

Terence Tao

Analysis I (Exercise)

Third Edition

GRAMM CYAO

Update: Tuesday 16th November, 2021

Contents

2	The natural numbers	1
2.1	The Peano axioms	1
2.2	Addition	2
2.3	Multiplication	6
3	Set theory	1
3.1	Fundamentals	1
3.2	Russell's paradox (Optional)	11
3.3	Functions	13
3.4	Images and inverse images	19
3.5	Cartesian products	25
3.6	Cardinality of sets	35
3.7	*Zermelo-Fraenkel axioms of set theory	41
3.7.1	ZFC axiom in Analysis I of Terence Tao	41
3.7.2	ZFC axiom in Set Theory of Ralf Schindler	43
4	Integers and rationals	47
4.1	The integers	47
4.2	The rationals	52
4.3	Absolute value and exponentiation	58
4.4	Gaps in the rational numbers	65

5	The real numbers	69
5.1	Cauchy sequences	69
5.2	Equivalent Cauchy sequences	70
5.3	The construction of the real numbers	71
5.4	Ordering the reals	75
5.5	The least upper bound property	80
5.6	Real exponentiation, part I	82
6	Limits of sequences	89
6.1	Convergence and limit laws	89
6.2	The extended real number system	96
6.3	Suprema and Infima of sequences	98
6.4	Limsup Liminf, and limit points	100
6.5	Some standard limits	107
6.6	Subsequences	108
6.7	Real exponentiation, part II	111
7	Series	113
7.1	Finite series	113
7.2	Infinite series	124
7.3	Sums of non-negative numbers	129
7.4	Rearrangement of series	132
7.5	The root and ratio tests	132
8	Infinite sets	135
8.1	Countability	135
8.2	Summation on infinite sets	143
8.3	Uncountable sets	150
8.4	The axiom of choice	153
8.5	Ordered sets	155
9	Continuous functions on \mathbf{R}	161
9.1	Subsets of the real line	161

9.2	The algebra of real-valued functions	168
9.3	Limiting values of functions	170
9.4	Continuous functions	174
9.5	Left and right limits	178
9.6	The maximum principle	180
9.7	The intermediate value theorem	180
9.8	Monotonic functions	181
9.9	Uniform continuity	184
9.10	Limits at infinity	188
10	Differentiation of functions	189
10.1	Basic definitions	189
10.2	Local maxima, local minima, and derivatives	196
10.3	Monotone functions and derivatives	199
10.4	Inverse functions and derivatives	201
10.5	L'Hôpital's rule	204
11	The Riemann integral	207
11.1	Partitions	207
11.2	Piecewise constant functions	209
11.3	Upper and lower Riemann integrals	214
11.4	Basic properties of the Riemann integral	216
11.5	Riemann integrability of continuous functions	223
11.6	Riemann integrability of monotone functions	224
11.7	A non-riemann integrable function	227
11.8	The Riemann-Stieltjes integral	228
11.9	The two fundamental theorems of calculus	235

CHAPTER II

Starting at the beginning: the natural numbers

2.1

The Peano axioms

Axiom 2.1. 0 is a natural number.

Axiom 2.2. If n is a natural number, then $n++$ is also a natural number.

Axiom 2.3. 0 is not the successor of any natural number; i.e., we have $n++ \neq 0$ for every natural number n .

Axiom 2.4. Different natural numbers must have different successors; i.e., if n, m are natural numbers and $n \neq m$, then $n++ \neq m++$. Equivalently, if $n++ = m++$, then we must have $n = m$.

Axiom 2.5 (Principle of mathematical induction). Let $P(n)$ be any property pertaining to a natural number n . Suppose that $P(0)$ is true, and suppose that whenever $P(n)$ is true, $P(n++)$ is also true. Then $P(n)$ is true for every natural number n .

2.2**Addition**

Why 2.2.1. Notice that we can prove easily, using Axioms 2.1, 2.2, and induction (Axiom 2.5), that the sum of two natural numbers is again a natural number (why?, p.25).

Proof. Let n and m be natural numbers, we prove that $n + m$ is a natural number. Use induction on n , and firstly we consider the base case $n = 0$ which is a natural number according to Axiom 2.1. By definition of addition, $0 + m = m$ is exactly a natural number. Now suppose inductively that $n + m$ is a natural number. We now have to show that $(n++) + m$ is a natural number. By definition of addition that $(n++) + m = (n + m)++$ and $n + m$ is a natural number. Then by Axiom 2.2, the successor of $n + m$, i.e. $(n + m)++$, is a natural number. This close the induction. \square

Exercise 2.2.1. Prove Proposition 2.2.5. (Hint: fix two of the variables and induct on the third.)

Proposition 2.2.5 (Addition is associative). For any natural numbers a, b, c , we have $(a + b) + c = a + (b + c)$.

Proof. We use induction on c and keeping a and b fixed. First consider the base case $c = 0$. As we have proved that the sum of two natural numbers is a natural number. By Definition 2.2.1 and Lemma 2.2.2, we have $(a + b) + 0 = a + b$ and $a + (b + 0) = a + b$. Associativity of addition be hold under the base case $c = 0$. Now we assume inductively that $(a + b) + c = a + (b + c)$; we have to show that $(a + b) + (c++) = a + (b + (c++))$. For the left side of equation, by Lemma 2.2.3, $(a + b) + (c++) = (a + b + c)++$. And for the right side, use Lemma 2.2.2 twice, we have $a + (b + (c++)) = a + ((b + c)++) = (a + b + c)++$. This close the induction. \square

Exercise 2.2.2. Prove Lemma 2.2.10. (Hint: use induction.)

Lemma 2.2.10. Let a be a positive number. Then there exists exactly one natural number b such that $b++ = a$.

Proof. We induct on b . Firstly, consider the base case $b = 0$, and we have $0++ = a$. We know a equals to 1 and it is not equal to 0 from Axiom 2.2 and Axiom 2.3, so 1 is a positive number. Now we suppose inductively that $b++ = a$ is a positive number. We need to show that $(b++)++$ also a positive number. From our assumption we have $(b++)++ = a++$. Since a is a positive number and not equal to 0, by Axiom 2.3, $a++$ not equals to 0 too. Thus $a++$ a positive number by Definition 2.2.7. This close our induction. \square

Exercise 2.2.3. Prove Proposition 2.2.12. (Hint: you will need many of the preceding propositions, corollaries, and lemmas.)

Proposition 2.2.12 (Basic properties of order for natural numbers).

Let a, b, c be natural numbers. Then

- (a) (Order is reflexive) $a \geq a$.
- (b) (Order is transitive) If $a \geq b$ and $b \geq c$, then $a \geq c$.
- (c) (Order is anti-symmetric) If $a \geq b$ and $b \geq a$, then $a = b$.
- (d) (Addition preserves order) $a \geq b$ if and only if $a + c \geq b + c$.
- (e) $a < b$ if and only if $a++ \leq b$.
- (f) $a < b$ if and only if $b = a + d$ for some positive number d .

Proof. (a) Suppose for contradiction that $a < a$. By Definition 2.2.11, we have $a \leq a$ and $a \neq a$. But this is impossible that $a \neq a$, thus $a \leq a$.

(b) Suppose that $a \geq b$ and $b \geq c$. By Definition 2.2.11, there exists two natural numbers n and m such that $a = b + n$ and $b = c + m$. Therefore, we have $a = (c + m) + n = c + (m + n)$ by associativity of addition. Because the sum of two natural number is still a natural number, then we successfully found a natural number $z = n + m$ such that $a = c + z$, and $a \geq c$ by

Definition 2.2.11.

