y - \varepsilon < y \leqslant x \leqslant y + \varepsilon$, as required.
  - If $x - y < 0$, i.e. if $x < y$, then $|x-y| = y - x < \varepsilon$ by hypothesis. This leads to $y - \varepsilon < x$, which is the first part of the result. Also, since $x < y$, we have $x < y < y + \varepsilon$. Combining all those results, we finally have $y - \varepsilon < x < y < y + \varepsilon$ as required.
- Conversely, suppose that $y - \varepsilon < x < y + \varepsilon$. Adding $(-y)$ to each part leads to $-\varepsilon < x - y < \varepsilon$. There are now three cases depending on the sign of $x - y$:
  - If $x - y > 0$, then $|x-y| = x - y < \varepsilon$ as required.
  - If $x - y < 0$, then $|x-y| = y - x$. But we know that $y - \varepsilon < x$, i.e. $y - x < \varepsilon$, as required.
  - If $x - y = 0$, then by definition, $|x-y| = 0 < \varepsilon$, as required.

Exercise 5.4.7. — *Let x and y be real numbers. Show that $x \leqslant y + \varepsilon$ for all real numbers $\varepsilon > 0$ if and only if $x \leqslant y$. Show that $|x-y| \leqslant \varepsilon$ for all real numbers $\varepsilon > 0$ if and only if $x = y$.*

1. Let's prove the first statement. It is obvious that if $x \leqslant y$, then $x \leqslant y + \varepsilon$ for all $\varepsilon > 0$. Let's thus prove the other direction. Let's use the contrapositive statement: suppose that $x > y$. Thus, $x - y > 0$. Now let δ be a positive real number defined by $\delta = \frac{x-y}{2} > 0$. We have:

$$\begin{aligned} y + \delta &= y + \frac{x}{2} + \frac{x}{2} \\ &= \frac{x}{2} + \frac{y}{2} \\ &< \frac{x}{2} + \frac{x}{2} = x \end{aligned}$$

Thus, if $x > y$, there exists $\delta > 0$ such that $x > y + \delta$. Using the contrapositive, we conclude that if $x \leqslant y + \varepsilon$ for all $\varepsilon > 0$, then $x \leqslant y$.

2. The second statement has a very similar proof. Once again, it is obvious that if $x = y$, then $|x - y| \leqslant \varepsilon$, since we have $|x - y| = 0$. For the other direction, let's use the contrapositive once again, and suppose that $x \neq y$. Thus, $|x-y| > 0$, and in particular, $|x - y| > \frac{|x-y|}{2} > 0$. Consequently, if $x \neq y$, there exists $\delta = \frac{|x-y|}{2} > 0$ such that $|x - y| > \delta$. The contrapositive means that if $|x - y| \leqslant \varepsilon$ for all $\varepsilon > 0$, then $x = y$, as required.

EXERCISE 5.4.8. — *Let $(a_n)_{n=1}^{\infty}$ be a Cauchy sequence of rationals, and let x be a real number. Show that if $a_n \leqslant x$ for all $n \geqslant 1$, then $LIM_{n\to\infty}a_n \leqslant x$. Similarly, show that if $a_n \geqslant x$ for all $n \geqslant 1$, then $LIM_{n\to\infty}a_n \geqslant x$.*

Suppose, for the sake of contradiction, that we have both $a_n \leqslant x$ for all $n \geqslant 1$ and $\text{LIM}_{n\to\infty}a_n > x$. Then, according to Proposition 5.4.14, there exists a rational q such that $x < q < \text{LIM}_{n\to\infty}a_n$. We have thus $a_n \leqslant x < q$ for all $n \geqslant 1$. According to Corollary 5.4.10 with the constant sequence of rationals $(q_n)_{n=:}^{\infty} = q$, we should have $\text{LIM}_{n\to\infty}a_n \leqslant q$.

Thus, we have both $\text{LIM}_{n\to\infty}a_n \leqslant q$ and $x < q < \text{LIM}_{n\to\infty}a_n$, which is a contradiction and closes the proof.

EXERCISE 5.5.1. — *Let E be a subset of the real numbers $\mathbb{R}$, and suppose that E has a least upper bound M which is a real number, i.e., $M = \sup(E)$. Let $-E$ be the set $-E := \{-x : x \in E\}$. Show that $-M$ is the greatest lower bound of $-E$, i.e., $-M = \inf(-E)$.*

According to the definition of the greatest lower bound, we have two separate things to show: first, that $-M$ is a lower bound for $-E$, and second, that any other lower bound for $-E$ is inferior to $-M$.

1. Let z be an element of $-E$. By definition, we have $z = -x$ for some $x \in E$. Since $x \in E$, we have $x \leqslant M$, i.e. $z = -x \geqslant -M$. This means that $-M$ is a lower bound for $-E$.

2. Let be $-A$ another lower bound for $-E$, and show that $-A \geqslant -M$. Let be $x \in E$. Thus, we have $-x \in -E$, and by definition, $-x \geqslant -A$. This can also be written $x \leqslant A$, meaning that A is an upper bound for E. But by definition of the least upper bound M, we have $A \leqslant M$, i.e. $-A \geqslant -M$ as required: this closes the proof.

EXERCISE 5.5.2. — *Let E be a non-empty subset of $\mathbb{R}$, let $n \geqslant 1$ be an integer, and let $L < K$ be integers. Suppose K/n is an upper bound for E, but that L/n is not an upper bound for E. Without using Theorem 5.5.9, show that there exists an integer $L < m \leqslant K$ such that m/n is an upper bound for E, but that $(m-1)/n$ is not an upper bound for E. (Hint: prove by contradiction, and use induction.)*

This is not an easy exercise, so that we begin by an informal draft of the proof. Tao's advice is to prove by contradiction, i.e. to suppose that the following statement holds:

($\mathcal{H}$): there exists no integer m such that $L < m \leqslant K$, where m/n is an upper bound for E, but $\frac{m-1}{n}$ is not.

What would be the contradiction if we accept this fact? Let's proceed by "descending induction". Let's start with $m = K$: we already know that K/n is an upper bound, thus according to ($\mathcal{H}$), $(K-1)/n$ is necessarily also an upper bound. Now let's continue with $m = K-1$: we can see that, necessarily, $(K-2)/n$ should also be an upper bound. And so on, until we finally reach (after $K-L-1$ steps) $m = L+1$, which will be still supposed to be an upper bound at this stage; but according to ($\mathcal{H}$), L/n should also be an upper bound, which would be a contradiction with the fundamental assumption of this exercise.

Thus, we have to combine, in some way, an induction reasoning with this hypothesis ($\mathcal{H}$) that we would like to reject.

Actually, if we suppose that ($\mathcal{H}$) is true, we can show by induction that for any natural j, $(K-j)/n$ is an upper bound for E:

- The base case $j = 0$ is straightforward: we already know that K/n is an upper bound.
- Now let's suppose inductively that $(K-j)/n$ is an upper bound, and let's show that $(K-j-1)/n$ is also an upper bound. We know that L/n is not an upper bound, so there must exist some $x_0 \in E$ such that $x_0 > L/n$. But since $(K-j)/n$ is an upper bound, we have the following inequalities: $L/n < x_0 \leqslant (K-j)/n \leqslant K/n$. In particular, this means that $L < K-j$; and this also means that $L < K-j \leqslant K$. Thus, we also have necessarily $(K-j-1)/n$ as an upper bound, otherwise the integer $(K-j)$ would be a contradiction for ($\mathcal{H}$). This closes the induction.

Now the contradiction appears formally: take the (positive) natural number $K-L$, and apply the previous statement proved by induction. We should have $(K-(K-L))/n = L/n$ as an upper bound for E, which is a contradiction with the main assumption of this exercise.

Exercise 5.5.3. — *Let E be a non-empty subset of $\mathbb{R}$, let $n \geqslant 1$ be an integer, and let m, m' be integers with the properties that m/n and m'/n are upper bounds for E, but $(m-1)/n$ and $(m'-1)/n$ are not upper bounds for E. Show that $m = m'$. This shows that the integer m constructed in Exercise 5.5.2 is unique.*

We will show successively that $m' \leqslant m$ and $m \leqslant m'$, which will imply that $m = m'$.

- Since $(m-1)/n$ is not an upper bound for E, there exists $x_0 \in E$ such that:
$$x_0 > (m-1)/n \tag{5.2}$$
- But since $x_0 \in E$ and m'/n is an upper bound, we have actually:
$$(m-1)/n < x_0 \leqslant m'/n \tag{5.3}$$
- Similarly, since $(m'-1)/n$ is not an upper bound for E, there exists $x_1 \in E$ such that:
$$x_1 > (m'-1)/n \tag{5.4}$$
- But since $x_1 \in E$ and m/n is an upper bound, we have actually:
$$(m'-1)/n < x_1 \leqslant m/n \tag{5.5}$$
- Thus, combining (5.3) and (5.5), we have both $m'-1 < m$ and $m-1 < m'$. But m and m' are integers: recall that for integers, $a-1 < b$ and $a \leqslant b$ are equivalent (see for instance Proposition 2.2.12 for the naturals). Thus we have both $m' \leqslant m$ and $m \leqslant m'$, as required.

Exercise 5.5.4. — *Let $q_1, q_2, q_3, \ldots$ be a sequence of rational numbers with the property that $|q_n - q_{n'}| \leqslant 1/M$ whenever $M \geqslant 1$ is an integer and $n, n' \geqslant M$. Show that $q_1, q_2, q_3, \ldots$ is a Cauchy sequence. Furthermore, if $S := LIM_{n\to\infty} q_n$, show that $|q_M - S| \leqslant 1/M$ for every $M \geqslant 1$. (Hint: use Exercise 5.4.8.)*

1. Let $\varepsilon > 0$ be a positive rational number. To show that $(q_n)_{n=1}^{\infty}$ is a Cauchy sequence, we must prove that there exists a natural number $N \geqslant 1$ such that $n, n' \geqslant N \implies |q_n - q_{n'}| \leqslant \varepsilon$.

   Let's apply the archimedean property[18]: since 1 and ε are both positive real numbers, there exists a natural number M such that $M\varepsilon \geqslant 1$, i.e. such that $1/M \leqslant \varepsilon$.

   Thus, for a given value of ε, taking this natural number M provides the required result. Indeed, if $n, n' \geqslant M$, we have $|q_n - q_{n'}| \leqslant 1/M$ by initial hypothesis, and $1/M \leqslant \varepsilon$ by archimedean property. Thus $|q_n - q_{n'}| \leqslant \varepsilon$, as required.

2. Let be $S = \mathrm{LIM}_{n\to\infty} q_n$. We want to show that $|q_M - S| \leqslant 1/M$ for all $M \geqslant 1$.

   Recall that $|q_n - q_{n'}| \leqslant 1/M$ for any $n, n' \geqslant M$ by initial hypothesis. By fixing $n' := M$ to a given value, we get in particular:

   $$\forall n \geqslant M \geqslant 1, \quad |q_n - q_M| \leqslant \frac{1}{M} \tag{5.6}$$

   This is equivalent to:

   $$\forall n \geqslant M \geqslant 1, \quad -\frac{1}{M} + q_M \leqslant q_n \leqslant \frac{1}{M} + q_M \tag{5.7}$$

   At this point, we cannot immediately apply the insight from Exercise 5.4.8 because of this "$n \geqslant M \geqslant 1$" part (we would like to have this property for "$n \geqslant 1$" instead). We can overcome this difficulty by defining a new "instrumental" sequence $(a_n)_{n=1}^{\infty}$:

   $$a_n = \begin{cases} q_M & \text{if } n \leqslant M \\ q_n & \text{if } n > M \end{cases} \tag{5.8}$$

   The sequences $(a_n)_{n=1}^{\infty}$ and $(q_n)_{n=1}^{\infty}$ are equivalent since they are eventually equal. Thus, $\mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} q_n$ by definition of a real number.

   And this time, we can adapt the insight from (5.7) by removing the problematic "$n \geqslant M$" part: we have actually, for all $n \geqslant 1$,

   $$\forall n \geqslant M \geqslant 1, \quad -\frac{1}{M} + q_M \leqslant a_n \leqslant \frac{1}{M} + q_M \tag{5.9}$$

   so that, according to Exercise 5.4.8, we have $-1/M + q_M \leqslant S \leqslant 1/M + q_M$, i.e. $|q_M - S| \leqslant 1/M$.

   This is valid for any given $M \geqslant 1$, and thus closes the proof.

[18] This has actually previously been shown in Exercise 5.4.4.

Exercise 5.5.5. — *Establish an analogue of Proposition 5.4.14, in which "rational" is replaced by "irrational".*

We have to prove that there exists an irrational number s between any real numbers x, y. Actually, we only know one concrete example of irrational number so far in Tao's text: the number $\sqrt{2}$ (see Proposition 4.4.4). We thus should use it in this proof.

Let be x, y two real numbers, and consider the two real numbers $x + \sqrt{2}$ and $y + \sqrt{2}$. According to Proposition 5.4.14, there exists a rational number q such that $x + \sqrt{2} < q < y + \sqrt{2}$; i.e. $x < q - \sqrt{2} < y$.

But it is easy to show that $q - \sqrt{2}$ cannot be a rational number: if it was a rational, then we would have $q - \sqrt{2} = q'$, with q' a rational number. And then, we would have $\sqrt{2} = q - q' \in \mathbb{Q}$, which would contradict Proposition 4.4.4.

Thus we can indeed construct an irrational number between any pair of real numbers x, y.

Exercise 5.6.1. — *Prove Lemma 5.6.6. (Hints: review the proof of Proposition 5.5.12; proofs by contradiction will be useful.)*

The claims to prove are:

(a) If $y = x^{1/n}$, then $y^n = x$.

To prove this one, we can basically go back to Proposition 5.5.12 and adapt its proof, as suggested by Terence Tao. The general sketch is as follows: by Definition 5.6.4, we have $y := \sup\{z \in \mathbb{R} \mid z \geqslant 0, z^n \leqslant x\}$. We should show that both the assertions $y^n > x$ and $y^n < x$ lead to contradictions with this definition. What could be such contradictions? For instance, it could be the fact that there exists a greater number than y in the set $\{z \in \mathbb{R} \mid z \geqslant 0, z^n \leqslant x\}$, i.e. that y is *not* an upper bound. To show that, we must find a small number $\zeta > 0$ such that $(y + \zeta) \in \{z \in \mathbb{R} \mid z \geqslant 0, z^n \leqslant x\}$, or in other words, $(y + \zeta)^n \leqslant x$. But another possible contradiction could be to show that y is an upper bound but not the *least* upper bound, i.e. that there exists a small number $0 < \varepsilon < 1$ such that $y - \varepsilon$ is also an upper bound of the set $\{z \in \mathbb{R} \mid z \geqslant 0, z^n \leqslant x\}$.

Starting with each of the two hypotheses $y^n > x$ and $y^n < x$, we'll try to obtain one of these two contradictions.

- First suppose that we have $y^n < x$. Here we will refrain from using the binomial theorem in the proof, but we will use it as an intuition. Let be a small real number $0 < \varepsilon < 1$. First, we will show that if $y^n < x$, then we can find a real number $M > 0$ such that $(y + \varepsilon)^n \leqslant y^n + M\varepsilon$. This would be easiest with the binomial theorem, but it can also by shown by induction. Indeed, the base case $n = 1$ is obvious, since $M = 1$ is okay. Now, let's suppose inductively that there exists a natural number M such that we have $(y + \varepsilon)^n \leqslant y^n + M\varepsilon$. Thus:

$$\begin{aligned}(y+\varepsilon)^{n+1} &\leqslant (y+\varepsilon)(y^n + M\varepsilon) \\ &\leqslant y^{n+1} + \varepsilon(My + y^n + M\varepsilon) \\ &\leqslant y^{n+1} + \varepsilon\underbrace{(My + y^n + M)}_{:=M'} \text{ (because } \varepsilon < 1)\end{aligned}$$

i.e. there exists also a natural number M' such that $(y + \varepsilon)^{n+1} \leqslant y^{n+1} + M'\varepsilon$, which closes the induction. And thus, we have: $(y + \varepsilon)^n \leqslant y^n + M\varepsilon < y^n < x$, for some $\varepsilon > 0$. This means that y is no longer the supremum of the set $\{z \in \mathbb{R} \mid z \geqslant 0, z^n \leqslant x\}$, which contradicts our initial hypothesis. Thus, $y^n < x$ leads to a contradiction.

- Then suppose that we have $y^n > x$. A similar approach applies. Let $0 < \varepsilon < 1$ be a small number. Although not detailed here, it is now quite easy to show in a similar fashion that there exists a real number M such that $(y - \varepsilon)^n \geqslant y^n - M\varepsilon$. But since $y^n > x$ (strictly), we know that there exists a small number $\zeta > 0$ such that $y^n > y^n - \zeta > x$. If we choose ε so that we have $M\varepsilon = \zeta$, then we get $(y - \varepsilon)^n > x$. This means that there exists a smaller number than y which is also an upper bound for $\{z \in \mathbb{R} \mid z \geqslant 0, z^n \leqslant x\}$, which contradicts the definition of y as the supremum of this set.

Thus, both the statements $y^n > x$ and $y^n < x$ are impossible, which allows us to conclude that $y^n = x$, as required.

Note that, in other words, we've just showed that for any non-negative real x, we have:

$$\left(x^{1/n}\right)^n = x \tag{5.10}$$

and this is thus something we can use in the next steps.

(b) Conversely, if $y^n = x$, then $y = x^{1/n}$. (Additional hint: use the previous result, and Proposition 4.3.12.)

Suppose that we have $y^n = x$. Since the n-th root is well-defined, the n-th roots of two equal numbers are also equal, i.e. $(y^n)^{1/n} = x^{1/n}$. Now we can use the insight from equation (5.10). In one hand, we have $\left[(y^n)^{1/n}\right]^n = y^n$; and this is equal to $\left(x^{1/n}\right)^n$. But according to Proposition 4.3.12 (and its counterpart for the real numbers), the equality $y^n = \left(x^{1/n}\right)^n$ implies $y = x^{1/n}$ as required.

Note that, in other words, we've just showed that for any non-negative real x, we have:

$$(x^n)^{1/n} = x \tag{5.11}$$

(c) $x^{1/n}$ is a non-negative number, and is positive iff x is positive.

First, $x^{1/n}$ is obviously non-negative, since it is defined as the supremum of a (non-empty) set of non-negative numbers.

Now let's prove that it is positive iff $x > 0$.

- If $x^{1/n} > 0$, then we have $\left(x^{1/n}\right)^n > 0^n$ according to (the counterpart for real numbers of) Proposition 4.3.12(b). But $0^n = 0$, and $\left(x^{1/n}\right)^n = x$ according to equation (5.10). Thus we have indeed $x > 0$.
- If $x > 0$, let's suppose for the sake of contradiction that $x^{1/n} = 0$. Then, still according to equation (5.10), we should have $x = \left(x^{1/n}\right)^n = 0$, which is a contradiction.

(d) We have $x > y \iff x^{1/n} > y^{1/n}$.

- First suppose that $x^{1/n} > y^{1/n}$. According to Proposition 4.3.12(b), we have: $\left(x^{1/n}\right)^n > \left(y^{1/n}\right)^n$. Also, by equation (5.10), $\left(x^{1/n}\right)^n = x$ and $\left(y^{1/n}\right)^n = y$, so we have indeed $x > y$. as required.
- Now suppose that $x > y$. Let's suppose for the sake of contradiction that we have $x^{1/n} \leqslant y^{1/n}$. We would thus use Proposition 4.3.12(b) and equation (5.10) again and see that $\left(x^{1/n}\right)^n \leqslant \left(y^{1/n}\right)^n$, i.e. that $x \leqslant y$, which is a contradiction. Thus we have necessarily $x^{1/n} > y^{1/n}$.

(e) If $x > 1$, then $x^{1/k}$ is a decreasing (i.e. $x^{1/k} < x^{1/l}$ whenever $k > l$) function of k. If $0 < x < 1$, then $x^{1/k}$ is an increasing function of k. If $x = 1$, then $x^{1/k}$ for all k. Here k ranges over the positive integers.

- Let be $x = 1$. We know that $1^k = 1$ for any positive integer k. Now, applying equation (5.11) leads to $1 = 1^{1/k}$ for any positive integer k, as required.
- Let be $x > 1$, and two positive integers $k > l$. We have to show that $x^{1/k} < x^{1/l}$. First, note that if $k > l$, then we have $k = l + p$ with p a positive integer (recall Definition 4.1.10). Let's suppose, for the sake of contradiction, that we have $x^{1/k} \geqslant x^{1/l}$. Thus, we should have $(x^{1/k})^{kl} \geqslant (x^{1/l})^{kl}$ according to Proposition 4.3.10(c), i.e. $x^l \geqslant x^k$ (we use equation (5.10) for this latest claim). But this last inequality could be written $x^l \geqslant x^{l+p}$, i.e. $1 \geqslant x^p$ by cancellation law. But this is a contradiction: if $x > 1$, we cannot have $x^p \leqslant 1$ with p a positive integer[19]. Thus, $x^{1/k}$ is indeed a decreasing function of k in this case.
- Let be $0 < x < 1$, and two positive integers $k > l$ (thus, we still have $k = l + p$ for a certain positive integer p). We have to show that $x^{1/k} > x^{1/l}$. A very similar proof applies. Let's suppose, for the sake of contradiction, that we have $x^{1/k} \leqslant x^{1/l}$. We should thus have $(x^{1/k})^{kl} \leqslant (x^{1/l})^{kl}$, i.e. $x^l \leqslant x^k$. This means that $x^l \leqslant x^{l+p}$, i.e. $1 \leqslant x^p$, which is impossible if $0 < x < 1$. This shows the contradiction and closes the proof.

(f) We have $(xy)^{1/n} = x^{1/n}y^{1/n}$.

- On the one hand, we have $\left[(xy)^{1/n}\right]^n = xy$ according to equation (5.10).
- On the other hand, we have $\left[x^{1/n}y^{1/n}\right]^n = (x^{1/n})^n(y^{1/n})^n = xy$, where we used successively Proposition 4.3.12(a) and equation (5.10).

Thus, both expressions are equal.

(g) We have $(x^{1/n})^{1/m} = x^{1/nm}$.

- On the one hand, we have $(x^{1/nm})^{nm} = x$ according to equation (5.10).
- On the other hand, $((x^{1/n})^{1/m})^{nm} = ((x^{1/n})^{1/m})^{mn} = \left[((x^{1/n})^{1/m})^m\right]^n = \left[x^{1/n}\right]^n = x$; where we used several times equation (5.10), and Proposition 4.3.12.

Thus, both expressions are equal.

[19] Just perform a quick inductive proof if needed.

Exercise 5.6.2. — *Prove Lemma 5.6.9.* Let $x, y > 0$ be positive reals, and let q, r be rationals. In the whole exercise, let's say that $q = a/b$ and $r = a'/b'$, with $b, b' > 0$. The claims to prove are:

(a) x^q is a positive real.

By Definition 5.6.7, $x^q = (x^{1/b})^a$. Since $x > 0$, we know by Lemma 5.6.6(c) that $x^{1/b}$ is a positive real. Thus, $x^q = (x^{1/b})^a$ is also positive (by the counterpart for reals of Proposition 4.3.12(b)).

(b) $x^{q+r} = x^q x^r$ and $(x^q)^r = x^{qr}$.

- We have $q + r = \frac{ab'+a'b}{bb'}$ (recall Definition 4.2.2). By Definition 5.6.7, $x^{q+r} = (x^{1/bb'})^{ab'+a'b}$. But we also can write $q = ab'/bb'$ and $r = a'b/bb'$. We know by Lemma 5.6.8 that our choices of numerators and denominators for q and r do not matter as regards x^q and x^r. Thus we also have $x^q x^r = (x^{1/bb'})^{ab'}(x^{1/bb'})^{a'b}$, which is equal to $(x^{1/bb'})^{ab'+a'b}$ by Proposition 4.3.12(a). Thus, $x^{q+r} = x^q x^r$.
- We have:

$$\begin{aligned} ((x^q)^r)^{bb'} &= ((((x^{1/b})^a)^{1/b'})^{a'})^{bb'} \\ &= (((x^{1/b})^a)^{1/b'})^{a'bb'} \\ &= ((((x^{1/b})^a)^{1/b'})^{b'})^{a'b} \\ &= ((x^{1/b})^a)^{a'b} \\ &= ((x^{1/b})^b)^{aa'} \\ &= x^{aa'} \end{aligned}$$

And, also:

$$\begin{aligned} (x^{qr})^{bb'} &= ((x^{1/bb'})^{aa'})^{bb'} \\ &= (x^{1/bb'})^{aa'bb'} \\ &= ((x^{1/bb'})^{bb'})^{aa'} \\ &= x^{aa'} \end{aligned}$$

Thus, both expressions are equal, and this implies $x^{qr} = (x^q)^r$ according to Proposition 4.3.12(c).

(c) $x^{-q} = 1/x^q$.

First, note that Definitions 4.3.11 and 5.6.2 only say that $x^{-n} = 1/x^n$ if n is a *positive* integer. But this is actually true for any integer n: if $n = 0$, then both x^{-n} and $1/x^n$ are equal to 1 and are thus equal; and if $n < 0$, then there exists a positive integer m such that $n = -m$, and we have $x^{-n} = x^m$, $1/x^n = 1/x^{-m} = 1/(1/x^m) = x^m$ by Definition 5.6.2. Thus in all cases,

$$x^{-n} = 1/x^n, \quad \forall n \in \mathbb{Z} \tag{5.12}$$

Now the claim is straightforward: if $q = a/b$ with $b > 0$, then $x^{-q} = (x^{1/b})^{-a} = 1/((x^{1/b})^a) = 1/x^q$ according to equation (5.12).

(d) If $q > 0$, then $x > y$ iff $x^q > y^q$.

First, note that if $q > 0$, then $q = a/b$ with both a, b as positive integers.

- If $x > y$, then $x^{1/b} > y^{1/b}$ according to Lemma 5.6.6(d). And then, $(x^{1/b})^a > (y^{1/b})^a$ according to (the counterpart for real numbers of) Proposition 4.3.12(b), i.e. $x^q > y^q$.
- If $x^q > y^q$, then by definition, $(x^{1/b})^a > (y^{1/b})^a$, and both terms are positive. Thus, by Lemma 5.6.6(d), we have $((x^{1/b})^a)^{1/a} > ((y^{1/b})^a)^{1/a}$, i.e. $x^{1/b} > y^{1/b}$. And, by Proposition 4.3.12(b), $x = (x^{1/b})^b > (y^{1/b})^b = y$, as required.

(e) If $x > 1$, then $x^q > x^r$ if and only if $q > r$. If $x < 1$, then $x^q > x^r$ if and only if $q < r$.

First, note that we have:

$$(x^{1/b})^a = (x^a)^{1/b} \tag{5.13}$$

because on the one hand, $((x^{1/b})^a)^b = (x^{1/b})^{ab} = (x^{1/b})^{ba} = ((x^{1/b})^b)^a = x^a$, where we used Proposition 4.3.12(a) twice. And on the other hand, $((x^a)^{1/b})^b = x^a$ by Lemma 5.6.6(a). Thus, we have $((x^{1/b})^a)^b = ((x^a)^{1/b})^b$, which implies $(x^{1/b})^a = (x^a)^{1/b}$ by Proposition 4.3.12(c). (This proof holds for $a \neq 0$, but equation (5.13) is obvious if $a = 0$.)

Now we go back to the main claims to prove.

- Let be $x > 1$. First suppose that $q > r$, and let's show that $x^q > x^r$. If $q > r$, then we have $a/b > a'/b'$, i.e. $ab'/bb' > a'b/bb'$, i.e. $ab' > a'b$. Since both ab' and $a'b$ are integers and $x > 1$, we thus have $x^{ab'} > x^{a'b}$. And, by Lemma 5.6.6(d), we have $(x^{ab'})^{1/bb'} > (x^{a'b})^{1/bb'}$, i.e. $x^{ab'/bb'} > x^{a'b/bb'}$, and finally $x^q > x^r$ by Lemma 5.6.8.

  Now suppose that $x^q > x^r$ (note that both of them are positive since $x > 1$), and let's show that $q > r$, i.e. that $ab' > a'b$. Since we can also write $q = ab'/bb'$ and $r = a'b/bb'$ (and this does not affect the result by Lemma 5.6.8), we have $x^{ab'} = (x^q)^{bb'} > (x^r)^{bb'} = x^{a'b}$. And, if we multiply both sides by $x^{-a'b}$, which is a positive number, we get $x^{ab'-a'b} > 1$. We see that, in this inequality, we obviously cannot have $ab' - ab' = 0$. We cannot have $ab' - a'b < 0$ either, because in this case, we would have $ab' - a'b = -n$ with n a positive integer. I.e., we would have $x^{-n} > 1$, i.e. $1/x^n > 1$, a fact which is incompatible with our initial hypothesis $x > 1$. Thus, we only have one possibility: $ab' - a'b > 0$, i.e. $q > r$.
- Let be $0 < x < 1$. Detailed proof is not given here, but is similar to the previous case.

(f) $(xy)^q = x^q y^q$.

We have:

$$\begin{aligned}
(xy)^q &= ((xy)^{1/b})^a \text{ (Definition 5.6.7)} \\
&= (x^{1/b} y^{1/b})^a \text{ (Lemma 5.6.6(f))} \\
&= (x^{1/b})^a (y^{1/b})^a \text{ (Proposition 4.3.12(a))} \\
&= x^q y^q
\end{aligned}$$

as required.

Exercise 5.6.3. — *If x is a real number, show that $|x| = (x^2)^{1/2}$.*

To show that $(x^2)^{1/2}$ is equal to $|x|$, we should consider three cases according to the definition of absolute values, and prove that $(x^2)^{1/2} = 0$ if $x = 0$, $(x^2)^{1/2} = x$ if $x > 0$, and $(x^2)^{1/2} = -x$ if $x < 0$.

- If $x = 0$, we have $x^2 = 0$, and thus $(x^2)^{1/2} = 0$ according to Lemma 5.6.6(c).
- If $x > 0$, we have $(x^2)^{1/2} = x$ according to equation (5.11), i.e. Lemma 5.6.6(b).
- If $x < 0$, then $-x > 0$. Also, we know that $x^2 = (-x)^2$ for every real x. Thus, $(x^2)^{1/2} = ((-x)^2)^{1/2} = -x$ as required, still by Lemma 5.6.6(b).

6. Limits of sequences

Exercise 6.1.1. — *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, such that $a_{n+1} > a_n$ for each natural number n. Prove that whenever n and m are natural numbers such that $m > n$, then we have $a_m > a_n$. (We refer to these sequences as increasing sequences.)*

Let be $k = m - n$, so that $k > 0$ represents the difference between the two indexes under comparison. Let's induct on k to prove that $a_{n+k} > a_n$ for all natural number n and any $k > 0$.

- The base case is $k = 1$, i.e. $m = n + 1$. We must show here that $a_{n+1} > a_n$, but this is true by our initial hypothesis.
- Now suppose inductively that this is true for a certain natural k, i.e. that we have $a_{n+k} > a_n$ for all n. We have to show that we have $a_{n+k+1} > a_n$ for all n. But we know by our initial hypothesis that $a_{n+k+1} > a_{n+k}$, thus we have $a_{n+k+1} > a_{n+k} > a_n$, i.e. by transitivity, $a_{n+k+1} > a_n$ as required. This closes the induction.

Exercise 6.1.2. — *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, and let L be a real number. Show that $(a_n)_{n=m}^{\infty}$ converges to L iff, given any real $\varepsilon > 0$, one can find an $N \geqslant m$ such that $|a_n - L| \leqslant \varepsilon$ for all $n \geqslant N$.*

- First suppose that $(a_n)_{n=m}^{\infty}$ converges to L. Let be ε a positive real number. By Definition 6.1.5, $(a_n)_{n=m}^{\infty}$ is eventually ε-close to L, i.e., there exists a natural $N \geqslant m$ such that $|a_n - L| \leqslant \varepsilon$, as required.
- Now show to converse implication. Let be ε a positive real. We know that we can find an $N \geqslant m$ such that $|a_n - L| \leqslant \varepsilon$ for all $n \geqslant N$, and thus $(a_n)_{n=m}^{\infty}$ is eventually close to ε. This is true for all ε, so that $(a_n)_{n=m}^{\infty}$ converges to L.

Exercise 6.1.3. — *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let c be a real number, and let $m' \geqslant m$ be an integer. Show that $(a_n)_{n=m}^{\infty}$ converges to c if and only if $(a_n)_{n=m'}^{\infty}$ converges to c.*

We will prove each part of this equivalence separately.

- First, suppose that $(a_n)_{n=m'}^{\infty}$ converges to c. This means that, for every $\varepsilon > 0$, there exists $N' \geqslant m'$ such that $n \geqslant N' \implies |a_n - c| \leqslant \varepsilon$. And we want to show that, still for every $\varepsilon > 0$, there exists $N \geqslant m$ such that $n \geqslant N \implies |a_n - c| \leqslant \varepsilon$. We claim that taking $N := N'$ is convenient. Indeed, since $m' \geqslant m$, we have $n \geqslant m$ as soon as we have $n \geqslant m'$, so that N' satisfies the required condition.
- Now suppose that $(a_n)_{n=m}^{\infty}$ converges to c. This means that:

  $$\forall \varepsilon > 0, \exists N \geqslant m \ : \ n \geqslant m \to |a_n - c| \leqslant \varepsilon \tag{6.1}$$

  So let be $\varepsilon > 0$, and let's take $N' := N + m'$. We have both $N' \geqslant N$ and $N' \geqslant m'$. Thus, if $n \geqslant N'$, we also have $n \geqslant N$, which implies $|a_n - c| \leqslant \varepsilon$ according to equation (6.1). Thus, we have indeed found a positive integer $N' \geqslant m'$ such that $n \geqslant N' \to |a_n - c| \leqslant \varepsilon$, i.e., we have proved that $(a_n)_{n=m'}^{\infty}$ converges to c.

Exercise 6.1.4. — *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let c be a real number, and let $k \geqslant 0$ be a non-negative integer. Show that $(a_n)_{n=m}^{\infty}$ converges to c if and only if $(a_{n+k})_{n=m}^{\infty}$ converges to c.*

This exercise is pretty similar to the previous one.[20]

- First, suppose that $(a_n)_{n=m}^{\infty}$ converges to c. Let $\varepsilon > 0$ be a real number. We must show that there exists $M \geqslant m$ such that $n + k \geqslant M \to |a_{n+k} - c| \leqslant \varepsilon$. We know that there exists $N \geqslant m$ such that $n \geqslant N \to |a_n - c| \leqslant \varepsilon$. But for any $k \leqslant 0$, we have $n + k \geqslant n$, so that as soon as $n \geqslant N$, we have $n + k \geqslant n \geqslant N$, and thus $|a_{n+k} - c| \leqslant \varepsilon$. So, choosing $M := N$ is suitable here, and we have showed the first part.
- Now prove the converse implication: suppose that $(a_{n+k})_{n=m}^{\infty}$ converges to c. We want to show that there exists $N \geqslant m$ such that $n \geqslant N \to |a_n - c| \leqslant \varepsilon$. We already know that there exists $M \geqslant m$ such that $n \geqslant M \to |a_{n+k} - c| \leqslant \varepsilon$. Let be $N := M + k$. We can see that if $n \geqslant N$, we have $n \geqslant M + k$, i.e. $n - k \geqslant M$, and thus $|a_{(n-k)+k} - c| \leqslant \varepsilon$ as required. So, $N := M + k$ is suitable, and $(a_n)_{n=m}^{\infty}$ converges to c.

Exercise 6.1.5. — *Prove Proposition 6.1.12.*

We must prove that convergent sequences are Cauchy sequences. Let $(a_n)_{n=m}^{\infty}$ be a convergent sequence to c, and let's prove that $(a_n)_{n=m}^{\infty}$ is a Cauchy sequence.

Let be $\varepsilon > 0$ a positive real number. By Definition 6.1.5, there exists a natural number $N \geqslant m$ such that $n \geqslant N \implies |a_n - c| \leqslant \varepsilon/2$.

Now let be $j, k \geqslant N$ two natural numbers. We have, by triangular inequality:

$$\begin{aligned} |a_j - a_k| &= |a_j - c + c - a_k| \\ &\leqslant |a_j - c| + |a_k - c| \\ &\leqslant \varepsilon/2 + \varepsilon/2 \\ &\leqslant \varepsilon \end{aligned}$$

Thus, we have showed that there exists $N \geqslant m$ such that, for any natural numbers $j, k \geqslant N$, we have $|a_j - a_k| \leqslant \varepsilon$. This means that $(a_n)_{n=m}^{\infty}$ is indeed a Cauchy sequence.

Exercise 6.1.6. — *Prove Proposition 6.1.15; i.e. that formal limits are genuine limits.*

Let $(a_n)_{n=m}^{\infty}$ be a Cauchy sequence of real numbers, and $L := \mathrm{LIM}_{n\to\infty} a_n$. We have to show that $(a_n)_{n=m}^{\infty}$ converges to L.

Now we assume, for the sake of contradiction, that $(a_n)_{n=m}^{\infty}$ is not eventually ε-close to L, i.e. that there exists $\varepsilon > 0$ such that:

$$\forall N \geqslant m, \exists n \geqslant N \ : \ |a_n - L| > \varepsilon \tag{6.2}$$

Precisely, let's consider this positive real ε in what follows. Also, recall that $(a_n)_{n=m}^{\infty}$ is a Cauchy sequence, so that:

$$\exists T \geqslant m \ : \ j, k \geqslant T \implies |a_j - a_k| \leqslant \varepsilon/2 \tag{6.3}$$

[20] Actually, it would be possible to use Exercise 6.1.3 here, but, quite unexpectedly, this would just make the proof a little harder, because it would require to use some notions (like composition of functions, or intervals) that were very rarely used until now. The present proof is easier to write at this point of our knowledge.

According to (6.2), there exists $t \geqslant T$ such that $|a_t - L| > \varepsilon$. And since $t \geqslant T$, let's also consider any $s \geqslant T$: according to (6.3), we have $|a_s - a_t| \leqslant \varepsilon/2$. In particular, let's now fix t so that we can rephrase this property as:

$$\forall s \geqslant T,\ |a_s - a_t| \leqslant \varepsilon/2 \quad ; \quad \text{i.e., } a_t - \varepsilon/2 \leqslant a_s \leqslant a_t + \varepsilon/2 \tag{6.4}$$

The inequality $|a_t - L| > \varepsilon$ opens two possible cases:

- if $a_t > L + \varepsilon$, then we have from (6.4):

$$L + \varepsilon - \varepsilon/2 < a_t - \varepsilon/2 \leqslant a_s$$

  so that $L + \varepsilon/2 \leqslant a_s$ for all $s \geqslant T$. According to Exercise 5.4.8, we should conclude that $L + \varepsilon/2 \leqslant L$, which is clearly a contradiction.

- if $a_t < L - \varepsilon$, then then we have from (6.4):

$$a_s \leqslant a_t + \varepsilon/2 \leqslant L - \varepsilon + \varepsilon/2$$

  so that $a_s \leqslant L - \varepsilon/2$ for all $s \geqslant T$. According to Exercise 5.4.8, we should conclude that $L \leqslant L - \varepsilon/2$, which is also a contradiction.

Thus, necessarily, $(a_n)_{n=m}^{\infty}$ is eventually ε-close to L for any $\varepsilon > 0$, i.e. $(a_n)_{n=m}^{\infty}$ converges to L.

Exercise 6.1.7. — *Show that Definition 6.1.16 is consistent with Definition 5.1.12 (i.e., prove an analogue of Proposition 6.1.4 for bounded sequences instead of Cauchy sequences).*

Let be $(a_n)_{n=m}^{\infty}$ a sequence of real[21] numbers.

- First suppose that $(a_n)_{n=m}^{\infty}$ is bounded according to Definition 5.1.12. It means that there exists a non-negative rational $M \geqslant 0$ such that $|a_n| \leqslant M$ for all $n \geqslant m$. We have to show that there exists a positive real $M' > 0$ such that $|a_n| \leqslant M'$ for all $n \geqslant m$. If we just take $M' := M + 1$, we have $M' > M$ and $M' \in \mathbb{R}$ because $M' \in \mathbb{Q}$. Thus, we have $|a_n| \leqslant M < M'$ for all $n \geqslant m$, as required. Thus, $(a_n)_{n=m}^{\infty}$ is bounded according to Definition 6.1.16.

- Now suppose that $(a_n)_{n=m}^{\infty}$ is bounded according to Definition 6.1.16. It means that there exists a positive real number $M' > 0$ such that $|a_n| \leqslant M'$ for all $n \geqslant m$. We have to show that there exists a non-negative rational number $M \geqslant 0$ such that $|a_n| \leqslant M$ for all $n \geqslant m$. According to Proposition 5.4.12, there exists a positive integer N such that $M' \leqslant N$. This implies that $|a_n| \leqslant M' \leqslant N$ for all $n \geqslant m$. And since N is a positive integer, it's also a non-negative rational. Thus, $(a_n)_{n=m}^{\infty}$ is bounded according to Definition 5.1.12.