(c) Suppose that $a \geq b$ and $b \geq a$. By Definition 2.2.11, there exists two natural number n and m such that $a = b + n$ and $b = a + m$. Then we have $a = (a + m) + n = a + (n + m)$, this equation holds if and only if $n + m = 0$ by Lemma 2.2.2. By Corollary 2.2.9, we have $n = 0$ and $m = 0$. Thus we have $a = b$ and $b = a$.

(d) Suppose that $a \geq b$. Then there is a natural number n let $a = b + n$. By cancellation law, commutativity and associativity of addition, we have $c + a = c + b + n$ implies that $a + c = (b + c) + n$. According to Definition 2.2.11, we have $a + c \geq b + c$. Now we suppose that $a + c \geq b + c$. In the same way, we have $a + c = (b + c) + n$ for some natural number n . By cancellation law, we have $a \geq b$, as desired.

(e) Suppose that $a < b$, by Definition 2.2.11, $b = a + n$ for some n and $a \neq b$. Then we can know that n is a positive number. By Lemma 2.2.10, there exists a natural number m such that $b = a + (m++)$. According to Definition 2.2.1 and Lemma 2.2.3, $b = a + (m++) = (a + m)++ = (a++) + m$, and $b \geq a++$. Otherwise, if $b \geq a++$ be hold, we have $b = (a++) + n$ for some n , and it is same possesses and we have $b > a$. So we proved two expresses is equivalent.

(f) If $a < b$ we have $b = a + d$ for some natural number, and $a \neq b$. Then we know that $d \neq 0$, so natural number d must be a positive number. We now suppose that $b = a + d$ for some positive number d , then d also a natural number which not equal to 0. Because $b = a$ if and only if $b = a + n$ for $n = 0$ be hold. So $b \neq a$, and $a < b$. \square

Exercise 2.2.4. *Justify the three statements marked (why?) in the proof of Proposition 2.2.13.*

Proof.

- When $a = 0$ we have $0 \leq b$ for all b (why?).

Because for all b there exists a natural number n which equals to b let $b = 0 + b$.

- If $a > b$, then $a ++ > b$ (why?).

If $a > b$, we have $a = b + n$ for some natural number and $b \neq a$. Because $a ++ = (b + n) ++ = b + (n ++)$ and $n ++$ still a natural number. So we have $a ++ > b$.

- If $a = b$, then $a ++ > b$ (why?).

If $a = b$, then $a = b + 0$, and $a ++ = (b + 0) ++ = b + (0 ++)$ $= b + 1$. Because $1 \neq 0$, $a ++ \neq b$. So $a ++ > b$. \square

Exercise 2.2.5. Prove Proposition 2.2.14. (Hint: define $Q(n)$ to be the property that $P(m)$ is true for all $m_0 \leq m < n$; note that $Q(n)$ is vacuously true when $n \leq m_0$.)

Proposition 2.2.14 (Strong principle of induction). 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 $m \geq m_0$, we have the following implication: if $P(m)$ is true for all natural numbers $m_0 \leq m' < m$, 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 $m \geq m_0$.)

Proof. We define $Q(n)$ to be the property that $P(m)$ is true for all $m_0 \leq m < n$. Now we just need to prove the truth of $Q(n)$. We use induction on n . Consider base case $n = 0$. $P(m)$ is vacuously true since there doesn't exist an $m_0 < 0$. Now we assume inductively that $P(m)$ is true for all $m_0 \leq m < n$. We have to prove that $P(m)$ is true for all $m_0 \leq m < n ++$. It holds as we have proved in Exercise 2.2.4. This closes the induction, and $Q(n)$ is true for all natural numbers n . Because m is a natural number, then we have $Q(m)$ is true when $P(m')$ is true for all $m_0 \leq m' < m$. Then we conclude that $P(m)$ is true for all natural numbers $m \geq m_0$. \square

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 $m \leq n$; this is known as the **principle of backwards induction**. (Hint: apply induction to the variable n .)

Proof. We use induction on n . First we consider the base case $n = 0$. $P(0)$ is true by hypothesis. Now we suppose inductively that $P(n)$ is true for all $m \leq n$. Then $P(n++)$ also is true for all $m \leq n < n++$. This closed the induction. \square

2.3

Multiplication

Exercise 2.3.1. Prove Lemma 2.3.2. (Hint: modify the proofs of Lemmas 2.2.2, 2.2.3 and Proposition 2.2.4.)

Lemma 2.3.2 (Multiplication is commutative). Let n, m be natural numbers. Then $n \times m = m \times n$.

Proof. We begin with some lemmas. Firstly we prove that for any natural number m and n , $m \times 0 = 0$ and $n \times (m++) = (n \times m) + n$.

We prove first statement. We use induction on m . The base case $0 \times 0 = 0$ follows since we know that $0 \times m = 0$ for any natural number m , and 0 is a natural number. Now suppose inductively that $m \times 0 = 0$. We wish to show that $(m++) \times 0 = 0$. But by definition of multiplication, $(n++) \times 0 = (n \times 0) + 0$, which is equal to 0 since $n \times 0 = 0$. This close the induction.

Now we prove second statement. We induct on n . We first consider the base case $n = 0$. In this case we have to prove $(0++) \times m = 1 \times m = (0 \times m) + m$. By definition of multiplication, $1 \times m = m$ and $n \times m = 0$, so both sides are equal to m and are thus equal to each other. Now we assume inductively that $n \times (m++) = (n \times m) + n$; we now have to show that $(n++) \times (m++) = ((n++) \times m) + (n++)$. The left-hand side is $(n \times (m++)) + (m++)$ by definition of multiplication, which is equal to $(n \times m) + m + n + 1$ by the inductive hypothesis. Similarly, we have

$(n++) \times m = (n \times m) + m$ by the definition of multiplication, and so the right-hand side is also equal to $(n \times m) + m + n + 1$. Thus both sides are equal to each other, and we have closed the induction.

Now we can begin to prove the commutativity of multiplication. We induct on n . First we do the base case $n = 0$, i.e. we show $0 \times m = m \times 0$. By the definition of multiplication, $0 \times m = 0$, while by the first statement we have proved, $m \times 0 = 0$. Thus the base case is done. Now suppose inductively that $n \times m = m \times n$, now we have to prove that $(n++) \times m = m \times (n++)$ to close the induction. By the definition of multiplication, $(n++) \times m = (n \times m) + m$. By the second statement we have proved, $m \times (n++) = (m \times n) + m$, but this is equal to $(n \times m) + m$ by the inductive hypothesis $n \times m = m \times n$. Thus $(n++) \times m = m \times (n++)$ and we have closed the induction. \square

Exercise 2.3.2. *Prove Lemma 2.3.3. (Hint: prove the second statement first.)*

Lemma 2.3.3 (Positive natural numbers have no zero divisors). Let n, m be natural numbers. Then $n \times m = 0$ if and only if at least one of n, m is equal to zero. In particular, if n and m are both positive, then nm is also positive.

Proof. First we prove the second statement. Suppose that a and b are two arbitrary natural numbers, then two positive numbers can be denoted as $n = a++$ and $m = b++$. Statement nm is positive is equivalent to $(a++) \times (b++) \neq 0$ by definition of positive. By definition of multiplication and its lemmas we proved in Exercise 2.3.1, $(a++) \times (b++) = (a \times b) + a + b + 1$, which is obviously not equal to zero. So nm is also not equal to zero and is a positive number.

Now we turn to look at the first statement. If both n and m not equal to zero, they are positive number by definition of positive. As we have proved former, it must be positive, i.e. $n \times m \neq 0$. If n or m is equal to zero, we have, as we have proved, $0 \times m = n \times 0 = 0 \times 0 = 0$, as desired. \square

Exercise 2.3.3. Prove Proposition 2.3.5. (Hint: modify the proof of Proposition 2.2.5 and use the distributive law.)

Proposition 2.3.5. (Multiplication is associative). For any natural numbers a, b, c , we have $(a \times b) \times c = a \times (b \times c)$.

Proof. We use induction on c and keeping a and c fixed. First consider the base case $c = 0$. Because multiplication of two natural number still a natural number by Lemma 2.3.3. So we have $(a \times b) \times 0 = 0$ and $a \times (b \times 0) = a \times 0 = 0$ by definition of multiplication and its inference. Associativity of multiplication be hold under the base case $c = 0$. Now we assume inductively that $(a \times b) \times c = a \times (b \times c)$; we have to show that $(a \times b) \times (c++) = a \times (b \times (c++))$. We firstly consider left side of equation. By reference of definition of multiplication, $(a \times b) \times (c++) = (a \times b \times c) + (a \times b)$. Then we look right side of equation, its process is same with left side and use distributive law, $a \times (b \times (c++)) = a \times ((b \times c) + b) = (a \times b \times c) + (a \times b)$. This close the induction. \square

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

Proof. We induct on a . First we consider the base case $a = 0$, we have $(0 + b)^2 = b^2$ and $0^2 + 2 \times 0 \times b + b^2 = b^2$. Thus the base case is done. Now suppose inductively that $(a + b)^2 = a^2 + 2ab + b^2$, and we have to prove that $((a++) + b)^2 = (a++)^2 + 2(a++)b + b^2$ to close the induction. We consider left side of equation by our inductive hypothesis

$$\begin{aligned} ((a++) + b)^2 &= ((a + b) + 1)^2 \\ &= (a + b)^2 + 2(a + b) + 1 \\ &= a^2 + 2ab + b^2 + 2(a + b) + 1. \end{aligned}$$

Then we consider the left side of equation

$$\begin{aligned} (a++)^2 + 2(a++)b + b^2 &= (a + 1)^2 + 2(a + 1)b + b^2 \\ &= a^2 + 2a + 1 + 2ab + 2b + b^2 \\ &= a^2 + 2ab + b^2 + 2(a + b) + 1. \end{aligned}$$

Then we look right side of equation is same with left side. This close the induction. \square

Exercise 2.3.5. *Prove Proposition 2.3.9. (Hint: fix q and induct on n .)*

Proposition 2.3.9 (Euclidean algorithm). Let n be a natural number, and let q be a positive number. Then there exist natural numbers m, r such that $0 \leq r < q$ and $n = mq + r$.

Proof. We use induction on n and keeping q fixed. Consider the base case $n = 0$. We can let $m = r = 0$ which is natural number, then $n = 0 = 0q + 0 = 0$. Now we suppose inductively that $n = mq + r$ and $0 \leq r < q$. We need to show that there exist natural numbers m, r let $n++ = mq + r$ and $0 \leq r < q$. First we consider that there exists a natural number r' that $0 \leq r = r' + 1 < q$, then $n++ = n + 1 = mq + r' + 1$ be satisfied by our inductive hypothesis. Then we turn to consider that $0 \leq r = r' + 1 = q$, let $m = m' + 1$ and $n + 1 = mq + q + r = q(m + 1) + r$. This close the induction. \square

CHAPTER III

Set theory

3.1

Fundamentals

Exercise 3.1.1. *Show that the definition of equality in Definition 3.1.4 is reflexive, symmetric, and transitive.*¹

Proof. First of all to prove that equality of sets is reflexive. Given any object $x \in X$. Because $x = x$ for any objects, so $X = X$ by definition of equality of sets. Then we prove that equality of sets is symmetric. Given two objects $x \in X$ and $y \in Y$ such that $X = Y$. We have $x = y$ by definition of equality of sets, then $y = x$ also true. So we have $Y = X$. We prove that equality of sets is transitive finally. Given three objects $x \in X, y \in Y$ and $z \in Z$ such that $X = Y$ and $Y = Z$. Then we have $x = y$ and $y = z$ such that $x = z$. So $X = Z$ by definition of equality of sets. \square

Proof. We prove by contradiction. Suppose there exist two empty sets \emptyset and \emptyset' which are not equal to each other, $\emptyset \neq \emptyset'$. By definition of empty

¹Definition 3.1.4 has to be given the status of an axiom (the axiom of extensionality) rather than a definition. Then Exercise 3.1.1 is now trivial and can be deleted. — Errata from Tao.

set, for any objects x such that $x \notin \emptyset$ and $x \notin \emptyset'$. Therefore, we know that $x \in \emptyset$ and $x \in \emptyset'$ for any objects x are always false. Then we can imply such conclusion that $\emptyset = \emptyset'$ is true by Definition 3.1.4 which become a vacuously true statement, a contradiction. (*Note: $\emptyset = \emptyset'$ is only situation for Definition 3.1.4. which is true, because is equivalent relation between two sets are qual and they have same elements. We must let the statements both side of “ \iff ” be false.*) \square

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

Proof. $\emptyset \neq \{\emptyset\}$. By Axiom 3.1, empty set \emptyset is a set and it also is an object. By Axiom 3.3, for any object $x \in \{\emptyset\}$ if and only if $x = \emptyset$. By Axiom 3.2, we have $x \notin \emptyset$. Thus there exists an element x lies in $\{\emptyset\}$ but not in \emptyset . By Definition 3.1.4, $\emptyset \neq \{\emptyset\}$.

$\{\emptyset\} \neq \{\{\emptyset\}\}$. By Axiom 3.3 and Definition 3.1.4, for $x \in \{\emptyset\}$ and $y \in \{\{\emptyset\}\}$, we have $x \neq y$ since $x = \emptyset$ and $y = \{\emptyset\}$. Hence, by Definition 3.1.4, $\{\emptyset\} \neq \{\{\emptyset\}\}$.

$\{\{\emptyset\}\} \neq \{\emptyset, \{\emptyset\}\}$. For every element y of $\{\{\emptyset\}\}$, by Axiom 3.3, we have $y = \{\emptyset\}$. Then we can find an element \emptyset of $\{\emptyset, \{\emptyset\}\}$ such that $y \neq \emptyset$. By Definition 3.1.4, $\{\{\emptyset\}\} \neq \{\emptyset, \{\emptyset\}\}$.

$\emptyset \neq \{\emptyset, \{\emptyset\}\}$. By Axiom 3.3, for any object $z \in \{\emptyset\}$ if and only if $z = \emptyset$ or $z = \{\emptyset\}$. But for either of them is not an element of \emptyset by Axiom 3.2. \square

Exercise 3.1.3. *Prove the remaining claims in Lemma 3.1.13.*

Lemma 3.1.13. If a and b are objects, then $\{a, b\} = \{a\} \cup \{b\}$. If A, B, C are sets, then the union operation is commutative (i.e., $A \cup B = B \cup A$) and associative (i.e., $(A \cup B) \cup C = A \cup (B \cup C)$). Also, we have $A \cup A = A \cup \emptyset = \emptyset \cup A = A$.

Proof. We first prove commutativity identity $A \cup B = B \cup A$. By Definition 3.1.4, we need to show that every element x of $A \cup B$ is an element of $B \cup A$,

and vice versa. Suppose that x is an element of $A \cup B$. By Axiom 3.4, this means that at least one of $x \in A$ or $x \in B$ is true. We divide into two cases. If $x \in A$, then by Axiom 3.4, $x \in B \cup A$ is true. If $x \in B$ instead, then $x \in B \cup A$ is still true. And a similar argument shows that every element of $B \cup A$ lies in $A \cup B$, and so $A \cup B = B \cup A$, as desired.