Exercise 6.1.8. — *Prove Theorem 6.1.19 about limit laws.*

We'll prove each statement successively.

[21] If $(a_n)_{n=m}^{\infty}$ is a sequence of rational numbers, there is literally nothing to prove, so that we can skip this case.

(a) Let be $\varepsilon > 0$. We have to prove that there exists a natural number $N \geqslant m$ such that for all $n \geqslant N$, we have $|(a_n + b_n) - (x + y)| \leqslant \varepsilon$.

Since $(a_n)_{n=m}^{\infty}$ converges to x, there exists $N_1 \geqslant m$ such that $n \geqslant N_1 \to |a_n - x| \leqslant \varepsilon/2$.

Similarly, there exists $N_2 \geqslant m$ such that $n \geqslant N_2 \to |b_n - y| \leqslant \varepsilon/2$.

Let be $N := \max(N_1, N_2)$. Thus, by triangular inequality, we have for all $n \geqslant N$:

$$\begin{aligned}
|(a_n + b_n) - (x + y)| &\leqslant |a_n + x| + |b_n + y| \\
&\leqslant \varepsilon/2 + \varepsilon/2 \\
&\leqslant \varepsilon
\end{aligned}$$

as required. This closes the proof.

(b) Let be $\varepsilon > 0$. We have to prove that there exists a natural number $N \geqslant m$ such that for all $n \geqslant N$, we have $|a_n b_n - xy| \leqslant \varepsilon$.

Let's start by some algebraic manipulations:

$$\begin{aligned}
|a_n b_n - xy| &= |a_n b_n - a_n y + a_n y - xy| \\
&= |a_n(b_n - y) + y(a_n - x)| \\
&\leqslant |a_n| \times |b_n - y| + |y| \times |a_n - x|
\end{aligned}$$

Hopefully, there exists an upper bound for each term of this last expression, at least eventually. Indeed:

- Since $(a_n)_{n=m}^{\infty}$ is convergent, it is bounded according to Corollary 6.1.17. Thus, there exists $M \geqslant 0$ such that $|a_n| \leqslant M$ for all $n \geqslant m$.
- Since $(b_n)_{n=m}^{\infty}$ converges to y, there exists an integer $N_b \geqslant m$ such that $n \geqslant N_b \to |b_n - y| \leqslant \frac{\varepsilon}{2M}$.
- Since $(a_n)_{n=m}^{\infty}$ converges to x, there exists an integer $N_a \geqslant m$ such that $n \geqslant N_b \to |a_n - x| \leqslant \frac{\varepsilon}{2(|y|+1)}$. (Note that we don't choose $\frac{\varepsilon}{2|y|}$, because we don't know whether $|y| \neq 0$ or not, so we need an additional precaution.)

Let be $N := \max(N_a, N_b)$. For $n \geqslant N$, we thus have:

$$\begin{aligned}
|a_n b_n - xy| &\leqslant |a_n| \times |b_n - y| + |y| \times |a_n - x| \\
&\leqslant M \times \frac{\varepsilon}{2M} + |y| \times \frac{\varepsilon}{2(|y| + 1)} \\
&\leqslant \varepsilon/2 + \varepsilon/2 \\
&\leqslant \varepsilon
\end{aligned}$$

as required. This closes the proof.

(c) Here, a direct proof would be possible (and short), for instance by using Proposition 4.3.3(d). But following Tao's hint, let's use the previous results of this exercise instead. Let's consider the constant sequence $(c_n)_{n=m}^{\infty} = c, c, c, \ldots$; so that the sequence $(ca_n)_{n=m}^{\infty}$ is actually the product of the two sequences $(a_n)_{n=m}^{\infty}$ and $(c_n)_{n=m}^{\infty}$. Obviously, $(c_n)_{n=m}^{\infty}$ converges to c, so that according to statement *(b)*, we have $\lim_{n\to\infty} ca_n = \lim_{n\to\infty} c_n a_n = \lim_{n\to\infty} c_n \lim_{n\to\infty} a_n = cx$, which closes the proof.

(d) According to *(c)* taking $c = -1$, we have $\lim_{n\to\infty}(-b_n) = -\lim_{n\to\infty} b_n = -y$. Furthermore, according to *(a)*, we have $\lim_{n\to\infty}(a_n + (-b_n)) = \lim_{n\to\infty} a_n + \lim_{n\to\infty}(-b_n) = x - y$, as requested.

(e) Following Tao's hint, we begin by proving an intermediate result, which is the following: if $(b_n)_{n=m}^{\infty}$ is a sequence of non-zero real numbers, and converges to $y \neq 0$, then $(b_n)_{n=m}^{\infty}$ is bounded away from zero. Note that, by the "second formula" of triangular inequality, we have:

$$|y| - |b_n| \leqslant ||y| - |b_n|| \leqslant |y - b_n| \tag{6.5}$$

Furthermore, since $(b_n)_{n=m}^{\infty}$ converges to y, there exists a positive integer $M \geqslant m$ such that, for all $n \geqslant M$, we have $|b_n - y| \leqslant |y|/2$. This result, along with equation (6.5), implies that $|b_n| \geqslant |y|/2$ for all $n \geqslant M$, so that $(b_n)_{n=m}^{\infty}$ is eventually bounded away from zero, for $n \geqslant M$. But since $b_n \neq 0$ for all n, just consider $c = \min(|b_m|, |b_{m+1}|, \ldots, |b_{M-1}|, |y|/2)$: we thus have $|b_n| \geqslant c$ for all $n \geqslant m$, and c is positive. Thus, $(b_n)_{n=m}^{\infty}$ is indeed bounded away from zero.

Note that saying "$(b_n)_{n=m}^{\infty}$ is bounded away from zero", or $|b_n| \geqslant c$ for all $n \geqslant m$, means that $|1/b_n| \leqslant 1/c$, i.e. that $(1/b_n)_{n=m}^{\infty}$ is bounded.

Let be $\varepsilon > 0$ a positive real number. Since $(b_n)_{n=m}^{\infty}$ converges to y, there exists a positive integer N such that for all $n \geqslant N$, we have $|y - b_n| \leqslant \varepsilon |y| c$ (which is a positive real number, because $|y| \neq 0$. Now let's consider:

$$\begin{aligned}
\left|\frac{1}{b_n} - \frac{1}{y}\right| &= \left|\frac{y - b_n}{y b_n}\right| \\
&= \left|\frac{1}{y}\right| \times \left|\frac{1}{b_n}\right| \times |y - b_n| \\
&\leqslant \frac{1}{|y|} \times \frac{1}{c} \times |y - b_n| \\
&\leqslant \frac{1}{|y|} \times \frac{1}{c} \times \varepsilon |y| c \quad \text{for all } n \geqslant N \\
&\leqslant \varepsilon
\end{aligned}$$

This means that, for any positive real ε, one can find a positive integer $N \geqslant m$ such that $\left|\frac{1}{b_n} - \frac{1}{y}\right| \leqslant \varepsilon$. Thus, $(1/b_n)_{n=m}^{\infty}$ converges to $1/y$.

(f) This is a direct consequence of the parts *(b)* and *(e)*. Indeed, by part *(b)*, we have: $\lim_{n\to\infty}(a_n/b_n) = \lim_{n\to\infty}(a_n \times 1/b_n) = \lim_{n\to\infty} a_n \times \lim_{n\to\infty} 1/b_n$. And by part *(e)*, we have $\lim_{n\to\infty} 1/b_n = 1/\lim_{n\to\infty} b_n$, which gives the property we wanted to show.

(g) We must show that $\lim_{n\to\infty} \max(a_n, b_n) = \max(x, y)$. We immediately see that we have two different cases: if $x \geqslant y$, then $\max(x, y) = x$ and then we must show that $\lim_{n\to\infty} \max(a_n, b_n) = x$; else if $x < y$, then $\max(x, y) = y$ and we must show that $\lim_{n\to\infty} \max(a_n, b_n) = y$. Let's consider those two cases separately. In what follows, let be $\varepsilon > 0$ a positive real.

- If $x \geqslant y$, we actually have to prove that there exists a positive integer $N \geqslant m$ such that for all $n \geqslant N$, we have $|\max(a_n, b_n) - x| \leqslant \varepsilon$. We already know that $(a_n)_{n=m}^{\infty}$ converges to x, thus:

$$\exists N_a \geqslant m \;:\; n \geqslant N_a \Longrightarrow |a_n - x| \leqslant \varepsilon \tag{6.6}$$

Similarly, since $(b_n)_{n=m}^{\infty}$ converges to y:

$$\exists N_b \geqslant m \;:\; n \geqslant N_b \Longrightarrow |b_n - y| \leqslant \varepsilon \tag{6.7}$$

Now let's consider $N := \max(N_a, N_b)$, which is a positive integer. For all $n \geqslant N$, we have both $x - \varepsilon \leqslant a_n \leqslant \varepsilon + x$ by (6.6), and $y - \varepsilon \leqslant b_n \leqslant \varepsilon + y \leqslant \varepsilon + x$ by (6.7). Combining those two relationships[22] leads to:

$$x - \varepsilon \leqslant a_n \leqslant \max(a_n, b_n) \leqslant \varepsilon + x$$

whose the most important part is $x - \varepsilon \leqslant \max(a_n, b_n) \leqslant x + \varepsilon$, which is equivalent to $|\max(a_n, b_n) - x| \leqslant \varepsilon$, as requested initially. This closes the proof for this first case.

- If $x < y$, the proof is very similar. We have to prove that there exists a positive integer $N \geqslant m$ such that for all $n \geqslant N$, we have $|\max(a_n, b_n) - y| \leqslant \varepsilon$. Once again, we can combine the previous results to get:

$$y - \varepsilon \leqslant b_n \leqslant \max(a_n, b_n) \leqslant \varepsilon + y$$

i.e. $|\max(a_n, b_n) - y| \leqslant \varepsilon$ as requested.

(h) Taking the inspiration in the way we defined the greatest lower bound from the least upper bound (Exercise 5.5.1), we could first note that $\min(a, b) = -\max(-a, -b)$. Then, this would be a direct consequence from part *(c)* and *(g)*.

Exercise 6.1.9. — *Explain why Theorem 6.1.19(f) fails when the limit of the denominator is* 0.

Theroem 6.1.19(f) says that if $(a_n)_{n=m}^{\infty}$ converges to x and $(b_n)_{n=m}^{\infty}$ converges to $y \neq 0$ (with $b_n \neq 0$ for all $n \geqslant m$), then $(a_n/b_n)_{n=m}^{\infty}$ converges to x/y. If $y = 0$, the number x/y simply does not exist, so that the statement is pointless.

Note however that the sequence $(a_n/b_n)_{n=m}^{\infty}$ *may* converge even if we have $\lim b_n = 0$: think of the situation $a_n = b_n = 1/n$, for instance.

Exercise 6.1.10. — *Show that the concept of equivalent Cauchy sequence, as defined in Definition 5.2.6, does not change if ε is required to be positive real instead of positive rational. More precisely, if $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are sequences of reals, show that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are eventually ε-close for every rational $\varepsilon > 0$ if and only if they are eventually ε-close for every real $\varepsilon > 0$.*

Following Tao's hint, we just adapt the proof of Proposition 6.1.4.

- First suppose that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are eventually ε-close for every real $\varepsilon > 0$. In particular, this means that they are ε-close for any *rational* $\varepsilon > 0$, so there is nothing to prove.

- Then suppose that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are eventually ε-close for every rational $\varepsilon > 0$. Now let $\varepsilon > 0$ be a *real* number. According to Proposition 5.4.12, there exists a positive

[22] We also use the property that if $x \leqslant a$ and $y \leqslant a$, then $\max(x, y) \leqslant a$.

rational q such that $q \leqslant \varepsilon$. Since q is a rational, $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are q-close, according to our initial hypothesis, i.e., $|a_n - b_n| \leqslant q \leqslant \varepsilon$ for all n greater than some positive integer N. In particular, this means that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are eventually ε-close, as requested.

EXERCISE 6.2.1. — *Prove Proposition 6.2.5.*

Hereafter, x, y, z are extended real numbers. Thus, they can be either real numbers, or $\pm\infty$.

(a) We have to show that $x \leqslant x$. We have three cases here. If x is a real number, this is an obvious statement. If $x = +\infty$, then by Definition 6.2.3(b), we have $x \leqslant +\infty$ for all extended real x, so that the claim $x \leqslant x$ is still true. Similarly, if $x = -\infty$, we have $-\infty \leqslant x$ for all extended real x according to Definition 6.2.3(c).

(b) We have to show the trichotomy of extended real numbers, i.e. that we always have exactly one of the statements $x < y$, $x = y$ or $x > y$. Both x and y can be either real numbers, $+\infty$ or $-\infty$, so that we have nine cases to study exhaustively.

(i) If both x, y are real numbers, then this is simply Proposition 5.4.7(a)[23].

(ii) If x is a real number and $y = +\infty$, then we have $x \leqslant y$ by Definition 6.2.3(b), and we also have $x \neq y$ by Definition 6.2.1. Thus, $x < y$ is true, and the two other statements $x = y$ and $x > y$ are false. (It is easy to see that $x > y$ is false, because it corresponds to none of the cases listed in Definition 6.2.3.)

(iii) If x is a real number and $y = -\infty$, then we have $y \leqslant x$ according to Definition 6.2.3(c), and $x \neq y$ according to Definition 6.2.1. Thus, $x > y$ is true; whereas $x = y$ is false, and $x < y$ is also false because it corresponds to none of the cases listed in Definition 6.2.3.

(iv) If $x = +\infty$ and y is a real number, then we have $x \geqslant y$ by Definition 6.2.3(b), and $x \neq y$ by Definition 6.2.1. Thus, $x > y$ is true; whereas $x = y$ is false, and $x < y$ is also false because it corresponds to none of the cases listed in Definition 6.2.3.

(v) If $x = +\infty$ and $y = +\infty$, then we have $x = y$. Thus, both statements $x < y$ and $x > y$ are incompatible with this one. It means that exactly one statement ($x = y$) is true in this case.

(vi) If $x = +\infty$ and $y = -\infty$, then we have $y \leqslant x$ by Definition 6.2.3(b-c). According to Definition 6.2.1, we also have $x \neq y$. Thus, $y > x$ is true; whereas $x = y$ is false, and $x > y$ is also false because it corresponds to none of the cases listed in Definition 6.2.3.

(vii) The three last cases where $x = -\infty$ and y is either $-\infty$, a real, or $+\infty$, can be shown in a similar fashion, and are "symmetrical" with all previous proofs.

(c) We have to show that if $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$. We'll choose wisely below some of the cases that really make sense in such a siuation.

- If $z = +\infty$, then according to Definition 6.2.3(b), $x \leqslant z$ is always true, regardless of the value of y.

[23] Actually, the statement (a) from Proposition 4.2.9, and applied to reals by Proposition 5.4.7, but I'll shorten this for the rest of the exercise.

- If $z = -\infty$, only possible situation makes sense: the situation where $x = y = z = -\infty$. Thus, the transitivity ($x \leqslant z$) is still true in this case.
- If z is a real number, then x, y cannot be $+\infty$. If $y = -\infty$, then necessarily $x = -\infty$, and we thus have $x \leqslant z$. If x, y are real numbers, then it is simply Proposition 5.4.7(c). And if y is a real number and $x = -\infty$, then $x \leqslant z$ by Definition 6.2.3(c).

(d) Finally we have to show that if $x \leqslant y$, then $-y \leqslant -x$. If $y = +\infty$, then $-y = -\infty$ and thus $-y \leqslant -x$ for all extended real $-x$, thus the statement is true in this case. If $y = -\infty$, then we have necessarily $x = -\infty$, and thus $-x = -y$, so that the statement still holds. Finally, if y is a real number, then the statement is simply Proposition 5.4.7 if x is real, and is obvious if $x = -\infty$.

Exercise 6.2.2. — *Prove Theorem 6.2.11.*

We have to prove the following statements:

(a) For every $x \in E$ we have $x \leqslant \sup(E)$ and $x \geqslant \inf(E)$.

- First, suppose that $+\infty \in E$. We thus have $\sup(E) = +\infty$ by Definition 6.2.6(b); and by Definition 6.2.3(b), $x \leqslant +\infty = \sup(E)$ for all $x \in E$, as required.
- Now suppose that $+\infty \notin E$, but $-\infty \in E$. There are two possible cases here. If $x = -\infty$, then $-\infty = x \leqslant y$ for all extended real$y$ by Definition 6.2.3(c); and in particular, with $y = \sup(E)$, we have $x \leqslant \sup(E)$ as required. On the other hand, if x is a real number, then $x \in E \backslash \{-\infty\}$, so that $x \leqslant \sup(E \backslash \{-\infty\}) := \sup(E)$ by Definition 6.2.6(c), as required.
- Finally, if E consists only of real numbers, we have $x \leqslant \sup(E)$ by Definition 5.5.10 as long as E has an upper bound; otherwise we have $\sup(E) = +\infty$ by Definition 6.2.6(b), and thus $x \leqslant \sup(E)$ bu Definition 6.2.3(b). Thus, in all cases, we have indeed $x \leqslant \sup(E)$.
- The other statement, $x \geqslant \inf(E)$, is a direct consequence. Indeed, if $x \in E$, then by definition, $-x \in -E$, and thus $-x \leqslant \sup(-E)$ according to our previous conclusions. And finally, $-x \leqslant \sup(-E)$ is equivalent to $x \geqslant -\sup(-E)$, i.e. $x \geqslant \inf(E)$, according to Proposition 6.2.5(d) and the definition of the greatest lower bound.

(b) Suppose that $M \in \overline{\mathbb{R}}$ is an upper bound for E, i.e., $x \leqslant M$ for all $x \in E$. Then we have $\sup(E) \leqslant M$.

- If $+\infty \in E$, then $\sup(E) = +\infty$ by Definition 6.2.6(b). But the only possible upper bound for E is $+\infty$, so that we have necessarily $\sup(E) = M = +\infty$. The statement is thus true is this case.
- If E consists only of real numbers, there are three sub-cases. If E is empty, the $\sup(E) = -\infty$, so that $\sup(E) \leqslant M$ regardless of the value of M. If E is non-empty and not bounded above, then $\sup(E) = +\infty$ and M can only be equal to $+\infty$, so that $\sup(E) = M$. And finally, if E is non-empty but bounded above, then the results comes from Definition 5.5.10.

- If $+\infty \notin E$ but $-\infty \in E$, then $\sup(E) := \sup(E\backslash\{-\infty\})$. But the set $E\backslash\{-\infty\}$ consists only of real numbers, so that we can basically go back to the previous case, to get $\sup(E) = \sup(E\backslash\{-\infty\}) \leqslant M$.

EXERCISE 6.3.1. — *Verify the claim in Example 6.3.4: Let* $a_n := 1/n$*; thus* $(a_n)_{n=1}^{\infty}$ *is the sequence* $1, 1/2, 1/3, \ldots$. *Then* $\sup(a_n)_{n=1}^{\infty} = 1$ *and* $\inf(a_n)_{n=1}^{\infty} = 0$.

Let's prove those two statements separately.

- First, note that 1 is an upper bound for $(a_n)_{n=1}^{\infty}$: since $n \geqslant 1$, we always have $1/n \leqslant 1$. Furthermore, let M be an upper bound for $(a_n)_{n=1}^{\infty}$. By definition, we must have $M \geqslant a_n$ for all $n \geqslant 1$; and in particular $M \geqslant a_1 = 1$. Thus, 1 is indeed the least upper bound of $(a_n)_{n=1}^{\infty}$.
- Second, 0 is a lower bound for $(a_n)_{n=1}^{\infty}$: we obviously have $1/n \geqslant 0$ for all $n \geqslant 1$. Let's suppose for the sake of contradiction that there exists a greater lower bound m of $(a_n)_{n=1}^{\infty}$. By definition, this means that $a_n \geqslant 0$ for all $n \geqslant 1$, and that $m > 0$. But according to the archimedean property with $\varepsilon = 1$, there exists a natural number N such that $mN \geqslant 1$, i.e. $m \geqslant 1/N$, that is to say $a_N \leqslant m$. This is a contradiction, since m is not a lower bound anymore. Thus, 0 is indeed the greatest lower bound of $(a_n)_{n=1}^{\infty}$.

EXERCISE 6.3.2. — *Prove Proposition 6.3.6.*

Let $E = \{a_n : n \geqslant m\}$, and $x = \sup(a_n)_{n=m}^{\infty}$. By Definition 6.3.1, we thus have $x = \sup(E)$. Let's prove all statements from the Proposition.

- Obviously, we have $a_n \leqslant x$ for all $n \geqslant m$, according to Theorem 6.2.11(a).
- Let be $M \in \overline{\mathbb{R}}$ an upper bound for $(a_n)_{n=m}^{\infty}$. This is equivalent to say that M is an upper bound for E. Thus, according to Theorem 6.2.11(b), we have $M \geqslant x$.
- Now let be $y \in \overline{\mathbb{R}}$ such that $y < x$. Suppose, for the sake of contradiction, that $y \geqslant a_n$ for all $n \geqslant m$. This means that y is an upper bound for $(a_n)_{n=m}^{\infty}$, and that $y < x$. This contradicts the conclusion from the previous bullet point of this exercise. Thus, there exists an $n \geqslant m$ such that $y < a_n$. And since $a_n \in E$, we also have $y < a_n \leqslant x$, as required.

EXERCISE 6.3.3. — *Prove Proposition 6.3.8. (increasing bounded sequences converge).*

Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers. The hypotheses are, first, that $a_{n+1} \geqslant a_n$ for all $n \geqslant m$; and second, that there exists a positive integer M such that $|a_n| \leqslant M$ for all $n \geqslant m$.

Let be $\ell = \sup(a_n)_{n=m}^{\infty}$. Since $(a_n)_{n=m}^{\infty}$ is bounded, ℓ is a real number. We will show that $(a_n)_{n=m}^{\infty}$ converges to ℓ.

Let be $\varepsilon > 0$ a real number. First, we already know that:

$$a_n \leqslant \ell \tag{6.8}$$

for all $n \geqslant m$. Thus, in particular, we have $a_n \leqslant \ell + \varepsilon$ for all $n \geqslant m$.

Also, according to the third statement of Proposition 6.3.6, if y is a real number such that $y < \ell$, there exists at least one element of $(a_n)_{n=m}^{\infty}$, say a_{n_0}, such that $y < a_{n_0} < \ell$. In particular, with $y = \ell - \varepsilon$, we thus have:

$$\ell - \varepsilon < a_{n_0} \leqslant a_n \tag{6.9}$$

But recall that the sequence $(a_n)_{n=m}^{\infty}$ is an increasing sequence: we have (as a quick induction shows) $a_n \geqslant a_{n_0}$ for all $n \geqslant n_0$. Combining the two equations (6.8) and (6.9), we get $\ell - \varepsilon < a_{n_0} \leqslant a_n \leqslant \ell \leqslant \ell + \varepsilon$ for all $n \geqslant n_0$.

Let's summarise:

$$\forall \varepsilon > 0, \exists n_0 \geqslant m \quad : \quad n \geqslant n_0 \Longrightarrow \ell - \varepsilon \leqslant a_n \leqslant \ell + \varepsilon \tag{6.10}$$

which means precisely that $(a_n)_{n=m}^{\infty}$ converges to ℓ, as required.

Exercise 6.3.4. — *Explain why Proposition 6.3.10 fails when $x > 1$. In fact, show that the sequence $(x^n)_{n=1}^{\infty}$ diverges when $x > 1$. Compare this with the argument in Example 1.2.3; can you now explain the flaws in the reasoning in that example?*

First, the proof of Proposition 6.3.10 supposes that the sequence $(x^n)_{n=1}^{\infty}$ is decreasing and has a lower bound of 0, which is not the case here: the sequence $(x^n)_{n=1}^{\infty}$ is increasing and has no real upper bound (or, let's say that its upper bound is $+\infty$). So the situation is not identical, and even not "symmetrical".

Actually, when $x > 1$, the sequence $(x^n)_{n=1}^{\infty}$ diverges to $+\infty$. Let's suppose, for the sake of contradiction, that it converges to a real number ℓ instead. Since $x > 1$, we have $0 < 1/x < 1$, so that by Proposition 6.3.10, $(1/x^n)_{n=1}^{\infty}$ is a sequence that converges to 0. Thus, if we denote $a_n = x^n \times 1/x^n$, the sequence $(a_n)_{n=1}^{\infty}$ is the product of two convergent sequences, so that if we apply the limit laws (Theorem 6.1.19(b)), we would have $\lim_{n\to\infty} a_n = \ell \times 0 = 0$. But this is impossible, since $a_n = 1$ for all $n \geqslant 1$. It is thus impossible for $(x^n)_{n=1}^{\infty}$ to be a convergent sequence if $x > 1$.

Finally, the issue in Example 1.2.3 is that there is a "hidden" assumption in it: it is not explicitly stated, but it is a strong one. In this example, we begin with the statement: "let L be the limit $L = \lim_{n\to\infty} x^n$", and then we apply the limit laws (Theorem 6.1.19) as if we were sure that $(x^n)_{n=1}^{\infty}$ is a convergent sequence, i.e. that L is a real number. But in the next steps of this example (where we take $x = -1$ or $x = 2$), this is absolutely not the case: instead this sequence diverges. Thus, it was wrong to apply the limit laws, and the whole reasoning is flawed. It is only correct in the case where $x \leqslant 1$, i.e. where $L = 0$ or $x = 1$, which is also a remark we make during this example.

Exercise 6.4.1. — *Prove Proposition 6.4.5.*

Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, which converges to c. There are two statements to prove.

- First, let's prove that c is a limit point for $(a_n)_{n=m}^{\infty}$. Intuitively this is obvious since being a limit point is somewhat weaker than being a limit, but show it rigorously this requires some manipulations on quantifiers.

  Let be $\varepsilon > 0$ a real number and $N \geqslant m$ a positive integer. We must show that there exists $k \geqslant N$ such that $|a_k - c| \leqslant \varepsilon$.

Actually, we know that $(a_n)_{n=m}^{\infty}$ converges to c. It means that there exists $N' \geqslant m$ such that, for *all* $n \geqslant N'$, we have $|a_n - c| \leqslant \varepsilon$.

Thus, any positive integer $n \geqslant \max(N, N')$ will do the trick, and in particular $n = \max(N, N')$: we will have $n \geqslant N$ and $|a_n - c|$ as required.

- Now let's prove that c is the *only* limit point of $(a_n)_{n=m}^{\infty}$. Here, drawing a picture will help a lot. Let's suppose, for the sake of contradiction, that there exists another limit point $c' \neq c$ for $(a_n)_{n=m}^{\infty}$. Let's consider the positive real number $\varepsilon = |c-c'|/3$ (we know it is positive, because $c - c' \neq 0$ by hypothesis).

  Since $(a_n)_{n=m}^{\infty}$ converges to c, there exists a positive integer $N \geqslant m$ such that $|a_n - c| \leqslant \varepsilon$ for all $n \geqslant N$. Also, since c' is a limit point, there exists $n_0 \geqslant N$ such that $|a_{n_0} - c'| \leqslant \varepsilon$. Thus:

$$\begin{aligned} |c - c'| &\leqslant |c - a_{n_0} + a_{n_0} - c'| \\ &\leqslant |a_{n_0} - c| + |a_{n_0} - c'| \\ &\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3} = \frac{2}{3}|c - c'| \end{aligned}$$

  This is a clear contradiction. Thus, $(a_n)_{n=m}^{\infty}$ has no other limit point than c.

Exercise 6.4.2. — *State and prove analogues of Exercises 6.1.3 and 6.1.4 for limit points, limit superior and limit inferior.*

As stated before, Exercises 6.1.3 and 6.1.4 provided very similar statements, so that we'll just prove here the analogues for Exercise 6.1.3. They can be stated as follows.

Let $m' \geqslant m$ two positive integers, and c a real number.

1. Statement for limit points: c is a limit point for $(a_n)_{n=m}^{\infty}$ iff c is a limit point for $(a_n)_{n=m'}^{\infty}$.

   Actually, c is a limit point for $(a_n)_{n=m}^{\infty}$ iff:

$$\forall \varepsilon > 0, \forall N \geqslant m, \exists n \geqslant N \;:\; |a_n - c| \leqslant \varepsilon \tag{6.11}$$

   Similarly, c is a limit point for $(a_n)_{n=m'}^{\infty}$ iff:

$$\forall \varepsilon > 0, \forall N' \geqslant m', \exists n \geqslant N' \;:\; |a_n - c| \leqslant \varepsilon \tag{6.12}$$

   - First suppose that c is a limit point for $(a_n)_{n=m}^{\infty}$, and let's show that it is a limit point for $(a_n)_{n=m'}^{\infty}$. Let be $\varepsilon > 0$ a positive real and $N' \geqslant m'$ a positive integer. Since $m' \geqslant m$, we also have $N' \geqslant m$. Thus, by (6.11), we know that there exists $n \geqslant N'$ such that $|a_n - c|$, as required. Thus, c is a limit point for $(a_n)_{n=m'}^{\infty}$.
   - Now suppose that c is a limit point for $(a_n)_{n=m'}^{\infty}$, and let's show that it is a limit point for $(a_n)_{n=m}^{\infty}$. Let be $\varepsilon > 0$ a positive real and $N \geqslant m$ a positive integer. We can distinguish two sub-cases here. If $N \geqslant m'$, then (6.12) indeed provides a $n \geqslant N$ such that $|a_n - c| \leqslant \varepsilon$, as required. Else, if $m \leqslant N < m'$, then according to (6.12), there exists an $n \geqslant m'$ such that $|a_n - c| \leqslant \varepsilon$. We thus have $m \leqslant N < m' \leqslant n$, i.e. in particular $N < n'$, and $|a_n - c| \leqslant \varepsilon$, as required. Thus, in both sub-cases, c is a limit point for $(a_n)_{n=m}^{\infty}$.

2. Statement for limit superior: ℓ is the limit superior of $(a_n)_{n=m}^{\infty}$ iff ℓ is the limit superior of $(a_n)_{n=m'}^{\infty}$.

   Actually, ℓ is the limit superior of $(a_n)_{n=m}^{\infty}$ iff:

$$\ell = \inf(a_N^+)_{N=m}^{\infty} \tag{6.13}$$

   Similarly, ℓ' is the limit superior of $(a_n)_{n=m'}^{\infty}$ iff:

$$\ell' = \inf(a_N^+)_{N=m'}^{\infty} \tag{6.14}$$

   First, note that the sequence $(a_N^+)_{N=m}^{\infty}$ is decreasing. In particular, thanks to Proposition 6.3.8, it means that its limit is equal to its greatest lower bound, i.e., $\lim_{N\to+\infty}(a_N^+) = \inf(a_N^+)_{N=m}^{\infty}$. But we already know, thanks to Exercises 6.1.3 and 6.1.4, that this limit does not depend on the starting index of the sequence. Thus, we indeed have $\ell = \ell'$.

3. The statement and proof for limit inferior are similar to the previous one, since the sequence $(a_N^-)_{N=m}^{\infty}$ is increasing.

Exercise 6.4.3. — *Prove parts (c), (d) and (e) of Proposition 6.4.12.*

Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, L^+ its limit superior and L^- its limit inferior. In the proofs below, we will make the additional hypothesis that L^- and L^+ are finite; the general case may need a painful split into cases depending on whether they are finite or not.

(c) Prove that $\inf(a_n)_{n=m}^{\infty} \leqslant L^- \leqslant L^+ \leqslant \sup(a_n)_{n=m}^{\infty}$.

We will prove each inequality separately.

- First let's prove that $\inf(a_n)_{n=m}^{\infty} \leqslant L^-$. By Definition 6.4.6, $L^- = \sup(a_N^-)_{N=m}^{\infty}$. Thus, we have $L^- \geqslant a_N^-$ for all $N \geqslant m$. In particular, for $N = m$, we have $L^- \geqslant a_m^- = \inf(a_n)_{n=m}^{\infty}$, as required.
- The proof for $L^+ \leqslant \sup(a_n)_{n=m}^{\infty}$ is similar. By Definition 6.4.6, $L^+ = \inf(a_N^+)_{N=m}^{\infty}$. Thus, we have $L^+ \leqslant a_N^+$ for all $N \geqslant m$. In particular, for $N = m$, we have $L^+ \leqslant a_m^+ = \sup(a_n)_{n=m}^{\infty}$, as required.
- Finally, let's prove that $L^- \leqslant L^+$. Let be $\varepsilon > 0$ a positive real. Since $L^+ + \varepsilon > L^+$, by Proposition 6.4.12(a), there exists a positive integer $N_1 \geqslant m$ such that $a_n < L^+ + \varepsilon$ whenever $n \geqslant N_1$. Similarly, there exists a positive integer $N_2 \geqslant m$ such that $a_n > L^- - \varepsilon$ whenever $n \geqslant N_2$. If we set $N := \max(N_1, N_2)$, then for all $n \geqslant N$, we get:

$$L^- - \varepsilon < a_n < L^+ + \varepsilon$$

  which leads to $L^- - L^+ < 2\varepsilon$ for all $\varepsilon > 0$. This implies that $L^- - L^+ \leqslant 0$ (otherwise we would have an obvious contradiction).

  Thus, we have indeed $L^- \leqslant L^+$, as expected[24]

[24]For an alternative proof that may not require the additional hypothesis that both L^- and L^+ are finite, argue by contradiction, and show that somehow, if $L^- > L^+$, we will get eventually something like $a_n < a_n$ later in the sequence.

(d) Prove that if c is any limit point of $(a_n)_{n=m}^{\infty}$, then we have $L^- \leqslant c \leqslant L^+$.

We only show here the inequality $c \leqslant L^+$; its counterpart $L^- \leqslant c$ can be proved the same way.

Let's suppose, for the sake of contradiction, that we have $c > L^+$. Let be $\varepsilon = \frac{c-L^+}{3}$, which is by hypothesis a positive real number. We have $L^+ + \varepsilon > L^+$. Thus, by Proposition 6.4.12(a), there exists $N \geqslant m$ such that:

$$a_n < L^+ + \varepsilon = \frac{2L^+ + c}{3} \text{ for all } n \geqslant N \tag{6.15}$$

But this is a contradiction with the fact that c is a limit point for $(a_n)_{n=m}^{\infty}$. Indeed, if c is a limit point, then there exists $n \geqslant N$ such that $|a_n - c| \leqslant \varepsilon$, i.e.: $c - \varepsilon < a_n < c + \varepsilon$, and in particular

$$a_n > c - \varepsilon = \frac{3c - c + L^+}{3} = \frac{2c + L^+}{3} \tag{6.16}$$

But equations (6.15) and (6.16) are incompatible, since they would lead to $a_n < \frac{2L^+ + c}{3} < \frac{2c + L^+}{3} < a_n$, i.e. $a_n < a_n$. Thus, our starting hypothesis $c \leqslant L^+$ is false, and we have necessarily $c > L^+$ as required.

(e) Prove that if L^+ is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$ (and similarly for L^-).

Once again, we only give the proof for L^+, the other one being very similar.

Let be $\varepsilon > 0$ a positive real, and $N \geqslant m$ a positive integer. Since $L^+ + \varepsilon > L^+$, according to Proposition 6.4.12(a), there exists $N' \geqslant m$ such that $a_n < L^+ + \varepsilon$ for all $n \geqslant N'$. In particular, if $n \geqslant M := \max(N, N')$, then we have $a_n < L^+ + \varepsilon$.

Also, we have $L^+ - \varepsilon < L^+$. Thus, according to Proposition 6.4.2(b), there exists $n \geqslant M$ such that $a_n > L^+ - \varepsilon$.

Thus, for any given $\varepsilon > 0$ and any given $N \geqslant m$, there exists $n \geqslant M \geqslant N$ such that $L^+ - \varepsilon < a_n < L^+ + \varepsilon$, i.e. $|a_n - L^+| \leqslant \varepsilon$. L^+ is thus a limit point, as required.

(f) Prove that $(a_n)_{n=m}^{\infty}$ converges to c iff $L^+ = L^- = c$.

- First, let's suppose that $L^+ = L^- = c$. According to Proposition 6.4.5, if $(a_n)_{n=m}^{\infty}$ is a convergent sequence, then its limit is a limit point. But according to Proposition 6.4.2(d), if $L^- = L^+ = c$, the only possible limit point for $(a_n)_{n=m}^{\infty}$ is c. However, we have only shown that *if $(a_n)_{n=m}^{\infty}$ is a convergent sequence, then it converges to* c, but we do not know whether or not $(a_n)_{n=m}^{\infty}$ is convergent.
  Let be $\varepsilon > 0$. According to Proposition 6.4.2(a), if $x = L^+ + \varepsilon = c + \varepsilon > L^+$, then there exists $N \geqslant m$ such that $a_n < c + \varepsilon$ for all $n \geqslant N$. Similarly, since $y = L^- - \varepsilon = c - \varepsilon < L^-$, then there exists $N' \geqslant m$ such that $a_n > c - \varepsilon$ for all $n \geqslant N'$.
  Thus, for all $n \geqslant \max(N, N')$, we have $c - \varepsilon < a_n < c + \varepsilon$, i.e. $|a_n - c| \leqslant \varepsilon$. This means that $(a_n)_{n=m}^{\infty}$ converges to c.
- Conversely, let's suppose that $(a_n)_{n=m}^{\infty}$ converges to c.
  Let be $\varepsilon > 0$. First, since $(an)_{n=m}^{\infty}$ converges to c, there exists $N \geqslant m$ such that $|a_n - c| \leqslant \varepsilon/2$ for all $n \geqslant N$. Also, since L^- is a limit point, there exists $k \geqslant N$ such that $|a_k - L^-| \leqslant \varepsilon/2$. Thus, we have both $|a_k - c| \leqslant \varepsilon/2$ and $|a_k - L^-| \leqslant \varepsilon/2$.

We then derive, by triangular inequality: $|L^- - c| \leqslant |L^- - a_n| + |a_n - c| \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon$. This means that $|L^- - c| \leqslant \varepsilon$ for all $\varepsilon > 0$, which is equivalent to $L^- = c$.

Similarly, we could prove that $c = L^+$.

Finally, we indeed have $L^- = c = L^+$ as required.

Exercise 6.4.4. — *Prove Lemma 6.4.13.*

Let be $(a_n)_{n=m}^\infty$ and $(b_n)_{n=m}^\infty$ two sequences of real numbers, such that $a_n \leqslant b_n$ for all $n \geqslant m$. We will prove the four statements of this lemma.

1. Prove that $\sup(a_n)_{n=m}^\infty \leqslant \sup(b_n)_{n=m}^\infty$.

   Let be $B = \sup(b_n)_{n=m}^\infty$. By definition, B is an upper bound for $(b_n)_{n=m}^\infty$, so that we have $B \geqslant b_n \geqslant a_n$ for all $n \geqslant m$. In particular, B is also an upper bound for $(a_n)_{n=m}^\infty$. And, by Definition 5.5.5 of a least upper bound, we thus have $B \geqslant \sup(a_n)_{n=m}^\infty$ as required.

2. Prove that $\inf(a_n)_{n=m}^\infty \leqslant \inf(b_n)_{n=m}^\infty$.

   A very similar argument applies. Let be $A = \sup(a_n)_{n=m}^\infty$. By definition, A is an lower bound for $(a_n)_{n=m}^\infty$, so that we have $A \leqslant a_n \leqslant b_n$ for all $n \geqslant m$. In particular, A is also a lower bound for $(b_n)_{n=m}^\infty$. We thus have $A \leqslant \inf(b_n)_{n=m}^\infty$ as required.

3. Prove that $\limsup a_n \leqslant \limsup b_n$.

   Let be, for any $N \geqslant m$, $a_N^+ = \sup(a_n)_{n=N}^\infty$, and $b_N^+ = \sup(b_n)_{n=N}^\infty$. According to the first point of this exercise, we have $a_N^+ \leqslant b_N^+$ for all $N \geqslant m$. According to the second point, we have $\inf(a_N^+)_{N=m}^\infty \leqslant \inf(b_N^+)_{N=m}^\infty$, i.e. $\limsup a_n \leqslant \limsup b_n$ as required.

4. Prove that $\liminf a_n \leqslant \liminf b_n$.

   Once again, the proof is similar to the previous point.

Exercise 6.4.5. — *Use Lemma 6.4.13 to prove Corollary 6.4.14.*

Let $(a_n)_{n=m}^\infty, (b_n)_{n=m}^\infty, (c_n)_{n=m}^\infty$ be sequences of real numbers such that $a_n \leqslant b_n \leqslant c_n$ for all $n \geqslant m$; and the sequences $(a_n)_{n=m}^\infty$ and $(c_n)_{n=m}^\infty$ converge to L.