Now we prove that $A \cup A = A \cup \emptyset = \emptyset \cup A = A$. We first prove $A \cup A = A$. By Axiom 3.4, for every element $x \in A \cup A$ lies in $x \in A$, and vice versa. Then we prove $A \cup \emptyset = A$. Because \emptyset is an empty set, we have $x \notin \emptyset$ for any x . If $x \in A \cup \emptyset$, we only have $x \in A$. So for every element of $A \cup \emptyset$ lies in A . And a similar argument shows that every element of A lies in $A \cup \emptyset$, and so $A \cup \emptyset = A$. Because the union operation is associative, then we have $A \cup \emptyset = \emptyset \cup A$. So summarize our conclusion, we have $A \cup A = A \cup \emptyset = \emptyset \cup A = A$, as desired. \square

Exercise 3.1.4. *Prove the remaining claims in Proposition 3.1.18.*

Proposition 3.1.18 (Sets are partially ordered by set inclusion). Let A, B, C be sets. If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$. If $A \subseteq B$ and $B \subseteq A$, then $A = B$. Finally, if $A \subsetneq B$ and $B \subsetneq C$ then $A \subsetneq C$.

Proof. We prove the second claim. Suppose that $A \subseteq B$ and $B \subseteq A$. We have to show that every element of A is an element of B . So, let us pick an arbitrary element x of A . Then, since $A \subseteq B$, x must be an element of B . Now we pick an another arbitrary element y of B . Then, since $B \subseteq A$, y must be an element of A . So every element of A is an element of B and vice versa, as desired. \square

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

Proof. Firstly we prove that $A \subseteq B$ is equivalent to $A \cup B = B$. We suppose that $x \in A \subseteq B$, then by Definition 3.1.15, for any object x , $x \in A$ implies $x \in B$. Now, if $x \in A \cup B$, by Axiom 3.4, this means at least one of $x \in A$

or $x \in B$ is true. We divide into two cases. If $x \in A$ which implies $x \in B$; and if $x \in B$ which is our desired. So every element of $A \cup B$ is an element of B . A similar argument shows that every element of B lies in $A \cup B$. So $A \subseteq B$ implies $A \cup B = B$. Then we suppose that $A \cup B = B$ is true. By Definition 3.1.4, for every element of $A \cup B$ is an element of B , and vice versa. If $x \in A$, by Axiom 3.4, it also an element of B . So if $x \in A$ for any object x , it implies $x \in B$. Therefore, $A \subseteq B$ is equivalent to $A \cup B = B$.

Then we prove that $A \subseteq B$ is equivalent to $A \cap B = A$. We still suppose that $x \in A \subseteq B$. We assume that $x \in A \cap B$, by Definition 3.1.23, it means both of $x \in A$ and $x \in B$ is true. It is obviously that every element of $A \cap B$ is an element of A . While if $x \in A$, by hypothesis, it implies $x \in B$. Then both of $x \in A$ and $x \in B$ is true. So $A \cap B = A$ is true by Definition 3.1.23. Thus $A \subseteq B$ implies $A \cap B = A$. Now we suppose that $A \cap B$ is equal to A . Then by Definition 3.1.4, for every element of $A \cap B$ is an element of A , and vice versa. If $x \in A$, then we have $x \in B$, which by Definition 3.1.15, $A \subseteq B$, and $A \cap B = A$ implies $A \subseteq B$. So they are equivalent.

We have proved that the equivalency of $A \subseteq B$ and $A \cap B = A$, $A \subseteq B$ and $A \cup B = B$. So $A \cap B = A$ is equivalent to $A \cup B = B$. \square

Exercise 3.1.6. *Prove Proposition 3.1.28. (Hint: one can use some of these claims to prove others. Some of the claims have also appeared previously in Lemma 3.1.13.)*

Proof. Let A, B, C be sets, and let X be a set containing A, B, C as subsets.

(a) (Minimal element) We have $A \cup \emptyset = A$ and $A \cap \emptyset = \emptyset$.

Prove (a). We have proved that $A \cup \emptyset = A$ in Exercise 3.1.3. So only task for us is to prove $A \cap \emptyset = \emptyset$. We suppose that for any object x , $x \in A \cap \emptyset$. By Definition 3.1.23, $x \in A$ and $x \in \emptyset$. But by definition of empty set, $x \notin \emptyset$. So, $x \notin A \cap \emptyset$ for any object x which must be an empty set. Because empty set is unique. So $A \cap \emptyset = \emptyset$.

(b) (Maximal element) We have $A \cup X = X$ and $A \cap X = A$.

Prove (b). Because X contains A , and we have proved that it equivalent to $A \cup X = X$ and $A \cap X = A$.

(c) (*Identity*) We have $A \cap A = A$ and $A \cup A = A$.

Prove (c). We have proved that $A \cup A = A$ in Exercise 3.1.3. We want to prove that $A \cap A = A$. By Definition 3.1.23, we have $x \in A$ and $x \in A$, which is equivalent to $x \in A$. Therefore, every element of $A \cap A$ is an element of A . A similar argument shows that every element of A is an element of $A \cap A$. So, $A \cap A$ is equal to A .

(d) (*Commutativity*) We have $A \cup B = B \cup A$ and $A \cap B = B \cap A$.

Prove (d). We have proved the commutativity identity of union operation. We just prove intersection case, $A \cap B = B \cap A$. Suppose that $x \in A \cap B$, then by Definition 3.1.23, this means that both of $x \in A$ and $x \in B$ is true. Thus is naturally that $x \in B$ and $x \in A$ also true. By Definition 3.1.23 again, we have $x \in B \cap A$, and vice versa. So we proved that every element of $A \cap B$ is an element of $B \cap A$ and vice versa.

(e) (*Associativity*) We have $(A \cup B) \cup C = A \cup (B \cup C)$ and $(A \cap B) \cap C = A \cap (B \cap C)$.

Prove (e). The first statement have proved in the Lemma 3.1.13. We prove that $(A \cap B) \cap C = A \cap (B \cap C)$. By Definition 3.1.4, we need to show that every element of $(A \cap B) \cap C$ is an element of $A \cap (B \cap C)$, and vice versa. So first suppose that x is an element of $(A \cap B) \cap C$. By Definition 3.1.23, this means that both $x \in A \cap B$ and $x \in C$ is true. We use Definition 3.1.23 again, then we have both $x \in A$ and $x \in B$ are true. So we know that by Definition 3.1.23 $x \in (B \cap C)$ and $x \in A$ are true. So $x \in A \cap (B \cap C)$. Thus in all case we see that every element of $(A \cap B) \cap C$ lies in $A \cap (B \cap C)$. A similar argument shows that every element of $A \cap (B \cap C)$ is an element of $(A \cap B) \cap C$, and so $(A \cap B) \cap C = A \cap (B \cap C)$ as desired.

(f) (*Distributivity*) we have $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Prove (f). We prove the first statement, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. We need to prove that, by Definition 3.1.4, every element of $A \cap (B \cup C)$ is an element of $(A \cap B) \cup (A \cap C)$, and vice versa. We suppose that x is an element of $A \cap (B \cup C)$. By Definition 3.1.23, this means both $x \in A$ and $x \in (B \cup C)$ is true; by Axiom 3.4, at least one of $x \in B$ or $x \in C$ is true. We divide into two cases. If $x \in B$, then we have $x \in A$ and $x \in B$ is true, by Definition 3.1.23, $x \in A \cap B$. While if $x \in C$ is true, by Definition 3.1.23 again, $x \in A \cap C$. Then we have at least one of $x \in A \cap B$ or $x \in A \cap C$ is true. So $x \in (A \cap B) \cup (A \cap C)$ by Axiom 3.4. Thus in all cases we see that every element of $A \cap (B \cup C)$ is an element of $(A \cap B) \cup (A \cap C)$. A similar argument shows that every element of $(A \cap B) \cup (A \cap C)$ is an element of $A \cap (B \cup C)$, so $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ as desired. For the second statement, it is similar process to prove, so we have $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

(g) (*Partition*) We have $A \cup (X \setminus A) = X$ and $A \cap (X \setminus A) = \emptyset$.