According to Corollary 6.4.14, we have $\limsup a_n \leqslant \limsup b_n \leqslant \limsup c_n$. But according to Proposition 6.4.12(f), when a sequence converges to a real number L, its limit superior is simply equal to its limit L. Thus, we have $L = \limsup a_n \leqslant \limsup b_n \leqslant \limsup c_n = L$, i.e. $\limsup b_n = L$.

The same statement apply for $\liminf b_n$, which is also equal to L.

Thus, still according to Proposition 6.4.12(f), $(b_n)_{n=m}^\infty$ converges to L, as required.

Exercise 6.4.6. — *Give an example of two bounded sequences $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ such that $a_n < b_n$ for all $n \geqslant 1$, but that $\sup(a_n)_{n=1}^\infty \nless (b_n)_{n=1}^\infty$. Explain why this does not contradict Lemma 6.4.13.*

The sequences defined by $a_n = -1/n$ and $b_n = 0$ for all $n \geqslant 1$ are suitable, since $-1/n < 0$ for all $n \geqslant 1$, but both least upper bounds are equal to 0.

This does not contradict Lemma 6.4.13 which deals with large inequalities; instead, it is even perfectly in accordance with Remark 5.4.11 in the previous chapter.

Exercise 6.4.7. — *Prove Corollary 6.4.17.*

This corollary says that $(a_n)_{n=m}^{\infty}$ converges to 0 iff $(|a_n|)_{n=m}^{\infty}$ converges to 0. Let's denote $b_n = |a_n|$ in what follows.

- First suppose that $\lim_{n\to\infty}|a_n| = 0$. We know that we have, for all $n \geqslant m$, $-|a_n| \leqslant a_n \leqslant |a_n|$. According to the squeeze test, we have indeed $\lim_{n\to\infty}a_n = 0$.
- Now suppose that $\lim_{n\to\infty}a_n = 0$. Let $\varepsilon > 0$ be a positive real number. We have $|b_n - 0| = |b_n| = ||a_n|| = |a_n| = |a_n - 0|$. We know, by definition, that there exists $N \geqslant m$ such that, for all $n \geqslant N$, we have $|a_n - 0| \leqslant \varepsilon$, and thus $|b_n - 0| \leqslant \varepsilon$. This means that $(b_n)_{n=m}^{\infty}$ converges to 0, as required.

Exercise 6.5.1. — *Show that $\lim_{n\to\infty}1/n^q = 0$ for any rational $q > 0$.*

Since q is a positive rational, let's suppose that $q := a/b$, with $a, b > 0$ two positive integers. With a little algebra, we can rewrite $1/n^q = (1/n)^q = (1/n)^{a/b} = ((1/n)^{1/b})^a$, using in particular Definition 5.6.7, and in a somewhat hidden manner, equation (5.13) from this document.

Using Corollary 6.5.1, we know that $\lim_{n\to\infty}1/n^{1/k} = 0$ for every integer $k \geqslant 1$. In particular, we have $\lim_{n\to\infty}(1/n)^{1/b} = \lim_{n\to\infty}(1/n^{1/b}) = 0$. Thus, using the limit laws (Theorem 6.1.19(b), iterated a times), we conclude that $\lim_{n\to\infty}1/n^q = \lim_{n\to\infty}(1/n^{1/b})^a = 0^a = 0$.

Exercise 6.5.2. — *Prove Lemma 6.5.2.*

Here, there are a lot of cases to consider.

- First, if $x = 1$, then $x^n = 1$ for all $n \geqslant 1$, so that $(x^n)_n$ is a constant sequence. Thus, $\lim_{n\to\infty}x^n = 1$.
- Similarly, if $x = 0$, then $\lim_{n\to\infty}x^n = 0$.
- If $-1 < x < 1$ and $x \neq 0$, we have $0 < |x| < 1$. This means, according to Proposition 6.3.10, that $\lim_{n\to\infty}|x|^n = 0$. But, we can note that we have, for all $n \geqslant 1$, the inequality $-|x^n| \leqslant x^n \leqslant |x|^n$. Since $\lim_{n\to\infty}|x|^n = \lim_{n\to\infty} -|x|^n = 0$, then according to the squeeze test, we also have $\lim_{n\to\infty}x^n = 0$.
- If $x = -1$, then the sequence $(x^n)_n$ is alternatively equal to 1 or -1. Consequently, it has two limits points, which are 1 and -1; it is even possible to show that $L^- = -1$ and $L^+ = 1$. Thus, it cannot converge, since a convergent sequence has only one limit point, and cannot have $L^- \neq L^+$.
- If $x > 1$, then Exercise 6.3.4 says that the sequence $(x^n)_n$ is divergent.
- Finally, if $x < -1$, then the exact same reasoning performed in Exercise 6.3.4 would also prove that $(x^n)_n$ is divergent.

Thus, as required, we have shown that $\lim_{n\to\infty}x^n$ exists and is equal to zero when $|x| < 1$, exists and is equal to 1 when $x = 1$, and diverges when $x = -1$ or when $|x| > 1$.

Exercise 6.5.3. — *Prove Lemma 6.5.3. (You may need to treat the cases $x \geqslant 1$ and $x < 1$ separately. You might wish to first use Lemma 6.5.2 to prove the preliminary result that for every $\varepsilon > 0$ and every real number $M > 0$, there exists an n such that $M^{1/n} \leqslant 1 + \varepsilon$.)*

We first prove the preliminary result suggested by Terence Tao. Let be $\varepsilon > 0$ a positive real number, so that $1 + \varepsilon > 1$. By Lemma 6.5.2, this implies that the sequence $(a_n)_{n=1}^{\infty}$ defined by $a_n = (1+\varepsilon)^n$ diverges. Furthermore, by Lemma 5.6.9(e), $(1+\varepsilon)^n$ is a strictly increasing function of n. It cannot be bounded because, by Proposition 6.3.8, it would imply that $(a_n)_{n=1}^{\infty}$ is a convergent sequence. Thus, $(a_n)_{n=1}^{\infty}$ is strictly increasing and not bounded above. In other words, for all $M > 0$, there exists a natural number N such that $(1+\varepsilon)^N \geqslant M$; and since we have an increasing function of n, we even get $(1+\varepsilon)^n \geqslant M$ for all $n \geqslant N$. This is equivalent to $(1+\varepsilon) \geqslant M^{1/n}$ by Lemma 5.6.6(d), which proves this preliminary result.

Let's go back to the main claim. We'll distinguish two cases. Let be $\varepsilon > 0$.

1. If $x \geqslant 1$, we have $x^{1/n} > 1$ for all $n \geqslant 1$ (because if we suppose, for the sake of contradiction, that $x^{1/n} < 1$, we would have $(x^{1/n})^n < 1^n$, i.e. $x < 1$). Thus, we have $0 < x^{1/n} - 1$ for all $n \geqslant 1$. Furthermore, according to the preliminary result, there exists an N such that $x^{1/n} - 1 \leqslant \varepsilon$ for all $n \geqslant N$. We thus have $-\varepsilon \leqslant 0 \leqslant x^{1/n} - 1 \leqslant \varepsilon$ for all $n \geqslant N$, i.e. $|x^{1/n} - 1| \leqslant \varepsilon$ for all $n \geqslant N$. This means that $\lim_{n\to\infty} x^{1/n} = 1$, as required.

2. If $x < 1$, we have $1/x > 1$. Thus, when considering $1/x$, we are back to the previous case, i.e. we have $\lim_{n\to\infty}(1/x)^{1/n} = 1$. But, by Lemma 5.6.9(b), we have $(1/x)^{1/n} = (x^{-1})^{1/n} = x^{-1/n} = 1/x^{1/n}$. Thus, we have $\lim_{n\to\infty} 1/x^{1/n} = 1$. Finally, by the limit laws (Theorem 6.1.19(e)), we get $\lim_{n\to\infty} x^{1/n} = 1/1 = 1$, as required.

Exercise 6.6.1. — *Prove Lemma 6.6.4.*

This lemma states that being a subsequence is reflexive and transitive.

- Let's show that $(a_n)_{n=1}^{\infty}$ is a subsequence of $(a_n)_{n=1}^{\infty}$. This is obvious: let's take the strictly increasing function $f(n) = n$ for all $n \in \mathbb{N}^*$: this will indeed give $a_n = a_{f(n)}$, i.e. $(a_n)_{n=1}^{\infty}$ as a subsequence of $(a_n)_{n=1}^{\infty}$.

- Now consider that $(b_n)_{n=1}^{\infty}$ is a subsequence of $(a_n)_{n=1}^{\infty}$, and that $(c_n)_{n=1}^{\infty}$ is a subsequence of $(b_n)_{n=1}^{\infty}$. It means that we have, for all $n \in \mathbb{N}^*$, $b_n = a_{f(n)}$ and $c_n = b_{g(n)}$, with f, g two strictly increasing functions from $\mathbb{N}^*$ to $\mathbb{N}^*$. The function $f \circ g$ is also strictly increasing, as the composition of two strictly increasing functions. Thus, we have $c_n = b_{g(n)} = a_{f \circ g(n)}$ for all $n \in \mathbb{N}^*$, i.e. $(c_n)_{n=1}^{\infty}$ as a subsequence of $(a_n)_{n=1}^{\infty}$.

Exercise 6.6.2. — *Can you find two sequences $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ which are not the same sequence, but such that each is a subsequence of the other?*

It might appear counterintuitive, but the answer is yes! One can think of the sequences defined by $a_n = (-1)^n$ and $b_n = (-1)^{n+1}$. We have, for any natural number n, $a_n = b_{n+1}$ and $b_n = a_{n+1}$. I.e., we have found a strictly increasing function $f : n \mapsto n+1$ such that $a_n = b_{f(n)}$ and $b_n = a_{f(n)}$ for all $n \in \mathbb{N}$; and these sequences are not equal.

Note that any example where $(b_n)_{n=0}^{\infty}$ is simply a "shift" of a periodic sequence $(a_n)_{n=0}^{\infty}$ would do the trick. For instance, the sequence $a_n = 0, 0, 1, 1, 0, 0, \ldots$ and the sequence $b_n = 1, 1, 0, 0, 1, 1, \ldots$, with the function $f : n \mapsto n+2$.

Exercise 6.6.3. — *Let $(a_n)_{n=0}^{\infty}$ be a sequence which is not bounded. Show that there exists a subsequence $(b_n)_{n=0}^{\infty}$ of $(a_n)_{n=0}^{\infty}$ such that $\lim_{n\to\infty} 1/b_n$ exists and is equal to zero. (Hint: for each natural number j, recursively introduce the quantity $n_j := \min\{n \in \mathbb{N} : |a_n| \geqslant j, n > n_j\}$,* *omitting the condition $n > n_{j-1}$ when $j = 0$. First explain why the set $\{n \in \mathbb{N} : |a_n| \geqslant j, n > n_j\}$ is non-empty; then set $b_j := a_{n_j}$. To ensure the existence and uniqueness of the minimum, one either needs to invoke the well ordering principle (which we have placed in Proposition 8.1.4, but whose proof does not rely on any material not already presented), or the least upper bound principle (Theorem 5.5.19).)*

For any natural number j, let's denote $A_j := \{n \in \mathbb{N} : |a_n| \geqslant j, n > n_{j-1}\}$, where n_j is recursively defined by:

$$\begin{cases} n_0 := \min\{n \in \mathbb{N} : |a_n| \geqslant 0\} \\ n_j := \min\{n \in \mathbb{N} : |a_n| \geqslant j; n > n_j\} \end{cases}$$

- For all natural number j, the set A_j is non-empty because $(a_n)_{n=0}^{\infty}$ is not bounded. Indeed, let's suppose for the sake of contradiction that there exists a natural number j such that $A_j = \varnothing$. It means that, for all natural numbers $n > n_j$, we have $|a_n| < j$. In such a case, $(a_n)_{n=j}^{\infty}$ is bounded by j, and thus $(a_n)_{n=0}^{\infty}$ is bounded by $M := \max(j, |a_0|, \dots, |a_{j-1}|)$. This contradicts our initial hypothesis that $(a_n)_{n=0}^{\infty}$ is not bounded. Thus, A_j is non-empty for all $j \in \mathbb{N}$.
- A_j also has a lower bound for all $j \in \mathbb{N}$. Indeed, by definition, n_{j-1} is a lower bound for A_j. So, A_j is always non-empty and has a lower bound. Thus, by Theorem 5.5.19, A_j has a greatest lower bound, that we will denote $\inf(A_j)$. (Furthermore, it is unique, by Theorem 5.5.18.)
- Let's show that $\inf(A_j) \in A_j$ for all $j \in \mathbb{N}$. Since $\inf(A_j)$ is the greatest lower bound, $\inf(A_j) + 1/2$ is not a lower bound for A_j (otherwise we would have an obvious contradiction). Thus, there exists $n \in A_j$ such that $\inf(A_j) \leqslant n < \inf(A_j) + 1/2$. Let's suppose, for the sake of contradiction, that there exists an $m \in A_j$ such that $m < n$. We would have $\inf(A_j) \leqslant m < n < \inf(A_j) + 1/2$, so that we could place two distinct integers within a range of width $1/2$, which is impossible[25]. Thus, we have $n \leqslant m$ for all $m \in A_j$. It means that n is also a lower bound for A_j, and thus $n \leqslant \inf(A_j)$. And since we have both $n \leqslant \inf(A_j)$ and $n \geqslant \inf(A_j)$, we finally get $\inf(A_j) = n \in A_j$, as required[26]. Thus, the numbers n_j are well-defined.
- Finally, let be $\varepsilon > 0$ a positive real number, and a function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(j) = n_j$. It is a strictly increasing function by definition of the numbers n_j. Let's consider the sequence $(b_n)_{n=0}^{\infty}$ defined by $b_j := a_{n_j} = a_{f(j)}$ for all $j \in \mathbb{N}$.

  By definition, we also have $|a_{n_j}| \geqslant j$ for all $j \in \mathbb{N}$, which is equivalent to $1/|a_{n_j}| \leqslant 1/j$.

  By the archimedean property (Corollary 5.4.13, and Exercise 5.4.4), there exists a natural number j such that $\varepsilon > 1/j > 0$.

  Finally, let be a natural number $k > j$: we thus have $|a_{n_k}| \geqslant k > j$, i.e. $1/|a_{n_k}| < 1/j$.

  Thus, unwrapping all these findings: for any $\varepsilon > 0$, there exists a natural number j such that, for any $k > j$, we have $1/|a_{n_k}| = |b_k - 0| \leqslant 1/j \leqslant \varepsilon$, i.e., $\lim_{n\to\infty} b_n = 0$, as required.

[25] One may also give a proof by contradiction by writing $x + 1/2 - x = \underbrace{x + 1/2 - m}_{>0} + \underbrace{m - n}_{>1} + \underbrace{n - x}_{>0}$, to prove this fact in this particular situation.

[26] We actually have proven a generalizable lemma: for any non-empty set $X \subset \mathbb{N}$, the greatest lower bound $\inf(X)$ is in X. This will be useful in Exercise 6.6.5 as well; and will be useful in Chapter 8.

Exercise 6.6.4. — *Prove Proposition 6.6.5. (Note that one of the two implications has a very short proof.)*

First let's prove a preliminary result: if $f : \mathbb{N} \to \mathbb{N}$ is a strictly increasing function, then $f(n) \geqslant n$ for all $n \in \mathbb{N}$. We can use induction on n:

- for the base case $n = 0$, we indeed have $f(0) \geqslant 0$ because $f(0) \in \mathbb{N}$;
- now suppose inductively that $f(n) \geqslant n$. Since f is strictly increasing, we have $f(n+1) > f(n) \geqslant n$, and thus, $f(n+1) > n$, i.e. $f(n+1) \geqslant n+1$. This closes the induction for this preliminary result.

For the exercise itself, we must prove that the two statements "The sequence $(a_n)_{n=0}^{\infty}$ converges to L" and "Every subsequence of $(a_n)_{n=0}^{\infty}$ converges to L" are equivalent.

- First suppose that every subsequence of $(a_n)_{n=0}^{\infty}$ converges to L. According to Lemma 6.6.4, $(a_n)_{n=0}^{\infty}$ is a subsequence of $(a_n)_{n=0}^{\infty}$, thus in particular, $(a_n)_{n=0}^{\infty}$ converges to L.
- Now suppose that the sequence $(a_n)_{n=0}^{\infty}$ converges to L. Let be $\varepsilon > 0$, and $f : \mathbb{N} \to \mathbb{N}$ a strictly increasing function such that $(a_{f(n)})_{n=0}^{\infty}$ is a subsequence of $(a_n)_{n=0}^{\infty}$.

  Since $(a_n)_{n=0}^{\infty}$ converges to L, there exists a natural number N such that $|a_n - L| \leqslant \varepsilon$ for all $n \geqslant N$. But, according to our preliminary result, we have $f(n) \geqslant n$, and thus by transitivity, $f(n) \geqslant N$. This means that $|a_{f(n)} - L| \leqslant \varepsilon$. Unfolding those findings: there exists an N such that $|a_{f(n)} - L| \leqslant \varepsilon$ for all $n \geqslant N$, i.e., $(a_{f(n)})_{n=0}^{\infty}$ converges to L.

Thus, both statements are indeed equivalent.

Exercise 6.6.5. — *Prove Proposition 6.6.6. (Hint: to show that (a) implies (b), define the numbers n_j for each natural number j by the formula $n_j := \min\{n > n_{j-1} : |a_n - L| \leqslant 1/j\}$, with the convention $n_0 := 0$, explaining why the set $\{n > n_{j-1} : |a_n - L| \leqslant 1/j\}$ is non-empty. Then consider the sequence a_{n_j}. To ensure the existence and uniqueness of the minimum, one either needs to invoke the well ordering principle (which we have placed in Proposition 8.1.4, but whose proof does not rely on any material not already presented), or the least upper bound principle (Theorem 5.5.19).)*

This exercise is pretty similar to Exercise 6.6.3, and we will re-use here a preliminary result shown there—see also note 26.

1. First we show that (b) implies (a). Let be $f : \mathbb{N} \to \mathbb{N}$ a strictly increasing function such that $(a_{f(n)})_{n=0}^{\infty}$ converges to L. We must show that L is a limit point of $(a_n)_{n=0}^{\infty}$, i.e. that:
$$\forall \varepsilon > 0,\ \forall N \geqslant 0,\ \exists n \geqslant N \ : \ |a_n - L| \leqslant \varepsilon \tag{6.17}$$
   So let be $\varepsilon > 0$ and $N \geqslant 0$. If $(a_{f(n)})_{n=0}^{\infty}$ converges to L, there exists $M \geqslant 0$ such that $n \geqslant M \to |a_{f(n)} - L| < \varepsilon$. We know that f is a strictly increasing function, so that $f(n) \geqslant n$ for all $n \in \mathbb{N}$, as previously shown in Exercise 6.6.4. Thus, if one considers $p := \max(M, N)$, we have:
   - $f(p) \geqslant p \geqslant N$

- $f(p) \geqslant p \geqslant M$, so that $|a_{f(p)} - L| \leqslant \varepsilon$

  Choosing $n := f(p) \geqslant N$, the condition in formula (6.17) is thus verified.

2. Now we prove that (a) implies (b).

   - First, we note that the set $\{n > n_{j-1} : |a_n - L| \leqslant 1/j\}$ is non-empty for all $j \geqslant 1$. For instance, with $j = 1$, this set is equal to $\{n > 0 : |a_n - L| \leqslant 1\}$: this set is non-empty since L is a limit point of $(a_n)_{n=0}^{\infty}$. This gives an intuition for the general case. Generally speaking, let's suppose for the sake of contradiction that the set $A_j := \{n > n_{j-1} : |a_n - L| \leqslant 1/j\}$ is non empty for one given $j \geqslant 1$. It means that we can define $\varepsilon := 1/j$ and $N := n_{j-1}$ such that, for all $n \geqslant N$, we have $|a_n - L| > \varepsilon$. This is precisely the negation of the fact that L is a limit point of $(a_n)_{n=0}^{\infty}$ (see formula (6.17)): we have a clear contradiction. Thus, this set A_j is non-empty for every natural number $j \geqslant 1$.
   - Thus, for all $j \geqslant 1$, this set A_j is non-empty, and it has a lower bound (for instance, 0 is obviously a lower bound, since A_j is a subset of $\mathbb{N}$). By Theorem 5.5.9, we thus know that this set has a (unique) greatest lower bound, which we can write $\inf(A_j)$. By note 26, we have $\inf(A_j) \in A_j$, i.e., $\inf(A_j) = \min(A_j) = n_j$ is well-defined.
   - Now let be $f : \mathbb{N} \to \mathbb{N}$ the strictly increasing function defined by $f(j) := n_j$; and let's consider the subsequence of $(a_n)_{n=0}^{\infty}$ defined by $a_{f(j)} = a_{n_j}$. Let be $\varepsilon > 0$. By the archimedean property (or Exercise 5.4.4), there exists a natural number $j \geqslant 1$ such that $\varepsilon > 1/j > 0$. We thus have a natural number n_j such that $|a_{n_j} - L| \leqslant 1/j < \varepsilon$. Furthermore, if $k > j$, we have $n_k > n_j$, and thus we have: $|a_{n_k} - L| \leqslant 1/k \leqslant 1/j < \varepsilon$. Unfolding all these findings:

     $$\forall \varepsilon > 0, \exists j \geqslant 1 \ : \ n \geqslant j \to |a_{f(j)} - L| \leqslant \varepsilon \tag{6.18}$$

     which means that the subsequence $(a_{f(n)})_{n=0}^{\infty}$ converges to 0, as required.

7. Series

EXERCISE 7.1.1. — *Prove Lemma 7.1.4. (Hint: you will need to use induction, but the base case might not necessarily be at 0.)*

Recall that, by Definition 7.1.1, a series $\sum_{i=m}^{n} a_i$ is recursively defined by $\sum_{i=m}^{n} a_i = 0$ for $n < m$, and $\sum_{i=m}^{n+1} a_i = \sum_{i=m}^{n} a_i + a_{n+1}$ for $m \geqslant n$.

The claims to prove are the following.

(a) For all integers $m \leqslant n < p$, we have $\sum_{i=m}^{n} a_i + \sum_{i=n+1}^{p} a_i = \sum_{i=m}^{p} a_i$.

Let's use induction on p, while keeping m, n fixed. Since we have $n < p$, the base case is $p = n + 1$. For $p = n + 1$, we have:

$$\sum_{i=m}^{n} a_i + \sum_{i=n+1}^{p} a_i = \sum_{i=m}^{n} a_i + \sum_{i=n+1}^{n+1} a_i = \sum_{i=m}^{n} a_i + a_{n+1} = \sum_{i=m}^{n+1} a_i = \sum_{i=m}^{p} a_i$$

as expected. The base case is done. Now let's suppose inductively that $\sum_{i=m}^{n} a_i + \sum_{i=n+1}^{p} a_i = \sum_{i=m}^{p} a_i$, and let's prove that this equality still holds for $p + 1$. We have:

$$\begin{aligned}
\sum_{i=m}^{n} a_i + \sum_{i=m}^{p+1} &= \sum_{i=m}^{n} a_i + \sum_{i=m}^{p} a_i + a_{p+1} \text{ (Definition 7.1.1)} \\
&= \sum_{i=m}^{p} a_i + a_{p+1} \text{ (induction hypothesis)} \\
&= \sum_{i=m}^{p+1} a_i \text{ (Definition 7.1.1)}
\end{aligned}$$

which closes the induction and proves the equality for all integers $m \leqslant n < p$.

(b) For k an integer, and $m \leqslant n$ two integers, we have $\sum_{i=m}^{n} a_i = \sum_{j=m+k}^{n+k} a_{j-k}$.

Let's use induction on n, the "offset" k being fixed. For the base case $n = m$, we have $\sum_{j=m+k}^{n+k} a_{j-k} = \sum_{j=m+k}^{m+k} a_{j-k} = a_{m+k-k} = a_m$. On the other hand, we also have $\sum_{i=m}^{n} a_i = \sum_{i=m}^{m} a_i = a_m$, which proves the base case.

Now let's suppose inductively that $\sum_{i=m}^{n} a_i = \sum_{j=m+k}^{n+k} a_{j-k}$. We thus have:

$$\begin{aligned}
\sum_{j=m+k}^{n+k+1} a_{j-k} &= \sum_{j=m+k}^{n+k} a_{j-k} + a_{n+1} \\
&= \sum_{i=m}^{n} a_i + a_{n+1} \text{ (induction hypothesis)} \\
&= \sum_{i=m}^{n+1} a_i
\end{aligned}$$

which closes the induction, the property being true for all $n \geqslant m$, with any arbitrary offset k.

(c) For $m \leqslant n$ two integers, we have $\sum_{i=m}^{n}(a_i + b_i) = (\sum_{i=m}^{n} a_i) + (\sum_{i=m}^{n} b_i)$.

Let's use induction on n. The base case is $m = n$ and is obvious, since we have on the one hand $\sum_{i=m}^{m}(a_i + b_i) = a_m + b_m$, and on the other hand $(\sum_{i=m}^{m} a_i) + (\sum_{i=m}^{m} b_i) = a_m + b_m$.

Now let's suppose inductively that $\sum_{i=m}^{n}(a_i + b_i) = (\sum_{i=m}^{n} a_i) + (\sum_{i=m}^{n} b_i)$. Then we have:

$$\begin{aligned}
\sum_{i=m}^{n+1}(a_i + b_i) &= \sum_{i=m}^{n}(a_i + b_i) + (a_{n+1} + b_{n+1}) \\
&= \left(\sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i\right) + a_{n+1} + b_{n+1} \text{ (induction hypothesis)} \\
&= \sum_{i=m}^{n} a_i + a_{n+1} + \sum_{i=m}^{n} b_i + b_{n+1} \\
&= \sum_{i=m}^{n+1} a_i + \sum_{i=m}^{n+1} b_i
\end{aligned}$$

which closes the induction. The property is thus true for all $n \geqslant m$.

(d) For $m \leqslant n$ integers, and c a real number, we have $\sum_{i=m}^{n} ca_i = c\,(\sum_{i=m}^{n} a_i)$.

Let's induct on n, first considering the base case $n = m$. We have $\sum_{i=m}^{m} ca_i = ca_m = c\,(\sum_{i=m}^{m} a_i)$ as required.

Now let's suppose inductively that $\sum_{i=m}^{n} ca_i = c\,(\sum_{i=m}^{n} a_i)$. Then we have:

$$\begin{aligned}
\sum_{i=m}^{n+1} ca_i &= \sum_{i=m}^{n} ca_i + ca_{n+1} \\
&= c\left(\sum_{i=m}^{n} a_i\right) + ca_{n+1} \text{ (induction hypothesis)} \\
&= c\left(\sum_{i=m}^{n} a_i + a_{n+1}\right) \\
&= c\sum_{i=m}^{n+1} a_i
\end{aligned}$$

which closes the induction. The property is thus true for all $n \geqslant m$.

(e) For $m \leqslant n$ integers, prove that $|\sum_{i=m}^{n} a_i| \leqslant \sum_{i=m}^{n} |a_i|$.

Let's use induction on n, first considering the base case $n = m$. We have $|\sum_{i=m}^{m} a_i| = |a_m| = \sum_{im}^{m} |a_i|$, as required.

Now let's suppose inductively that we have $|\sum_{i=m}^n a_i| \leqslant \sum_{i=m}^n |a_i|$. Then we have:

$$\begin{aligned}
\left|\sum_{i=m}^{n+1} a_i\right| &:= \left|\sum_{i=m}^{n} a_i + a_{n+1}\right| \\
&\leqslant \left|\sum_{i=m}^{n} a_i\right| + |a_{n+1}| \text{ (Proposition 4.3.3(b))} \\
&\leqslant \sum_{i=m}^{n} |a_i| + |a_{n+1}| \text{ (induction hypothesis)} \\
&\leqslant \sum_{i=m}^{n+1} |a_i|
\end{aligned}$$

which closes the induction. The property is thus true for all $n \geqslant m$.

(f) For $m \leqslant n$ integers, prove that, if $a_i \leqslant b_i$ for all $m \leqslant i \leqslant n$, we have $\sum_{i=m}^n \leqslant \sum_{i=m}^n b_i$.

Once again, let's use induction on n, starting with the base case $n = m$. In this case, we have $\sum_{i=m}^n a_i := a_m \leqslant b_m := \sum_{i=m}^n b_i$, as required.

Now let's suppose inductively that we have $\sum_{i=m}^n \leqslant \sum_{i=m}^n b_i$. Then, if $a_{n+1} \leqslant b_{n+1}$, we have:

$$\begin{aligned}
\sum_{i=m}^{n+1} a_i &:= \sum_{i=m}^{n} a_i + a_{n+1} \\
&\leqslant \sum_{i=m}^{n} a_i + b_{n+1} \\
&\leqslant \sum_{i=m}^{n} b_i + b_{n+1} \text{ (induction hypothesis)} \\
&\leqslant \sum_{i=m}^{n+1} b_i
\end{aligned}$$

which closes the induction. The property is thus true for all $n \geqslant m$.

EXERCISE 7.1.2. — *Prove Proposition 7.1.11.*

First recall the main definition: if X is a finite set with n elements, $f : X \to \mathbb{R}$ a function, and g a bijection from $[\![1, n]\!]$ to X, then we have $\sum_{x \in X} f(x) := \sum_{i=1}^n f(g(i))$.

The statements to prove are:

(a) If $X = \varnothing$ and f is the empty function, we have $\sum_{x \in X} f(x) = 0$.

In this case, X has $n = 0$ elements, and the empty function g is a bijection between X and $[\![1, 0]\!]$, so that by definition, we have $\sum_{x \in X} f(x) = \sum_{i=1}^0 f(g(i)) = 0$.

(b) If $X = \{x_0\}$, we have $\sum_{x \in X} f(x) = f(x_0)$.

In this case, X has $n = 1$ element, and $g : \{1\} \to X$ such that $g(1) = x_0$ is a bijection. Thus, by definition, we have $\sum_{x \in X} f(x) = \sum_{i=1}^1 f(g(i)) = f(g(1)) = f(x_0)$.

(c) If X is a finite set and $g : Y \to X$ a bijection, then $\sum_{x \in X} f(x) = \sum_{y \in Y} f(g(y))$.

First note that since g is a bijection, Y also has n elements.

On the one hand, since X has n elements, there exists a bijection $h : [\![1, n]\!] \to X$ such that $\sum_{x \in X} f(x) = \sum_{i=1}^{n} f(h(i))$.

On the other hand, let be $k : [\![1, n]\!] \to Y$ a function defined by $k = g^{-1} \circ h$. Since k is the composition of two bijections, it is itself a bijection (cf. Exercise 3.3.2). We thus have : $\sum_{y \in Y} f(g(y)) = \sum_{i=1}^{n} f \circ g \circ g^{-1} \circ h(i) = \sum_{i=1}^{n} f(h(i))$.

Thus, we have indeed $\sum_{x \in X} f(x) = \sum_{y \in Y} f(g(y))$.

(d) For $n \leqslant m$ two integers and $X = \{i \in \mathbb{Z} : n \leqslant i \leqslant m\}$, we have $\sum_{i=n}^{m} a_i = \sum_{i \in X} a_i$.

Let be $f : [\![n, m]\!] \to X$ defined by $f(i) = a_i$ a function; so that we have $\sum_{i \in X} a_i = \sum_{i \in X} f(i)$. First of all, note that X has $m - n + 1$ elements. Furthermore, let be $g : [\![1, m-n+1]\!] \to X$ the bijection defined by $g(i) = i+n-1$ (we could show rigorously that this function is bijective, but this is straightforward). We thus have, by definition: $\sum_{i \in X} a_i = \sum_{i \in X} f(i) = \sum_{i=1}^{m-n+1} f(g(i)) = \sum_{i=1}^{m-n+1} f(i + n - 1) = \sum_{i=1}^{m-n+1} a_{i+n-1}$.

Now we can use Lemma 7.1.4(b), to get $\sum_{i=1}^{m-n+1} a_{i+n-1} = \sum_{i=n}^{m} a_i$, so that we have finally showed that $\sum_{i \in X} a_i = \sum_{i=n}^{m} a_i$.

(e) If X, Y are sets such that $X \cap Y = \varnothing$, and $f : X \cup Y \to \mathbb{R}$, we have $\sum_{z \in X \cup Y} f(z) = \sum_{x \in X} f(x) + \sum_{y \in Y} f(y)$.

Suppose that X has n elements and Y has m elements. Thus, there exist two bijections $g : [\![1, n]\!] \to X$ and $h : [\![1, m]\!] \to Y$. Thus, by definition, we have $\sum_{x \in X} f(x) = \sum_{i=1}^{n} f(g(i))$ and $\sum_{y \in Y} f(y) = \sum_{i=1}^{m} f(h(i))$.

By Lemma 7.1.4(b), we have $\sum_{i=1}^{m} f(h(i)) = \sum_{i=n+1}^{n+m} f(h(i - n))$.

Now let's construct a new function $k : [\![1, n+m]\!] \to X \cup Y$, defined by: $k(i) = g(i) \in X$ if $1 \leqslant i \leqslant n$ and $k(i) = h(i) \in Y$ if $n + 1 \leqslant i \leqslant n + m$. In particular, $k(i) \in X \cup Y$ for all $i \in [\![1, n + m]\!]$. Let's prove that k is a bijection.

- Let's suppose that there exists $z \in X \cup Y$ such that $k(i) \neq z$ for all $i \in [\![1, n + m]\!]$. If $z \in X$, this means in particular that there exists one $z \in X$ such that $g(i) \neq z$ for all $i \in [\![1, n]\!]$, a contradiction. A similar contradiction follows if $z \in Y$. Thus, k is surjective.
- Now suppose that there exist two $i, j \in [\![1, n + m]\!]$ such that $k(i) = k(j)$. If both $i, j \in [\![1, n]\!]$, it is a contradiction with the fact that g is injective; a similar contradiction follows for h if both $i, j \in [\![n + 1, m]\!]$. Finally, if $i \in [\![1, n]\!]$ and $j \in [\![n+1, m]\!]$ (or the converse), it is a contradiction with the fact that $X \cap Y = \varnothing$. Thus, k is injective.

Thus, we have:

$$\begin{aligned}
\sum_{x\in X\cup Y} f(x) &= \sum_{i=1}^{n+m} f(k(i)) \\
&= \sum_{i=1}^{n} f(k(i)) + \sum_{i=n+1}^{n+m} f(k(i)) \text{ (by Lemma 7.1.4(a))} \\
&= \sum_{i=1}^{n} f(g(i)) + \sum_{i=n+1}^{n+m} f(h(i)) \\
&= \sum_{x\in X} f(x) + \sum_{y\in Y} f(y)
\end{aligned}$$

as required, which closes the proof.

(f) If X is a finite set, then we have $\sum_{x\in X}(f(x)+g(x)) = \sum_{x\in X} f(x) + \sum_{x\in X} g(x)$.

Suppose that X has n elements, so that there exists a bijection $h : [\![1,n]\!] \to X$. Then we have:

$$\begin{aligned}
\sum_{x\in X} f(x)+g(x) &:= \sum_{i=1}^{n} f(h(i)) + g(h(i)) \\
&= \sum_{i=1}^{n} f(h(i)) + \sum_{i=1}^{n} g(h(i)) \text{ (Lemma 7.1.4(e))} \\
&:= \sum_{x\in X} f(x) + \sum_{x\in X} g(x)
\end{aligned}$$

(g) If c is a real number, the we have $\sum_{x\in X} cf(x) = c\sum_{x\in X} f(x)$.

If we define a function g as $g(x) := cf(x)$, and h a bijection $[\![1,n]\!] \to X$, we have:

$$\begin{aligned}
\sum_{x\in X} cf(x) = \sum_{x\in X} g(x) = \sum_{i=1}^{n} g(h(i)) &= \sum_{i=1}^{n} cf(h(i)) \\
&= c\times \sum_{i=1}^{n} f(h(i)) \text{ (Lemma 7.1.4(d))} \\
&= c\times \sum_{x\in X} f(x)
\end{aligned}$$

(h) If f, g are functions such that $f(x) \leqslant g(x)$ for all $i \in \mathbb{R}$, then we have $\sum_{x\in X} f(x) \leqslant \sum_{x\in X} g(x)$.

If X has n elements, let $h : [\![1,n]\!] \to X$ be a bijection. Then we have:

$$\begin{aligned}
\sum_{x\in X} f(x) &= \sum_{i=1}^{n} f(h(i)) \\
&\leqslant \sum_{i=1}^{n} g(h(i)) \text{ (Lemma 7.1.4(f))} \\
&= \sum_{x\in X} g(x)
\end{aligned}$$

(i) Triangle inequality: we have $|\sum_{x\in X} f(x)| \leqslant \sum_{x\in X} |f(x)|$.

If we consider the function $k : X \to \mathbb{R}$ such that $k(x) = |f(x)|$ for all $x \in X$; and $h : [\![1, n]\!] \to X$ a bijection, we have:

$$\begin{aligned}
\left|\sum_{x\in X} f(x)\right| &= \left|\sum_{i=1}^{n} f(h(i))\right| \\
&\leqslant \sum_{i=1}^{n} |f(h(i))| \text{ (Lemma 7.1.4(e))} \\
&= \sum_{i=1}^{n} k(h(i)) = \sum_{x\in X} k(x) \\
&= \sum_{x\in X} |f(x)|
\end{aligned}$$

Exercise 7.1.4. — *Define the factorial function $n!$ for natural numbers n by the recursive definition $0! := 1$ and $(n+1)! := n! \times (n+1)$. If x and y are real numbers, prove the binomial formula $(x+y)^n = \sum_{j=0}^{n} \frac{n!}{j!(n-j)!} x^j y^{n-j}$ for all natural numbers n. (Hint: induct on n.)*

First we begin with a small preliminary result, which is that for all n, j natural numbers, we have $\frac{(n+1)!}{j!(n+1-j)!} = \frac{n!}{(j-1)!(n-j+1)!} + \frac{n!}{j!(n-j)!}$. Indeed:

$$\begin{aligned}
\frac{n!}{(j-1)!(n-j+1)!} + \frac{n!}{j!(n-j)!} &= \frac{n! \times j}{j!(n-j+1)!} + \frac{n! \times (n-j+1)}{j!(n-j+1)!} \\
&= \frac{n! \times (n+1)}{j!(n-j+1)!} \\
&= \frac{(n+1)!}{j!(n+1-j)!}
\end{aligned}$$

Now we can go back to the main proof, for which we will induct on n.

- Let's start with the base case $n = 0$. On the one hand, by Definition 5.6.1, we have $(x+y)^0 = 1$. On the other hand, we have $\sum_{j=0}^{0} \frac{n!}{j!(n-j)!} x^j y^{n-j} = \frac{0!}{0! \times 0!} x^0 y^0 = 1$, so that the base case is done.
- Now suppose inductively that we have $(x+y)^n = \sum_{j=0}^{n} \frac{n!}{j!(n-j)!} x^j y^{n-j}$, and let's prove that this property is still true for $n+1$. We have:

$$
\begin{aligned}
(x+y)^{n+1} &= (x+y)\times(x+y)^n \\
&= x\left(\sum_{j=0}^{n}\frac{n!}{j!(n-j)!}x^j y^{n-j}\right) + y\left(\sum_{j=0}^{n}\frac{n!}{j!(n-j)!}x^j y^{n-j}\right) \quad \text{(by induction hypothesis)} \\
&= \left(\sum_{j=0}^{n}\frac{n!}{j!(n-j)!}x^{j+1} y^{n-j}\right) + \left(\sum_{j=0}^{n}\frac{n!}{j!(n-j)!}x^j y^{n-j+1}\right) \\
&= \left(x^{n+1} + \sum_{j=0}^{n-1}\frac{n!}{j!(n-j)!}x^{j+1} y^{n-j}\right) + \left(y^{n+1} + \sum_{j=1}^{n}\frac{n!}{j!(n-j)!}x^j y^{n-j+1}\right) \\
&= \left(x^{n+1} + \sum_{j=1}^{n}\frac{n!}{(j-1)!(n-j+1)!}x^{j} y^{n-j+1}\right) + \left(y^{n+1} + \sum_{j=1}^{n}\frac{n!}{j!(n-j)!}x^j y^{n-j+1}\right) \\
&= x^{n+1} + \sum_{j=1}^{n}\left(\frac{n!}{(j-1)!(n-j+1)!} + \frac{n!}{j!(n-j)!}\right)x^j y^{n-j+1} + y^{n+1} \\
&= x^{n+1} + \sum_{j=1}^{n}\frac{(n+1)!}{j!(n+1-j)!}x^j y^{n-j+1} + y^{n+1} \\
&= \sum_{j=0}^{n+1}\frac{(n+1)!}{j!(n+1-j)!}x^j y^{n-j+1}
\end{aligned}
$$

so that the property is true for all natural number n.

Exercise 7.1.5. — *Let X be a finite set, let m be an integer, and for each $x \in X$ let $(a_n(x))_{n=m}^{\infty}$ be a convergent sequence of real numbers. Show that the sequence $(\sum_{x\in X} a_n(x))_{n=m}^{\infty}$ is convergent, and that $\lim_{n\to\infty}\sum_{x\in X} a_n(x) = \sum_{x\in X}\lim_{n\to\infty} a_n(x)$.*

Let's suppose that X has r elements. Note that, in particular, the property is equivalent to $\lim_{n\to\infty}\sum_{i=1}^{r} a_n(h(i)) = \sum_{i=1}^{r}\lim_{n\to\infty} a_n(h(i))$ for any bijection $h : [\![1, r]\!] \to X$.

Let's induct on r, the cardinality of X.

- We start with the base case $r = 0$, i.e. $X = \varnothing$. In this case, by Proposition 7.1.11(a), we have immediately $\sum_{x\in X} a_n(x) = 0$, so that $\lim_{n\to\infty}\sum_{x\in X} a_n(x) = 0$ as the limit of a constant sequence. Similarly, we have $\sum_{x\in X}\lim_{n\to\infty} a_n(x) = 0$, as a sum over an empty set X. Thus, the base case is done.

- Now suppose inductively that $\lim_{n\to\infty}\sum_{x\in X} a_n(x) = \sum_{x\in X}\lim_{n\to\infty} a_n(x)$ when X has r elements, and let's show that this property is still true when X has $r+1$ elements.

  Let be $h : [\![1, r+1]\!] \to X$ a bijection. By Definition 7.1.6, we have:

$$
\begin{aligned}
\sum_{x\in X} a_n(x) &= \sum_{i=1}^{r+1} a_n(h(i)) \\
&= \sum_{i=1}^{r} a_n(h(i)) + a_n(h(r+1))
\end{aligned}
$$

By the induction hypothesis, $\sum_{i=1}^{r} a_n(h(i))$ is a convergent sequence; and $a_n(h(r+1))$ is a convergent sequence by the initial hypothesis; so that we can apply Theorem 6.1.19(a):

$$\begin{aligned}
\lim_{n\to\infty} \sum_{x\in X} a_n(x) &= \lim_{n\to\infty} \left(\sum_{i=1}^{r} a_n(h(i)) + a_n(h(r+1)) \right) \\
&= \lim_{n\to\infty} \sum_{i=1}^{r} a_n(h(i)) + \lim_{n\to\infty} a_n(h(r+1)) \\
&= \sum_{i=1}^{r} \lim_{n\to\infty} a_n(h(i)) + \lim_{n\to\infty} a_n(h(r+1)) \text{ (by induction hypothesis)} \\
&= \sum_{i=1}^{r+1} \lim_{n\to\infty} a_n(h(i)) \text{ (Definition 7.1.1)} \\
&= \sum_{x\in X} \lim_{n\to\infty} a_n(x) \text{ (by Proposition 7.1.11(c))}
\end{aligned}$$

so that the property is also true when X has $n+1$ elements. This closes the induction.

Exercise 7.2.2. — *Prove Proposition 7.2.5.*

We have to prove that the formal series $\sum_{n=m}^{\infty} a_n$ converges iff, for any $\varepsilon > 0$, there exists an integer $N \geqslant m$ such that $\left|\sum_{n=p}^{q} a_n\right| \leqslant \varepsilon$ for all $p, q \geqslant N$.

First, note that the second statement is just another way to say that the partial sum $S_n = \sum_{i=m}^{n} a_n$ is a Cauchy sequence. Indeed, for $q \geqslant p \geqslant N$, we have $|S_q - S_p| = |\sum_{n=m}^{q} a_n - \sum_{n=m}^{p} a_n| = \left|\sum_{n=p}^{q} a_n\right|$ (by Lemma 7.1.4(a)). Thus, the equivalence to prove is simply the fact that $\sum_{n=m}^{\infty} a_n$ converges iff the sequence of its partial sums $(S_N)_{N=m}^{\infty}$ is a Cauchy sequence. Proposition 6.1.12 and Theorem 6.4.18 immediately provide this equivalence.

Exercise 7.2.3. — *Prove Corollary 7.2.6.*

Let's suppose that $\sum_{n=m}^{\infty} a_n$ is a convergent sequence. Thus, according to Proposition 7.2.5 (see also previous exercise), for all $\varepsilon > 0$, there exists $N \geqslant 0$ such as $p, q \geqslant N \implies \left|\sum_{n=p}^{q} a_n\right| \leqslant \varepsilon$.

In particular, let's take $p = q$, and we have $|a_p| \leqslant \varepsilon$ for all $p \geqslant N$, which says precisely that $(a_n)_{n=m}^{\infty}$ converges to 0.

Exercise 7.2.5. — *Prove Proposition 7.2.14.*

We have to prove several statements here:

(a) Let's consider the partial sums $S_N := \sum_{n=m}^{N} a_n$ and $T_N := \sum_{n=m}^{N} b_n$. Saying that the formal series $\sum_{n=m}^{\infty} a_n$ and $\sum_{n=m}^{\infty} b_n$ are convergent means that the sequences of their partials sums, $(S_N)_{N=m}^{\infty}$ and $(T_N)_{N=m}^{\infty}$ are convergent, towards x and y respectively.

By Theorem 6.1.19(a), the sequence $(S_N+T_N)_{N=m}^{\infty}$ converges to $x+y$. But by definition, we have $S_N + T_N = \sum_{n=m}^{N} a_n + \sum_{n=m}^{N} b_n = \sum_{n=m}^{N} a_n + b_n$ by Lemma 7.1.4(c). It means that the partial sum of the formal series $\sum_{n=m}^{\infty} a_n + b_n$ converges to $x + y$, so that we have proved the result.

(b) The proof would be very similar to the previous one; just use Lemma 7.1.4(d) and Theorem 6.1.19(c) instead.

(c) For all integers $N \geqslant m$ and $k \geqslant 0$, we know by Lemma 7.1.4(a) that we have $\sum_{n=m}^{N} a_n = \sum_{n=m}^{m+k-1} a_n + \sum_{n=m+k}^{N} a_n$. If we set $S_N := \sum_{n=m}^{N} a_n$; $x := \sum_{n=m}^{m+k-1} a_n$; and $T_N := \sum_{n=m+k}^{N} a_n$, we have for all $N \geqslant m$, the equality $S_N = x + T_N$. As a very general fact (not shown in the book, but easy to prove), the sequence $(S_N)_{N=m}^{\infty}$ converges to L iff the sequence $(T_N)_{N=m}^{\infty}$ converges to $L - x$. The statement follows.

(d) Let be $\varepsilon > 0$ a positive real number. Since the formal series $\sum_{n=m}^{\infty} a_n$ converges to x, there exists a positive integer M such that $|\sum_{n=m}^{N} a_n - x| \leqslant \varepsilon$ for all $N \geqslant M$. By Lemma 7.1.4(b), this is equivalent to $|\sum_{n=m+k}^{N+k} a_{n-k} - x| \leqslant \varepsilon$, for k a positive integer and for all $N \geqslant M$.

Thus, there exists a positive integer $M' := M + k$ such that, for all $N \geqslant M'$, we have $|\sum_{n=m+k}^{N} a_{n-k} - x| \leqslant \varepsilon$, which means that $\sum_{n=m+k}^{\infty} a_{n-k}$ also converges to x.

Exercise 7.2.6. — *Prove Lemma 7.2.15.*

Let be $(a_n)_{n=0}^{\infty}$ a sequence that converges to 0, and let's consider the formal series $\sum_{n=0}^{\infty}(a_n - a_{n+1})$. Its partials sums are, for any integer $N \geqslant 0$, $S_N := \sum_{n=0}^{N}(a_n - a_{n+1}) = a_0 - a_{N+1}$. This can be shown by an induction on $N \geqslant 0$:

- Let's start with the base case $N = 0$. In this case, we have $S_0 = \sum_{n=0}^{0}(a_n - a_{n+1}) = a_0 - a1$, as expected.

- Now suppose inductively that the property for a given positive integer N, and let's show that it is still true for $N + 1$. We have:

$$\begin{aligned} S_{N+1} &:= \sum_{n=0}^{N+1}(a_n - a_{n+1}) \\ &= \sum_{n=0}^{N}(a_n - a_{n+1}) + (a_{N+1} - a_{N+2}) \text{ (Definition 7.1.1)} \\ &= a_0 - a_{N+1} + a_{N+1} - a_{N+2} \text{ (induction hypothesis)} \\ &= a_0 - a_{N+2} \end{aligned}$$

which closes the induction, so that the property is true for all $N \geqslant 0$.

Now, let be $\varepsilon > 0$ a positive real number. Recall that $(a_n)_{n=0}^{\infty}$ converges to 0, so that there exists a positive integer $M \geqslant 0$ such that $|a_n| \leqslant \varepsilon$ for all $n \geqslant M$. On the other hand, we have, for all $N \geqslant 0$, $|S_N - a_0| = |a_0 - a_{N+1} - a_0| = |a_{N+1}|$. Thus, there exists a positive integer $M' := M - 1$ such that, for all $N \geqslant M'$, we have $|S_N - a_0| \leqslant \varepsilon$. This means that $(S_N)_{N=0}^{\infty}$ converges to a_0, i.e. that the formal series $\sum_{n=0}^{\infty}(a_n - a_{n+1})$ converges to a_0, as required.

Exercise 7.3.1. — *Use Proposition 7.3.1 to prove Corollary 7.3.2.*

First, an important remark: if we have $|a_n| \leqslant b_n$ for all $n \geqslant m$, we have in particular $b_n \geqslant 0$ for all $n \geqslant m$, so that the formal series $\sum_{n=m}^{\infty} b_n$ is a series of non-negative numbers. Since it is convergent by hypothesis, it is thus possible to apply Proposition 7.3.1 to this series: there exists a real number M such that $T_N := \sum_{n=m}^{N} b_n \leqslant M$ for all $N \geqslant m$.

- First, let's show the first part of Corollary 7.3.2, that is to say that $\sum_{n=m}^{\infty} a_n$ is absolutely convergent as soon as $\sum_{n=m}^{\infty} b_n$ is convergent.

  Since we have $|a_n| \leqslant b_n$ for all $n \geqslant m$ by hypothesis, we have also for all $N \geqslant m$ the inequality $\sum_{n=m}^{N} |a_n| \leqslant \sum_{n=m}^{N} b_n \leqslant M$ by Lemma 7.1.4(f), so that the formal series $\sum_{n=m}^{\infty} |a_n|$ is itself convergent by Proposition 7.3.1 again. This proves the first statement.

- Now let's prove that $|\sum_{n=m}^{\infty} a_n| \leqslant \sum_{n=m}^{\infty} |a_n|$. This is actually simply Proposition 7.2.9, so there is nothing to add here.

- Finally, let's prove that $\sum_{n=m}^{\infty} |a_n| \leqslant \sum_{n=m}^{\infty} b_n$. By Proposition 7.2.14(a)-(b), we know that, since $\sum_{n=m}^{\infty} |a_n|$ and $\sum_{n=m}^{\infty} b_n$ are convergent, then $\sum_{n=m}^{\infty} b_n - |a_n|$ is convergent, so that its partial sum $V_N := \sum_{n=m}^{N} b_n - |a_n|$ converges to some limit L. Also, we already know that $b_n - |a_n| \geqslant 0$ for all $n \geqslant m$, so that $V_N \geqslant 0$ for all $N \geqslant m$, and thus $L \geqslant 0$ by Proposition 5.4.9.

  To summarize: $\sum_{n=m}^{\infty}(b_n - |a_n|) = \sum_{n=m}^{\infty} b_n - \sum_{n=m}^{\infty} |a_n| = L \geqslant 0$, and thus in particular, $\sum_{n=m}^{\infty} b_n \geqslant \sum_{n=m}^{\infty} |a_n|$, as required.

  This closes the proof.

Exercise 7.3.2. — *Prove Lemma 7.3.3.*

Let be x a real number and let's consider the formal series $\sum_{n=0}^{\infty} x^n$.

- If $|x| \geqslant 1$, then $(x^n)_{n=0}^{\infty}$ does not converge to 0, by Lemma 6.5.2. Thus, by the zero test (Corollary 7.2.6), the formal series $\sum_{n=0}^{\infty} x^n$ is not convergent.

- If $|x| < 1$, then let's prove that for all $N \geqslant 0$, the partial sum S_N is equal to $\sum_{n=0}^{N} x^n$. Let's induct on N, starting with the base case $N = 0$. In this case, we have $S_0 = \sum_{n=0}^{0} x^n = 1$; and on the other side, we have $\frac{1-x^{N+1}}{1-x} = \frac{1-x}{1-x} = 1$, so the base case is done.

  Now let's suppose inductively that $S_N = \frac{1-x^{N+1}}{1-x}$ and let's prove that this property still holds for S_{N+1}. On the one hand we have:

$$\begin{aligned} S_{N+1} &:= \sum_{n=0}^{N} x^n + x^{N+1} \\ &= \frac{1-x^{N+1}}{1-x} + x^{N+1} \text{ (induction hypothesis)} \\ &= \frac{1-x^{N+1} + (1-x)x^{N+1}}{1-x} \\ &= \frac{1-x^{N+2}}{1-x} \end{aligned}$$

  as required, so that the property holds for all $N \geqslant 0$: we always have $S_N = \frac{1-x^{N+1}}{1-x}$.

  But, by Lemma 6.5.2 (combined with Exercise 6.1.4), we know that $(x^{n+1})_{n=0}^{\infty}$ converges to 0; so that by the convergence laws, $(S_N)_{N=0}^{\infty}$ converges to $\frac{1}{1-x}$.

  However, this only shows that $(S_N)_{N=0}^{\infty}$ is conditionally convergent. Actually, it is absolutely convergent. Indeed, according to Proposition 4.3.10(d), we have $|x^n| = |x|^n$, and the series $\sum_{n=0}^{\infty} |x|^n$ is convergent, by the previous result. This closes the proof.

8. Infinite sets

EXERCISE 8.1.1. — *Let X be a set. Show that X is infinite if and only if there exists a proper subset $Y \subsetneq X$ of X which has the same cardinality as X. (This exercise requires the axiom of choice, Axiom 8.1)*

This exercise is about the characterization of Dedekind-infinite sets.

- First suppose that X and Y have the same cardinality, i.e. that there exists a bijection $f : Y \to X$ between X and a proper subset $Y \subsetneq X$. For the sake of contradiction, let's suppose that X is finite. Then, by definition[27], there exists $n \in \mathbb{N}$ such that $\#X = n$. On the one hand, since Y is a proper subset of X, by Proposition 3.6.14 (c), we must have $\#Y < \#X$, i.e. $\#Y < n$. On the other hand, since f is a bijection, X and Y have the same cardinality, i.e. $\#Y = \#X = n$. Thus we have a contradiction; X cannot be finite.

- Conversely, suppose that X is infinite, and let's show that we can construct a bijection $f : X \to Y$ from X to a proper subset $Y \subsetneq X$.

  We begin by defining recursively an infinite collection $A_0, A_1, A_2, \dots$ of non-empty subsets of X. Since X is infinite, in particular it is non-empty; thus by the lemma of single choice we can pick a $x_0 \in X$. We define $A_0 := \{a_0\}$. Now since X is infinite, $X \backslash A_0$ is clearly non-empty, and thus we can also pick a $a_1 \in X \backslash A_0$. We can thus define the non-empty subset $A_1 := A_0 \cup \{a_1\}$. Recursively, let's suppose that A_n has been defined and is a finite non-empty subset of X. Thus, the set $X \backslash \bigcup_{i=0}^{n} A_i$ is non-empty, so that we can pick a a_{n+1} in it, and then form the set $A_{n+1} := A_n \cup \{a_{n+1}\}$. The axiom of (countable) choice ensures that we can repeat this recursive algorithm infinitely many times, to finally get the set $A := \bigcup_{n \in \mathbb{N}} A_n$.

  Now let's define the proper subset $Y := X \backslash \{a_0\}$, and the following function $f : X \to Y$

  $$\begin{cases} f(x) = x & \text{if } x \in X \backslash A \\ f(x) = a_{n+1} & \text{if } x \in A, x = a_n \end{cases}$$

  One can clearly see that f is a surjection, and also an injection. This closes the proof.

EXERCISE 8.1.2. — *Prove Proposition 8.1.4. (Hint: you can either use induction, or use the principle of infinite descent, Exercise 4.4.2, or use the least upper bound (or greatest lower bound) principle, Theorem 5.5.9.) Does the well-ordering principle work if we replace the natural numbers by the integers? What if we replace the natural numbers by the positive rationals? Explain.*

Let be $X \subset \mathbb{N}$ such that $\#X = p$; let's show that there exists $n \in X$ such that $n \leqslant m$ for all $m \in X$.

We could indeed show this property by using the greatest lower bound principle: we have showed in Exercise 6.6.3 that, for every subset $X \subset \mathbb{N}$, the greatest lower bound $\inf(X)$ belongs to X. But we will give here another proof, using induction on p.

[27] This is Definition 3.6.10, i.e. the definition of Cantor-finite sets.

- Let's start with the base case $n = 1$ (note that the base case is not $n = 0$, since X is supposed to be non-empty). In such a case, X is a singleton, i.e. there exists $n \in \mathbb{N}$ such that $X = \{n\}$. By reflexivity, we have of course $n \leqslant n$, i.e. we have $n \leqslant m$ for all $m \in X$, as required.

- Now suppose inductively that the property is true for a cardinality of p, and let's show that it is still true if $\#X = p + 1$. By Lemma 3.6.9, there exists $x \in \mathbb{N}$ and a subset $Y \subset \mathbb{N}$ with $\#Y = p$, such that $X = Y \cup \{x\}$. By the induction hypothesis, there exists $n \in Y$ such that $n \leqslant m$ for all $m \in Y$.

  Now, let be $n' := \min(n, x)$. We thus have $n' \leqslant n \leqslant m$ for all $m \in Y$ (and thus, in particular, $n' \leqslant m$); and $n' \leqslant x$ by definition. Thus, we have $n' \leqslant m$ for all $m \in Y \cup \{x\}$, that is to say $n' \leqslant m$ for all $m \in X$, as required. This closes the induction.

Note that the well ordering principle does not apply for the integers, because a non-empty set of integers can have arbitrarily small negative numbers: consider for instance the set $A = \{\dots, -3, -2, -1, 0\}$. This is not applicable to the positive rationals either: one can think of the set $B = \{1/n : n \in \mathbb{N}^\star\}$, which has an infimum of 0, but has no smallest element in B.

Exercise 8.1.3. — *Fill in the gaps marked (?) in Proposition 8.1.5.*

Those gaps only state a few results that we will better show below.

1. Show that the set $\{x \in X : x \neq a_m \text{ for all } m < n\}$ is infinite.

   By definition, we have $X = \{a_0, \cdots, a_{n-1}\} \cup \{x \in X : x \neq a_m \text{ for all } m < n\}$. Obviously, the set $\{a_0, \cdots, a_{n-1}\}$ is finite. Let's suppose for the sake of contradiction that $\{x \in X : x \neq a_m \text{ for all } m < n\}$ is also finite. Then, by Proposition 3.6.14(b), X would also be finite as the union of two finite sets, which is a contradiction.

2. Show that $(a_n)_{n=0}^{\infty}$ is strictly increasing.

   By definition, we have:

   $$\begin{aligned} a_n &:= \min\{x \in X : x \neq a_m \text{ for all } m < n\} \\ a_{n+1} &:= \min\{x \in X : x \neq a_m \text{ for all } m < n+1\} \end{aligned}$$

   Let's write $A_n := \{x \in X : x \neq a_m \text{ for all } m < n\}$; we thus have $A_{n+1} \subset A_n$ for all natural number n. Consequently, $\min A_{n+1} \geqslant \min A_n$ for all n; i.e. $a_{n+1} \geqslant a_n$ for all n. Furthermore, by definition, $a_{n+1} \neq a_n$, so that we finally have $a_{n+1} > a_n$ for all n, i.e. $(a_n)_{n=0}^{\infty}$ is a strictly increasing sequence.

3. Show that $a_n \neq a_m$ for all $n \neq m$.

   Suppose, for the sake of contradiction, that there exists two distinct natural numbers n, m such that $a_n = a_m$. This would be a contradiction with the fact that $(a_n)_{n=0}^{\infty}$ is a strictly increasing sequence.

4. Show that $a_n \in X$ for all $n \in \mathbb{N}$.

   By Proposition 8.1.4 (well-ordering principle), the minimum of a subset of natural numbers is well-defined: $\min X = \inf X \in X$ if $X \subset \mathbb{N}$.

5. Obvious statement, simple rephrasing of the definition.

6. Show that $a_n \geqslant n$ for all n.

   This can be shown by a simple induction: $(a_n)_{n=0}^{\infty}$ is a strictly increasing sequence of natural numbers, so that $a_0 \geqslant 0$ (base case), and if we suppose inductively that $a_n \geqslant n$, we thus have $a_{n+1} > a_n \geqslant n$, i.e. $a_{n+1} \geqslant n+1$. This closes the induction.

Exercise 8.1.4. — *Prove Proposition 8.1.8.*

Let Y be a set, and $f : \mathbb{N} \to Y$ a function (non necessarily bijective). We have to show that $f(\mathbb{N})$ is at most countable.

Let's define the set $A := \{n \in \mathbb{N} : f(m) \neq f(n) \text{ for all } 0 \leqslant m < n\}$, intuitively the set of natural numbers n such that $f(n)$ does not appear in the sequence $f(0), \cdots, f(n-1)$. We write $f_{|A}$ the restriction of f to A, and let's show that $f_{|A}$ is a bijection from A to $f(\mathbb{N})$.

- First, $f_{|A}$ is injective: let's suppose that we have two natural numbers $a, b \in A$ such that $f(a) = f(b)$. By definition of A, we must have $f(m) \neq f(a)$ for all $m < a$. Since $f(a) = f(b)$, we necessarily have $b \geqslant a$. Similarly, we must have $f(m) \neq f(b)$ for all $m < b$, which implies $a \geqslant b$. The two inequalities $a \geqslant b$ and $b \geqslant a$ imply $a = b$, which shows that $f_{|A}$ is injective.

- Now, let's prove that $f_{|A} : A \to f(\mathbb{N})$ is surjective. Let be $y \in f(\mathbb{N})$, and let's suppose, for the sake of contradiction, that $f(a) \neq y$ for all $a \in A$.

  By definition of $f(\mathbb{N})$, there exists $a_1 \in \mathbb{N}$ such that $f(a_1) = y$. Since we suppose $f(a) \neq y$ for all $a \in A$, we have $a_1 \notin A$. By definition of the set A, it means that there exists a natural number $a_2 < a_1$ such that $f(a_2) = f(a_1) = y$, and still $a_2 \notin A$. Similarly, there exists a natural number $a_3 < a_2$ such that $a_3 \notin A$ and $f(a_3) = f(a_2) = f(a_1) = y$. Actually, we are constructing like this a sequence $(a_n)_{n=1}^{\infty}$ of natural numbers which is in infinite descent. But this is impossible (see Exercise 4.4.2). Thus, for every $y \in f(\mathbb{N})$, we necessarily have a $a \in A$ such that $f(a) = y$; i.e., $f_{|A} : A \to f(\mathbb{N})$ is surjective.

Thus, $f_{|A} : A \to f(\mathbb{N})$ is bijective. By Corollary 8.1.6, every subset of $\mathbb{N}$ is at most countable, so that A (which is clearly a subset of $\mathbb{N}$) is at most countable. And since $f(\mathbb{N})$ has the same cardinality as A, this shows that $f(\mathbb{N})$ is at most countable and closes the proof.

Exercise 8.1.5. — *Use Proposition 8.1.8 to prove Corollary 8.1.9.*

We have to show the following claim: if X is a countable set, and $f : X \to Y$ a function, then $f(X)$ is at most countable (i.e., any image of a countable set is itself countable).

By definition, if X is a countable set, there exists a bijective function $g : \mathbb{N} \to X$. Let's consider the function $h = f \circ g$, which is a function from $\mathbb{N}$ to Y. We will show that $f(X) = h(\mathbb{N})$.

- First, $f(X) \subseteq h(\mathbb{N})$. Indeed, let be $y \in f(X)$. By definition, there exists $x \in X$ such that $y = f(x)$. But since g is bijective, there exists $n \in \mathbb{N}$ such that $y = f(g(n))$, i.e., there exists $n \in \mathbb{N}$ such that $y = h(n)$. Thus, $y \in h(\mathbb{N})$.

- Furthermore, $h(\mathbb{N}) \subseteq f(X)$. Let be $y \in h(\mathbb{N})$. By definition, there exists $n \in \mathbb{N}$ such that $y = h(n) = f \circ g(n) = f(g(n))$. But since $g(n) \in X$, we have $y = f(g(n)) \in f(X)$ as required.

Thus, we have indeed $f(X) = h(\mathbb{N})$, and since $h(\mathbb{N})$ is a countable set by Proposition 8.1.8, $f(X)$ is indeed countable.

EXERCISE 8.1.6. — *Let A be a set. Show that A is at most countable if and only if there exists an injective map $f : A \to \mathbb{N}$.*

First suppose that there exists an injective function $f : A \to \mathbb{N}$. By definition, f is thus a bijection between A and $f(A)$, and those two sets thus have the same cardinality. But since we have $f(A) \subset \mathbb{N}$, the set $f(A)$ is at most countable by Corollary 8.1.6. Thus, A is at most countable, as required.

Now suppose that A is at most countable. We have two options here:

1. If A is countably infinite, then there exists a bijection $f : A \to \mathbb{N}$, and in particular, f is thus injective.

2. If A is finite, then A has a finite number of elements—say n elements. Consequently, there exists a bijective function $g : A \to [\![1, n]\!]$. Now let be $f : A \to \mathbb{N}$ such that $f(a) = g(a)$ for all $a \in A$. In particular, f is injective—but not necessarily surjective.

In both cases, we have found an injective map $f : A \to \mathbb{N}$ as required. This closes the proof.

EXERCISE 8.1.7. — *Prove Proposition 8.1.10.*

We have to prove that if X and Y are countable, then $X \cup Y$ is countable. We follow the hint given by Terence Tao for this proof.

Since X and Y are countable, there exist two bijections $f : \mathbb{N} \to X$ and $g : \mathbb{N} \to Y$. Let be $h : \mathbb{N} \to X \cup Y$ a function defined by $h(2n) = f(n)$ and $h(2n + 1) = g(n)$ for all natural numbers n. Let's show that we have $h(\mathbb{N}) = X \cup Y$.

- First, the fact that $h(\mathbb{N}) \subset X \cup Y$ is obvious: for all $m \in \mathbb{N}$, $h(m)$ belongs either to X or Y depending on whether m is odd or even, but in both cases, $h(m) \in X \cup Y$.

- Now let's prove that $X \cup Y \subset h(\mathbb{N})$. Let be $z \in X \cup Y$. If $z \in X$, then there exists $n \in \mathbb{N}$ such that $f(n) = z$, and thus $h(2n) = z$. This means that $z \in h(\mathbb{N})$. A similar argument applies if $z \in Y$. Thus, we have indeed $X \cup Y \subset h(\mathbb{N})$.

Those two properties show that $X \cup Y = h(\mathbb{N})$. According to Proposition 8.1.8 or Corollary 8.1.9, $h(\mathbb{N})$ is at most countable, as the image of the countable set $\mathbb{N}$. Thus, $h(\mathbb{N})$ can be either finite or countable. Let's suppose that it is a finite set. In this case, $X \cup Y$ is a finite set. By Proposition 3.6.14(c), X is also finite, as a subset of a finite set. This is a clear contradiction with our initial hypothesis that X is countably infinite. Thus, $h(\mathbb{N}) = X \cup Y$ is countable, as required.

EXERCISE 8.1.8. — *Use Corollary 8.1.13 to prove Corollary 8.1.14.*

We must show that any cartesian product of two countable sets is itself countable, i.e., if X and Y are countable, then $X \times Y$ is countable.

By definition, there exist two bijections $f : X \to \mathbb{N}$ and $g : Y \to \mathbb{N}$. Now let be the function $h : X \times Y \to \mathbb{N} \times \mathbb{N}$ defined by $h(x, y) = (f(x), g(y))$ for all $x \in X$ and $y \in Y$. We will show that h is also bijective.

- h is injective: if we suppose that $h(x,y) = h(x',y')$, then we have $(f(x), g(y)) = (f(x'), g(y'))$, i.e. $f(x) = f(x')$ and $g(y) = g(y')$. Since f and g are bijective, this implies $x = x'$ and $y = y'$, so that h is injective.

- h is also surjective because f and g are surjective: for all $n \in \mathbb{N}$, there exists $x \in X$ such that $n = f(x)$; and similarly, for all $m \in \mathbb{N}$ there exists $y \in Y$ such that $m = g(y)$. Thus, for all $(n,m) \in \mathbb{N} \times \mathbb{N}$ there exists $(x,y) \in X \times Y$ such that $h(x,y) = (n,m)$, i.e., h is surjective.

Thus, $X \times Y$ and $\mathbb{N} \times \mathbb{N}$ have the same cardinality. By Corollary 8.1.13, this means that $X \times Y$ is countable.

Exercise 8.1.9. — *Suppose that I is an at most countable set, and for each $\alpha \in I$, let A_α be an at most countable set. Show that the set $\bigcup_{\alpha \in I} A_\alpha$ is also at most countable. In particular, countable unions of countable sets are countable. (This exercise requires the axiom of choice, see Section 8.4.)*

This statement, although quite intuitive, is actually tricky to prove rigorously. There are a bunch of things that make the proof even trickier; for instance:

- the sets I and A_α are said to be *at most* countable, i.e. either finite or countably infinite: do we have to handle those cases separately?

- intuitively and informally, we have actually a denumerable sequence of sets, and in each set, we can also count the elements. For instance, we have a set $A_{\alpha_1} = \{a_{11}, a_{12}, \ldots\}$, a set $A_{\alpha_2} = \{a_{21}, a_{22}, \ldots\}$, and so on. Each element of $\bigcup_\alpha A_\alpha$ thus has two indices (the index of its set, and the index of its place in this set), so that we can think of a map between $\mathbb{N} \times \mathbb{N}$ and $\bigcup_\alpha A_\alpha$. But the sets A_α are not supposed to be disjoint, so that we do not really see how this map could be necessarily bijective (a same object x can belong to several A_α, and thus, several pairs (n,k) can provide an $x = a_{nk}$).

We thus need to think to a way to overcome these problems. A first remark can address (at least part of) them: if there exists an injection $f : A \to C$ between a set A and a countable set C, then A is at most countable (see also Exercise 3.6.7). Indeed, f is a bijection between A and $f(A)$, so that they have the same cardinality; and $f(A)$ is at most countable because it is a subset of C (Corollary 8.1.7). Alternatively, if A is at most countable, there exists a bijection between A and a subset of $\mathbb{N}$.

Using the notion of injection instead of bijection seems to be the way to go. So, let's begin the main proof!

- Since I is at most countable, there exists a bijection $g : N \to I$, where N is a subset of the natural numbers. Thus, $(A_\alpha)_{\alpha \in I} = (A_{g(m)})_{m \in N}$.

- Also, since each set $A_{g(m)}$ is at most countable, there exists an injective function $f_m : A_{g(m)} \to \mathbb{N}$ for each given $m \in N$. So, for each $m \in N$, let be $\mathcal{F}_m$ the set of all possible injections from $A_{g(m)}$ to $\mathbb{N}$. By the axiom of choice, we can choose simultaneously an injection in each of these sets (although we do not know which one exactly), so that we end up with an at most countable set of injections $\{f_m\}_{m \in N}$.

- Now let's consider $x \in \bigcup_{\alpha \in I} A_\alpha = \bigcup_{m \in N} A_{g(m)}$. As we said previously, x can belong to one or several sets in this union. Thus, let's consider the set $\{m \in N \,:\, x \in A_{g(m)}\}$. It's a subset of $\mathbb{N}$, so that by the well-ordering principle (Proposition 8.1.4), there exists exactly one minimal element n in this set.

- Now let's define $\theta : \bigcup_{m \in N} A_{g(m)} \to \mathbb{N} \times \mathbb{N}$ the function such that, for all $x \in \bigcup_{m \in N} A_{g(m)}$, we have $\theta(x) = (n, f_n(x))$ with n the minimal element defined above. θ is injective since n is uniquely defined and f_n is injective. Thus, $\bigcup_{m \in N} A_{g(m)}$ is at most countable, as required.

Exercise 8.2.1. — *Prove Lemma 8.2.3.*

We have to prove that, if X is a countable set, then $\sum_{x \in X} f(x)$ is absolutely convergent iff $\sup\{\sum_{x \in A} |f(x)| \,:\, A \subseteq X,\ A \text{ finite}\}$.

- First suppose that $\sup\{\sum_{x \in A} |f(x)| \,:\, A \subseteq X,\ A \text{ finite}\} < \infty$, and let's show that the series $\sum_{x \in X} f(x)$ is absolutely convergent.

  A preliminary remark: since X is a countable set, there exists a bijection $g : \mathbb{N} \to X$. Let be N a natural number. Since g is a bijection from $\mathbb{N}$ to X, the set $A_N := g(\llbracket 0, N \rrbracket)$ is a (finite) subset of X for any value of N. Furthermore, since the restriction $g : \llbracket 0, N \rrbracket \to A_N$ is bijective, we have $\sum_{x \in A_N} |f(x)| = \sum_{n=0}^{N} |f(g(n))|$. Thus, by our initial hypothesis,

$$\sup\left\{\sum_{n=0}^{N} |f(g(n))| \,:\, N \in \mathbb{N}\right\} = \sup\left\{\sum_{x \in A_N} |f(x)| \,:\, N \in \mathbb{N}\right\} \tag{8.1}$$

$$\leqslant \sup\left\{\sum_{x \in A} |f(x)| \,:\, A \subseteq X,\ A \text{ finite}\right\} < \infty \tag{8.2}$$

  Let's denote $S_N := \sum_{n=0}^{N} |f(g(n))|$. The sequence $(S_N)_{N=0}^{\infty}$ is a sequence of partial sums of non-negative numbers, so that it is increasing, and thus converges iff it is bounded (Proposition 6.3.8). By equations (8.1)–(8.2), we have $\sup(S_N)_{N=0}^{\infty} < \infty$, which means that $(S_N)_{N=0}^{\infty}$ converges. It means that $\sum_{n=0}^{\infty} |f(g(n))|$ converges for some bijection $g : \mathbb{N} \to X$, i.e. that $\sum_{x \in X} f(x)$ is absolutely convergent by Definition 8.2.1.

- Now suppose that $\sum_{x \in X} f(x)$ is absolutely convergent and let's show that we have $\sup\{\sum_{x \in A} |f(x)| \,:\, A \subseteq X,\ A \text{ finite}\} < \infty$.

  Since $\sum_{x \in X} f(x)$ is absolutely convergent, then by Definition 8.2.1, there exists some bijection $g : \mathbb{N} \to X$ such that $\sum_{n=0}^{\infty} f(g(n))$ is absolutely convergent. But by Proposition 7.4.1, we know that if one such bijection g exists, then the series $\sum_{n=0}^{\infty} f(h(n))$ is also absolutely convergent for any other bijection $h : \mathbb{N} \to X$.

  So, let's choose a bijection h that suits us. Let A be a finite (non-empty) subset of X having N elements; we define $h : \mathbb{N} \to X$ a bijection such that $h(\llbracket 0, N-1 \rrbracket) = A$. We thus have (by Proposition 7.2.14(c)):

$$\sum_{n=0}^{\infty} |f(h(n))| = \sum_{n=0}^{N-1} |f(h(n))| + \sum_{n=N}^{\infty} |f(h(n))| \tag{8.3}$$

which is equivalent to

$$\sum_{x\in X}|f(x)| = \sum_{x\in A}|f(x)| + \sum_{n=N}^{\infty}|f(h(n))| \tag{8.4}$$

And since $\sum_{n=N}^{\infty}|f(h(n))|$ converges (Proposition 7.2.14(d)) to a positive real number (Proposition 5.4.9), we get: $\sum_{x\in A}|f(x)| \leqslant \sum_{x\in X}|f(x)|$ for any subset A of X.

Finally, our initial hypothesis was that $\sum_{x\in X} f(x)$ is absolutely convergent, i.e. that $\sum_{x\in X}|f(x)|$ is a (positive) real number M. Thus, there exists $M \in \mathbb{R}$ such that $\sum_{x\in A}|f(x)| \leqslant M$, which is equivalent to $\sup\{\sum_{x\in A}|f(x)| \, : \, A \subseteq X, \ A \text{ finite}\} < \infty$, as required. This closes the proof.

Exercise 8.2.2. — *Prove Lemma 8.2.5.*

Let's follow the hint, and first consider the number:

$$M := \sup\left\{\sum_{x\in A}|f(x)| \, : \, A \subseteq X, A \text{ finite}\right\} \tag{8.5}$$

which is (by Definition 8.2.4) a finite real number since $\sum_{x\in X} f(x)$ is supposed to be absolutely convergent; and for each positive integer n, the set:

$$A_n = \{x \in X \, : \, |f(x)| > 1/n\}$$

- Let's show that all the sets A_n are finite. Suppose, for the sake of contradiction, that there exists a natural number n such that A_n is infinite. Since it is an infinite set, in particular there exists a subset finite $A \subset A_n$ such that $\#A = 2Mn$. Thus we have $\sum_{x\in A}|f(x)| > \sum_{x\in A} 1/n = 2M$. I.e., we have found a finite subset $A \subset A_n \subseteq X$ such that:
$$\sum_{x\in A}|f(x)| > \sup\left\{\sum_{x\in A}|f(x)| \, : \, A \subseteq X, A \text{ finite}\right\},$$
an obvious contradiction. Thus, A_n is finite for all $n > 0$.

- Since A_n is a finite subset of X, then by equation (8.5), we have $\sum_{x\in A_n}|f(x)| \leqslant M$. But also, we have by definition of A_n (and Proposition 7.1.11(h)) that $\sum_{x\in A_n} 1/n < \sum_{x\in A_n}|f(x)|$, so that by transitivity:
$$M > \sum_{x\in A_n} 1/n = \#A_n \times (1/n)$$
and thus $\#A_n < Mn$ for all natural number $n > 0$.

- Now, we show that
$$A := \{x \in X \, : \, f(x) \neq 0\} = \bigcup_{n=1}^{\infty} A_n \tag{8.6}$$
On the one hand, if $x \in A$, then $f(x) \neq 0$, and in particular $|f(x)| > 0$. By Exercise 5.4.4, there exists a positive integer m such that $|f(x)| > 1/m > 0$, i.e. such that $x \in A_m$.

Conversely, if $x \in \bigcup_{n=1}^{\infty} A_n$, then by definition there exists a positive integer m such that $x \in A_m$, so that $|f(x)| > /m$, and thus $x \in A$.

Both sets are thus equal.

- Finally, by Exercise 8.1.9, a countable union of countable set is countable. Thus, $\bigcup_{n=1}^{\infty} A_n$ is countable, i.e. $\{x \in X : f(x) \neq 0\}$ is countable. This closes the proof.

Exercise 8.2.3. — *Prove Proposition 8.2.6.*

All statements can be deduced in a similar fashion from the usual series laws given in Chapter 7, so that we won't prove all of them in painful details below; we just give an example of proof for one of them. Remember that since X is possibly uncountable here, only Definition 8.2.4 and Lemma 8.2.5 can be used to prove the statements. However, Lemma 8.2.5 allows to reduce them into a case where the sum is computed on an at most countable set, so that Propositions 7.1.11 and 7.2.14 apply (more or less) immediately.

For what follows, let's define $F := \{x \in X : f(x) \neq 0\}$ and $G := \{x \in X : g(x) \neq 0\}$. By Lemma 8.2.5, both F and G are at most countable.

Also, let's define $M_f := \sup\{\sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite}\}$, by Definition 8.2.4 we know that M_f is a finite real number; similarly we define $M_g := \sup\{\sum_{x \in A} |g(x)| : A \subseteq X, A \text{ finite}\}$ which is also a finite real number.

(b) Let be c a real number. First a general remark: if $\sum_{x \in X} f(x)$ is absolutely convergent, then $\sum_{x \in X} cf(x)$ is also absolutely convergent. Indeed, we know that for all finite subset $A \subseteq X$, we have $\sum_{x \in A} |f(x)| \leqslant M_f$, so that by Proposition 7.1.11(g) we have $\sum_{x \in A} |cf(x)| \leqslant |c| \sum_{x \in A} |f(x)| \leqslant |c| M_f$. It means that the set $\{\sum_{x \in A} |cf(x)| : A \subseteq X, A \text{ finite}\}$ is bounded, i.e. that $\sum_{x \in X} cf(x)$ is absolutely convergent. Now consider several cases.