Prove (g). We prove the first statement, $A \cup (X \setminus A) = X$. We need to prove that every element of $A \cup (X \setminus A)$ is an element of X . Assume that $x \in A \cup (X \setminus A)$. By Axiom 3.4 and Definition 3.1.27, this means that at least one of $x \in A$ or both $x \in X$ and $x \notin A$ are true is true. If $x \in A$ is true, because $A \subseteq X$, so we have $x \in X$. If $x \in X$ and $x \notin A$, then we still have $x \in X$. Thus in all cases we have every element of $A \cup (X \setminus A)$ is an element of X . A similar argument shows that every element of X is an element of $A \cup (X \setminus A)$. So $A \cup (X \setminus A) = X$. Now we prove the second statement, $A \cap (X \setminus A) = \emptyset$. We need to prove that there doesn't exist any element in $A \cap (X \setminus A)$. By Definition 3.1.23 and Definition 3.1.27, we have all of $x \in A$, $x \in X$ and $x \notin A$ is true. But it is impossible that both $x \in A$ and $x \notin A$ be true. So we can't find any object which satisfy the condition $x \in A \cap (X \setminus A)$. Then it is an empty set.

(h) (*De Morgan laws*) We have $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$ and $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Prove (h). We first prove that $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$. By Definition 3.1.27 and Axiom 3.4, for any object x , $x \in X$ and at least one of $x \notin A$ and $x \notin B$ is true. When $x \notin A$, we have $x \in X$ and $x \notin A$; while when $x \notin B$, then we have $x \in X$ and $x \notin B$ is true. By Definition 3.1.23, $x \in (X \setminus A) \cap (X \setminus B)$. A similar argument shows that $(X \setminus A) \cap (X \setminus B)$ implies that $X \setminus (A \cup B)$. So $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$. It same way to prove the second statement. \square

Exercise 3.1.7. Let A, B, C be sets. Show that $A \cap B \subseteq A$ and $A \cap B \subseteq B$. Furthermore, show that $C \subseteq A$ and $C \subseteq B$ if and only if $C \subseteq A \cap B$. In a similar spirit, show that $A \subseteq A \cup B$ and $B \subseteq A \cup B$, and furthermore that $A \subseteq C$ and $B \subseteq C$ if and only if $A \cup B \subseteq C$.

Proof. We prove that every element of $A \cap B$ is an element of A . Suppose that $x \in A \cap B$, by Definition 3.1.23, for any x is an element of A and B . Then we have every element of $A \cap B$ is an element of A . Similarly, we also have it is an element of B . So by Definition 3.1.15, $A \cap B \subseteq A$ and $A \cap B \subseteq B$ are true.

Then we turn to prove second claim. Assume that both of $C \subseteq A$ and $C \subseteq B$ are true. By Definition 3.1.15, $x \in C \implies x \in A$ and $x \in C \implies x \in B$. Which means $x \in C \implies x \in A \wedge x \in B$, By Definition 3.1.23, $x \in C \implies x \in A \cap B$ and again, by Definition 3.1.15, $C \subseteq A \cap B$. It is similar to prove that if for any element of x , $x \in C \subseteq A \cap B$ we have $x \in C \subseteq A$ and $x \in C \subseteq B$. Which is our desired.

According to Addition Rule in logics, if $x \in A$ is true, then we must have at least one of $x \in A$ or $x \in B$ is true, and vice versa. Then by Axiom 3.4 and Definition 3.1.15, the third claim be proved.

Finally, we prove the fourth claim. We suppose that for every element of x , both of $x \in A \subseteq C$ and $x \in B \subseteq C$ are true. By Axiom 3.4, at least one of $x \in A$ and $x \in B$ is true, then we have $x \in C$. And by Definition

3.1.15, $x \in A \cup B \subseteq C$. Now we need to show that every element of A and B is an element of C . Case of A or B , one of them is empty set is satisfied as we have proved in Exercise 3.1.6. We use contradiction. Suppose that there exists an object x_0 such that $x_0 \in B$ implies that $x_0 \notin C$ which B is a non-empty set. Then there exists an object $x_0 \in A \cup B$ which does not lie in C . It is a contradiction. So every element of B is an element of C . A similar argument shows that every element of A is an element of C , as desired. \square

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

Proof. We prove the first claim. We need to prove that every element of $A \cap (A \cup B)$ is an element of A . We suppose $x \in A \cap (A \cup B)$. By Definition 3.1.23, both of $x \in A$ and $x \in A \cup B$ are true. We see the second part, by Axiom 3.4, at least one of $x \in A$ or $x \in B$ is true. Because we have known that $x \in A$ must be true. So we have $x \in A$. While if we suppose that $x \in A$. Then at least one of $x \in A$ or $x \in B$ holds and, by Axiom 3.4, we have $x \in A \cup B$. By Definition 3.1.23 again, we have $x \in A \cap (A \cup B)$. So every element of $A \cap (A \cup B)$ is an element of A .

Then we prove the second claim. We use distributivity and where we replace C by A , then we have $A \cup (A \cap B) = (A \cup B) \cap (A \cup A)$. Because $A \cup A = A$ as we proved in Exercise 3.1.3 (or by Proposition 3.1.28). Then we found that $A \cup (A \cap B) = A \cap (A \cup B)$ which is equal to A , as desired. \square

Exercise 3.1.9. Let A, B, X be sets such that $A \cup B = X$ and $A \cap B = \emptyset$. Show that $A = X \setminus B$ and $B = X \setminus A$.

Proof. We have to prove that for any element $x \in X$, if it is an element of A , it would not be an element of B and vice versa. We use contradiction. Suppose that there exists such an element $x_0 \in X$ such that both of $x \in A$ and $x \in B$ are true. According to our hypothesis that $A \cap B = \emptyset$, by Definition 3.1.23, there doesn't exist any element x such that both of $x \in A$ and $x \in B$ are true. But because of $A \cup B = X$, by Axiom 3.4, we know that at least one of $x \in A$ or $x \in B$ is true. So we can conclude that $x \notin A$ when $x \in B$

and vice versa. And we also have every element of A or B is an element of X . Now we see the first claim that $A = X \setminus B$. If we suppose that $x \in A$, we must have $x \notin B$ and $x \in X$. By Definition 3.1.27, $x \in X \setminus B$. While if we suppose that $x \in X \setminus B$, by Definition 3.1.27, $x \in X$ and $x \notin B$, then we have $x \in A$. So every element of A is an element of $X \setminus B$. A similar argument shows that every element of B is an element of $X \setminus A$, as desired. \square

Exercise 3.1.10. *Let A and B be sets. Show that the three sets $A \setminus B$, $A \cap B$, and $B \setminus A$ are disjoint, and that their union is $A \cup B$.*

Proof. We first prove that $(A \setminus B) \cap (A \cap B) = \emptyset$. We need to show that for any element of $A \setminus B$ is not an element of $A \cap B$. Suppose that $x \in A \setminus B$. By Definition 3.1.27, both of $x \in A$ and $x \notin B$ hold, which contradictive with $x \in A \cap B$ by Definition 3.1.23. So there doesn't exist any element x such that $x \in (A \setminus B) \cap (A \cap B)$, and which is an empty set. An similar argument shows that there doesn't exist any element x such that $x \in (B \setminus A) \cap (A \cap B)$, and which is an empty set. And it is obviously that $(A \setminus B) \cap (B \setminus A) = \emptyset$.