If $c = 0$ there is almost nothing to prove. On the one hand, we have $c \times \sum_{x \in X} f(x) = 0$ (because $\sum_{x \in X} f(x)$ is a finite real number); on the other hand we have $\sum_{x \in X} cf(x) := \sum_{\{x \in X : cf(x) = 0\}} f(x) = \sum_{\varnothing} f(x) = 0$ by Proposition 7.1.11(a). Thus, we have $\sum_{x \in X} cf(x) = 0 = c \sum_{x \in X} f(x)$, and the claim follows.

Now, suppose instead that $c \neq 0$. Note that, in this case, $f(x) = 0 \iff cf(x) = 0$, so that:

$$\{x \in X : f(x) \neq 0\} = \{x \in X : cf(x) \neq 0\} \tag{8.7}$$

Also, we know that $c \times \sum_{x \in X} f(x) := c \times \sum_{x \in F} f(x)$, the set F being at most countable.

- If F is finite, then by Proposition 7.1.11(g), we have $c \times \sum_{x \in F} f(x) = \sum_{x \in F} cf(x)$. Thus, we have $c \times \sum_{x \in X} f(x) := c \times \sum_{x \in F} f(x) = \sum_{x \in F} cf(x) =: \sum_{x \in X} cf(x)$ (using equation (8.7) under the hood for the last equality), as expected.
- If F is countably infinite, since we have already shown that $\sum_{x \in X} cf(x) := \sum_{x \in F} cf(x)$ is absolutely convergent, we can apply Definition 8.2.1: there exists a bijection $g : \mathbb{N} \to F$ such that $\sum_{n=0}^{\infty} cf(g(n))$ is absolutely convergent. By Proposition 7.2.14(b), we finally have:

$$\sum_{x \in F} cf(x) := \sum_{n=0}^{\infty} cf(g(n)) = c \sum_{n=0}^{\infty} f(g(n)) =: c \sum_{x \in F} f(x)$$

as expected.

EXERCISE 8.2.4. — *Prove Lemma 8.2.7.*

Let be $\sum_{n=0}^{\infty} a_n$ a series which is conditionally but not absolutely convergent; and let's define the two sets $A_+ := \{n \in \mathbb{N} : a_n \geqslant 0\}$ and $A_- := \{n \in \mathbb{N} : a_n < 0\}$. We have to prove that neither $\sum_{n \in A_+} a_n$ nor $\sum_{n \in A_-} a_n$ are conditionally convergent.

We'll use a proof by contradiction.

- First, we are going to suppose something which is a little too strong: let's suppose that *both* $\sum_{n \in A_+} a_n$ and $\sum_{n \in A_-} a_n$ are conditionally convergent[28].

  In this case, $\sum_{n \in A_+} a_n$ is absolutely convergent (because for the positive series, conditional and absolute convergence are the same thing). Also, the series $\sum_{n \in A_-} (-a_n)$ is convergent by Proposition 7.2.14(b); and as it is a positive series, it is absolutely convergent; so that $\sum_{n \in A_-} a_n$ itself is absolutely convergent by Proposition 8.2.6(b). Thus, for both $\sum_{n \in A_+} a_n$ and $\sum_{n \in A_-} a_n$, conditional convergence implies absolute convergence.

  Since $\mathbb{N} = A_+ \sqcup A_-$, we have by Proposition 8.2.6(c) that $\sum_{n \in A_+} a_n + \sum_{n \in A_+} a_n = \sum_{n \in A} a_n$ is absolutely convergent; a contradiction.

  Thus, $\sum_{n \in A_+} a_n$ and $\sum_{n \in A_-} a_n$ cannot be both conditionally convergent.

- However, we have not really proved Lemma 8.2.7 when doing this. Instead, we need to prove that *even only one of them* cannot be conditionally convergent.

  Actually, it turns out that when one of them is convergent, the other one is also convergent, which closes the proof.

  Indeed, let's suppose that $\sum_{n \in A_+} a_n$ is (conditionally, and thus absolutely) convergent. Let's define the (positive) series $\sum_{n=0}^{\infty} b_n$, where $b_n = 0$ whenever $a_n < 0$, and $b_n = a_n$ whenever $a_n \geqslant 0$.

  We obviously have $\sum_{n \in A_+} b_n = \sum_{n \in A_+} a_n$ (so that $\sum_{n \in A_+} b_n$ is absolutely convergent); and $\sum_{n \in A_-} b_n = 0$ (so that it is also absolutely convergent). Thus, by Proposition 8.2.6(c), $\sum_{n=0}^{\infty} b_n$ is absolutely convergent, and is equal to $\sum_{n \in A_+} a_n$.

  Finally, since $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are both (at least conditionally) convergent, then $\sum_{n=0}^{\infty} (a_n - b_n)$ is also convergent. Since we always have $a_n - b_n \leqslant 0$ for all natural number n, conditional convergence implies absolute convergence, so that by Proposition 8.2.6(c), both $\sum_{n \in A_+} (a_n - b_n) = 0$ and $\sum_{n \in A_-} (a_n - b_n) = \sum_{n \in A_-} a_n$ are absolutely convergent.

  In particular, we have shown that whenever $\sum_{n \in A_+} a_n$ is convergent, the other series $\sum_{n \in A_-} a_n$ is also convergent, as expected. This closes the proof.

EXERCISE 8.3.1. — *Let X be a finite set of cardinality n. Show that 2^X is a finite set of cardinality 2^n.*

Let's use induction on n, the number of elements of X.

- For the base case, if $n = 0$, X is simply the empty set. In such a case, $2^{\varnothing} = \{\varnothing\}$ has indeed one single element, i.e. has a cardinality of $2^0 = 1$, and is thus finite.

[28] This hypothesis is too strong because the proper negation which should be the starting point for our proof by contradiction would be: "one of $\sum_{n \in A_+} a_n$ or $\sum_{n \in A_-} a_n$ is conditionally convergent". We'll fix that later.

- Now suppose inductively that 2^X has 2^n elements if X has n elements, and let's show that this property still holds for $n+1$.

  Suppose that $\#X = n+1$, and let's pick $x \in X$. If we denote $A = X \backslash \{x\}$, then we have $X = A \cup \{x\}$, with $\#A = n$. Note that any subset of X consists either of a subset of A, or of the pairwise union between $\{x\}$ and a subset of A (obviously, any such set is a subset of X; and conversely, any subset of X may or may not contain x, and is otherwise composed of elements of A).

  Thus, we have $2^X = \{S : S \in 2^A\} \sqcup \{S \cup \{x\} : S \in 2^A\}$. By the induction hypothesis, 2^A has cardinality 2^n, so that, by Proposition 3.6.14(b), we have $\#X = 2^n + 2^n = 2^{n+1}$, as expected.

Exercise 8.5.2. — *Give examples of a set X and a relation $\leqslant_X$ such that the relation $\leqslant_X$ is...*

We have the cases below:

(a) Reflexive and anti-symmetric but not transitive. Consider $X = \mathbb{N}$, and $\leqslant_X$ defined by $n \leqslant_X m$ iff $0 \leqslant n-m \leqslant 1$. This relation is obviously reflexive. It is also anti-symmetric, since $n \leqslant_X m$ and $m \leqslant_X n$ imply that we have both "$n = m$ or $n = 1+m$" and "$n = m$ or $m = 1+n$", which lets $n = m$ for the only possibility. But it is not transitive since $1 \leqslant_X 2$ and $2 \leqslant_X 3$ but we do not have $1 \leqslant_X 3$.

(b) Reflexive and transitive but not anti-symmetric. Consider $X = \mathbb{R}^2$ and $\leqslant_X$ defined by $(x,y) \leqslant_X (x',y')$ iff $x \leqslant x'$. It is obviously reflexive and transitive, but not anti-symmetric since $(2,3) \leqslant_X (2,4)$ and $(2,4) \leqslant_X (2,3)$ but we do not have $(2,3) = (2,4)$.

(c) Anti-symmetric and transitive but not reflexive. Consider $X = \mathbb{R}$ and $\leqslant_X$ defined by the usual strict inequality $<$. Obviously, it is transitive. It might not immediately clear why it is also anti-symmetric, however: how could it happen that we have $x = y$ after both statements $x < y$ and $y < x$, since each of them implies in particular that $x \neq y$? By trichotomy of order, we can never have both of them at the same time, so that the implication $(x < y)$ and $(y < x) \implies x = y$ is *vacuously* true. And thus, $<$ is indeed anti-symmetric.

Exercise 8.5.3. — *Given two positive integers $n, m \in \mathbb{N} - \{0\}$, we say that n divides m, and write $n|m$, if there exists a positive integer a such that $m = na$. Show that the set $\mathbb{N} - \{0\}$ with the ordering relation $|$ is a partially ordered set but not a totally ordered one.*

To show that this defines a partially order set, we just have to prove that the relation $|$ is reflexive, anti-symmetric and transitive.

- It is obviously reflexive because we have $n = n \times 1$, and 1 is a positive integer.

- It is anti-symmetric because if we suppose that $n|m$, we must have $m = na$ for some positive integer a. Similarly, if we have $m|n$, we must have $n = mb$ for some positive integer b. Gathering these two statements, we get $n = n(ab)$, i.e. $ab = 1$ by cancellation law (Corollary 2.3.7). The only possibility for $ab = 1$ on the natural numbers is that both $a = 1$ and $b = 1$, as a quick proof by contradiction would show. Thus, we have $n = mb = m$, as required.

- It is also transitive because if $n|m$ and $m|p$, we have $m = na$ and $p = mb$ for some natural numbers a, b. Thus, $p = n(ab)$, with ab a natural number, so that $n|p$ as required.

However, this does not define a totally ordered set, since we have neither $2|3$ nor $3|2$.

EXERCISE 8.5.4. — *Show that the set of positive real numbers, $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$, has no minimal element.*

This is actually a direct consequence of Exercise 5.4.4, and thus of the Archimedean property of $\mathbb{R}$. Suppose that there exists $m \in \mathbb{R}_+$ such that $m = \min(\mathbb{R}_+)$. Obviously, we must have $m < 1$, since there exists elements such as $1/2 \in \mathbb{R}_+$. Let's apply the Archimedean property with $x = 1$: there exists a positive integer M such that $Mm > 1$, i.e. such that $1/M < m$. But $1/M$ is a positive rational number, thus a positive real number, which is a contradiction.

EXERCISE 8.5.5. — *Let $f : X \to Y$ be a function from one set X to another set Y. Suppose that Y is partially ordered with some ordering relation $\leqslant_Y$. Define a relation $\leqslant_X$ on X by defining $x \leqslant_X x'$ if and only if $f(x) <_Y f(x')$ or $x = x'$. Show that this relation $\leqslant_X$ turns X into a partially ordered set. If we know in addition that the relation $\leqslant_Y$ makes Y totally ordered, does this mean that the relation $\leqslant_X$ makes X totally ordered also? If not, what additional assumption needs to be made on f in order to ensure that $\leqslant_X$ makes X totally ordered?*

First we prove that $\leqslant_X$ makes X a partially ordered set, by showing that the three properties of ordering relations hold for $\leqslant_X$:

- Reflexivity is obviously okay: $x \leqslant_X x$ is true iff $f(x) <_Y f(x)$ or $x = x$; one of these two statements is clearly true.

- $\leqslant_X$ is anti-symmetric: suppose that we have both $x \leqslant_X y$ and $y \leqslant_X x$ for $x, y \in X$. Since $x \leqslant_X y$, we have either $x = y$ (and in this case we are done) or $f(x) <_Y f(y)$. Let's suppose that $f(x) <_Y f(y)$. By the second hypothesis, we have $y \leqslant_X x$, so that we have either $y = x$ (and in this case we are done) or $f(y) <_Y f(x)$. But since we already have supposed that $f(x) <_Y f(y)$, we cannot have $f(y) <_Y f(x)$. Thus, the only possibility is $x = y$, as expected.

- $\leqslant_X$ is also transitive: if we suppose that we have both $x \leqslant_X y$ and $y \leqslant_X z$, we have actually [$x = y$ or $f(x) <_Y f(y)$] and [$y = z$ or $f(y) <_Y f(z)$]; and we must show that this implies [$x = z$ or $f(x) <_Y f(z)$]. A cumbersome but easy distinction into four pairs of cases would show that $\leqslant_X$ is transitive.

However, the hypothesis that Y is totally ordered does not make X a totally ordered set. For instance, let's take $X = Y = \mathbb{R}$ and f the constant function $f(x) = 0\, \forall x \in \mathbb{R}$. In this case, we have neither $2 \leqslant_X 3$ (because $2 = 3$ is false, and $f(2) < f(3)$ is false), nor $3 \leqslant_X 2$ (the same remark applies). So, at least, f must not be constant. But it's still not enough! Indeed, let's take $f(x) = x^2$. We have neither $-1 \leqslant_X 1$ (because $-1 = 1$ is false, and $f(-1) < f(1)$ is false) nor $1 \leqslant_X -1$ (same remark). Thus, f must be at least injective.

Exercise 8.5.7. — *Let X be a partially ordered set, and let Y be a totally ordered subset of X. Show that Y can have at most one maximum and at most one minimum.*

Let's suppose that there exist two distinct elements $m \neq m'$ in Y such that both m and m' are minimal elements. Since Y is totally ordered, we have either $m \leqslant m'$, or $m' \leqslant m$. Suppose that $m \leqslant m'$. By Definition 8.5.5, since m' is a minimal element, it is impossible that $m < m'$; thus $m = m'$ is the only possibility. The same conclusion applies if we suppose that $m' \leqslant m$, so we are done.

A similar argument shows that Y can have at most one maximal element.

Exercise 8.5.8. — *Show that every finite non-empty subset of a totally ordered set has a minimum and a maximum. (Hint: use induction.) Conclude in particular that every finite totally ordered set is well-ordered.*

Let be X a totally ordered set, and Y a finite non-empty subset of X. Since Y is finite, we may say that it has n elements. Let's induct on n.

- For the base case, let's suppose that $n = 1$ (we don't start from 0 since Y is supposed to be non-empty). It means that Y is a singleton set, i.e., $Y = \{x\}$ for some $x \in X$. We can say that x is a minimal element, since there is no $y \in Y$ such that $y < x$ (because there is simply no other element in Y). Similarly, x is a maximal element, so the base case is done.

- Now suppose inductively that the property holds for any subset Y which has n elements, and let's prove that it is still true when Y is supposed to have $n+1$ elements. By Lemma 3.6.9, if we suppose $\#Y = n + 1$, we can write $Y = A \cup \{x\}$ with $\#A = n$. By the induction hypothesis, A has a minimum m and a maximum M. If we have $x < m$, then x is a minimum for Y; else (i.e., if $m \leqslant x$) m is still a minimum for Y. Similarly, if we have $x > M$, then x is a maximum for Y; else (i.e., if $M \geqslant x$) M is still a maximum for Y. In all cases, Y has a minimum and a maximum, as expected. This closes the induction, and proves that every finite non-empty subset of a totally ordered set has a minimum and a maximum.

If X is a finite totally ordered set, all non-empty subsets $Y \subseteq X$ are finite. Thus, we have even proved that every finite totally ordered set is well-ordered, in the sense of Definition 8.5.8.

Exercise 8.5.9. — *Let X be a totally ordered set such that every non-empty subset of X has both a minimum and a maximum. Show that X is finite.*

Suppose for sake of contradiction that X is infinite. By the initial hypothesis, there exists a minimal element $x_0 \in X$. Since X is infinite, the set $X - \{x_0\}$ is non-empty; let be x_1 its minimum. We have $x_0 \leqslant x_1$ because $x_1 \in X$ and $x_0 = \min(X)$, but we also have $x_0 \neq x_1$ because $x_1 \in X - \{x_0\}$ by definition. Thus, we have $x_0 < x_1$. Now let's consider the subset of X defined by $X - \{x_0, x_1\}$: it it still a non-empty subset (otherwise X could not be infinite), and the same argument as above would show that its minimum x_2 (wich certainly exists) is such that $x_0 < x_1 < x_2$. We can thus repeat these steps indefinitely to construct an increasing sequence $x_0 < x_1 < x_2 < \cdots$.

More formally, we can define recursively $x_0 := \min(X)$ and, for each natural number n, $x_{n+1} := \min(X - \cup_{i=0}^{n}\{x_i\})$; so that $(x_n)_{n=0}^{\infty}$ is an increasing sequence.

By definition, the set $\{x_0, x_1, \cdots\}$ is an infinite set, is thus a non-empty subset of X, but has no maximal element. This contradicts our initial hypothesis. Thus, X cannot be infinite: it is a finite set.

Exercise 8.5.10. — *Prove Proposition 8.5.10.*

Let be X a totally ordered set, and P a property such that, for any $n \in X$, we have the following implication: $[P(m)$ is true for all $m \in X$ with $m < n] \Rightarrow [P(n)$ is true$]$. We have to prove that $P(n)$ is true for all $n \in X$.

First, note that, if we denote $0 := \min(X)$ (and this minium necessarily exists since X is totally ordered), $P(0)$ must be true. Indeed, the statement "$P(m)$ is true for all $m \in X$ with $m < 0$" is vacuously true, so that $P(0)$ must be true.

Now suppose, for the sake of contradiction, that the set Y defined by

$$Y := \{n \in X \,:\, P(m) \text{ is false for some } m \in X \text{ with } m \leqslant n\}$$

is non-empty. Since X is totally ordered, there exists a minimal element M in Y, i.e. there exists some $M := \min(Y)$. In particular, M is the lowest element of X for which $P(M)$ is false.

It is impossible that $M = 0$, because it would imply that $P(0)$ is false, which has been excluded. Thus, we have $M \neq 0$, or more precisely, $M > 0$.

Since $M > 0$, for all elements $m \in X$ such that $0 \leqslant m < M$ (there is at least one such element, which is 0), $P(m)$ is true. By our initial hypothesis, it would thus imply that $P(M)$ is true; a contradiction.

Thus, Y is empty, and $P(n)$ is true for all $n \in X$.

9. Continuous functions on $\mathbb{R}$

EXERCISE 9.1.1. — *Let X be any subset of the real line, and let Y be a set such that $X \subseteq Y \subseteq \overline{X}$. Show that $\overline{Y} = \overline{X}$.*

We will show that we have $\overline{X} \subseteq \overline{Y}$, and $\overline{Y} \subseteq \overline{X}$. Let be $\varepsilon > 0$ a positive real number.

- Let be $x' \in \overline{X}$. By definition, there exists $x \in X$ such that $|x - x'| \leqslant \varepsilon$. But since $X \subseteq Y$, we also have $x \in Y$. Thus, for any arbitrary $\varepsilon > 0$, we have indeed an $x \in Y$ such that $|x' - x| \leqslant \varepsilon$, which means that $x' \in \overline{Y}$. So, $\overline{X} \subseteq \overline{Y}$.

- Let be $y' \in \overline{Y}$. By definition, there exists $y \in Y$ such that $|y' - y| \leqslant \varepsilon/2$. But since $Y \subseteq \overline{X}$, we have $y \in \overline{X}$. It means that we can find an $x \in X$ such that $|y - x| \leqslant \varepsilon/2$. Thus, we have found an $x \in X$ such that $|y' - x| \leqslant |y' - y| + |y - x| \leqslant \varepsilon/2 + \varepsilon/2 \leqslant \varepsilon$, for any arbitrary $\varepsilon > 0$. It means that $y' \in \overline{X}$; so that $\overline{Y} \subseteq \overline{X}$.

- We have thus proved that $\overline{X} = \overline{Y}$, as expected.

EXERCISE 9.1.2. — *Prove Lemma 9.1.11.*

Let X and Y be subsets of $\mathbb{R}$.

- Prove that $X \subseteq \overline{X}$. If $x \in X$, then for any $\varepsilon > 0$ we always have an $y \in X$ such that $|x - y| \leqslant \varepsilon$: let's take $y = x$, so that $|x - y| = 0 \leqslant \varepsilon$.

- Prove that $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$. Let be $\varepsilon > 0$ a real number, and $z' \in \overline{X \cap Y}$. By definition, there exists $z \in X \cap Y$ such that $|z' - z| \leqslant \varepsilon$. We have $z \in X$ and $z \in Y$. Thus, for an arbitrary $\varepsilon > 0$, we have found a $z \in X$ such that $|z' - z| \leqslant \varepsilon$, i.e. $z' \in \overline{X}$. Similarly, $z' \in \overline{Y}$. Thus, $z' \in \overline{X} \cap \overline{Y}$. Thus, we have proved $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$ as required.

- Prove that $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$. Let be $\varepsilon > 0$ a positive real number, and $z' \in \overline{X} \cup \overline{Y}$. We have either $z' \in \overline{X}$ or $z' \in \overline{Y}$. Suppose that $z' \in \overline{X}$. By definition, there exists $x \in X$ such that $|z' - x| \leqslant \varepsilon$. But since $X \subseteq X \cup Y$, we have found an $x \in X \cup Y$ such that $|z' - x| \leqslant \varepsilon$, and thus we conclude that $z' \in \overline{X \cup Y}$. The same conclusion applies if we suppose instead that $z' \in \overline{Y}$. Thus, we have indeed $\overline{X} \cup \overline{Y} \subseteq \overline{X \cup Y}$.

  Now let be $z' \in \overline{X \cup Y}$. Suppose that $z' \notin \overline{X} \cup \overline{Y}$; this means that $z' \notin \overline{X}$ and $z' \notin \overline{Y}$. Thus, there exists $\delta > 0$ such that $|z' - x| > \delta$ for all $x \in X$. Similarly, there exists $\gamma > 0$ such that $|z' - y| > \gamma$ for all $y \in Y$. We conclude that, for $\varepsilon := \min(\delta, \gamma)$, we have $|z' - x| > \varepsilon$ for all $x \in X$ and $|z' - y| > \varepsilon$ for all $y \in Y$; i.e. that $|z' - z| > \varepsilon$ for all $z \in X \cup Y$. Thus, $z \notin \overline{X \cup Y}$, which obviously contradicts our initial hypothesis. It means that whenever $z' \in \overline{X \cup Y}$, we have $z' \in \overline{X} \cup \overline{Y}$, i.e. $\overline{X \cup Y} \subseteq \overline{X} \cup \overline{Y}$ as required.

- Prove that if $X \subseteq Y$, then $\overline{X} \subseteq \overline{Y}$. Let be $x' \in \overline{X}$, and $\varepsilon > 0$ a real number. By definition, there exists $x \in X$ such that $|x' - x| \leqslant \varepsilon$. But since $X \subseteq Y$, we also have $x \in Y$. Thus, for any $\varepsilon > 0$, we have an $x \in Y$ such that $|x' - x| \leqslant \varepsilon$, which means that $x' \in \overline{Y}$.

EXERCISE 9.1.3. — *Prove Lemma 9.1.13*

We will prove all statements in a slightly different order, and will use previous results from this chapter.

First, note that, by Definition 9.1.8, any adherent point to a subset $X \subseteq \mathbb{R}$ is necessarily a real number. It means that, for any set $X \subseteq \mathbb{R}$, we have $\overline{X} \subseteq \mathbb{R}$. This gives our first statement:

(i) The closure of $\mathbb{R}$ is $\mathbb{R}$. Indeed, we have $\overline{\mathbb{R}} \subseteq \mathbb{R}$ by our remark above, and we have $\mathbb{R} \subseteq \overline{\mathbb{R}}$ by Lemma 9.1.11. The claim follows.

(ii) Now we prove that the closure of $\varnothing$ is $\varnothing$. Suppose, for the sake of contradiction, that $\overline{\varnothing}$ is not empty. Then there exists, by Definition 9.1.8, a real number $x \in \overline{\varnothing}$. So, for a given $\varepsilon > 0$, we necessarily have a $z \in \varnothing$ such that $|x - z| \leqslant \varepsilon$, but there exists no such z. Thus, $\overline{\varnothing}$ is empty.

(iii) The closure of $\mathbb{N}$ is $\mathbb{N}$. Indeed, by Lemma 9.1.1, we already know that $\mathbb{N} \subseteq \overline{\mathbb{N}}$, so that we just have to show that $\overline{\mathbb{N}} \subseteq \mathbb{N}$. Let be $n \in \overline{\mathbb{N}}$. Suppose, for the sake of contradiction, that $n \notin \mathbb{N}$, i.e. that $n \in \mathbb{R} \backslash \mathbb{N}$. By Exercise 5.4.3, there exists an integer m such that $m \leqslant n < m + 1$. Let be $d_1 = |m - n| > 0$, which is a positive real number since $n \notin \mathbb{N}$ and $m \in \mathbb{N}$. Similarly, let be $d_2 := |n - (m + 1)| > 0$. Now, let be $d := \min(d_1, d_2)/2$, which is also a positive real number. Obviously, n is not d-adherent to $\mathbb{N}$, so that $n \notin \overline{\mathbb{N}}$, a contradiction. Thus, we have indeed $\mathbb{N} = \overline{\mathbb{N}}$.

(iv) The proof for $\mathbb{Z} = \overline{\mathbb{Z}}$ is similar.

(v) Finally, let's show that $\overline{\mathbb{Q}} = \mathbb{R}$. As for the previous cases, we only have to show that $\mathbb{R} \subseteq \overline{\mathbb{Q}}$, since we already know that $\overline{\mathbb{Q}} \subseteq \mathbb{R}$. So, let be $z \in \mathbb{R}$, and let's show that $z \in \overline{\mathbb{Q}}$. Let be $\varepsilon > 0$ a positive real number. By Proposition 5.4.14, there always exists a rational $y \in \mathbb{Q}$ such that $z - \varepsilon < y < z + \varepsilon$, i.e. such that $|y - z| < \varepsilon$. Thus, z is ε-adhrent to $\mathbb{Q}$ for any $\varepsilon > 0$. This means that $z \in \overline{\mathbb{Q}}$, as expected.

Exercise 9.1.4. — *Give an example of two subsets X, Y of the real line such that $\overline{X \cap Y} \neq \overline{X} \cap \overline{Y}$.*

We already know, from Lemma 9.1.11, that $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$. Thus, we have to find $X, Y \subseteq \mathbb{R}$ such that $\overline{X} \cap \overline{Y}$ fails to be a subset of $\overline{X \cap Y}$.

Let be $X =]1, 2[$ and $Y =]2, 3[$. We have (by Lemma 9.1.12) $\overline{X} = [1, 2]$ and $\overline{Y} = [2, 3]$, so that $\overline{X} \cap \overline{Y} = \{2\}$. On the other hand, we have $X \cap Y = \varnothing$, so that $\overline{X \cap Y} = \varnothing$ (by Lemma 9.1.13). Thus, we indeed have $\overline{X \cap Y} \neq \overline{X} \cap \overline{Y}$, as expected.

Exercise 9.1.5. — *Prove Lemma 9.1.14.*

Let X be a subset of $\mathbb{R}$, and let be $\ell \in \mathbb{R}$. We have to prove that the following two statements are equivalent: (a) There exists a sequence $(a_n)_{n=1}^{\infty}$ whose all elements lie in X and which converges to ℓ. (b) $\ell \in \overline{X}$.

- First, we prove that (a) implies (b). Suppose that we have a sequence $(a_n)_{n=1}^{\infty}$ that converges to $\ell \in \mathbb{R}$ and such that $a_n \in X$ for all $n \geqslant 1$, and let be $\varepsilon > 0$. By definition, there exists a positive integer N such that $|a_n - \ell| \leqslant \varepsilon$ for all $n \geqslant N$; and in particular, ℓ is ε-adherent to X since $|a_N - \ell| \leqslant \varepsilon$ with $a_N \in X$. Thus, $\ell \in \overline{X}$.

- Now prove that (b) implies (a). Let be $\ell \in \overline{X}$. For any given positive integer n, let be the set $A_n := \{x \in X \,:\, |x - \ell| \leqslant 1/n\}$. By definition, A_n is non empty for all $n \geqslant 1$, i.e.

there exists a real number $a_n \in A_n$ for all $n \geqslant 1$. By the axiom of choice, we can thus define a sequence $(a_n)_{n=1}^{\infty}$ where $a_n \in A_n$ for all $n \geqslant 1$. This means that for all $n \geqslant 1$, we have $\ell - 1/n \leqslant a_n \leqslant \ell + 1/n$. By the squeeze test, $(a_n)_{n=1}^{\infty}$ converges to ℓ.

- Thus, both statements are equivalent, as expected.

Exercise 9.1.6. — *Let X be a subset of $\mathbb{R}$. Show that $\overline{X}$ is closed. Furthermore, show that if Y is any closed set that contains X, then Y also contains $\overline{X}$.*

First we show that $\overline{X}$ is closed. We already know, by Lemma 9.1.11, that we have $\overline{X} \subseteq \overline{\overline{X}}$. Thus, we juste have to show that $\overline{\overline{X}} \subseteq \overline{X}$. Let be $x'' \in \overline{\overline{X}}$, and $\varepsilon > 0$ any positive real number. By definition, there exists $x' \in \overline{X}$ such that we have $|x'' - x'| \leqslant \varepsilon/2$. And since $x' \in \overline{X}$, then by definition, there exists $x \in X$ such that we have $|x' - x| \leqslant \varepsilon/2$. Thus, by triangular inequality, we have $|x'' - x| \leqslant |x'' - x'| + |x' - x| \leqslant 2\varepsilon/2 \leqslant \varepsilon$. Thus, for any $\varepsilon > 0$, there exists $x \in X$ such that $|x'' - x| \leqslant \varepsilon$, which show that $x'' \in \overline{X}$, and thus that $\overline{\overline{X}} \subseteq \overline{X}$, as expected.

Now let be Y a closed set, i.e. a set such that $Y = \overline{Y}$. By Lemma 9.1.11, if $X \subseteq Y$, then we have $\overline{X} \subseteq \overline{Y}$. But since $\overline{Y} = Y$, we have $\overline{X} \subseteq Y$, as expected. Thus, $\overline{Y}$ is the smallest closed set that contains X.

Exercise 9.1.7. — *Let $n > 1$ be a positive integer, and let $X_1, \ldots, X_n$ be closed subsets of $\mathbb{R}$. Show that $X_1 \cup \ldots \cup X_n$ is also closed.*

Let's use induction on n. To show the base case, $n = 2$, we must show that $X_1 \cup X_2$ is closed whenever X_1 and X_2 are closed. By Lemma 9.1.11, we have $\overline{X_1 \cup X_2} = \overline{X_1} \cup \overline{X_2}$, and this is equal to $X_1 \cup X_2$ since X_1, X_2 are closed. Thus the base case is done.

Now suppose inductively that $X_1 \cup \ldots \cup X_n$ is closed for some positive integer n, and let's show that $X_1 \cup \ldots \cup X_{n+1}$ is also closed. We have:

$$\begin{aligned}
\overline{X_1 \cup \ldots \cup X_n \cup X_{n+1}} &= \overline{(X_1 \cup \ldots \cup X_n) \cup X_{n+1}} \\
&= (\overline{X_1 \cup \ldots \cup X_n}) \cup \overline{X_{n+1}} \text{ (Lemma 9.1.11)} \\
&= (\overline{X_1} \cup \ldots \cup \overline{X_n}) \cup \overline{X_{n+1}} \text{ (by induction hypothesis)} \\
&= \overline{X_1} \cup \ldots \cup \overline{X_n} \cup \overline{X_{n+1}} \\
&= X_1 \cup \ldots \cup X_n \cup X_{n+1} \text{ (because all } X_i\text{'s are closed)}
\end{aligned}$$

This closes the induction.

Exercise 9.1.8. — *Let I be a set (possibly infinite), and for each $\alpha \in I$ let X_α be a closed subset of $\mathbb{R}$. Show that the intersection $\bigcap_{\alpha \in I} X_\alpha$ is also closed.*

Recall that $\bigcap_{\alpha \in I} X_\alpha$ is the set such that, for any real x, we have $x \in \bigcap_{\alpha \in I} X_\alpha$ iff $x \in X_\alpha$ for all $\alpha \in I$.

We already know by Lemma 9.1.11 that we have $\bigcap_{\alpha \in I} X_\alpha \subseteq \overline{\bigcap_{\alpha \in I} X_\alpha}$; thus, to show that $\bigcap_{\alpha \in I} X_\alpha$ is closed, we just have to show that $\overline{\bigcap_{\alpha \in I} X_\alpha} \subseteq \bigcap_{\alpha \in I} X_\alpha$.

Let be $x' \in \overline{\bigcap_{\alpha \in I} X_\alpha}$, and $\varepsilon > 0$ any positive real. By definition, there exists $x \in \bigcap_{\alpha \in I} X_\alpha$ such that $|x' - x| \leqslant \varepsilon$. By definition of an intersection, we have $x \in X_\alpha$ for all $\alpha \in I$, so that we also have $x' \in \overline{X_\alpha}$ for all $\alpha \in I$. Thus, $x' \in \bigcap_{\alpha \in I} \overline{X_\alpha}$. But since each X_α is closed, we finally get $x' \in \bigcap_{\alpha \in I} X_\alpha$, as expected.

Exercise 9.1.9. — *Let X be a subset of the real line, and x be a real number. Show that every adherent point of X is either a limit point or an isolated point of X, but cannot be both. Conversely, show that every limit point and every isolated point of X is an adherent point of X.*

Let's begin by the last assertion. First we prove that any isolated point of X is an adherent point of X. Let be $x \in X$ an isolated point. The claim is obvious since, by Lemma 9.1.11, we have $X \subseteq \overline{X}$, so that $x \in \overline{X}$. Also, we prove that any limit point of X is adherent to X. If x is a limit point of X, then $x \in \overline{X - \{x\}}$. It means that, for all $\varepsilon > 0$, there exists some $y \in X - \{x\}$ such that $|x - y| \leqslant \varepsilon$. But since $X - \{x\} \subseteq X$, there exists actually some $y \in X$ such that $|x - y| \leqslant \varepsilon$ for all $\varepsilon > 0$, which means that $x \in \overline{X}$ as expected.

Now suppose that $x \in \overline{X}$ and let's show that it is either a limit point or an isolated point. Having $x \in \overline{X}$ means that:

$$\forall \varepsilon > 0, \exists y \in X \,:\, |x - y| \leqslant \varepsilon$$

There are two cases which are mutually exclusive: either the $y \in X$ we can find for any given value of ε always lies in $X - \{x\}$, or it is not the case. In the first case, we have:

$$\forall \varepsilon > 0, \exists y \in X - \{x\} \,:\, |x - y| \leqslant \varepsilon$$

i.e., x is a limit point of X. In the other case, let's take the negation to get:

$$\exists \varepsilon > 0 \,:\, \forall y \in X - \{x\}, |x - y| > \varepsilon$$

i.e., x is an isolated point of X. The claim follows.

Exercise 9.1.10. — *If X is a non-empty subset of $\mathbb{R}$, show that X is bounded if and only if $\inf(X)$ and $\sup(X)$ are finite.*

First suppose that X is bounded. Thus, by definition, there exists a positive real number M such that $X \subseteq [-M, M]$. Thus, we have $\sup(X) \leqslant M$ (because M is an upper bound of X, and $\sup(X)$ is the least upper bound), so that $\sup(X)$ is finite. Similarly for $\inf(X) \geqslant -M$.

Now suppose that both $\inf(X)$ and $\sup(X)$ are finite. We thus have $\inf(X) = a$ and $\sup(X) = b$ for some real numbers a, b. By definition, we have $a \leqslant x \leqslant b$ for all $x \in X$. It means that $X \subseteq [a, b]$. Let's pick $M := \max(|a|, |b|)$; we thus have $X \subseteq [-M, M]$ as expected.

Exercise 9.1.11. — *Show that if X is a bounded subset of $\mathbb{R}$, then the closure $\overline{X}$ is also bounded.*

If X is bounded, we have $X \subseteq [-M, M]$ for some positive real number M. The interval $[-M, M]$ is a closed subset, thus its closure is also $[-M, M]$. Now let's use Exercise 9.1.6, with $Y := [-M, M]$ being the closed subset that contains X; we thus have $X \subseteq \overline{X} \subseteq [-M, M]$. It means that $\overline{X}$ is bounded, as expected.

Exercise 9.1.13. — *Prove Theorem 9.1.24 (Heine-Borel theorem).*

We have to prove that these two statements are equivalent for any subset X of $\mathbb{R}$: (a) X is closed and bounded; (b) from any sequence of elements of X, one can get a subsequence that converges to a real number $L \in X$.

- First we prove that (a) implies (b). Let be X a subset of $\mathbb{R}$ which is closed and bounded, and let be $(a_n)_{n=0}^{\infty}$ a sequence such that $a_n \in X$ for all $n \geqslant 0$. Since X is bounded, there exists a real number $M > 0$ such that $X \subseteq [-M, M]$. The sequence $(a_n)_{n=0}^{\infty}$ is bounded: we have necessarily $|a_n| \leqslant M$ (because $a_n \in X$) for all $n \geqslant 0$. Since $(a_n)_{n=0}^{\infty}$ is bounded, there exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges to a real number $L \in \mathbb{R}$ (Bolzano-Weierstrass theorem). But since X is closed, we have $L \in X$ by Corollary 9.1.17. The claim follows.

- Now we prove that (b) implies (a). We start from (b), and suppose for the sake of contradiction that X is not closed. Thus, there exists a sequence $(a_n)_{n=0}^{\infty}$ such that $a_n \in X$ for all $n \geqslant 0$, which converges to a real number $L \notin X$ (by negation of Corollary 9.1.17). But according to (b), from the sequence $(a_n)_{n=0}^{\infty}$, we can extract a subsequence that converges to $L' \in X$. It is a contradiction, since it implies that $L \neq L'$, which is impossible by Proposition 6.6.5. Thus, X is closed.

  Furthermore, X is bounded. Indeed,

Exercise 9.1.14. — *Show that any finite subset of $\mathbb{R}$ is closed and bounded.*

Let be A_n a finite subset of $\mathbb{R}$. We can use induction on n, the number of elements of A_n.

- For the base case, A_1 is simply a singleton set; say that $A_1 := \{a_1\}$ with $a_1 \in \mathbb{R}$. Obviously, A_1 is closed, because a_1 is an adherent point of A_1, and no other real can be ε-adherent to A_1 for all $\varepsilon > 0$ (if $x \neq a_1$, then x is not ε-adherent to A_1 for $\varepsilon = |x - a_1|/2$, for instance). Furthermore, A_1 is bounded, since $|x| \leqslant a_1$ for all $x \in A_1$.

- Now suppose inductively that any set A_n is closed and bounded if A_n has n elements, and let's show that any set A_{n+1} with $n+1$ elements is also closed and bounded. If A_{n+1} has $n+1$ elements, then we have $A_{n+1} = A_n \cup \{a_{n+1}\}$, where A_n has n elements (e.g., Proposition 3.6.14 (a)). But, by the base case, $\{a_{n+1}\}$ is closed and bounded; and by the induction hypothesis, A_n is closed and bounded. Thus, by Exercise 9.1.7, $A_n \cup \{a_{n+1}\}$ is also closed. This union set is also bounded: if A_n is bounded by M, then A_{n+1} is bounded by $\max(M, |a_{n+1}|)$.

  Now let's show that X is bounded. Suppose for the sake of contradiction that X is not bounded. In particular, either X has no upper bound, or no lower bound (or both). Suppose that X has no upper bound; the proof is similar if it has no lower bound. Saying that X is not bounded above means that, for all $n \geqslant 0$, the set $A_n := \{x \in X : x > n\}$ is not empty. Thus, by the axiom of (countable) choice, we can define a sequence $(a_n)_{n=0}^{\infty}$ such that $a_n \in A_n$ for all $n \geqslant 0$. In particular, we have $a_n > n$ for all $n \geqslant 0$. But this sequence has no limit point: indeed, saying that $(a_n)_{n=0}^{\infty}$ has a finite limit point L means that:

  $$\forall \varepsilon > 0, \forall N \geqslant 0, \exists n \geqslant N : L - \varepsilon \leqslant a_n \leqslant L + \varepsilon \tag{9.1}$$

  But this is impossible: let's take $\varepsilon = 1$, and any $n > L + 1$; we thus have $a_n > L + 1$ by definition, which is not compatible with (9.1). And according to Proposition 6.6.6, if $(a_n)_{n=0}^{\infty}$ has no limit point, then we cannot extract a convergent subsequence from $(a_n)_{n=0}^{\infty}$; a contradiction with our initial hypothesis (a). This closes the proof.

Exercise 9.2.1. — *Let f, g, h be functions from $\mathbb{R}$ to $\mathbb{R}$: which of the following statements are true?*

The first statement, $(f+g)\circ h = (f\circ h)+(g\circ h)$, is true. Indeed, we have for all $x \in \mathbb{R}$:

$$\begin{aligned}((f+g)\circ h)(x) &= (f+g)(h(x))\\ &= f(h(x))+g(h(x))\\ &= (f\circ h)(x)+(g\circ h)(x)\end{aligned}$$

The second statement, $f\circ(g+h) = (f\circ g)+(f\circ h)$, is false. Indeed, let's consider $f(x) = x^2$, $g(x) = x$ and $h(x) = -x$. On the one hand we have $f\circ(g+h)(1) = f(1-1) = f(0) = 0$; and the other hand we have $(f\circ g + f\circ h)(1) = f(1)+f(-1) = 1+1 = 2$.

The third statement, $(f+g)\times h = (f\times h)+(g\times h)$, is true. Indeed, we have for all $x \in \mathbb{R}$:

$$\begin{aligned}((f+g)\times h)(x) &= (f+g)(x)\times h(x)\\ &= (f(x)+g(x))\times h(x)\\ &= f(x)h(x)+g(x)h(x)\end{aligned}$$

The last statement, $f\times(g+h) = (f\times g)+(f\times h)$, is true. Indeed, we have for all $x \in \mathbb{R}$:

$$\begin{aligned}(f\times(g+h))(x) &= f(x)\times(g+h)(x)\\ &= f(x)\times(g(x)+h(x))\\ &= f(x)g(x)+f(x)h(x)\end{aligned}$$

Exercise 9.3.1. — *Prove Proposition 9.3.9.*

We must prove that these two statements are equivalent: (a) f converges to L at x_0 in E; (b) for every sequence $(a_n)_{n=0}^{\infty}$ which consists entirely of elements of E and converges to x_0, the sequence $(f(a_n))_{n=0}^{\infty}$ converges to L.

- First we prove that (a) implies (b). Suppose that f converges to L at x_0, and let be $(a_n)_{n=0}^{\infty}$ a sequence of elements of E that converges to x_0. We must show that:

  $$\forall \varepsilon > 0, \exists N \geqslant 0 \,:\, n \geqslant N \Longrightarrow |f(a_n)-L| \leqslant \varepsilon \tag{9.2}$$

  Let be $\varepsilon > 0$. We know that there exists a $\delta > 0$ such that $|x-x_0| < \delta \Rightarrow |f(x)-L| \leqslant \varepsilon$. Consider this positive real δ. Since $(a_n)_{n=0}^{\infty}$ converges to x_0, there exists a natural number N such that $n \geqslant N \Rightarrow |a_n - x_0| \leqslant \delta$. Thus, for this number N, we have:

  $$n \geqslant N \Longrightarrow |a_n - x_0| \leqslant \delta \Longrightarrow |f(a_n)-L| \leqslant \varepsilon$$

  so that this natural number N is suitable for the property in equation (9.2). This closes the first part of the proof.

- Conversely, let's prove that (b) implies (a). Suppose for the sake of contradiction that we have (b), but f does not converges to L at x_0. It means that:

  $$\exists \varepsilon > 0 \,:\, \forall \delta > 0, \exists x \in E, |x-x_0| < \delta \text{ and } |f(x)-L| > \varepsilon \tag{9.3}$$

  And in particular:

  $$\exists \varepsilon > 0 \,:\, \forall n \geqslant 0, \exists a_n \in E, |a_n - x_0| < \frac{1}{n+1} \text{ and } |f(a_n)-L| > \varepsilon \tag{9.4}$$

Consider this positive real $\varepsilon > 0$. Using equation (9.4), we build a sequence[29] $(a_n)_{n=0}^{\infty}$ of elements of E such that, for all $n \geqslant 0$, we have $x_0 - \frac{1}{n+1} < a_n < x_0 + \frac{1}{n+1}$. According to the squeeze test, it implies that $(a_n)_{n=0}^{\infty}$ converges to x_0. But at the same time, $f(a_n)$ is never ε-close to L, still according to equation (9.4). This is a contradiction with the hypothesis (b). The claim follows.

Exercise 9.3.2. — *Prove the remaining claims in Proposition 9.3.14.*

All the claims can be proved in a very similar fashion, and each of them is itself very similar to the proof of the first claim given in the book. We will just prove another one, namely that fg has a limit LM at x_0 in E.

Since x_0 is adherent to E, there exists a sequence $(a_n)_{n=0}^{\infty}$ of elements of E that converges to x_0 (Lemma 9.1.14). By Proposition 9.3.9, we thus know that $(f(a_n))_{n=0}^{\infty}$ converges to L. Similarly, $(g(a_n))_{n=0}^{\infty}$ converges to M. By Theorem 6.1.19(b), we have

$$\lim_{n\to\infty}((f \times g)(a_n))_{n=0}^{\infty} = \lim_{n\to\infty}(f(a_n))_{n=0}^{\infty} \times \lim_{n\to\infty}(g(a_n))_{n=0}^{\infty} = LM.$$

Thus, since $(a_n)_{n=0}^{\infty}$ is an arbitrary sequence, we have showed that for any sequence $(a_n)_{n=0}^{\infty}$ of elements of E that converges to x_0, the sequence $((fg)(a_n))_{n=0}^{\infty}$ converges to LM. Thus, by Proposition 9.3.9, fg has a limit LM at x_0 in E.

Exercise 9.3.3. — *Prove Lemma 9.3.18.*

Let be $\delta > 0$ a positive real number. We have the following equivalent statements:

$$\begin{aligned}
&\lim_{x\to x_0\,;\, x\in E\cap]x_0-\delta,x_0+\delta[} f(x) = L \\
&\iff (\forall\varepsilon > 0, \exists\alpha > 0 \,:\, x \in \{z \in E\cap]x_0 - \delta, x_0 + \delta[\,:\, |z - x_0| < \alpha\} \Rightarrow |f(x) - L| < \varepsilon) \\
&\iff (\forall\varepsilon > 0, \exists\alpha > 0 \,:\, x \in \{z \in E\cap]x_0 - \min(\delta,\alpha), x_0 + \min(\delta,\alpha)[\} \Rightarrow |f(x) - L| < \varepsilon) \\
&\iff (\forall\varepsilon > 0, \exists\gamma > 0 \,:\, x \in \{z \in E\cap]x_0 - \gamma, x_0 + \gamma[\} \Rightarrow |f(x) - L| < \varepsilon) \\
&\iff \lim_{x\to x_0\,;\, x\in E} f(x) = L
\end{aligned}$$

Exercise 9.4.1. — *Prove Proposition 9.4.7.*

We will prove the following implications: (a) implies (c), (c) implies (d), (d) implies (b), and (b) implies (a).

- First we prove that (a) implies (c). Let be $\varepsilon > 0$ a positive real. If f is continuous at x_0, then by Definition 9.4.1, we have $\lim_{x\to x_0} f(x) = f(x_0)$. Thus, for any $\varepsilon' > 0$, there exists $\delta > 0$ such that $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| \leqslant \varepsilon'$. Let's take $\varepsilon' := \varepsilon/2$, and we have $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| \leqslant \varepsilon' < \varepsilon$, as expected for (c).
- Then, it is simply obvious that (c) implies (d), since (c) is simply one of the possible cases of (d).

[29] We "secretly" use the axiom of choice here.

- Now we show that (d) implies (b). Let be $(a_n)_{n=0}^{\infty}$ a sequence that converges to x_0. Also, let $\varepsilon > 0$ be a positive real number. Then, by (d), there exists $\delta > 0$ such that $|x - x_0| \leqslant \delta \Rightarrow |f(x) - f(x_0)| \leqslant \varepsilon$. Let's consider this number $\delta > 0$. Since $(a_n)_{n=0}^{\infty}$ converges to x_0, there exists $N \geqslant 0$ such that $n \geqslant N \Rightarrow |a_n - x_0| \leqslant \delta$.

  Gathering all those implications, for any arbitrary $\varepsilon > 0$, we have found a natural number $N \geqslant 0$ such that, if $n \geqslant N$, then we have $|a_n - x_0| \leqslant \delta$, and thus $|f(a_n) - f(x_0)| \leqslant \varepsilon$. It means that $(f(a_n))_{n=0}^{\infty}$ converges to $f(x_0)$, as expected.

- And finally, we have to prove that (b) implies (a). This is a direct application of Proposition 9.3.9.

Exercise 9.4.2. — *Let X be a subset of $\mathbb{R}$, and let $c \in \mathbb{R}$. Show that the constant function $f : X \to \mathbb{R}$ defined by $f(x) := c$ is continuous, and show that the identity function $g : X \to \mathbb{R}$ defined by $g(x) := x$ is also continuous.*

The remarks made in the textbook page 224, along with the Examples 9.4.2 and 9.4.3, would suffice to show this property. But we will ignore those paragraphs and give a proper proof, mainly using Proposition 9.3.9.

- First we show that the constant function $f(x) = c$ is continuous. Let be x_0 any real number. Let be $(a_n)_{n=0}^{\infty}$ any sequence of real numbers that converges to x_0. By Proposition 9.3.9, the sequence $(f(a_n))_{n=0}^{\infty}$ must converge to L, which is the limit $L := \lim_{x \to x_0} f(x)$. But the sequence $(f(a_n))_{n=0}^{\infty}$ is the constant sequence $c, c, c, \ldots$, which necessarily converges to c. Thus, by Proposition 9.3.9, we have $L = c$. This shows that f is continuous at any real number x_0.

- Now we show that the identity function $f(x) = x$ is continuous. Let be x_0 any real number. Let be $(a_n)_{n=0}^{\infty}$ any sequence of real numbers that converges to x_0. By Proposition 9.3.9, the sequence $(f(a_n))_{n=0}^{\infty}$ must converge to L, which is the limit $L := \lim_{x \to x_0} f(x)$. But the sequence $(f(a_n))_{n=0}^{\infty}$ is the sequence $(a_n)_{n=0}^{\infty}$, which is known to converge to x_0. Thus, by Proposition 9.3.9, we have $L = x_0$. This shows that f is continuous at any real number x_0.

Exercise 9.4.3. — *Prove Proposition 9.4.10. (Hint: you can use Lemma 6.5.3, combined with the squeeze test (Corollary 6.4.14) and Proposition 6.7.3.)*

Let be $a > 0$ a positive real, and we must show that $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := a^x$ is continuous. In other words, we must show that, for any real number x_0, we have:

$$\lim_{x \to x_0} a^x = a^{x_0}, \tag{9.5}$$

By Proposition 9.4.7, one way to show that is to prove that for any sequence $(r_n)_{n=0}^{\infty}$ of *real* numbers that converges to x_0, we have

$$\lim_{n \to \infty} a^{r_n} = a^{x_0}. \tag{9.6}$$

This result is not immediate because we just know similar results for sequences of *rational* numbers (e.g., Definition 6.7.2).

In what follows, as in the proof of Lemma 6.7.1, we should consider three distinct cases: $a > 1$, $a = 1$ and $a < 1$. We will only prove the first one, because the second one is obvious and the third one is very similar.

Thus, let be $(r_n)_{n=0}^{\infty}$ a sequence of *real* numbers that converges to x_0. We must show (9.6).

By Proposition 5.4.14, there exists a *rational* number α_n in the interval $]r_n - 1/n, r_n[$ for all $n > 0$. Similarly, there exists a *rational* number β_n in the interval $]r_n, r_n + 1/n[$ for all $n > 0$. Thus, using the axiom of countable choice, we can build two sequences of rational numbers, $(\alpha_n)_{n=0}^{\infty}$ and $(\beta_n)_{n=0}^{\infty}$, such that we have

$$r_n - 1/n < \alpha_n < r_n < \beta_n < r_n + 1/n \tag{9.7}$$

for all $n > 0$ (there is a tiny problem for $n = 0$, but we could have chosen $1/(n+1)$ instead of $1/n$ to solve it). Also, it is obvious that $(\alpha_n)_{n=0}^{\infty}$, $(r_n)_{n=0}^{\infty}$ and $(\beta_n)_{n=0}^{\infty}$ are all equivalent sequences, and thus they all converge to x_0.

Thus, starting from (9.7) and using Lemma 5.6.9(e) applied to real numbers, we have in particular, since $a > 1$ by hypothesis:

$$a^{\alpha_n} < a^{r_n} < a^{\beta_n} \tag{9.8}$$

for all $n > 0$. But, by Definition 6.7.2, we know that $\lim_{n\to\infty} a^{\alpha_n} = a^{x_0}$ and $\lim_{n\to\infty} a^{\beta_n} = a^{x_0}$. Thus, by the squeeze test, we also have $\lim_{n\to\infty} a^{r_n} = a^{x_0}$, as expected. This closes the proof.

Exercise 9.4.4. — *Prove Proposition 9.4.11. (Hint: from limit laws (Proposition 9.3.14) one can show that* $\lim_{x\to 1} x^n = 1$ *for all integers n. From this and the squeeze test (Corollary 6.4.14) deduce that* $\lim_{x\to 1} x^p = 1$ *for all real numbers p. Finally, apply Proposition 6.7.3.)*

We follow the hint to show that:

1. $\lim_{x\to 1} x^n = 1$ for all integers n. If $n = 0$, the statement is obvious. If n is a positive integer, then a quick induction shows the result: for the base case $n = 1$, we know that $\lim_{x\to 1} x = 1$ by Example 9.4.3 (for instance), and then we get the result for all $n > 0$ since $x^{n+1} = x^n \times x$, using the limit laws. A similar argument applies if $n < 0$, because we know (for instance by the limit laws) that $\lim_{x\to 1} 1/x = 1$, and that $1/x^{n+1} = (1/x) \times (1/x^n)$.

2. (TODO) $\lim_{x\to 1} x^p = 1$ for all real numbers p.

3. (TODO) Finally, we can conclude that the function $f :]0, \infty[\to \mathbb{R}$ defined by $f(x) := x^p$ is continuous. We must show that $\lim_{x\to a} x^p = a^p$ for all $a \in \mathbb{R}$. Let be $\varepsilon > 0$. By the previous result, there exists $\alpha > 0$ such that $|x - 1| < \alpha \Rightarrow |x^p - 1| < \varepsilon/a^p$. Multiplying this last inequality by a^p (which preserves order since $a > 0$), we get:

$$\forall \varepsilon > 0, \exists \alpha > 0 \,:\, |x - 1| < \alpha \Rightarrow |x^p - | < \varepsilon$$

Exercise 9.4.5. — *Prove Proposition 9.4.13.*

To show the result, by Proposition 9.4.7(b), we have to prove that, for any sequence $(a_n)_{n=0}^{\infty}$ of elements of X that converges to x_0, the sequence $(g \circ f(a_n))_{n=0}^{\infty}$ converges to $g \circ f(x_0)$.

Thus, let be $(a_n)_{n=0}^{\infty}$ a sequence of elements of X that converges to x_0. Because f is continuous at x_0, the sequence $(f(a_n))_{n=0}^{\infty}$ converges to $f(x_0)$. And since g is continuous at $f(x_0)$, the sequence $(g(f(a_n)))_{n=0}^{\infty}$ converges to $g(f(x_0))$.

Thus, for any sequence $(a_n)_{n=0}^{\infty}$ of elements of X that converges to x_0, the sequence $(g \circ f(a_n))_{n=0}^{\infty}$ converges to $g \circ f(x_0)$, i.e., by Proposition 9.4.7, $g \circ f$ is continuous at x_0, as expected.

Exercise 9.4.6. — *Let X be a subset of $\mathbb{R}$, and let $f : X \to \mathbb{R}$ be a continuous function. If Y is a subset of X, show that the restriction $f|_Y : Y \to \mathbb{R}$ of f to Y is also a continuous function.*

Let be $x_0 \in X$. Since f is continuous on X, it is continuous in x_0, and then we have:

$$\forall \varepsilon > 0, \exists \delta > 0 \, : \, x \in]x_0 - \delta, x_0 + \delta[\cap X \Longrightarrow |f(x) - f(x_0)| < \varepsilon \tag{9.9}$$

Now let be $y_0 \in Y$. In particular, $y_0 \in X$. Let be the $\delta > 0$ given by equation 9.9. Then for any $y \in]y_0 - \delta, y_0 + \delta[\cap Y$, we also have $y \in]y_0 - \delta, y_0 + \delta[\cap X$, so that we have indeed $|f(y) - f(y_0)| < \varepsilon$ still be equation 9.9. Since y_0 is arbitrary, $f|_Y$ is continuous.

Exercise 9.5.1. — *Let E be a subset of $\mathbb{R}$, let $f : E \to \mathbb{R}$ be a function, and let x_0 be an adherent point of E. Write down a definition of what it would mean for the limit $lim_{x \to x_0; x \in E} f(x)$ to exist and equal $+\infty$ or $-\infty$. If $f : \mathbb{R} - \{0\} \to \mathbb{R}$ is the function $f(x) := 1/x$, use your definition to conclude $f(0+) = +\infty$ and $f(0-) = -\infty$.*

We can state the following definition. We have $\lim_{x \to x_0; x > x_0} f(x) = +\infty$ iff for any positive real $A > 0$, there exists $\delta > 0$ such that $f(x) > A$ for all $x \in]x_0, x_0 + \delta[$. We have $\lim_{x \to x_0; x < x_0} f(x) = +\infty$ iff for any positive real $A > 0$, there exists $\delta > 0$ such that $f(x) > A$ for all $x \in]x_0 - \delta, x_0[$. We have $\lim_{x \to x_0} f(x) = +\infty$ iff $\lim_{x \to x_0; x > x_0} f(x) = +\infty = \lim_{x \to x_0; x < x_0} f(x)$.

(A similar definition can be written for $-\infty$.)

In particular, let's consider the function $f(x) := 1/x$ defined on $\mathbb{R} - \{0\}$. Let be $A > 0$ a positive real number. Then, let be $\delta := 1/(A+1) > 0$. For any $x \in]0, \delta[$, we have $|x - 0| = |x| < 1/(A+1)$, i.e. $x < 1/(A+1)$, and thus $1/x > A + 1 > A$. Thus, $f(0+) = \infty$ according to the above definition.

A similar proof can be given to show that $f(0-) = -\infty$.

Exercise 9.6.1. — *Give examples of: (a) a function $f :]1, 2[\to \mathbb{R}$ which is continuous and bounded, attains its minimum somewhere, but does not attain its maximum anywhere; (b) a function $f : [0, \infty[\to \mathbb{R}$ which is continuous and bounded, attains its maximum somewhere, but does not attain its minimum anywhere; (c) a function $f : [-1, 1] \to \mathbb{R}$ which is bounded but does not attain its minimum anywhere or its maximum anywhere; (d) a function $f : [-1, 1] \to \mathbb{R}$ which has no upper bound and no lower bound. Explain why none of the examples you construct violate the maximum principle.*

We can find the following examples:

(a) $f(x) = |x - 1.5|$ is an example of such a function. It does not contradict the maximum principle since $]1, 2[$ is not closed.

(b) $f(x) = 1/(x+1)$ is an example. It does not contradict the maximum principle since $]0, +\infty[$ is not closed.

(c) $f(x)$ defined by $f(x) = |x|$ if $x \notin \{-1, 0, 1\}$ and $f(x) = 0.5$ if $x \in \{-1, 0, 1\}$ is an example. It does not contradict the maximum principle since it is not continuous.

(d) $f(x)$ defined by $f(x) = 1/x$ if $x \neq 0$ and $f(0) = 0$ is an example. It does not contradict the maximum principle since it is not continuous.

Exercise 9.7.1. — *Prove Corollary 9.7.4.*

Since f is continuous on $[a, b]$, then by the maximum principle (Proposition 9.6.7), there exists $x_{max} \in [a, b]$ such that $f(x_{max}) = M$, and $x_{min} \in [a, b]$ such that $f(x_{min}) = m$.

We can suppose without loss of generality that $x_{min} < x_{max}$; because the case $x_{max} < x_{min}$ can be proved similarly, and there is almost nothing to prove if $x_{min} = x_{max}$ (f would be constant, so picking any $c \in [a, b]$ would be sufficient).

By Exercise 9.4.6, since f is continuous on $[a, b]$, then f is also continuous on $[x_{min}, x_{max}]$. We can thus apply the intermediate value theorem on this interval to get the result: for any $y \in [m, M]$, there exists indeed a $c \in [x_{min}, x_{max}]$ such that $f(c) = y$.

Also, we know that for any $y \in f([a, b])$, we have both $y \leqslant M$ and $y \geqslant m$ by definition, so that $f([a, b]) \subseteq [m, M]$. And by the result we have just shown, we also know that $[m, M] \subseteq f([a, b])$, thus we are done.

Exercise 9.7.2. — *Let $f : [0, 1] \to [0, 1]$ be a continuous function. Show that there exists a real number x in $[0, 1]$ such that $f(x) = x$. This point x is known as a fixed point of f.*

Let be the function $g : [0, 1] \to \mathbb{R}$ defined by $g(x) := f(x) - x$.

- This function g is continuous on $[0, 1]$, since f is continuous by hypothesis, the identity function $x \mapsto x$ is also continuous (Exercice 9.4.2), and the difference of two continuous functions is still continuous (Proposition 9.4.9).
- Furthermore, note that $f(x) = x$ iff $g(x) = 0$. We thus have to prove that there exists a $x \in [0, 1]$ such that $g(x) = 0$.
- We have $g(0) = f(0) \in [0, 1]$; and $g(1) = f(1) - 1 \in [-1, 0]$. Thus, we have in particular $g(1) \leqslant 0 \leqslant g(0)$, and g is continuous. The intermediate value theorem thus says that there exists a $x \in [0, 1]$ such that $g(x) = 0$, as desired.

Exercise 9.8.1. — *Explain why the maximum principle remains true if the hypothesis that f is continuous is replaced with f being monotone, or with f being strictly monotone. (You can use the same explanation for both cases.)*

Let be $f : [a, b] \to \mathbb{R}$ an increasing function. By definition, we have $f(x) \leqslant f(y)$ for all $x, y \in [a, b]$ such that $x \leqslant y$. In particular, since $x \leqslant b$ for all $x \in [a, b]$, we have $f(x) \leqslant f(b)$ for all $x \in [a, b]$. And thus, since $b \in [a, b]$, f attains its maximum at b on $[a, b]$. Similarly, it attains its minimum at a.

The proof is similar if f is supposed to be strictly increasing, or decreasing, or strictly decreasing.

Exercise 9.8.2. — *Give an example to show that the intermediate value theorem becomes false if the hypothesis that f is continuous is replaced with f being monotone, or with f being strictly monotone. (You can use the same counterexample for both cases.)*

Let's simply consider the (strictly increasing function $f : [0, 2] \to \mathbb{R}$ defined by $f(x) = x$ if $x \in [0, 1[$ and $f(x) = x + 1$ if $x \in [1, 2]$. Then, for instance, we have $f(0) = 0$ and $f(2) = 3$, but there is no real $x \in [0, 2]$ such that $f(x) = 3/2$.

Exercise 9.8.3. — *Let $a < b$ be real numbers, and let $f : [a, b] \to \mathbb{R}$ be a function which is both continuous and one-to-one. Show that f is strictly monotone.*

(This one may be more a detailed sketch of a proof rather than an actual rigorous proof.) We divide the proof into three cases:

1. If we suppose $f(a) = f(b)$, then we have immediately a contradiction, since f is not injective. This case is thus simply impossible. (More generally, f cannot be constant on any interval $[c, d] \subseteq [a, b]$, since it would obviously not be injective in such a case. Thus, f can only be strictly increasing or strictly decreasing on such intervals.)

2. If we suppose $f(a) > f(b)$, let's suppose for the sake of contradiction that f is not strictly monotone increasing on $[a, b]$. Actually, f must at least be strictly increasing on a subset of $[a, b]$, since $f(a) > f(b)$ (otherwise we have an obvious contradiction), but this hypothesis means that f is not strictly increasing on the whole interval $[a, b]$. Thus, there must exist three real numbers c, d, e, such that $a \leqslant c < d < e \leqslant b$, and we have yet another time several possible cases:

   - f is strictly increasing on $[c, d]$ and strictly decreasing on $[d, e]$, i.e., we have $f(c) < f(d)$ and $f(d) > f(e)$. One can show that[30] $m := \frac{f(c)+f(e)}{2}$ belongs to both intervals $[f(c), f(d)]$ and $[f(d), f(e)]$. Thus, applying the intermediate value theorem on each interval $[c, d]$ and $[d, e]$ shows that there exists $x_1 \in [c, d]$, $x_2 \in [d, e]$ such that $f(x_1) = f(x_2) = m$, although $x_1 \neq x_2$. (The only possibility for x_1 and x_2 to be equal to each other would be that $x_1 = x_2 = d$, which would imply that f is constant on both $[c, d]$ and $[d, e]$, a contradiction.) Thus f would not be injective, a contradiction.
   - A similar argument applies if f is strictly decreasing on $[c, d]$ and then increasing on $[d, e]$.

3. If we suppose $f(a) < f(b)$, a direct adaptation of the second case shows immediately that f must be strictly monotone decreasing.

Exercise 9.9.1. — *Prove Lemma 9.9.7.*

We have to prove that two sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent if and only if $\lim_{n\to\infty}(a_n - b_n) = 0$.

- If these two sequences are equivalent (in the sense of Definition 9.9.5), then let be $\varepsilon > 0$. By definition, there exists $N \geqslant 1$ such that, for all $n \geqslant N$, we have $|a_n - b_n| \leqslant \varepsilon$, i.e., $|(a_n - b_n) - 0| \leqslant \varepsilon$. This means precisely that $\lim_{n\to\infty}(a_n - b_n) = 0$.

[30] Actually, this is yet another division into cases, depending on whether $f(c) > f(e)$ or $f(c) \leqslant f(e)$, but this is easily shown, and even more easily understandable by drawing a picture.

- All these steps can be reversed to show that if $\lim_{n\to\infty}(a_n - b_n) = 0$, then $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent.

Exercise 9.9.2. — *Prove Proposition 9.9.8.*

Recall the two statements:

(a) f is uniformly continuous on X

(b) For any two equivalent sequences $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ of elements of X, the sequences $(f(x_n))_{n=1}^{\infty}$ and $(f(y_n))_{n=1}^{\infty}$ are also equivalent.

Let's prove each implication separately.

- First, we show that (a) implies (b). Let be $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ two equivalent sequences of elements of X. Let be $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta > 0$ such that $|f(x) - f(y)| \leqslant \varepsilon$ whenever $|x - y| \leqslant \delta$ (for $x, y \in X$).

  But since $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ are equivalent, there exists $N \geqslant 1$ such that $n \geqslant N$ implies $|x_n - y_n| \leqslant \delta$.

  Thus, if $n \geqslant N$, we have $|x_n - y_n| \leqslant \delta$, and thus $|f(x_n) - f(y_n)| \leqslant \varepsilon$. This means that $(f(x_n))_{n=1}^{\infty}$ and $(f(y_n))_{n=1}^{\infty}$ are equivalent.

- Now we show that (b) implies (a). We will use contradiction here[31]. Thus, we suppose that (b) is true, but that f is not uniformly continuous. For f, not being uniformly continuous means that:

$$\exists \varepsilon > 0, \exists y_0 \in X \,:\, \forall \delta \leqslant 0,\ |f(x) - f(y_0)| \geqslant \varepsilon,\ \text{although } |x - y_0| < \delta \tag{9.10}$$

  If we consider a sequence $(x_n)_{n=1}^{\infty}$ of elements of X that converges to y_0, then the sequences $(x_n)_{n=1}^{\infty}$ and $(y_0)_{n=1}^{\infty}$ are equivalent. This is thus a clear contradiction with equation (9.10).

Exercise 9.9.3. — *Prove Proposition 9.9.12.*

We must prove that uniformly continuous functions map Cauchy sequences to Cauchy sequences. Let be an arbitrary $\varepsilon > 0$. First, note that if $f : X \to \mathbb{R}$ is uniformly continuous, then we have

$$\exists \delta > 0 \,:\, (x, y \in I, |x - y| < \delta) \Longrightarrow |f(x) - f(y)| < \varepsilon \tag{9.11}$$

But since $(x_n)_{n=0}^{\infty}$ is supposed to be a Cauchy sequence, then we have

$$\exists N \geqslant 0 \,:\, j, k \geqslant N \Longrightarrow |x_j - x_k| < \delta \tag{9.12}$$

Thus combining equations (9.11) and (9.12), if we suppose that $j, k \geqslant N$, then we have $|x_j - x_k| < \delta$, and thus $|f(x_j) - f(x_k)| < \varepsilon$. Since this is true for any arbitrary $\varepsilon > 0$, this shows that $(f(x_n))_{n=0}^{\infty}$ is a Cauchy sequence, as required.

[31] Note that we use the same approach as in Exercise 9.3.1.

Exercise 9.9.4. — *Use Proposition 9.9.12 to prove Corollary 9.9.14.*

First, recall that the limit $\lim_{x\to x_0;x\in X} f(x)$ exists and is equal to a real number L iff the sequence $(f(a_n))_{n=0}^{\infty}$ converges to L for *any* sequence $(a_n)_{n=0}^{\infty}$ that converges to x_0 (Proposition 9.3.9).

Since x_0 is an adherent point of X, there exists a sequence $(a_n)_{n=0}^{\infty}$ entirely composed of elements on X that converges to x_0 (Lemma 9.1.14). Since it is a convergent sequence, it is also a Cauchy sequence, because $\mathbb{R}$ is complete (Theorem 6.4.18). Thus, by Proposition 9.9.12, $(f(a_n))_{n=0}^{\infty}$ is a Cauchy sequence; and by Theorem 6.4.18 again, it is thus a convergent sequence. Say that it converges to L.

An important remark: we cannot say yet that the proof is over, and that we have $\lim_{x\to x_0;x\in X} f(x) = L$! We need one last step: let's consider $(b_n)_{n=0}^{\infty}$ *any other* sequence that converges to x_0. Thus, by Lemma 9.9.1, $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are equivalent; and by Proposition 9.9.8, $(f(a_n))_{n=0}^{\infty}$ and $(f(b_n))_{n=0}^{\infty}$ are equivalent. Thus, $(f(b_n))_{n=0}^{\infty}$ also converges to L. This, time, is shows indeed that $\lim_{x\to x_0;x\in X} f(x) = L$, as expected.

Exercise 9.9.5. — *Prove Proposition 9.9.15.*

Suppose for the sake of contradiction that the set $f(E)$ is not bounded. Thus, for all real number $M \geqslant 0$, there exists $x \in E$ such that $|f(x)| \geqslant M$. Since this is true of all real number, it is in particular true for any *natural number* $n \geqslant 0$, which implies that the set

$$X_n := \{x \in E \,:\, |f(x)| \geqslant n\}$$

is non-empty for all natural number $n \geqslant 0$. By the axiom of choice, we can construct a sequence $(x_n)_{n=0}^{\infty}$ such that $x_n \in X_n$ for all $n \geqslant 0$ (and in particular, $x_n \in E$).

Since E is bounded by hypothesis, the sequence $(x_n)_{n=0}^{\infty}$ is bounded. By the Bolzano-Weierstrass theorem (Theorem 6.6.8), there exists a subsequence $(x_{n_j})_{j=0}^{\infty}$ of $(x_n)_{n=0}^{\infty}$ that converges to some real number L (note that we have $n_j \geqslant j$ for all j, as a quick induction shows). By Lemma 9.1.14, L is thus an adherent point of E. By Corollary 9.9.14, the limit $\lim_{x\to L;x\in E} f(x)$ exists and is finite.

In particular, it implies that $\lim_{j\to\infty} f(x_{n_j})$ exists and is finite. In other words, $(f(x_{n_j}))_{j=0}^{\infty}$ is a convergent sequence, and in particular, is a bounded sequence. But this contradicts the definition of the x_{n_j}, since we have $|f(x_{n_j})| \geqslant n_j \geqslant j$ for all j, by construction.

10. Differentiability

Exercise 10.1.2. — *Prove Proposition 10.1.7.*

We have to prove the principle of Newton's approximation, i.e. that (a) f is differentiable at x_0 iff (b) we have a linear approximation of f at x_0:

$$\forall \varepsilon > 0, \exists \delta > 0 \,:\, |x - x_0| \leqslant \delta \Longrightarrow |f(x) - (f(x_0) + L(x - x_0))| \leqslant \varepsilon |x - x_0| \tag{10.1}$$

- First consider the case $x \neq x_0$. We know that the following statements are logically equivalent:

$$\begin{aligned} & f \text{ is differentiable at } x_0 \in X \\ \Longleftrightarrow & \lim_{x \to x_0; x \in X - \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = L \\ \Longleftrightarrow & \forall \varepsilon > 0, \exists \delta > 0 \,:\, |x - x_0| \leqslant \delta \Longrightarrow \left| \frac{f(x) - f(x_0)}{x - x_0} - L \right| \leqslant \varepsilon \\ \Longleftrightarrow & \forall \varepsilon > 0, \exists \delta > 0 \,:\, |x - x_0| \leqslant \delta \Longrightarrow |f(x) - f(x_0) - L(x - x_0)| \leqslant \varepsilon |x - x_0| \end{aligned}$$

- If $x = x_0$, there is almost nothing to prove, since all the terms implied here are null.

This closes the proof in both cases.

Exercise 10.1.3. — *Prove Proposition 10.1.10.*

We have to show that differentiability implies continuity. Let be $f : X \to \mathbb{R}$ a function which is continuous at $x_0 \in X$. It means that we have

$$\lim_{x \to x_0; x \in X - \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = L \tag{10.2}$$

where L is a real number. Let be $\varepsilon > 0$. We have to prove that there exists a $\delta > 0$ such that $|x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \varepsilon$.

Since f is supposed to be differentiable at x_0, by (10.2) we know that there exists a $\alpha > 0$ such that $|x - x_0| < \alpha \Longrightarrow \left| \frac{f(x) - f(x_0)}{x - x_0} - L \right| < \varepsilon$.

In other words (this is Proposition 10.1.7),

$$\exists \alpha > 0, \, |x - x_0| < \alpha \Longrightarrow |f(x) - f(x_0) - L(x - x_0)| < \varepsilon |x - x_0|. \tag{10.3}$$

Thus, by triangular inequality, we have

$$\exists \alpha > 0, \, |x - x_0| < \alpha \Longrightarrow |f(x) - f(x_0)| < |x - x_0|(|L| + \varepsilon) \tag{10.4}$$

Now, let be $\delta := \min(\alpha, \frac{\varepsilon}{|L| + \varepsilon})$. If $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| < \varepsilon$ as expected. This closes the proof.

Exercise 10.1.4. — *Prove Theorem 10.1.13.*

Let's prove the different statements of this theorem.

(a) If $f(x) = c$ for all $x \in \mathbb{R}$, then at any point $x_0 \in \mathbb{R}$, the quotient $\frac{f(x)-f(x_0)}{x-x_0}$ is equal to 0. Thus, we have in particular $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{x \to x_0} 0 = 0$. This means that $f'(x_0) = 0$ for all $x_0 \in \mathbb{R}$, as expected.

(b) Similarly, if $f(x) = x$ for all $x \in \mathbb{R}$, then at any point $x_0 \in \mathbb{R}$, the quotient $\frac{f(x)-f(x_0)}{x-x_0}$ is equal to 1. Thus, the same argument leads to $f'(x_0) = 1$ for all $x_0 \in \mathbb{R}$.

(c) We have:

$$\begin{aligned}\frac{(f+g)(x)-(f+g)(x_0)}{x-x_0} &= \frac{f(x)+g(x)-f(x_0)-g(x_0)}{x-x_0}\\ &= \frac{f(x)-f(x_0)}{x-x_0} + \frac{g(x)-g(x_0)}{x-x_0}\end{aligned}$$

and thus, in particular, since both terms of the right-hand side converge by hypothesis, we have by Proposition 9.3.14:

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

as expected.

(d) We have:

$$\begin{aligned}\frac{(fg)(x)-(fg)(x_0)}{x-x_0} &= \frac{f(x)g(x)-f(x_0)g(x_0)}{x-x_0}\\ &= \frac{f(x)g(x)-f(x_0)g(x)+f(x_0)g(x)-f(x_0)g(x_0)}{x-x_0}\\ &= g(x)\frac{f(x)-f(x_0)}{x-x_0} + f(x_0)\frac{g(x)-g(x_0)}{x-x_0}\end{aligned}$$

and since f, g are differentiable (and thus continuous) by hypothesis, we have by Proposition 9.3.14

$$\begin{aligned}\lim_{x \to x_0} \frac{(fg)(x)-(fg)(x_0)}{x-x_0} &= \lim_{x \to x_0} \left(g(x)\frac{f(x)-f(x_0)}{x-x_0} + f(x_0)\frac{g(x)-g(x_0)}{x-x_0} \right)\\ &= \lim_{x \to x_0} \left(g(x)\frac{f(x)-f(x_0)}{x-x_0} \right) + \left(f(x_0)\frac{g(x)-g(x_0)}{x-x_0} \right)\\ &= g(x_0)f'(x_0) + f(x_0) + g'(x_0)\end{aligned}$$

as expected.

(e) Obvious using limit laws and/or Proposition 9.3.14.

(f) Obvious by combining (c) and (e), setting $c := -1$.

(g) We have:

$$\frac{1/g(x)-1/g(x_0)}{x-x_0} = \frac{g(x_0)-g(x)}{g(x)g(x_0)(x-x_0)} = \frac{1}{g(x)g(x_0)} \times \frac{g(x_0)-g(x)}{x-x_0}$$

Since g is differentiable (and thus continuous), we have $\lim_{x\to x_0} \frac{1}{g(x)g(x_0)} = \frac{1}{g^2(x_0)}$; and we also have $\lim_{x\to x_0} \frac{g(x)-g(x_0)}{x-x_0} = g'(x_0)$. Thus, since both terms converge, we can apply Proposition 9.3.14 to get:

$$\begin{aligned}
\lim_{x\to x_0} \frac{1/g(x) - 1/g(x_0)}{x - x_0} &= \lim_{x\to x_0} \left(\frac{1}{g(x)g(x_0)} \times \frac{g(x_0) - g(x)}{x - x_0} \right) \\
&= \lim_{x\to x_0} \frac{1}{g(x)g(x_0)} \times \lim_{x\to x_0} \frac{g(x_0) - g(x)}{x - x_0} \\
&= \frac{-g'(x_0)}{g^2(x_0)}
\end{aligned}$$

as expected.

(h) Obvious by combining (g) and (d).

Exercise 10.1.5. — *Let n be a natural number, and let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) := x^n$. Show that f is differentiable on $\mathbb{R}$ and $f'(x) = nx^{n-1}$ for all $x \in \mathbb{R}$.*

Let's use induction on n.

- The case $n = 0$ corresponds to the constant function $f(x) = 1$; see Exercise 10.1.4(a).
- We really start at $n = 1$; in this case we have $f(x) = x$. We have shown in Exercise 10.1.4(b) that $f'(x) = 1 = 1 \times x^0$, as expected.
- Now suppose inductively that the property is true for some natural number $n \in \mathbb{N}$, and let's show that it is still true for $n + 1$. If we set $f(x) := x^{n+1}$, we can write $f(x) = x \times x^n$. By the induction hypothesis, $x \to x^n$ is differentiable on $\mathbb{R}$; and we know that the identity function $x \to x$ is also differentiable on $\mathbb{R}$. Thus f is differentiable on $\mathbb{R}$, and we can use the product rule:

  $$f'(x) = 1 \times x^n + x \times nx^{n-1} = (n + 1)x^n$$

  as expected. This closes the proof.

Exercise 10.1.6. — *Let n be a negative integer, and let $f : \mathbb{R} - \{0\} \to \mathbb{R}$ be the function $f(x) := x^n$. Show that f is differentiable on $\mathbb{R}$ and $f'(x) = nx^{n-1}$ for all $x \in \mathbb{R} - \{0\}$.*

Here, we suppose that $n < 0$. Thus, if we set $m := -n$, then m is a positive integer, and we have $f(x) = 1/x^m$ by Definition 5.6.2.

We know that $x \to x^m$ is differentiable on $\mathbb{R}$, and thus on $\mathbb{R} - \{0\}$. By Theorem 10.1.13(g), the function $x \to 1/x^m$ is differentiable on $\mathbb{R} - \{0\}$ and we have:

$$f'(x) = (1/x^m)' = -\frac{mx^{m-1}}{x^{2m}} = -mx^{-m-1} = nx^{n-1},$$

(using also Exercise 10.1.5,) as expected.