Then we prove that $(A \setminus B) \cup (A \cap B) \cup (B \setminus A) = A \cup B$. We suppose that $x \in (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$, by Axiom 3.4, at least one of $x \in A \setminus B$ or $x \in (A \cap B)$ or $x \in B \setminus A$ is true. We divide into three cases. If $x \in A \setminus B$, by Definition 3.1.27, both of $x \in A$ and $x \notin B$ is true, then at least one of $x \in A$ or $x \in B$ also true. While if $x \in A \cap B$, by definition 3.1.23, both of $x \in A$ and $x \in B$ are true, so naturally at least one of $x \in A$ or $x \in B$ is true. If $x \in B \setminus A$, by Definition 3.1.27, both of $x \in B$ and $x \notin A$ is true, then at least one of $x \in A$ or $x \in B$ also true. So every element of $(A \setminus B) \cup (A \cap B) \cup (B \setminus A)$ is an element of $A \cup B$. A similar argument shows that every element of $A \cup B$ is an element of $(A \setminus B) \cup (A \cap B) \cup (B \setminus A)$. This is our desired. \square

Exercise 3.1.11. *Show that the axiom of replacement implies the axiom of specification.*

Proof. Let A be a set, and for each $x \in A$, let $P(x)$ be a property pertaining

to x . Define $Q(x, y)$ by

$$x = y \iff P(x) \text{ is true,}$$

for all x and y . Then for any $x \in X$ there is $y = x$ such that $Q(x, y)$ is hold, this implies that, by Axiom 3.6, for any $y = x$

$$(y \in A \text{ and } P(y) \text{ is true}) \iff y \in \{x \in A : P(x) \text{ is true}\}.$$

□

3.2**Russell's paradox (Optional)**

Exercise 3.2.1. *Show that the universal specification axiom, Axiom 3.8, if assumed to be true, would imply Axioms 3.2, 3.3, 3.4, 3.5, and 3.6. (If we assume that all natural numbers are objects, we also obtain Axiom 3.7.) Thus, this axiom, if permitted, would simplify the foundations of set theory tremendously (and can be viewed as one basis for an intuitive model of set theory known as “naive set theory”). Unfortunately, as we have seen, Axiom 3.8 is “too good to be true”!*

Proof.

- Axiom 3.2 (Empty set). Let $P(x)$ be the property that $x \neq x$. Then by Axiom 3.8,

$$y \in \{x : x \neq x \text{ is true}\} \iff y \neq y \text{ is true.}$$

Since there is no y such that $y \neq y$. Thus the set above is empty.

- Axiom 3.3 (Singleton sets and pair sets). Let $P(x)$ be the property that $x = a$. Then by Axiom 3.8,

$$y \in \{x : x = a \text{ is true}\} \iff y = a \text{ is true.}$$

And let $P(x)$ be the property that $x = a$ or $x = b$. Then by Axiom 3.8,

$$y \in \{x : x = a \text{ or } x = b \text{ is true}\} \iff x = a \text{ or } y = b \text{ is true.}$$

- Axiom 3.4 (Pairwise union). Let $P(x)$ be the property that $x \in A$ or $x \in B$. Then by Axiom 3.8,

$$y \in \{x : x \in A \text{ or } x \in B \text{ is true}\} \iff y \in A \text{ or } y \in B \text{ is true.}$$

- Axiom 3.5 (Axiom of specification). Let $P(x)$ be the property that $x \in A$ and $P(x)$ is true. Then by Axiom 3.8,

$$y \in \{x : x \in A \text{ and } P(x) \text{ is true}\} \iff y \in A \text{ and } P(y) \text{ is true.}$$

- Axiom 3.6 (Replacement). Let $P(y)$ be the property that there exists $x \in A$ such that $P(x, y)$ is true. Then by Axiom 3.8,

$$\begin{aligned} z \in \{y : P(y) \text{ is true}\} &= \{x : P(x, y) \text{ is true for some } x \in A\} \\ &\iff P(x, z) \text{ is true for some } x. \end{aligned}$$

- Axiom 3.7 (Infinite). Let $P(x)$ be the property that

$$(x = 0 \iff 0 \in \mathbf{N}) \wedge \forall x (x = n \iff (x \in \mathbf{N} \iff x ++ \in \mathbf{N})).$$

□

Exercise 3.2.2. Use the axiom of regularity (and the singleton set axiom) to show that if A is a set, then $A \notin A$. Furthermore, show that if A and B are two sets, then either $A \notin B$ or $B \notin A$ (or both).

Proof. Suppose that A is a non-empty set, and A is an element of A . By Axiom 3.3, we have $A \in \{A\}$. Then by Axiom 3.9, A is either not a set, or is disjoint from $\{A\}$. So we know that $A \cap \{A\} = \emptyset$. But since $A \in A$ and $A \in \{A\}$, this implies that $A \cap \{A\} = A$, a contradiction. While if A is an empty set. By Axiom 3.2, $A \notin A$.

Furthermore, suppose that A, B are two sets and we both have $A \in B$ and $B \in A$. By Axiom 3.3, we have $A, B \in \{A, B\}$. By Axiom 3.9, we have $A \cap \{A, B\} = \emptyset$ and $B \cap \{A, B\} = \emptyset$, but this is contradictive with $A \in B, B \in A$ and $A, B \in \{A, B\}$. □

Exercise 3.2.3. Show (assuming the other axioms of set theory) that the universal specification axiom, Axiom 3.8, is equivalent to an axiom postulating the existence of a “universal set” Ω consisting of all objects (i.e., for all objects x , we have $x \in \Omega$). In other words, if Axiom 3.8 is true, then a universal set exists, and conversely, if a universal set exists, then Axiom 3.8 is true. (This may explain why Axiom 3.8 is called the axiom of universal specification). Note that if a universal set Ω existed, then we would have $\Omega \in \Omega$ by Axiom 3.1, contradicting Exercise 3.2.2. Thus the axiom of foundation specifically rules out the axiom of universal specification.

Proof. Suppose that Axiom 3.8 is true, then we have

$$y \in \{x : x \in \Omega\}.$$

Conversely, if a universal set Ω exists, by Axiom 3.5, we have

$$y \in \{x \in \Omega : P(x) \text{ is true}\} \implies y \in \{x : P(x) \text{ is true}\}.$$

Hence we obtain Axiom 3.8. □

3.3

Functions

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 $f, \tilde{f} : X \rightarrow Y$ and $g, \tilde{g} : Y \rightarrow Z$ are functions such that $f = \tilde{f}$ and $g = \tilde{g}$, then $g \circ f = \tilde{g} \circ \tilde{f}$.²*

Proof. By Definition 3.3.7, $f(x)$ is an element of Y , so it also is an object. For all $x \in X$, if we have an object $f(x)$, by Reflexive axiom, $f(x) = f(x)$. So by Definition 3.3.7 again, $f = f$ as our desired.

Prove symmetry of equality. Given that $f = g$ for all $x \in X$. By Definition 3.3.7, this means that $f(x) = g(x)$. Then we have $g(x) = f(x)$ by Transitive axiom, then by Definition 3.3.7 again, $g = f$.

Then we prove transitivity of equality. Suppose that we have $f = g$ and $g = h$ for all $x \in X$. By Definition 3.3.7, this means that $f(x) = g(x)$ and $g(x) = h(x)$, So $f(x) = h(x)$ by Symmetric axiom, and by Definition 3.3.7 again, $f = h$.