Exercise 10.1.7. — *Prove Theorem 10.1.5.*

This theorem defines the *chain rule* for functions of a real variable. Because each hypothesis in this theorem is really crucial, let's recall them:

Theorem (Chain rule). *Let X, Y be subsets of $\mathbb{R}$, let $x_0 \in X$ be a limit point of X, and let $y_0 \in Y$ be a limit point of Y. Let $f : X \to Y$ be a function such that $f(x_0) = y_0$, and such that f is differentiable at x_0. Suppose that $g : Y \to \mathbb{R}$ is a function which is differentiable at y_0. Then the function $g \circ f : X \to \mathbb{R}$ is differentiable at x_0, and $(g \circ f)'(x_0) = g'(y_0)f'(x_0)$.*

We know that f is differentiable at x_0, so that if we set $L_1 := f'(x_0)$ we have the following Newton's approximation (Proposition 10.1.7):

$$\forall \varepsilon > 0, \exists \delta > 0 \ : \ |x - x_0| \leqslant \delta \Longrightarrow |f(x) - f(x_0) - L_1(x - x_0)| \leqslant \varepsilon |x - x_0| \tag{10.5}$$

and, using the triangular inequality, this also implies that

$$\forall \varepsilon > 0, \exists \delta > 0 \ : \ |x - x_0| \leqslant \delta \Longrightarrow |f(x) - f(x_0)| \leqslant |L_1| \times |x - x_0| + \varepsilon |x - x_0|. \tag{10.6}$$

In particular, since f is differentiable at x_0, it is also continuous at x_0 (Proposition 10.1.10), thus we have

$$\forall \varepsilon > 0, \exists \delta > 0 \ : \ |x - x_0| \leqslant \delta \Longrightarrow |f(x) - f(x_0)| \leqslant \varepsilon. \tag{10.7}$$

Finally, we know that g is differentiable at $y_0 := f(x_0)$, so that if we set $L_2 := g'(y_0)$, we have the following Newton's approximation:

$$\forall \varepsilon > 0, \exists \delta > 0 \ : \ |y - y_0| \leqslant \delta \Longrightarrow |g(y) - g(y_0) - L_2(y - y_0)| \leqslant \varepsilon |y - y_0|. \tag{10.8}$$

By combining these three facts, we want to show that we have the following Newton's approximation for $g \circ f$:

$$\forall \varepsilon > 0, \exists \delta > 0 \ : \ |x - x_0| \leqslant \delta \Longrightarrow |g \circ f(x) - g \circ f(x_0) - L_2 L_1 (x - x_0)| \leqslant \varepsilon |x - x_0|. \tag{10.9}$$

Thus, let be $\varepsilon > 0$ a positive real number, and let's define another positive real number ε' such that $\varepsilon \leqslant \varepsilon'(\varepsilon' + |L_1| + |L_2|)$ — the rationale for introducing this quantity will be clearer a bit later in the proof. [32]

By equation (10.8), for this choice of $\varepsilon' > 0$, there exists a $\delta_1 > 0$ such that we have $|g(y) - g(y_0) - L_2(y - y_0)| \leqslant \varepsilon' |y - y_0|$ whenever ($y \in Y$ and) $|y - y_0| \leqslant \delta_1$.

In this result, consider the part $|y - y_0|$. Actually, we know that if we set $\varepsilon := \delta_1$ in equation (10.7), then there exists a real number $\delta_2 > 0$ such that $|f(x) - f(x_0)| \leqslant \delta_1$ whenever ($x \in X$ and) $|x - x_0| \leqslant \delta_2$.

Now, still for our initial choice of an $\varepsilon' > 0$, we know by equation (10.5) that there exists a $\delta_3 > 0$ such that $|f(x) - f(x_0) - L_1(x - x_0)| \leqslant \varepsilon' |x - x_0|$ whenever ($x \in X$ and) $|x - x_0| \leqslant \delta_3$.

Now we set $\delta := \min(\delta_2, \delta_3)$, and we unwrap all these facts. If $|x - x_0| \leqslant \delta$, then we have $|f(x) - f(x_0)| \leqslant \delta_1$, and thus,

$$|g(f(x)) - g(f(x_0)) - L_2(f(x) - f(x_0))| \leqslant \varepsilon' |f(x) - f(x_0)| \tag{10.10}$$

This is still quite far from what we want, i.e. equation (10.9). In particular, we still don't have the $L_1 L_2 (x - x_0)$ quantity that we want. One way to overcome this is to force this

[32] The existence of this ε' could be proved by more advanced knowledge in algebra, but it may be proved in another (painful but more elementary) way. Indeed, the number $(1 + |L_1| + |L_2|)$ is positive, and so is ε. Thus, by the Archimedean principle (Corollary 5.4.13), there exists a positive integer M such that $M\varepsilon > (1 + |L_1| + |L_2|)$, i.e. such that $\varepsilon > (\frac{1}{M} + \frac{1}{M}|L_1| + \frac{1}{M}|L_2|)$. And since we have $\frac{1}{M} \geqslant \frac{1}{M^2}$ for any positive integer M, then we have $\varepsilon > (\frac{1}{M^2} + \frac{1}{M}|L_1| + \frac{1}{M}|L_2|)$. Let's take $\varepsilon' = \frac{1}{M}$ and we are done.

quantity into the previous equation, i.e. to apply the "middle-man trick" (see hint of Exercise 10.1.4), and then to apply the triangular inequality:

$$\begin{aligned}
&|g(f(x)) - g(f(x_0)) - L_2(f(x) - f(x_0))| \\
&= |g(f(x)) - g(f(x_0)) - L_2(f(x) - f(x_0) - L_1(x - x_0)) - L_2L_1(x - x_0)| \\
&\geqslant |g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| - |L_2(f(x) - f(x_0) - L_1(x - x_0))|
\end{aligned}$$

We are now much closer to equation (10.9)! We can now combine all the inequalities to say that, if $|x - x_0| \leqslant \delta$, then we have:

$$\begin{aligned}
&|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| - |L_2(f(x) - f(x_0) - L_1(x - x_0))| \leqslant \varepsilon'|f(x) - f(x_0)| \\
\Longrightarrow&|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| \leqslant |L_2(f(x) - f(x_0) - L_1(x - x_0))| + \varepsilon'|f(x) - f(x_0)| \\
\Longrightarrow&|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| \leqslant |L_2| \times |f(x) - f(x_0) - L_1(x - x_0)| + \varepsilon'|f(x) - f(x_0)| \\
\Longrightarrow&|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| \leqslant |L_2|\varepsilon'|x - x_0| + \varepsilon'|f(x) - f(x_0)| \text{ (cf. eq. (10.5))} \\
\Longrightarrow&|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| \leqslant |L_2|\varepsilon'|x - x_0| + \varepsilon'(|L_1||x - x_0| + \varepsilon'|x - x_0|) \text{ (cf. eq. (10.6))} \\
\Longrightarrow&|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| \leqslant \varepsilon'(\varepsilon' + |L_1| + |L_2|)|x - x_0|
\end{aligned}$$

Thus, by definition of ε', we finally conclude that for our initial arbitrary $\varepsilon > 0$, there exists a $\delta > 0$ such that, whenever $x \in X$ and $|x - x_0| \leqslant \delta$, we have:

$$|g(f(x)) - g(f(x_0)) - L_2L_1(x - x_0)| \leqslant \varepsilon|x - x_0|$$

We have thus shown that $g \circ f$ has a Newton's approximation at x_0, i.e. that $g \circ f$ is differentiable at x_0 and that $(g \circ f)'(x_0) = L_2L_1 = g'(f(x_0))f'(x_0)$, as expected.

Exercise 10.2.1. — *Prove Proposition 10.2.6.*

We will prove the proposition for the case of a local maximum; the proof is similar for a local minimum. First, if f attains a local maximum at x_0, there exists a $\delta > 0$ such that:

$$\forall x \in]a, b[\cap]x_0 - \delta, x_0 + \delta[, \ f(x_0) \geqslant f(x)$$

Now, recall f is differentiable at x_0, so that we have $f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$. Since this limit does exist, the left limit and the right limit must match (if they don't, a quick proof by contradiction would show that f is not differentiable at x_0); i.e. we must have:

$$\lim_{x \to x_0;\, x \in]x_0 - \delta, x_0[} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0;\, x \in]x_0, x_0 + \delta[} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

In both cases, we have $f(x) \leqslant f(x_0)$ since x_0 is a local maximum, so that the numerator is negative. But in the case of the left limit, we have $x - x_0 \leqslant 0$, so that

$$f'(x_0) = \lim_{x \to x_0;\, x \in]x_0 - \delta, x_0[} \frac{f(x) - f(x_0)}{x - x_0} \geqslant 0$$

(taking the limit of a positive quantity leads to a positive limit), and similarly, for the right limit, we have

$$f'(x_0) = \lim_{x \to x_0;\, x \in]x_0, x_0 + \delta[} \frac{f(x) - f(x_0)}{x - x_0} \leqslant 0$$

But since left and right limits are equal, the only possibility is that they are both equal to zero, so that the limit $f'(x_0)$ itself is equal to zero. Thus, we have indeed $f'(x_0) = 0$, as expected.

Exercise 10.2.2. — *Give an example of a function $f :]-1, 1[\to \mathbb{R}$ which is continuous and attains a global maximum at 0, but which is not differentiable at 0. Explain why this does not contradict Proposition 10.2.6.*

The function f defined by $f(x) = x + 1$ if $x \in]-1, 0]$ and $f(x) = -x + 1$ if $x \in]0, 1[$ attains a global maximum of 1 at $x = 0$, but is not differentiable at 0 (since the left limit $\lim_{x\to 0^-} \frac{f(x)-f(0)}{x-0}$ is equal to 1 whereas the right limit $\lim_{x\to 0^+} \frac{f(x)-f(0)}{x-0}$ is equal to -1).

This does not contradict Proposition 10.2.6 because this proposition only deals with functions that are supposed to be differentiable.

Exercise 10.2.3. — *Give an example of a function $f :]-1, 1[\to \mathbb{R}$ which is differentiable, and whose derivative equals 0 at 0, but such that 0 is neither a local minimum nor a local maximum. Explain why this does not contradict Proposition 10.2.6.*

The function f defined by $f(x) := x^3$ is an example of such a function. Indeed, by Exercise 10.1.5, f is differentiable and we have $f'(x) = 3x^2$; so that $f'(0) = 0$. However, f is increasing on $]0, 1[$ (see Lemma 5.6.9(d)); and f is also increasing on $]-1, 0]$ because if we have $x < y \leqslant 0$, then we have $x^2 > y^2 \geqslant 0$, and thus $x^3 < y^3 \leqslant 0$. Thus, f is increasing on the whole interval $]-1, 1[$, so that 0 cannot be a local extremum.

This does not contradict Proposition 10.2.6 because this proposition only says that *if f attains a local extremum at x_0, then $f'(x_0) = 0$*; it does not say that "if $f'(x_0) = 0$, then f attains a local extremum at x_0".

Exercise 10.2.4. — *Prove Theorem 10.2.7.*

We suppose that $g : [a, b] \to \mathbb{R}$ is a continuous function, differentiable on $]a, b[$, and such that $g(a) = g(b)$. We must show that there exists an $x \in]a, b[$ such that $g'(x) = 0$.

Intuitively, if $g(a) = g(b)$, either f is totally "flat" (i.e., constant), or f is not constant and must be increasing and then decreasing (or decreasing and then increasing) somewhere to go back back to its starting point. Let's consider those cases separately.

- If g is constant, then $g'(x) = 0$ for all $x \in]a, b[$ (see Theorem 10.1.13(a)). Thus the claim is trivial: any $x \in]a, b[$ is a good pick.

- If g is not constant on $[a, b]$, then there exists a $x_0 \in]a, b[$ such that $x_0 \neq g(a)$ (and also $x_0 \neq g(b)$). Since g is continuous on $[a, b]$, then by the maximum principle (Proposition 9.6.7), there exists a $x_{max} \in [a, b]$ and a $x_{min} \in [a, b]$ that are respectively the global maximum and minimum of g on $[a, b]$. Once again, there are several (sub-)cases:

  - If $x_{max} \in]a, b[$, then by Proposition 10.2.6, we have $g'(x_{max}) = 0$: the claim follows.
  - If $x_{max} = a$, then g attains a local maximum at $x = a$ (and thus also at $x = b$). Since g is supposed to be non-constant, we have $x_{max} \neq x_{min}$. In particular, we thus have $x_{min} \neq a$ and $x_{min} \neq b$, i.e. $x_{min} \in]a, b[$. Using Proposition 10.2.6, the claim follows.

Exercise 10.2.5. — *Use Theorem 10.2.7 to prove Corollary 10.2.9. (Hint: consider a function of the form $f(x) - cx$ for some carefully chosen real number c.)*

Our goal here is to choose the number c so that the function $g : [a, b] \to \mathbb{R}$ defined by $g(x) := f(x) - cx$ will obey to the requirements of Rolle's theorem. Of course, since f is continuous on $[a, b]$ and differentiable on $]a, b[$, g is also continuous and differentiable on the same intervals. Thus, we only have to choose c so that we have $g(a) = g(b)$.

We have respectively $g(a) = f(a) - ca$, and $g(b) = f(b) - c(b)$, so that $g(a) = g(b)$ iff $c = (f(b) - f(a))/(b - a)$. Thus, let's define the function g by $g(x) := f(x) - \frac{f(b)-f(a)}{b-a}x$. We can apply Rolle's theorem to this function g, i.e., we know that there exists $x_0 \in]a, b[$ such that $g'(x_0) = 0$.

But we have $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$. Thus, for this number x_0, we have actually $f'(x_0) = \frac{f(b)-f(a)}{b-a}$, as expected.

Exercise 10.2.6. — *Let $M > 0$, and let $f : [a, b] \to \mathbb{R}$ be a function which is continuous on $[a, b]$ and differentiable on $]a, b[$, and such that $|f'(x)| \leqslant M$ for all $x \in]a, b[$. Show that for any $x, y \in [a, b]$, we have the inequality $|f(x) - f(y)| \leqslant M|x - y|$.*

The proposition is obvious when $x = y$ since both sides of the inequality are equal to 0; so that we can suppose that we have $a \leqslant x < y \leqslant b$ (if $y > x$, we just have to switch their roles in the arguments below).

Since f is continuous on $[a, b]$ and differentiable on $]a, b[$, then in particular its restriction $f|_{[x,y]}$ is continuous on $[x, y]$ and differentiable on $]x, y[$. We can thus apply the mean value theorem to this restriction $f|_{[x,y]}$; it tells us that there exists a real number $c \in]x, y[$ such that $f'(c) = \frac{f(x)-f(y)}{x-y}$. But since f' is bounded by M, we have $|f'(c)| \leqslant M$, i.e. $|f(x) - f(y)| \leqslant M|x - y|$, as expected.

Exercise 10.2.7. — *Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that f' is bounded. Show that f is uniformly continuous.*

First, saying that f' is bounded means that there exists a real number $M > 0$ such that $|f'(x)| \leqslant M$ for all $x \in \mathbb{R}$. Since $f'$is bounded, then from the previous exercise (10.2.6) we know that we have $|f(x) - f(y)| \leqslant M|x - y|$ for all $x, y \in \mathbb{R}$.

Now let be $\varepsilon > 0$. If we set $\delta := \varepsilon/M$, then for all $x, y \in \mathbb{R}$ such that $|x - y| \leqslant \delta$, we have

$$|f(x) - f(y)| \leqslant M|x - y| \leqslant \varepsilon$$

so that f is uniformly continuous, as expected.

Exercise 10.3.1. — *Prove Proposition 10.3.1.*

Let $f : X \to \mathbb{R}$ be a function which is differentiable at $x_0 \in X$. First suppose that f is monotone increasing, and let's show that we have $f'(x_0) \geqslant 0$.

Since f is differentiable at x_0, then we have

$$\lim_{x \to x_0;\, x \in X - \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

But f is monotone increasing, so that $\frac{f(x)-f(x_0)}{x-x_0}$ is always a non-negative quantity (if $x < x_0$ then $f(x) \leqslant f(x_0)$; and if $x > x_0$ then $f(x) \geqslant f(x_0)$). Thus, since the limit of a non-negative quantity cannot be negative, we have $f'(x_0) \geqslant 0$, as expected.

A similar argument show that if f is monotone decreasing, then $f'(x_0) \leqslant 0$.

Exercise 10.3.2. — *Give an example of a function $f :]-1,1[\to \mathbb{R}$ which is continuous and monotone increasing, but which is not differentiable at 0. Explain why this does not contradict Proposition 10.3.1.*

Consider the function $f :]-1,1[\to \mathbb{R}$ defined by $f(x) = x$ if $x \in]-1,0]$ and $f(x) = 2x$ if $x \in]0,1[$. This function f is continuous at 0, because its left and right limits at 0 are both null:

$$f(0-) = \lim_{x \to 0;\, x \in]-1,0[} f(x) = \lim_{x \to 0;\, x \in]-1,0[} x = 0$$
$$f(0+) = \lim_{x \to 0;\, x \in]0,1[} f(x) = \lim_{x \to 0;\, x \in]0,1[} 2x = 0$$

and since $f(0) = 0$, this shows that f is continuous at 0. However, f is not differentiable at 0, because

$$\lim_{x \to 0;\, x \in]-1,0[} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0;\, x \in]-1,0[} \frac{x}{x} = 1$$

whereas

$$\lim_{x \to 0;\, x \in]0,1[} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0;\, x \in]0,1[} \frac{2x}{x} = 2.$$

Thus, the right and left limits do not match, i.e., the limit $\lim_{x \to 0;\, x \in]-1,1[- \{0\}} \frac{f(x)-f(0)}{x-0}$ is not defined, and f is not differentiable at 0.

This example does not contradict Proposition 10.3.1, because it is only applicable when f is differentiable at a given x_0, which is not the case here (see also Remark 10.3.2).

Exercise 10.3.3. — *Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ which is strictly monotone increasing and differentiable, but whose derivative at 0 is zero. Explain why this does not contradict Proposition 10.3.1 or Proposition 10.3.3.*

There is a canonical example for that, given by $f(x) := x^3$. It is clearly strictly increasing, and differentiable at 0, with $f'(0) = 3 \times 0^2 = 0$ (see also Exercise 10.1.6). This is an example of an *inflection point*; and this does not contradict Proposition 10.3.1 (which allows $f'(x_0)$ to be zero when the function is increasing) nor Proposition 10.3.3 (which states a converse statement).

Exercise 10.3.4. — *Prove Proposition 10.3.3.*

Let $f : [a,b] \to \mathbb{R}$ be a differentiable function. We will prove only the first statement, i.e. that if $f'(x) > 0$ for all $x \in [a,b]$, then f is strictly monotone increasing. After that, the two other statements can be easily proven in a similar fashion.

Let be $x, y \in [a,b]$ such that $x < y$. To show that f is strictly increasing, we must show that we have $f(x) < f(y)$. Since f is supposed to be differentiable, it is in particular differentiable on $]x,y[$. We can thus apply the mean value theorem: there exists a real number $c \in]x,y[$ such that $f'(c) = \frac{f(y)-f(x)}{y-x}$. But since f' is positive, we have $\frac{f(y)-f(x)}{y-x} > 0$. Since the denominator is positive, the numerator must also be positive, so that we get $f(y) > f(x)$, as expected.

EXERCISE 10.3.5. — *Give an example of a subset $X \in \mathbb{R}$ and a function $f : X \to \mathbb{R}$ which is differentiable on X, is such that $f'(x) > 0$ for all $x \in X$, but f is not strictly monotone increasing. (Hint: the conditions here are subtly different from those in Proposition 10.3.3. What is the difference, and how can one exploit that difference to obtain the example?)*

The main difference with Proposition 10.3.3 is that, here, we do not consider a function f defined on an *interval*, but on an arbitrary subset of the real line. In particular, we can choose X as a subset that is not connected (see Definition 11.1.1). For instance, let's consider $X := [1, 2] \cup [3, 4]$ such that $f(x) := x$ when $x \in [1, 2]$ and $f(x) := x - 3$ when $x \in [3, 4]$. f is continuous and differentiable; and we have $f'(x) = 1 > 0$ for all $x \in X$. However, f is not strictly increasing, since we have $f(2) = 2$ and $f(3) = 0$.

11. The Riemann integral

Exercise 11.1.1. — *Prove Lemma 11.1.4.*

Here, proving that (b) implies (a) is slightly easier than the converse implication. We will thus begin by this implication.

- If X is a bounded interval, then there exist $a \leqslant b$ such that X is an interval of the form (a, b), where the parentheses stand for either opening or closing brackets. In either of these four cases, we clearly have $X \subset [-M, M]$ if we set $M := \max(|a|, |b|) + 1$, so that X is bounded (Definition 9.1.22).

  Furthermore, X is connected by definition of an interval (Definition 9.1.1). Indeed, let's consider two elements $x < y$ in X, and z another real such that $x < z < y$. By transitivity of order, we clearly have $z \in (a, b)$, as expected.

- Now suppose that X is bounded and connected, and let's show that it is a bounded interval. If X is empty, the claim follows, because of the convention adopted in the definitions. If X is not empty, consider $a := \inf(X)$ and $b := \sup(X)$ (they exist as real numbers because X is a bounded, non empty subset of the real line). By definition of the infimum and supremum, we have $X \subseteq [a, b]$.

  Now let be $z \in]a, b[$, i.e., $a < z < b$. By definition of the infimum of X, there exists an $x \in X$ such that $a < x < z$; otherwise a would not be the infimum of X (there would exist a smaller minorant). Similarly, there exists a $y \in X$ such that $z < y < b$. Thus, we have $x < z < y$, and since X is supposed to be connected, we have $z \in X$. We have just show that $]a, b[\subseteq X$.

  Thus, we have $]a, b[\subseteq X \subseteq [a, b]$; i.e., X can only be one of the four sets $[a, b]$, $[a, b[$, $]a, b]$ and $]a, b[$. In each case, X is a bounded interval, as expected.

Exercise 11.1.2. — *Prove Corollary 11.1.6.*

Let be I, J two bounded intervals; we must show that $I \cap J$ is also a bounded interval. By Lemma 11.1.4, we know that I, J are both bounded connected sets, and it will suffice to show that $I \cap J$ is also a bounded and connected set.

- Since I is bounded, we have $I \subset [-M, M]$ for some $M > 0$. Similarly, we have $J \subset [-N, N]$ for some $N > 0$. Thus, if we set $P := \max(M, N)$, we have clearly $I \cap J \subset [-P, P]$, so that $I \cap J$ is bounded.
- Let be $x < y$ in $I \cap J$, and another element z such that $x < z < y$. Let's show that $z \in I \cap J$. Since $x, y, z \in I \cap J$, then in particular $x, y, z \in I$, which is a connected set. Thus, $z \in I$. A similar argument show that $z \in J$. Thus, $z \in I \cap J$, as expected, which show that $I \cap J$ is connected.

Since $I \cap J$ is both bounded and connected, it is an interval, as expected.

Exercise 11.1.3. — *Let I be a bounded interval of the form $I =]a, b[$ or $I = [a, b[$ for some real numbers $a < b$. Let $I_1, \ldots, I_n$ be a partition of I. Prove that one of the intervals I_j in this partition is of the form $I_j =]c, b[$ or $I_j = [c, b[$ for some $a \leqslant c \leqslant b$. (Hint: prove by contradiction. First show that if I_j is not of the form $]c, b[$ or $[c, b[$ for any $a \leqslant c \leqslant b$, then* $\sup I_j$ *is strictly less than b.)*

First note that for any interval I of real numbers of the form $[b, t[$ $[b, t]$, $]b, t]$ or $]b, t[$, we have $\sup(I) = t$ and $\inf(I) = b$. (Indeed, t is an upper bound for each of these intervals, and for any $\varepsilon > 0$, $t - \varepsilon$ is not an upper bound. Similarly for b.)

Thus, since I_j is an element of a partition of I, we must have $\sup(I_j) \leqslant b$ (otherwise, if $\sup(I_j) > b$, we could find an element of I_j which is greater than b, i.e. an element of I_j that is not an element of I, which is not possible in a partition).

Thus, for all I_j in the partition, we have $\sup(I_j) \leqslant b$. If I_j is not of the form $]c, b[$ or $[c, b[$ for any $a \leqslant c \leqslant b$, then obviously $\sup(I_j) \neq b$, and thus $\sup(I_j) < b$.

Now let's suppose, for the sake of contradiction, that no interval I_j in $I_1, \ldots, I_n$ is of the form $I_j =]c, b[$ or $I_j =]c, b]$. Thus we have $\sup(I_j) < b$ for all $1 \leqslant j \leqslant n$. Since a partition is a *finite* collection of intervals, we can define $m := \max_{1 \leqslant j \leqslant n}(\sup I_j)$, and we will have $m < b$. In particular, the interval $]m, b[$ is non empty, is a subset of I, but the elements of $]m, b[$ cannot be found in any I_j in the partition. This is a contradiction.

Exercise 11.1.4. — *Prove Lemma 11.1.18.*

Let I be an interval, $\mathbf{P}$ and $\mathbf{P}'$ be partitions of I, and let's show that $\mathbf{P} \# \mathbf{P}'$ is a partition of I that is finer than both $\mathbf{P}$ and $\mathbf{P}'$. First, recall that $\mathbf{P} \# \mathbf{P}' := \{J \cap K \;:\; J \in \mathbf{P}, K \in \mathbf{P}'\}$.

- Let be $x \in I$. Since $\mathbf{P}$ is a partition of I there exists exactly one $J \in \mathbf{P}$ such that $x \in J$. Similarly, since $\mathbf{P}'$ is a partition of I, there exists exactly one $K \in \mathbf{P}'$ such that $x \in K$. Thus, we have $x \in J \cap K$, which is an element of $\mathbf{P} \# \mathbf{P}'$. This element is unique, otherwise this would obviously contradict the unicity of J and K. Thus, $\mathbf{P} \# \mathbf{P}'$ is indeed a partition of I.
- Let be $S \in \mathbf{P} \# \mathbf{P}'$. By definition, we have $S = J \cap K$ for $J \in \mathbf{P}$ and $K \in \mathbf{P}'$. Thus, in particular, $S \subseteq J$; this shows that $\mathbf{P} \# P'$ is finer than $\mathbf{P}$. Similarly, $S \subseteq K$; this show that $\mathbf{P} \# \mathbf{P}'$ is finer than $\mathbf{P}'$.

Exercise 11.2.1. — *Prove Lemma 11.2.7.*

Consider an interval $J' \in \mathbf{P}'$. Since $\mathbf{P}'$ is a partition that is finer than $\mathbf{P}$, then by Definition 11.1.14, there exists an interval $J \in P$ such that $J' \subseteq J$. But f is supposed to be piecewise constant with respect to $\mathbf{P}$, so that $f|_J$ is a constant function, i.e., there exists a constant c_j such that $f(x) = c_j$ for all $x \in J$. Thus, in particular, $f(x) = c_j$ for all $x \in J'$, so that $f|_{J'}$ is constant.

This shows that f is also piecewise constant with respect to $\mathbf{P}'$.

Exercise 11.2.2. — *Prove Lemma 11.2.8. (Hint: use Lemmas 11.1.18 and 11.2.7 to make f and g piecewise constant with respect to the same partition of I.)*

Consider that f is piecewise constant with respect to some partition $\mathbf{P}$ of I, and g is piecewise constant with respect to some partition $\mathbf{P}'$ of I.

By Lemma 11.1.18, $\mathbf{P}\#\mathbf{P}'$ is also a partition of I and is finer than both $\mathbf{P}$ and $\mathbf{P}'$. Furthermore, by Lemma 11.2.7, we know that f and g are thus piecewise constant with respect to $\mathbf{P}\#\mathbf{P}'$.

For this partition $\mathbf{P}\#\mathbf{P}'$, it thus clear that $f+g$ is piecewise constant. Indeed, for any $K \in \mathbf{P}\#\mathbf{P}'$, we have $f|_K = c_f$ and $g|_K = c_g$, so that $(f+g)|_K = c_f + c_g$. A similar argument applies for $f-g$, $\max(f,g)$, fg, and for f/g if g vanishes nowhere on I.

Exercise 11.2.3. — *Prove Proposition 11.2.13.*

Let I be a bounded interval, and let $f : I \to \mathbb{R}$ be a function. Suppose that $\mathbf{P}$ and $\mathbf{P}'$ are partitions of I such that f is piecewise constant both with respect to $\mathbf{P}$ and $\mathbf{P}'$. We have to show that $p.c. \int_{[\mathbf{P}]} f = p.c. \int_{[\mathbf{P}']} f$.

- Let's begin with a preliminary result. Consider $\mathbf{P}$ and $\mathbf{P}'$ two partitions of an interval I. We claim that, for any $J \in \mathbf{P}$, the set $S_J := \{J \cap K \,:\, K \in \mathbf{P}'\}$ is a partition of J (of course, we can switch the rôles of P and P' and this statement will still be true). Indeed, let be $x \in J$ (we consider that J is non empty, otherwise the result is trivial). In particular, $x \in I$. Since $\mathbf{P}'$ is a partition of I, there exists exactly one $K \in P'$ such that $x \in K$. Thus, $x \in J \cap K$, and since K is unique, so is $J \cap K$. Thus, S_J is indeed a partition of J.

- Another preliminary result : as a consequence, for any $J \in \mathbf{P}$, we have:

$$\begin{aligned}
\sum_{K \in \mathbf{P}'} |J \cap K| &= \sum_{X \in S_J} |X| && \text{(by Proposition 7.1.11(c))} \\
&= |J| && \text{(by Theorem 11.1.13)}
\end{aligned}$$

- Now let's begin the actual proof. First, note that, by Lemma 11.2.7, f is piecewise constant with respect to $\mathbf{P}\#\mathbf{P}'$, so that $p.c. \int_{[\mathbf{P}\#\mathbf{P}']} f$ is well-defined.

- We thus have:

$$\begin{aligned}
p.c. \int_{[\mathbf{P}\#\mathbf{P}']} f &= \sum_{X \in \mathbf{P}\#\mathbf{P}'} c_X |X| && \text{(by Definition 11.2.9)} \\
&= \sum_{(J,K) \in \mathbf{P} \times \mathbf{P}'} c_{J\cap K} |J \cap K| && \text{(by Definition 11.1.14)} \\
&= \sum_{J \in P} \sum_{K \in \mathbf{P}'} c_{J \cap K} |J \cap K| && \text{(by Corollary 7.1.14)} \\
&= \sum_{J \in P} \sum_{K \in \mathbf{P}'} c_J |J \cap K| && \text{(because } c_J = c_{J\cap K} \text{ for all } K \in \mathbf{P}'\text{)} \\
&= \sum_{J \in P} c_J \sum_{K \in \mathbf{P}'} |J \cap K| && \text{(by linearity of finite sums)} \\
&= \sum_{J \in P} c_J |J| && \text{(see preliminary result above)} \\
&= p.c. \int_{[\mathbf{P}]} f && \text{(by Definition 11.2.9)}
\end{aligned}$$

Note that when we used Fubini's theorem for finite series (Corollary 7.1.14) above, we could have written the double summation in the reverse order, i.e., $\sum_{K\in P'}\sum_{J\in\mathbf{P}} c_{J\cap K}|J\cap K|$. It is easy to see that, in this case, we would have got $p.c. \int_{[\mathbf{P}']} f$ in the end. Thus, we have shown that

$$p.c. \int_{[\mathbf{P}\#\mathbf{P}']} f = p.c. \int_{[\mathbf{P}]} f = p.c. \int_{[\mathbf{P}']} f$$

and we have in particular the equality of Proposition 11.2.13.

Exercise 11.2.4. — *Prove Theorem 11.2.16.*

Let be two piecewise constant functions $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$, and let $\mathbf{P}$ be a given partition of I. As a general remark (see Proposition 11.2.13 or Exercise 11.2.3 above), $p.c. \int_I f$ and $p.c. \int_I g$ do not depend on the choice of the partition of I, so that we can consider that $p.c. \int_I f = p.c. \int_{[\mathbf{P}]} f$ and similarly for g.

(a) Show that $p.c. \int_I (f+g) = p.c. \int_I f + p.c. \int_I g$. By Definition 11.2.9, we have $p.c. \int_I f = p.c. \int_{[\mathbf{P}]} f = \sum_{J\in\mathbf{P}} c_J|J|$, where c_J is the constant value of f on J. Similarly, $p.c. \int_I g = p.c. \int_{[\mathbf{P}]} g = \sum_{J\in\mathbf{P}} d_J|J|$, where d_J is the constant value of g on J. Since these two summations are finite, we have (by linearity of finite sums, Proposition 7.1.11(f))

$$p.c. \int_I f + p.c. \int_I g = \sum_{J\in\mathbf{P}} (c_J + d_J)|J|.$$

Now, observe that $f+g$ is also piecewise constant on I (Lemma 11.2.8) and that the constant value of $f+g$ on each element $J \in \mathbf{P}$ is obviously $c_J + d_J$ (with previous notations). Thus, we also have, by Definition 11.2.9,

$$p.c. \int_I (f+g) = p.c. \int_{[\mathbf{P}]} (f+g) = \sum_{J\in\mathbf{P}} (c_J + d_J)|J|.$$

This shows that we have indeed $p.c. \int_I (f+g) = p.c. \int_I f + p.c. \int_I g$.

(b) Show that, for any real number c, we have $p.c. \int_I (cf) = c(p.c. \int_I f)$. The exact same argument applies, replacing just Proposition 7.1.11(f) by Proposition 7.1.11(g):

$$p.c. \int_I (cf) := p.c. \int_{[\mathbf{P}]} (cf) = \sum_{J\in\mathbf{P}} (cc_J)|J| = c\sum_{J\in\mathbf{P}} c_J|J| = c\int_{[\mathbf{P}]} f =: c\int_I f$$

(c) Show that $p.c. \int_I (f-g) = p.c. \int_I f - p.c. \int_I g$. This is a direct consequence of (b), choosing $c = -1$, and then (a).

(d) Show that if $f(x) \geqslant 0$ for all $x \in I$, then $p.c. \int_I f \geqslant 0$. If $f(x) \geqslant 0$, then all constant values c_J are positive. Since the lengths $|J|$ are also clearly positive, we have immediately $p.c. \int_I f := p.c. \int_{[\mathbf{P}]} f = \sum_{J\in\mathbf{P}} c_J|J| \geqslant 0$, as expected.

(e) Show that if $f(x) \geqslant g(x)$ for all $x \in I$, then $p.c. \int_I f \geqslant \int_I g$. First note that $f(x) \geqslant g(x)$ iff $f(x) - g(x) \geqslant 0$. Thus, by the previous result (d), we have $p.c. \int_I (f-g) \geqslant 0$. By the previous result (c), this can also be written $p.c. \int_I f - p.c. \int_I g \geqslant 0$. This shows that we have indeed $p.c. \int_I f \geqslant \int_I g$.

(f) Show that if $f(x) = c$ for all $x \in I$, then we have $p.c. \int_I f = c|I|$. In this case, we $c_J = c$ for all $J \in \mathbf{P}$ with the previous notations. Thus we have

$$\begin{aligned}
p.c. \int_I f &= p.c. \int_{[\mathbf{P}]} f && \text{(Definition 11.2.14)} \\
&= \sum_{J \in \mathbf{P}} c|J| && \text{(Definition 11.2.9)} \\
&= c \sum_{J \in \mathbf{P}} |J| && \text{(Proposition 7.1.11(f))} \\
&= c|I| && \text{(Theorem 11.1.13)}
\end{aligned}$$

as expected.

(g) Here we have to show that if J is a bounded interval containing I and $F : J \to \mathbb{R}$ is the function defined by $F(x) := f(x)$ if $x \in I$ and $f(x) := 0$ if $x \notin I$, then F is piecewise constant on J, and $p.c. \int_J F = p.c. \int_I f$.

First, we want to say something like "$\{I, J \backslash I\}$ is a partition of I". This may seem trivial; however it is not really trivial using Tao's definition of a partition. Indeed, Definition 11.1.10 requires that all elements in a partition be bounded intervals. If we really want to be picky about that, we can split into cases. Indeed, since I is supposed to be a bounded interval, we have $I = (a, b)$, where the parentheses stand for either closing or opening brackets. Thus, J can be either of the form (c, b), (a, d) or (c, d), where $c \leqslant a < b \leqslant d$. In either case, $J \backslash I$ is a bounded interval, or the reunion of two bounded intervals. Thus, there clearly exists a partition $\mathbf{P}''$ of J, defined by $\mathbf{P}'' := \mathbf{P} \cup \mathbf{P}'$, where $\mathbf{P}$ is a partition of I and $\mathbf{P}'$ is a partition of $J \backslash I$.

Thus, F is clearly piecewise constant relative to $\mathbf{P}''$, since it is piecewise constant on $\mathbf{P}$ by hypothesis, and F is constant equal to 0 on each interval in $\mathbf{P}'$. Thus, F is piecewise constant on J.

We thus have

$$\begin{aligned}
p.c. \int_J F &= p.c. \int_{[\mathbf{P}'']} F && \text{(Definition 11.2.14)} \\
&= p.c. \int_{[\mathbf{P} \cup P']} F \\
&= \sum_{K \in \mathbf{P} \cup \mathbf{P}'} c_K |K| && \text{(Definition 11.2.9)} \\
&= \sum_{K \in \mathbf{P}} c_K |K| + \sum_{K \in \mathbf{P}'} c_K |K| && \text{(Proposition 7.1.11(e))} \\
&= \sum_{K \in \mathbf{P}} c_K |K| + \sum_{K \in \mathbf{P}'} 0 \times |K| \\
&= \sum_{K \in \mathbf{P}} c_K |K| \\
&= p.c. \int_I f
\end{aligned}$$

as expected.

(h) We have to show that if $\{J, K\}$ is a partition of I into two intervals J and K, then the functions $f|_J : J \to \mathbb{R}$ and $f|_K : K \to \mathbb{R}$ are piecewise constant on J and K respectively, and we have $p.c. \int_I f = p.c. \int_J f|_J + p.c. \int_K f|_K$.

f is supposed to be piecewise constant on I, so if we consider a partition $\mathbf{P}$ of I, then f is piecewise constant with respect to $\mathbf{P}$. We can see easily that $\mathbf{P}_J := \{J \cap L : L \in \mathbf{P}\}$ is a partition of J, and $\mathbf{P}_K := \{K \cap L : L \in \mathbf{P}\}$ is a partition of K. Thus, $\mathbf{P}_J \cup \mathbf{P}_K$ is a partition of I, and f is piecewise constant with respect to this partition. Thus, in particular, $f|_J$ is piecewise constant with respect to $\mathbf{P}_J$, and similarly with $f|_K$ and $\mathbf{P}_K$.

Inspired by the previous result (g), let's define the function $F : I \to \mathbb{R}$ by $F(x) := f|_J(x)$ if $x \in J$ and $F(x) := 0$ otherwise. Similarly, we define $G : I \to \mathbb{R}$ by $G(x) := f|_K(x)$ if $x \in K$ and $G(x) := 0$ otherwise. Clearly, we have $f(x) = F(x) + G(x)$ for all $x \in I$. Thus, we have:

$$\begin{aligned} p.c. \int_I f &= p.c. \int_I (F + G) \\ &= p.c. \int_I F + p.c. \int_I G && \text{(Theorem 11.2.16(a))} \\ &= p.c. \int_J f|_J + p.c. \int_K f|_K && \text{(Theorem 11.2.16(g))} \end{aligned}$$

as expected.

Exercise 11.3.1. — *Let $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$ and $h : I \to \mathbb{R}$ be functions. Show that if f majorizes g and g majorizes h, then f majorizes h. Show that if f and g majorize each other, then they must be equal.*

If f majorizes g and g majorizes h, we have $f(x) \geqslant g(x)$ and $g(x) \geqslant h(x)$ for any given $x \in I$. By transitivity of order on the real numbers, we thus have $f(x) \geqslant h(x)$. It is true for any given $x \in I$, thus f majorizes h.

Similarly, if f majorizes g and g majorizes f, then we have $f(x) \leqslant g(x)$ and $f(x) \geqslant g(x)$ for any given $x \in I$. Thus, we have $f(x) = g(x)$ by antisymmetry of order on the reals. Thus, we have indeed $f = g$.

Exercise 11.3.2. — *Let $f : I \to \mathbb{R}$, $g : I \to \mathbb{R}$ and $h : I \to \mathbb{R}$ be functions. If f majorizes g, is it true that $f + h$ majorizes $g + h$? Is is true that $f \cdot h$ majorizes $g \cdot h$? If c is a real number, is it true that cf majorizes cg?*

If f majorizes g, then $f(x) \geqslant g(x)$ for any given $x \in I$. Thus, whatever may be $h(x)$, we have $f(x) + h(x) \geqslant g(x) + h(x)$. This is true for any $x \in I$, thus $f + h$ majorizes $g + h$.

However, fh does not majorize gh. Indeed, since we give no restriction on h, this function might be negative at some places. The most trivial example would be the constant function $h(x) = c$ for some real number $c < 0$. In this case, for all $x \in I$, we have $f(x) \geqslant g(x)$ but $f(x)h(x) \leqslant g(x)h(x)$.

In particular, it shows at the same time that if we set $c := -1$ (or any other negative value), cf does not majorize cg.

Exercise 11.3.3. — *Prove Lemma 11.3.7.*

The claim we must prove is that, if f is piecewise constant on I, it is Riemann-integrable, abd its Riemann integral (defined in Definition 11.3.4) is equal to the piecewise constant integral (defined in Definition 11.2.14).

Recall that, in Definition 11.3.1, the inequalities are *large*: we say that a function f majorizes g on I iff $f(x) \geqslant g(x)$ for all $x \in I$. In particular, it means that the binary relation "majorizes" is reflexive, i.e., that any function f majorizes itself, since $f(x) \geqslant f(x)$ for all $x \in I$. We have a similar remark for the relation "minorizes".

With this remark in mind, the Lemma 11.3.7 is trivial:

- f majorizes f, thus by Definition 11.3.2, we have $\overline{\int}_I f \leqslant p.c. \int_I f$.
- f minorizes f, thus by Definition 11.3.2, we have $\underline{\int}_I f \geqslant p.c. \int_I f$.

Thus we have $\overline{\int}_I f \leqslant p.c. \int_I f \leqslant \underline{\int}_I f$. But we must also have $\overline{\int}_I f \geqslant \underline{\int}_I f$ by Lemma 11.3.3. The only possibility for those two statements to be simultaneously true is that $\overline{\int}_I f = p.c. \int_I f = \underline{\int}_I f$, i.e., f is Riemann-integrable, and its Riemann integral is equal to its piecewise constant integral.

Exercise 11.3.4. — *Prove Lemma 11.3.11.*

We will only prove the first statement of this lemma, i.e. that for any piecewise constant function g (w.r.t. a partition $\mathbf{P}$ of I) that majorizes f on I, we have

$$p.c. \int_I g \geqslant U(f, \mathbf{P}) ;$$

the second statement can be proved in the same way.

Let be $\mathbf{P}$ a partition of I. We know that the piecewise constant integral does not depend on the choice of the partition (Proposition 11.2.13), thus we have

$$p.c. \int_I g = p.c. \int_{[\mathbf{P}]} g = \sum_{J \in \mathbf{P}} c_J |J| = \sum_{J \in \mathbf{P};\, J \neq \varnothing} c_J |J| \tag{11.1}$$

where c_J is the constant value of g any interval $J \in \mathbf{P}$. (The last equality comes from the fact that $|J| = 0$ if $|J| = \varnothing$, so that the restriction on the last summation excludes nothing but null elements, and lets the whole sum identical.)

Since g majorizes f on I, for all $J \in \mathbf{P}$, and for all x in this interval J, we have $f(x) \leqslant c_J$, i.e. $f(x) \leqslant c_J$. In particular, $c_J \geqslant \sup_{x \in J} f(x)$ on each interval $J \in \mathbf{P}$. Thus we have

$$\sum_{J \in \mathbf{P};\, J \neq \varnothing} c_J |J| \geqslant \sum_{J \in \mathbf{P};\, J \neq \varnothing} (\sup_{x \in J} f(x)) |J| = U(f, \mathbf{P}) \tag{11.2}$$

Combining equations (11.1) and (11.2), we get $p.c. \int_I g \geqslant U(f, \mathbf{P})$, as expected.

Exercise 11.3.5. — *Prove Proposition 11.3.12.*

As for the last exercise, we will only prove the first statement, i.e. that $\overline{\int}_I f = \inf\{U(f, \mathbf{P}) : \mathbf{P}$ is a partition of $I\}$. The second statement can be proved in a similar fashion.

- First, let's prove that $\overline{\int}_I f \geqslant \inf\{U(f,\mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$.

  Let be $\mathbf{P}$ any partition of I, and g any function piecewise constant with respect to $\mathbf{P}$ that majorizes f on I. Thus we have *p.c.* $\int_I g \geqslant U(f,\mathbf{P})$, by Lemma 11.3.11.

  It means that $U(f,\mathbf{P})$ is a lower bound for the set $\{p.c. \int_I g \mid g \text{ is p.c. and majorizes } f\}$. Thus, by definition of an infimum (taken on functions g), we have

$$U(f,\mathbf{P}) \geqslant \inf\left\{ p.c. \int_I g \;\middle|\; g \text{ is p.c. and majorizes } f \right\} =: \overline{\int}_I f$$

  This is true for any given partition $\mathbf{P}$ of I. Thus, it means that $\overline{\int}_I f$ is a lower bound for the set $\{U(f,\mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$. By definition of an infimum (taken on partitions $\mathbf{P}$), we have

$$\overline{\int}_I f \geqslant \inf\left\{U(f,\mathbf{P}) : \mathbf{P} \text{ is a partition of } I\right\}$$

  as expected.

- Now, let's prove that $\overline{\int}_I f \leqslant \inf\{U(f,\mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$.

  Let be $\mathbf{P}$ any given partition of I. Let be a function $S : I \to \mathbb{R}$ defined as follows. For any interval $J \in \mathbf{P}$, we set $S(x) := \sup_{x \in J} f(x)$. Since f is bounded on I, the function S is well defined (the supremum on J is always a real number). Furthermore, by construction, S is piecewise constant on I with respect to $\mathbf{P}$, and S majorizes f on I. Thus, by Definition 11.3.2 of the upper Riemann integral,

$$\overline{\int}_I f \leqslant p.c. \int_I S(x) = p.c. \int_{[\mathbf{P}]} S(x) = \sum_{J \in \mathbf{P};\, J \neq \varnothing} (\sup_{x \in J} f(x))|J| = U(f,\mathbf{P}).$$

  Since the inequality $\overline{\int}_I f \leqslant U(f,\mathbf{P})$ is true for any partition $\mathbf{P}$ of I, the quantity $\overline{\int}_I f$ is a lower bound for the set $\{U(f,\mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$. Thus, by definition of an infimum (taken on partitions $\mathbf{P}$), we have

$$\overline{\int}_I f \leqslant \inf\{U(f,\mathbf{P}) \;:\; \mathbf{P} \text{ is a partition of } I\}$$

  as expected.

- Thus, we have $\overline{\int}_I f = \inf\{U(f,\mathbf{P}) : \mathbf{P} \text{ is a partition of } I\}$, as expected.

Exercise 11.4.1. — *Prove Theorem 11.4.1.*

Here we have to prove the properties of the (general) Riemann integral, using the properties of the piecewise constant integral developed earlier. It might not be so easy to see where to start, so that we will recall some definitions and derive some preliminary results that will be useful for the whole exercise.

- First, since f and g are supposed to be Riemann integrable, we thus have $\int_I f = \overline{\int}_I f = \underline{\int}_I f$, and similarly for g.

- By definition, we have $\overline{\int_I} f := \inf\{p.c. \int_I g : g \text{ is p.c. and majorizes } f\}$. In particular, by definition of an infimum, it means that for all $\varepsilon > 0$, there exists a piecewise constant function $\overline{f}$ that majorizes f such that

$$\overline{\int_I} f + \varepsilon > p.c. \int_I \overline{f} = \int_I \overline{f} \tag{11.3}$$

But since f is Riemann-integrable, we have $\int_I f = \overline{\int_I} f$; and since $\overline{f}$ is piecewise constant and majorizes f, we have $\int_I \overline{f} \geqslant \overline{\int_I} f$. Thus, equation (11.3) can be extended as

$$\int_I f + \varepsilon = \overline{\int_I} f + \varepsilon > \int_I \overline{f} \geqslant \overline{\int_I} f = \int_I f \tag{11.4}$$

- A similar argument for the lower integral shows that, for all $\varepsilon > 0$, there exists a piecewise constant function $\underline{f}$ minorizing f such that

$$\int_I f - \varepsilon = \underline{\int_I} f - \varepsilon < \int_I \underline{f} \leqslant \underline{\int_I} f = \int_I f \tag{11.5}$$

- Putting together equations (11.4)–(11.5), we finally get that for all $\varepsilon > 0$, there exists two piecewise constant functions $\overline{f}$ and $\underline{f}$ majorizing and minorizing f respectively, such that

$$\int_I f - \varepsilon < \int_I \underline{f} \leqslant \int_I f \leqslant \int_I \overline{f} < \int_I f + \varepsilon \tag{11.6}$$

- Of course, we have the exact analogue for g with

$$\int_I g - \varepsilon < \int_I \underline{g} \leqslant \int_I g \leqslant \int_I \overline{g} < \int_I g + \varepsilon \tag{11.7}$$

We are now ready to begin the actual proofs of the theorem.

(a) First let's show that $f+g$ is Riemann-integrable. This requires to show that $\overline{\int_I}(f+g) = \underline{\int_I}(f+g)$. Let be $\varepsilon > 0$ a positive real. With previous notations of equations (11.6)–(11.7), we know that $\overline{f} + \overline{g}$ majorizes $f+g$, so that we have

$$\begin{aligned}
\overline{\int_I}(f+g) &\leqslant p.c. \int_I (\overline{f} + \overline{g}) && \text{(Definition 11.3.2)} \\
&\leqslant p.c. \int_I \overline{f} + p.c. \int_I \overline{g} && \text{(Theorem 11.2.16(a))} \\
&< \int_I f + \varepsilon + \int_I g + \varepsilon && \text{(Equations (11.6)–(11.7))}
\end{aligned}$$

and, similarly,

$$\underline{\int_I}(f+g) > \int_I f + \int_I g - 2\varepsilon$$

so that we finally get

$$\int_I f + \int_I g - 2\varepsilon < \underline{\int_I}(f+g) \leqslant \overline{\int_I}(f+g) < \int_I f + \int_I g + 2\varepsilon \tag{11.8}$$

In particular, for all $\varepsilon > 0$,

$$-2\varepsilon < \underline{\int_I}(f+g) - \left(\int_I f + \int_I g\right) < 2\varepsilon$$

which means that $\underline{\int_I}(f+g) = (\int_I f + \int_I g)$.

Similarly, for all $\varepsilon > 0$, we have

$$-2\varepsilon < \overline{\int_I}(f+g) - \left(\int_I f + \int_I g\right) < 2\varepsilon$$

which means that $\overline{\int_I}(f+g) = (\int_I f + \int_I g)$.

Thus, we have $\overline{\int_I}(f+g) = \underline{\int_I}(f+g) = \int_I f + \int_I g$, which implies both that $f+g$ is Riemann-integrable and that $\int_I(f+g) = \int_I + \int_I g$.

(b) We have to show that for any real number c, the function cf is Riemann-integrable, and we have $\int_I(cf) = c(\int_I f)$. We will have to split into cases, depending on the value of c.

- The easiest case is $c = 0$. In this case, cf is the constant function 0. In particular, since cf is constant, it is piecewise constant on I, so it is Riemann-integrable (by Lemma 11.3.7). Thus we have

$$\int_I cf = p.c. \int_I cf = c\left(p.c. \int_I f\right) = c\left(\int_I f\right)$$

since we can apply the properties of the piecewise constant integral (more precisely, Theorem 11.2.16(b)).

- Now consider the case $c > 0$. Since c is positive, we have the property of preservation of order; i.e., if $\overline{f}$ majorizes f, then $c\overline{f}$ majorizes cf. Similarly, if $\underline{f}$ minorizes f, then $c\underline{f}$ minorizes cf (where $\underline{f}$ and $\overline{f}$ correspond to any fixed arbitrary $\varepsilon > 0$ in equation (11.6)).

Since $c\overline{f}$ is a piecewise constant function that majorizes cf, we have

$$\overline{\int_I} cf \leqslant p.c. \int_I c\overline{f} = c\left(p.c. \int_I \overline{f}\right) < c\left(\int_I f + \varepsilon\right) \tag{11.9}$$

where we used Theorem 11.2.16(b) and equation (11.6).

Similarly, since $c\underline{f}$ is a piecewise constant function that minorizes cf, we have

$$\underline{\int_I} cf \geqslant p.c. \int_I c\underline{f} = c\left(p.c. \int_I \underline{f}\right) > c\left(\int_I f - \varepsilon\right) \tag{11.10}$$

still using Theorem 11.2.16(b) and equation (11.6).

Thus, combining the two previous results (and using the fact that $\underline{\int_I} cf \leqslant \overline{\int_I} cf$, by Lemma 11.3.3), we have

$$c\int_I f - c\varepsilon < \underline{\int_I} cf \leqslant \overline{\int_I} cf < c\int_I f + c\varepsilon.$$

In particular, $c\int_I f - c\varepsilon < \underline{\int_I} cf < c\int_I f + c\varepsilon$ for any arbitrary $\varepsilon > 0$. Thus, $c\int_I f = \underline{\int_I} cf$. A similar argument shows that $c\int_I f = \overline{\int_I} cf$. This means that cf is Riemann-integrable, and that $\int_I cf = c\int_I f$, as expected.

- Now consider the case $c = -1$. In this case (still for an arbitrary $\varepsilon > 0$ in equation (11.6)), if $\overline{f}$ majorizes f, we have $c\overline{f} = -\overline{f} \leqslant -f$, i.e., $-\overline{f}$ minorizes $-f$. Similarly, $-\underline{f}$ majorizes $-f$.
  Since $-\underline{f}$ majorizes $-f$, we have
  $$\overline{\int_I} - f \leqslant p.c. \int_I -\underline{f} = -p.c. \int_I \underline{f} < -\int_I f + \varepsilon$$
  where we have used Theorem 11.2.16(b) and equation (11.6) (reversing the inequality "$\int_I f - \varepsilon < \int_I \underline{f}$" as "$-\int_I f + \varepsilon > -\int_I \underline{f}$").
  Similarly, we have
  $$-\int_I f - \varepsilon < p.c. \int_I \overline{f} \leqslant \underline{\int_I} - f$$
  Combining the previous results, we obtain
  $$-\int_I f - \varepsilon < \underline{\int_I} - f \leqslant \overline{\int_I} - f \leqslant -\int_I f + \varepsilon$$
  and since this is true for any arbitrary $\varepsilon > 0$, we have $\overline{\int_I} - f = \underline{\int_I} - f = -\int_I f$, so that $-f$ is Riemann-integrable, and $\int_I -f := -\int_I f$, as expected.
- Finally, consider the general case $c < 0$. We just have to combine the results of the two previous cases. If $c < 0$, we have $-c > 0$, so that $-cf$ is integrable, by the second case proved earlier. And by the third case proved above, $-(-cf) = cf$ is integrable. Furthermore, still using the two previous cases,
  $$\int_I cf = \int_I (-1) \times (-cf) = -1 \times \int_I (-cf) = (-1) \times (-c) \int_I f = c \int_I f$$
  as expected.

(c) Simply combine (a) and (b) with $c = -1$.

(d) Saying that $f(x) \geqslant 0$ for all $x \in I$ means that the constant (and thus piecewise constant) function 0 minorizes f on I. Thus, by definition, $\underline{\int_I} f \geqslant p.c. \int_I 0 = 0$ (for instance by Theorem 11.2.16(e), or by the previous statement (b) with $c = 0$). But since f is supposed to be Riemann-integrable, we have $\int_I f = \underline{\int_I} f$. The statement follows.

(e) Direct consequence of (d) and (c).

(f) If $f(x) = c$ for all $x \in I$, then f is constant, and thus piecewise constant on I. By Lemma 11.3.3, f is Riemann-integrable, and $\int_I f = p.c. \int_I f = c|I|$, where the last equality comes from Theorem 11.2.16(f).

(g) Let be $\overline{f}$ a piecewise function majorizing f on I. If we define the function $\overline{F} : J \to \mathbb{R}$ by $\overline{F}(x) = \overline{f}(x)$ when $x \in I$, and $\overline{F}(x) = 0$ otherwise, then it is clear that $\overline{F}$ is a piecewise function majorizing f on J. Thus for any given $\varepsilon > 0$, there exist such functions $\overline{f}, \overline{F}$, with
$$\overline{\int_J} F \leqslant p.c. \int_J \overline{F} = p.c. \int_I \overline{F} = p.c. \int_I \overline{f} < \int_I f + \varepsilon$$
(where we used Theorem 11.2.16(g) and equation (11.6)).

Similarly, with obvious notations, there exist two functions $\underline{f}, \underline{F}$, with

$$\underline{\int_J} F \geqslant p.c. \int_J \underline{F} = p.c. \int_I \underline{F} = p.c. \int_I \underline{f} > \int_I f - \varepsilon$$

Combining these results, we have, for all $\varepsilon > 0$,

$$\int_I f - \varepsilon < \underline{\int_J} F \leqslant \overline{\int_J} F < \int_I f + \varepsilon$$

which means that F is Riemann-integrable, and $\int_J F = \int_I f$, as expected.

(h) If we suppose that $f|_J$ and $f|_K$ are Riemann integrable, the result can be immediately deduced from parts (a) and (g). Indeed, let's take $F : I \to \mathbb{R}$ the function defined by $F(x) = f|_J(x)$ if $x \in J$ and $F(x) := 0$ otherwise; and similarly, the function $G : I \to \mathbb{R}$ defined by $G(x) = f|_K(x)$ if $x \in K$ and $G(x) = 0$ otherwise. Thus, it is clear that we have $f = F + G$. Using previous results, we have

$$\begin{aligned} \int_I f &= \int_I F + \int_I G && \text{(by (a))} \\ &= \int_J f|_J + \int_K f|_K && \text{(by (g))} \end{aligned}$$

as expected.

The painful part is precisely to show that $f|_J$ is Riemann integrable on J, and $f|_K$ is Riemann integrable on K. So, we have to prove that $\overline{\int_J} f|_J = \underline{\int_J} f|_J$, and $\overline{\int_K} f|_K = \underline{\int_K} f|_K$.

Let be $\varepsilon > 0$, and two piecewise continuous functions $\overline{f}$ and $\underline{f}$ that obey to equation (11.6). First, note that if $\overline{f}$ majorizes f on I, then $\overline{f}|_J$ majorizes $f|_J$ on J (and is of course piecewise continuous on J). Similarly, $\underline{f}|_J$ minorizes $f|_J$ on J, $\overline{f}|_K$ majorizes $f|_K$ on K, and $\underline{f}|_K$ minorizes $f|_K$ on K. Thus, we have in particular

$$p.c. \int_J \underline{f}|_J \leqslant \underline{\int_J} f|_J \leqslant \overline{\int_J} f|_J \leqslant p.c. \int_J \overline{f}|_J \tag{11.11}$$

and of course, a similar bound on K.

Since they are all piecewise continuous functions (on I, J or K respectively), we have by Theorem 11.2.16(h)

$$\begin{aligned} p.c. \int_I \overline{f} &= p.c. \int_J \overline{f}|_J \; + p.c. \int_K \overline{f}|_K \\ p.c. \int_I \underline{f} &= p.c. \int_J \underline{f}|_J \; + p.c. \int_K \underline{f}|_K \end{aligned}$$

Substituting this in equation (11.6), we obtain

$$\int_I f - \varepsilon < \left(p.c. \int_J \underline{f}|_J + p.c. \int_K \underline{f}|_K\right) \leqslant \left(p.c. \int_J \overline{f}|_J + p.c. \int_K \overline{f}|_K\right) < \int_I f + \varepsilon$$

so that in particular, $(p.c. \int_J \overline{f}|_J + p.c. \int_K \overline{f}|_K)$ and $(p.c. \int_J \underline{f}|_J + p.c. \int_K \underline{f}|_K)$ lie in an interval of length 2ε and thus are at most distant of 2ε of each other[33]

In other words, we have

$$0 \leqslant \left(p.c. \int_J \overline{f}|_J + p.c. \int_K \overline{f}|_K\right) - \left(p.c. \int_J \underline{f}|_J + p.c. \int_K \underline{f}|_K\right) \leqslant 2\varepsilon$$

or, even better, reordering the terms in the middle:

$$0 \leqslant \left(p.c. \int_J \overline{f}|_J - p.c. \int_J \underline{f}|_J\right) + \left(p.c. \int_K \overline{f}|_K + p.c. \int_K \underline{f}|_K\right) \leqslant 2\varepsilon \tag{11.12}$$

Since each term of the sum of the middle is non-negative (by Lemma 11.3.3), each term is itself non-negative and lesser than 2ε. In conclusion, we have

$$\begin{aligned} 0 \leqslant p.c. \int_J \overline{f}|_J - p.c. \int_J \underline{f}|_J \leqslant 2\varepsilon \\ 0 \leqslant p.c. \int_K \overline{f}|_K - p.c. \int_K \underline{f}|_K \leqslant 2\varepsilon \end{aligned}$$

for an arbitrary $\varepsilon > 0$ previously fixed. Thus, if we come back to equation (11.11), this imply that $\underline{\int}_J f|_J$ and $\overline{\int}_J f|_J$ lie in an interval of length at most 2ε, so that they are themselves distant at most of 2ε (see footnote 33). This is true for any arbitrary $\varepsilon > 0$ previously fixed, so that $f|_J$ is Riemann integrable on J.

Similarly, $f|_K$ is Riemann integrable on K.

Exercise 11.4.2. — *Let I be a bounded interval, let $f : I \to \mathbb{R}$ be a Riemann integrable function, and let $\mathbf{P}$ be a partition of I. Show that $\int_I f = \sum_{J \in \mathbf{P}} \int_J f$.*

Our best friends here will be Theorem 11.4.1(h) and Remark 11.4.2. The later is especially useful, since it is not crucial to split into cases depending on whether the intervals in $\mathbf{P}$ are open or closed (as it was the case in the proof of Theorem 11.1.13); we know that for all real numbers $a < b$ we have $\int_{]a,b[} f = \int_{[a,b[} f = \int_{]a,b]} f = \int_{[a,b]} f$.

As previously in this document, we will write (a, b) any interval of the form $]a, b[$, $]a, b]$, $[a, b[$ or $[a, b]$ (i.e., the parentheses stand for either open or closed brackets, and this notation makes a tacit split into cases that does not change anything for the integrals themselves).

Let be $I = (a, b)$ a bounded interval. We use induction on the cardinality $n := \#\mathbf{P}$.

- If $n = 2$, the result is simply Theorem 11.4.1(h).
- Suppose inductively that we have showed the result for $n \geqslant 2$, and let's prove that it is still true for $n + 1$. As in the proof of Theorem 11.1.13, if $\mathbf{P} = \{I_1, I_2, \dots, I_n, I_{n+1}\}$, then one of the intervals I_j is of the form (c, b) with $a < c \leqslant b$ (and is possibly just the singleton $\{b\}$ if $c = b$ and the brackets are closed). Without any loss of generality, we can suppose that this interval is I_{n+1}. Then, $\mathbf{P'} := \mathbf{P} - \{I_{n+1}\}$ is a partition of (a, c)

[33] Formally, if we have $x - \varepsilon < y \leqslant z < x + \varepsilon$, then $0 \leqslant z - y \leqslant 2\varepsilon$ (why? Use triangular inequality).

and has cardinality n. Thus, we have:

$$\begin{aligned}
\int_I f &= \int_{(a,c)} f + \int_{(c,b)} f && \text{(Theorem 11.4.1(h))} \\
&= \sum_{J\in\mathbf{P}'} \int_J f + \int_{(c,b)} f && \text{(by induction hypothesis)} \\
&= \sum_{J\in\mathbf{P}} \int_J f && \text{(because } \mathbf{P} = \mathbf{P}' \cup \{(c,b)\}\text{)}
\end{aligned}$$

which closes the induction and shows the property, as expected.

Exercise 11.5.1. — *Let $a < b$ be real numbers, and let $f : [a,b] \to \mathbb{R}$ be a continuous, non-negative function (so $f(x) \geqslant 0$ for all $x \in [a,b]$). Suppose that $\int_{[a,b]} f = 0$. Show that $f(x) = 0$ for all $x \in [a,b]$.*

Suppose for the sake of contradiction that there exists an $x_0 \in [a,b]$ such that $f(x_0) > 0$. (For simplicity, we will assume that this x_0 lies in $]a,b[$; but the proof is essentially the same if $x_0 = a$ or $x_0 = b$. The only change will consist in a modification of equation (11.13).)

Let be $\varepsilon := f(x_0)/2$; by hypothesis, we have $\varepsilon > 0$. Since f is continuous on $[a,b]$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ whenever ($x \in [a,b]$ and) $|x - x_0| < \delta$.

In particular, we have

$$(x \in [a,b],\ |x - x_0| < \delta) \implies 0 < \frac{f(x_0)}{2} < f(x) < 3\frac{f(x_0)}{2}$$

By Theorem 11.4.1(h), we have

$$\int_{[a,b]} f = \int_{[a,x_0-\delta]} f + \int_{]x_0-\delta,x_0+\delta[} f + \int_{]x_0+\delta,b]} f \tag{11.13}$$

and we know that all terms are (at least) non-negative in the right-hand-side:

- $f|_{[a,x_0-\delta]} \geqslant 0$ by hypothesis, so that $\int_{[a,x_0-\delta]} f \geqslant 0$ by Theorem 11.4.1(d);
- the constant function $x \mapsto f(x_0)/2$ minorizes f on $]x_0-\delta, x_0+\delta[$, so that, by Theorem 11.4.1(b,e), $\int_{]x_0-\delta,x_0+\delta[} f \geqslant 2\delta f(x_0)/2 > 0$;
- $f|_{]x_0+\delta,b]} \geqslant 0$ by hypothesis, so that $\int_{]x_0-\delta,b]} f \geqslant 0$ by Theorem 11.4.1(d).

Thus we have $\int_{[a,b]} f > 0$, a contradiction. This shows that f is the constant function 0 on $[a,b]$.

Exercise 11.5.2. — *Prove Proposition 11.5.6. (Hint: use Theorem 11.4.1(a) and (g).)*

Let I be a bounded interval, and $f : I \to \mathbb{R}$ a function that is both piecewise continuous and bounded. We have to show that f is Riemann integrable.

Since f is piecewise continuous, there exists a partition $\mathbf{P}$ of I such that $f|_J$ is continuous for all $J \in \mathbf{P}$. Thus we can define, for all $J \in \mathbf{P}$, a function $g_J : I \to \mathbb{R}$ (integrable on I by Proposition 11.5.3) by:

$$\begin{cases} g_J(x) := f|_J(x) & \text{if } x \in J \\ g_J(x) := 0 & \text{if } x \in I - J. \end{cases}$$

By Theorem 11.4.1(g), we know that, for all $J \in \mathbf{P}$, g_J is Riemann-integrable on I. Furthermore, we have clearly $f(x) = \sum_{J\in\mathbf{P}} g_J(x)$ for all $x \in I$. Thus, by Theorem 11.4.1(a), f is Riemann-integrable on I (and we have $\int_I f = \sum_{J\in P} \int_I g_J$), as expected.

Exercise 11.6.1. — *Prove Corollary 11.6.3.*

Let I be a bounded interval, and let $f : I \to \mathbb{R}$ be both monotone and bounded. We have to show that f is Riemann-integrable on I.

First, note that there are several trivial particular cases. Since I is bounded, it can be the empty set or a singleton $\{a\}$. In both cases the claim is trivial, with $\int_I f = 0$, by Remark 11.3.8. But I can also be a closed interval $[a, b]$: in this case, the claim to prove is simply Proposition 11.6.1. Thus, we just have to prove the claim when $I =]a, b[$, $]a, b]$ or $[a, b[$, with $a < b$ two real numbers.

Since f is supposed to be bounded on I, there exists a positive real number $M > 0$ such that $-M \leqslant f(x) \leqslant M$ for all $x \in I$.

Let be $0 < \varepsilon < (b-a)/2$ a small number. Thus, the closed interval $[a+\varepsilon, b-\varepsilon]$ is a (proper) subset of I. And, by Proposition 11.6.1, we know that the restriction of f to $[a+\varepsilon, b-\varepsilon]$ is Riemann-integrable on this interval, since it is monotone.

Let be $h : [a+\varepsilon, b-\varepsilon] \to \mathbb{R}$ a piecewise constant function majorizing f. By definition, we have

$$\int_{[a+\varepsilon,b-\varepsilon]} f \leqslant \int_{[a+\varepsilon,b-\varepsilon]} h < \int_{[a+\varepsilon,b-\varepsilon]} f + \varepsilon$$

Now we can define a new function $\tilde{h} : I \to \mathbb{R}$ by

$$\begin{cases} \tilde{h}(x) = h(x) & \text{if } x \in I \\ \tilde{h}(x) = M & \text{if } x \notin I \end{cases}$$

The function $\tilde{h}$ is clearly piecewise constant on I, and it clearly majorizes f on I. Thus, by definition of an upper integral, we have

$$\begin{aligned} \overline{\int_I} f \leqslant \int_I \tilde{h} &:= \varepsilon M + \int_{[a+\varepsilon,b-\varepsilon]} \tilde{h} + \varepsilon M \\ &\leqslant \int_{[a+\varepsilon,b-\varepsilon]} \tilde{h} + 2\varepsilon M \\ &\leqslant \int_{[a+\varepsilon,b-\varepsilon]} f + (1+2M)\varepsilon \end{aligned}$$

A similar argument would show that

$$\underline{\int_I} f \geqslant \int_{[a+\varepsilon,b-\varepsilon]} f - (1+2M)\varepsilon$$

thus leading to

$$\overline{\int_I} f - \underline{\int_I} f \leqslant (2+4M)\varepsilon.$$

Since this is true for all $\varepsilon > 0$, it shows that $\overline{\int_I} f = \underline{\int_I} f$, i.e. that f is Riemann-integrable on I, as expected.

Exercise 11.6.2. — *Formulate a reasonable notion of a piecewise monotone function, and then show that all bounded piecewise monotone functions are Riemann integrable.*

We can define a piecewise monotone function as follows. Let be $f : I \to \mathbb{R}$ a function, with I a bounded interval. We say that f is piecewise monotone on I iff there exists a partition $\mathbf{P}$ of I such that, for all $J \in \mathbf{P}$, the restriction $f|_J$ is monotone on J.

Now let's prove that all bounded piecewise monotone functions are Riemann integrable. Let be $f : I \to \mathbb{R}$ a bounded piecewise monotone function, and $\mathbf{P}$ a partition of I. For all $J \in \mathbf{P}$, we define a function $g_J : J \to \mathbb{R}$ as follows:

$$\begin{cases} g_J(x) := f|_J(x) & \text{if } x \in J \\ g_J(x) := 0 & \text{if } x \in I - J \end{cases}$$

Note that, since each $f|_J$ is piecewise monotone and bounded, then by Corollary 11.6.3, for all $J \in \mathbf{P}$, $f|_J$ is Riemann integrable on J. By Theorem 11.4.1(g), it implies that the functions g_J defined above are Riemann-integrable on I.

Furthermore, we have clearly

$$f = \sum_{J \in \mathbf{P}} g_J$$

so that, by Theorem 11.4.1(a), f is Riemann-integrable on I.

Exercise 11.6.3. — *Prove Proposition 11.6.4.*

Let be $f : [0, +\infty[\to \mathbb{R}$ a positive, decreasing function. We have to show that $\sum_{n=0}^{\infty} f(n)$ converges iff $\sup_{N>0} \int_{[0,N]} f$ is finite.

First observe that since f is decreasing and positive, it is bounded, and is thus Riemann-integrable on any interval of the form $[0, N]$ by Corollary 11.6.3.

Also, we have $f(n) \geqslant f(x) \geqslant f(n+1)$ for all x in an interval $[n, n+1[$. This motivates the following argument:

- Let be $\mathbf{P} := \{[n, n+1[: n \in \mathbb{N}\}$ a partition of $[0, +\infty[$.
- Let be the constant piecewise function $g : [0, +\infty[\to \mathbb{R}$ defined by $g(x) := f(n)$ when $x \in [n, n+1[$. The function g clearly majorizes f on $[0, +\infty[$.
- Let be the constant piecewise function $h : [0, +\infty[\to \mathbb{R}$ defined by $h(x) := f(n+1)$ when $x \in [n, n+1[$. The function g clearly minorizes f on $[0, +\infty[$.
- Thus, on any interval $[n, n+1[$ with $n \in \mathbb{N}$, we have:

  $$f(n) = \int_{[n,n+1[} g(x) \geqslant \int_{[n,n+1[} f \geqslant \int_{[n,n+1[} h = f(n+1)$$

  We can use Theorem 11.4.1(g) to sum over such intervals, to get:

  $$\sum_{n=0}^{N-1} f(n) \geqslant \int_{[0,N[} f \geqslant \sum_{n=0}^{N-1} f(n+1)$$

  i.e., using Remark 11.4.2, and Lemma 7.1.4,

  $$\sum_{n=0}^{N-1} f(n) \geqslant \int_{[0,N]} f \geqslant \sum_{n=1}^{N} f(n) \tag{11.14}$$

- In (11.14), if $\sum_{n=0}^{\infty} f(n)$ converges, then (since is equal to $\sup_N \sum_{n=0}^{N-1} f(n)$) this sum is an upper bound for $\int_{[0,N]} f$, so that the supremum $\sup_{n>0} \int_{[0,N]} f$ is necessarily finite.

- Conversely, if the supremum $\sup_{N>\int_{[0,N]} f} \int_{[0,N]} f$ is finite, then it is a (finite) upper bound for the $\sum_{n=1}^{N} f(n)$. Thus, the series $\sum_{n=0}^{\infty} f(n)$ converges, by Proposition 7.3.1.

Exercise 11.6.4. — *Give examples to show that both directions of the integral test break down if f is not assumed to be monotone decreasing.*

Consider the following two examples[34]:

- Let be the positive function $f : [0, +\infty[\to \mathbb{R}$ defined by:

$$\begin{cases} f(x) = 0 & \text{if } x \in [n, n + 1/4], n \in \mathbb{N} \\ f(x) = 1 & \text{if } x \in [n + 1/4, n + 3/4], n \in \mathbb{N} \\ f(x) = 0 & \text{if } x \in [n + 3/4, n + 1], n \in \mathbb{N} \end{cases}$$

  In particular, we have $f(n) = 0$ for all $n \in \mathbb{N}$; and $\int_{[n,n+1]} f = 1/2$ for all $n \in \mathbb{N}$. Thus, we clearly have $\sum_{n=0}^{\infty} f(n) = 0$, i.e. the series converges; but $\sup_{N>0} \int_{[0,N]} f = \sup_{N>0} N/2 = +\infty$.

- Let be the positive function $g : [0, +\infty[\to \mathbb{R}$ defined by:

$$\begin{cases} f(x) = 1 & \text{if } x \in \mathbb{N} \\ f(x) = 0 & \text{otherwise} \end{cases}$$

  For all $n \in \mathbb{N}$, we clearly have $\int_{[n,n+1]} f = 0$, and thus $\int_{[0,N]} f = 0$. Consequently, we also have $\sup_{N>0} \int_{[0,N]} f = 0$.

  On the other hand, we have $\sum_{n=0}^{N} f(n) = N$, so that $\sum_{n=0}^{\infty} f(n) = \infty$. Thus, the series diverges, even though the sup over the integrals $\int_{[0,N]} f$ is finite.

Exercise 11.8.1. — *Prove Lemma 11.8.4. (Hint: modify the proof of Theorem 11.1.13.)*

Let be X a closed interval, $I \subseteq X$ a closed interval, and $\alpha : X \to \mathbb{R}$ a function which is either monotone increasing or continuous, and $\mathbf{P}$ a partition of I. We have to show that

$$\sum_{J \in \mathbf{P}} \alpha[J] = \alpha[I] \tag{11.15}$$

As in the proof of Theorem 11.1.13, we use induction on $n := \#\mathbf{P}$.

- The base case $n = 0$ is trivial because in this case, $\mathbf{P}$ is empty, and thus I must be itself the empty set. In such a case, both sides of (11.15) are equal to 0.

  Similarly, if $n = 1$, then I must be a singleton $\{a\}$. In such a case, both sides of (11.15) are equal to $\alpha[\{a\}]$, so that the property is once again trivial.

- Now suppose inductively that the property is true when $\#\mathbf{P} = n$, and let's prove that it is still true when $\#\mathbf{P} = n + 1$.

  First, note that once again, the property is trivial when $I = \varnothing$ (since all elements of $\mathbf{P}$ are then the empty set themselves), and when $I = \{a\}$ (since they can only be $\{a\}$ or the empty set). Thus, we now consider that I is of the form $[a, b]$, $]a, b]$, $[a, b[$ or $]a, b[$ (with $a < b$ two real numbers).

[34] Note that we must still choose piecewise constant functions here, since we still do not know how to compute the integral of a more general function.

Now we split into two sub-cases:

1. If $b \in I$, then I is either $[a, b]$ or $]a, b]$. Furthermore, since $b \in I$, there exists exactly one $K \in \mathbf{P}$ such that $b \in K$. This interval K must be either $[c, b]$, $]c, b]$ or $\{b\}$, with $a \leqslant c \leqslant b$.

   To avoid splitting the proof into an very deep imbrication of sub-cases, let's consider that $I = [a, b]$ and $K = [c, b]$, but the proof is very similar in the other cases. Thus, $I - K = [a, c[$. We thus have:

$$\begin{aligned}
\alpha[I] &:= \Big(\lim_{x\to a^+} \alpha(x) - \alpha(a)\Big) + \Big(\lim_{x\to b^-} \alpha(x) - \lim_{x\to a^+} \alpha(x)\Big) + \Big(\alpha(b) - \lim_{x\to b^-} \alpha(x)\Big) \\
&= \alpha(b) - \alpha(a) \\
\alpha[I-K] &:= \Big(\lim_{x\to a^+} \alpha(x) - \alpha(a)\Big) + \Big(\lim_{x\to c^-} \alpha(x) - \lim_{x\to a^+} \alpha(x)\Big) \\
&= \lim_{x\to c^-} \alpha(x) - \alpha(a) \\
\alpha[K] &:= \Big(\lim_{x\to c^+} \alpha(x) - \lim_{x\to c^-} \alpha(x)\Big) + \Big(\lim_{x\to b^-} \alpha(x) - \lim_{x\to c^+} \alpha(x)\Big) + \Big(\alpha(b) - \lim_{x\to b^-} \alpha(x)\Big) \\
&= -\lim_{x\to c^-} \alpha(x) + \alpha(b)
\end{aligned}$$

   when α is supposed to be monotone, or more directly, using the simplified formula,

$$\begin{aligned}
\alpha[I] &:= \alpha(b) - \alpha(a) \\
\alpha[K] &:= \alpha(b) - \alpha(c) \\
\alpha[I-K] &:= \alpha(c) - \alpha(a)
\end{aligned}$$

   when α is supposed to be continuous and not necessarily monotone.

   In both cases, we can see that we have

$$\alpha[I] = \alpha[I-K] + \alpha[K]. \tag{11.16}$$

   And since $\mathbf{P} - \{K\}$ is a partition of $I - K$ of cardinality n, the induction hypothesis applies, and we get

$$\alpha[I] = \alpha[K] + \alpha[I-K] = \alpha[K] + \sum_{J \in \mathbf{P} - \{K\}} \alpha[J] = \sum_{J \in \mathbf{P}} \alpha[J]$$

   as expected. This closes the induction.

2. If $b \notin I$, then $I = [a, b[$ or $]a, b[$, and thus there exists exactly one $K \in \mathbf{P}$ which is equal to $]c, b[$ or $[c, b[$. We just have to follow the same approach as in the previous case.

Exercise 11.8.2. — *State and prove a version of Proposition 11.2.13 for the Riemann-Stieltjes integral.*

This proposition could be stated as follows:

Let be I a bounded interval, and $\mathbf{P}, \mathbf{P}'$ be partitions of I. Let be $f : I \to \mathbb{R}$ a function which is piecewise constant with respect to $\mathbf{P}$ and $\mathbf{P}'$, and let be $\alpha : I \to \mathbb{R}$ a continuous or increasing function. Then we have $p.c. \int_{[\mathbf{P}]} f \, d\alpha = p.c. \int_{[\mathbf{P}']} f \, d\alpha$.

To prove this statement, we just have to modify Exercise 11.2.13 (we will reuse here the notations are results introduced in this former exercise).

First let's consider $\mathbf{P}\#\mathbf{P'}$, which is a partition of I finer than both $\mathbf{P}$ and $\mathbf{P'}$ (Lemma 11.1.8). Furthermore, f is piecewise constant on $\mathbf{P}\#\mathbf{P'}$ (Lemma 11.2.7), so that $p.c. \int_{[\mathbf{P}\#\mathbf{P'}]} f \, d\alpha$ is well-defined.

We thus have:

$$\begin{aligned}
p.c. \int_{[\mathbf{P}\#\mathbf{P'}]} f \, d\alpha &:= \sum_{J \in \mathbf{P}\#\mathbf{P'}} c_J \, \alpha[J] && \text{(Definition 11.8.5)} \\
&= \sum_{(K,L) \in \mathbf{P} \times \mathbf{P'}} c_{K \cap L} \, \alpha[K \cap L] && \text{(Definition 11.1.6)} \\
&= \sum_{K \in \mathbf{P}} \sum_{L \in \mathbf{P'}} c_{K \cap L} \, \alpha[K \cap L] && \text{(Corollary 7.1.14)} \\
&= \sum_{K \in \mathbf{P}} c_K \Big(\sum_{L \in \mathbf{P'}} \alpha[K \cap L] \Big) && \text{(because } c_K = c_{K \cap L}\text{)} \\
&= \sum_{K \in \mathbf{P}} c_K (\alpha[K]) && \text{(Lemma 11.8.4)} \\
&=: p.c. \int_{[\mathbf{P}]} f \, d\alpha
\end{aligned}$$

The same argument would show that $p.c. \int_{[\mathbf{P}\#\mathbf{P'}]} f \, d\alpha = p.c. \int_{[\mathbf{P'}]} f \, d\alpha$, which closes the proof.