Finally we prove the substitution property. First we suppose that $f = \tilde{f}$ and $g = \tilde{g}$ for all $x \in X, y \in Y$, where $f, \tilde{f} : X \rightarrow Y$ and $g, \tilde{g} : Y \rightarrow Z$. By Definition 3.3.7, $f(x) = \tilde{f}(x)$ and $g(x) = \tilde{g}(x)$. It is easy to

²Of course, these statements are immediate from the axioms of equality in Appendix A.7 applied directly to the functions in question, but the point of the exercise is to show that they can also be established by instead applying the axioms of equality to elements of the domain and range of these functions, rather than to the functions itself. — Errata from Tao.

verify that $g \circ f$ and $\tilde{g} \circ \tilde{f}$ are defined and have same domain and range: $g \circ f : X \rightarrow Z, \tilde{g} \circ \tilde{f} : X \rightarrow Z$. For any $x \in X$ we have $f(x) = \tilde{f}(x)$, so we have $(g \circ f)(x) = g(f(x)) = g(\tilde{f}(x))$ for all $x \in X$. And for all $\tilde{f}(x) \in Y$, $g(x) = \tilde{g}(x)$. So $(g \circ f)(x) = g(\tilde{f}(x)) = \tilde{g}(\tilde{f}(x)) = (\tilde{g} \circ \tilde{f})(x)$. By Definition 3.3.10, $g \circ f = \tilde{g} \circ \tilde{f}$. \square

Exercise 3.3.2. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ 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 $g \circ f$.

Proof. We prove the first claim. Suppose that f and g are injective function. By Definition 3.3.14, for all $x \in X$ and $y \in Y$, $f(x) = f(x') \implies x = x'$ and $g(y) = g(y') \implies y = y'$. Composition function $g \circ f$ is defined: $g \circ f : X \rightarrow Z$. Let $y = f(x), y' = f(x')$, then $g(f(x)) = g(f(x')) \implies f(x) = f(x')$, hence $x = x'$. So by Definition 3.3.10 and Definition 3.3.14, $(g \circ f)(x) = (g \circ f)(x') \implies x = x'$ and $g \circ f$ is injective.

Then we suppose that f and g are surjective. We need to prove that for every $z \in Z$, there exists $x \in X$ such that $(g \circ f)(x) = z$. We can find $x_0 \in X$ such that $f(x_0) = y_0$ for every $y \in Y$, and for every $z \in Z$, there exists $y_0 \in Y$ such that $g(y_0) = z_0$. So for every $z \in Z$ there exists $x \in X$ such that $g(f(x)) = z$, as desired. \square

Remark 3.3.2. This follows that $g \circ f$ is bijective if both f and g are bijective.

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

Proof. The empty function is always injective. We use contradiction, suppose that the empty function is not injective. By Definition 3.3.14, there exists $x, x' \in \emptyset$ such that $f(x) = f(x')$ and $x \neq x'$. It is impossible for that we can't find any x and x' from empty set.

The empty set is surjective if $Y = \emptyset$. Because if X is empty and Y is not, there is no $x \in X$ such that whatever statement holds. If they're both

empty, the statement is vacuously true, since there are no $y \in Y$.

Therefore we see that the empty function is bijective if both X and Y are empty set. \square

Exercise 3.3.4. *In this section we give some cancellation laws for composition. Let $f : X \rightarrow Y$, $\tilde{f} : X \rightarrow Y$, $g : Y \rightarrow Z$, and $\tilde{g} : Y \rightarrow Z$ be functions. Show that if $g \circ f = g \circ \tilde{f}$ and g is injective, then $f = \tilde{f}$. Is the same statement true if g is not injective? Show that if $g \circ f = \tilde{g} \circ f$ and f is surjective, then $g = \tilde{g}$. Is the same statement true if f is not surjective?*

Proof. Prove the first statement. Suppose that $g \circ f = g \circ \tilde{f}$ and g is injective. Which means that $g(f(x)) = g(\tilde{f}(x))$ for all $x \in X$, and for all $y, y' \in Y$, $g(y) = g(y') \implies y = y'$ by Definition 3.3.7, Definition 3.3.10 and Definition 3.3.14. Because $f(x)$ and $\tilde{f}(x)$ are elements of Y , let $y = f(x)$ and $y' = \tilde{f}(x)$, we have $f(x) = \tilde{f}(x)$ for all $x \in X$. So by Definition 3.3.7, $f = \tilde{f}$.

If we suppose that g is not injective, then there exist $y, y' \in Y$ such that $g(y) = g(y')$ and $y \neq y'$. For all $x \in X$ we have $g(f(x)) = g(\tilde{f}(x))$. Because $f(x)$ and $\tilde{f}(x)$ are elements of Y , so they are not equal to each other: $f(x) \neq \tilde{f}(x)$ for some $x \in X$ is true. So $f \neq \tilde{f}$.

Then we prove the second statement. Suppose that $g \circ f = \tilde{g} \circ f$ and f is surjective. Which means that $g(f(x)) = \tilde{g}(f(x))$ for every $x \in X$, and for every $y \in Y$, there exists $x \in X$ such that $f(x) = y$. So for arbitrary $y \in Y$ we can find an element $x_0 \in X$ such that $f(x_0) = y$, so $g(y) = \tilde{g}(y)$ for every $y \in Y$. By Definition 3.3.7, $g = \tilde{g}$.

If f is not surjective, for every $x \in X$, there exists $y_0 \in Y$ such that $f(x) \neq y_0$. In this case composition is not defined. \square

Exercise 3.3.5. *Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ 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 also be surjective?*

Proof. We prove the first claim. We use contradiction. Suppose that f is

not injective, this means that there exists $x, x' \in X$ such that $f(x) = f(x')$ and $x \neq x'$. Since $g \circ f$ is injective, by Definition 3.3.14, $x \neq x'$ implies $g(f(x)) \neq g(f(x'))$. This is contradictory that $g(f(x)) \neq g(f(x'))$ and $f(x) = f(x')$ by Definition 3.3.1. So f is injective.

g is not necessary to be injective. There is a counterexample: Let \mathbf{R} denote the real numbers and let $\overline{\mathbf{R}}_+$ denote the non-negative real numbers. Suppose that $f : \overline{\mathbf{R}}_+ \rightarrow \mathbf{R}$ and $g : \mathbf{R} \rightarrow \mathbf{R}$, then $g \circ f : \overline{\mathbf{R}}_+ \rightarrow \mathbf{R}$. Define f, g by

$$f(x) = x^2, \quad g(y) = y^2.$$

We can see that f is injective and g is surjective. Then $g \circ f$ is defined by

$$(g \circ f)(x) = x^4.$$

Where is injective.

Now we prove the second claim. Use contradiction. Suppose that g is not surjective, this means that for all $y \in Y$, there exists $z \in Z$ such that $g(y) \neq z$. But $g \circ f$ is surjective, for all $z \in Z$, there exists $x \in X$ such that $g(f(x)) = z$. We can find such a $x_0 \in X$, satisfy that $g(f(x_0)) = z_0$. If we let $y_0 = f(x_0)$, then there exists y_0 such that $g(y_0) = z_0$, a contradiction.

g is not necessary to be injective. There is a counterexample: Suppose that $f : \overline{\mathbf{R}}_+ \rightarrow \mathbf{R}$ and $g : \mathbf{R} \rightarrow \overline{\mathbf{R}}_+$, then $g \circ f : \overline{\mathbf{R}}_+ \rightarrow \overline{\mathbf{R}}_+$. Define f, g by

$$f(x) = x^2, \quad g(y) = y^2.$$

We can see that f is injective and g is surjective. Then $g \circ f$ is defined by

$$(g \circ f)(x) = x^4.$$

Where is surjective. □

Remark 3.3.5. We summarize conclusions of Exercise 3.3.2, Exercise 3.3.5 here: the necessary condition for $g \circ f$ is injective or surjective is that f is injective or g is surjective.

Exercise 3.3.6. Let $f : X \rightarrow Y$ be a bijective function, and let $f^{-1} : Y \rightarrow X$ 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 (thus $(f^{-1})^{-1} = f$).

Definition (Inverse). If f is bijective, then for every $y \in Y$, there is exactly one x such that $f(x) = y$ (there is at least one because of surjectivity, and at most one because of injectivity). This value of x is denoted $f^{-1}(y)$; thus f^{-1} is a function from Y to X . We call f^{-1} the *inverse* of f .

Proof. We prove the first statement. Since f is bijection, for every $y \in Y$, there exists exactly one $x \in X$ such that $f(x) = y$. Then, by definition, for inverse $f^{-1} : Y \rightarrow X$, we have $f^{-1}(f(x)) = f^{-1}(y) = x$ for all $x \in X$.

Then we prove the second statement. By definition, there exists exactly one $x \in X$ such that $f(x) = y$. Hence $f(f^{-1}(y)) = f(x) = y$ for all $y \in Y$.

Now we can show that f^{-1} is invertible, i.e., it is injective, if f is injection. f^{-1} is injective: for all $y, y' \in Y$

$$f^{-1}(y) = f^{-1}(y') \implies f(f^{-1}(y)) = f(f^{-1}(y')) \implies y = y'.$$

f^{-1} is surjective: for every $x \in X$, there exists $y \in Y$ such that $y = f^{-1}(f(x)) = x$. Thus, by Definition 3.3.20, f^{-1} is invertible, and we have $(f^{-1})^{-1} = f$. \square

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

Proof. By Exercise, $g \circ f$ is bijective when both f and g are bijective.

Suppose that $g \circ f : X \rightarrow Z$ is a bijective function, and $(g \circ f)^{-1} : Z \rightarrow X$ be its inverse. We need to show that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. By Exercise 3.3.6, we have

$$\begin{aligned} (f^{-1} \circ g^{-1})((g \circ f)(x)) &= ((f^{-1} \circ g^{-1}) \circ (g \circ f))(x) \\ &= f^{-1}((g^{-1} \circ g)(f(x))) \\ &= f^{-1}(f(x)) = x, \end{aligned}$$

and

$$\begin{aligned} (g \circ f)((f^{-1} \circ g^{-1})(x)) &= ((g \circ f) \circ (f^{-1} \circ g^{-1}))(x) \\ &= g((f \circ f^{-1})(g^{-1}(x))) \\ &= g(g^{-1}(x)) = y. \end{aligned}$$

Hence $(f^{-1} \circ g^{-1})^{-1} = g \circ f = ((g \circ f)^{-1})^{-1}$, and we have $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. \square

Exercise 3.3.8. If X is a subset of Y , let $\iota_{X \rightarrow Y} : X \rightarrow Y$ be the inclusion map from X to Y , defined by mapping $x \mapsto x$ for all $x \in X$, i.e., $\iota_{X \rightarrow Y}(x) := x$ for all $x \in X$. The map $\iota_{X \rightarrow X}$ is in particular called the **identity map** on X .

- (a) Show that if $X \subseteq Y \subseteq Z$ then $\iota_{Y \rightarrow Z} \circ \iota_{X \rightarrow Y} = \iota_{X \rightarrow Z}$.
- (b) Show that if $f : A \rightarrow B$ is any function, then $f = f \circ \iota_{A \rightarrow A} = \iota_{B \rightarrow B} \circ f$.
- (c) Show that, if $f : A \rightarrow B$ is a bijective function, then $f \circ f^{-1} = \iota_{B \rightarrow B}$ and $f^{-1} \circ f = \iota_{A \rightarrow A}$.
- (d) Show that if X and Y are disjoint sets, and $f : X \rightarrow Z$ and $g : Y \rightarrow Z$ are functions, then there is a unique function $h : X \cup Y \rightarrow Z$ such that $h \circ \iota_{X \rightarrow X \cup Y} = f$ and $h \circ \iota_{Y \rightarrow X \cup Y} = g$.

Proof. Prove (a). For all $x \in X$, we have

$$(\iota_{Y \rightarrow Z} \circ \iota_{X \rightarrow Y})(x) = (\iota_{Y \rightarrow Z})(\iota_{X \rightarrow Y}(x)) = (\iota_{Y \rightarrow Z})(x) = x,$$

and

$$(\iota_{X \rightarrow Z})(x) = x.$$

Prove (b). For all $a \in A$ and $f(a) \in B$, we have

$$(f \circ \iota_{A \rightarrow A})(a) = f((\iota_{A \rightarrow A})(a)) = f(a),$$

and

$$(\iota_{B \rightarrow B} \circ f)(a) = (\iota_{B \rightarrow B})(f(a)) = f(a).$$

Prove (c). This immediately follows from Exercise 3.3.6 and the definition of identity map.

Prove (d). For all $w \in X \cup Y$, we define $h : X \cup Y \rightarrow Z$ by following rule, if $w \in X$, then $h(w) = f(w)$. While if $w \in Y$, then $h(w) = g(w)$. Then we have

$$(h \circ \iota_{X \rightarrow X \cup Y})(w) = h((\iota_{X \rightarrow X \cup Y})(w)) = f(w).$$

and

$$(h \circ \iota_{Y \rightarrow X \cup Y})(w) = h((\iota_{Y \rightarrow X \cup Y})(w)) = g(w).$$

□

3.4

Images and inverse images

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

Proof. By Definition 3.4.1, the forward image of V under f^{-1} is defined by $f^{-1}(V) = \{f^{-1}(y) : y \in V\}$. Suppose $x \in \{f^{-1}(y) : y \in V\}$. By Definition 3.4.1, there exists some $y_0 \in V$ such that $x_0 = f^{-1}(y_0)$ which is an element of X . Because f is bijective, for every $y \in V$ there exists $x_0 \in X$ such that $y_0 = f(x_0)$. So $f(x_0) \in V$ and then $x \in \{x \in X : f(x) \in V\}$. That's we proved $\{f^{-1}(y) : y \in V\} \subseteq \{x \in X : f(x) \in V\}$. By Definition 3.4.4, the inverse image of V under f is defined by $f^{-1}(V) = \{x \in X : f(x) \in V\}$. Suppose that $x \in \{x \in X : f(x) \in V\}$. Let $y = f(x)$. Because f^{-1} is bijective, then there exists $y \in V$ such that $x = f^{-1}(y)$. So $x \in \{f^{-1}(y) : y \in V\}$ by Definition 3.4.1. and we have $\{x \in X : f(x) \in V\} \subseteq \{f^{-1}(y) : y \in V\}$. So these two sets are equal to each other. □

Exercise 3.4.2. Let $f : X \rightarrow Y$ be a function from one set X to another set Y , 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 ?

Proof. Prove S is a subset of $f^{-1}(f(S))$. Suppose that $x \in S$, by Definition 3.4.1, this means that $f(x) \in f(S)$, and it is the subset of Y . By Definition 3.4.4, $x \in f^{-1}(f(S))$. Therefore, for every element of S is an element of $f^{-1}(f(S))$, and $S \subseteq f^{-1}(f(S))$. Now we going to prove $f^{-1}(f(S)) = S$ since f is injective. Suppose that f is injective, and let $x \in f^{-1}(f(S))$.

Then by Definition 3.4.4, $f(x) \in f(S)$, so $f(x) = f(y)$ for some $y \in S$ (by definition of function). Since f is injective, it follows that $x = y$, hence $x \in S$. So $f^{-1}(f(S)) \subseteq S$. Thus when f is injective, $f^{-1}(f(S)) = S$.

Prove $f(f^{-1}(U))$ is a subset of U . Suppose that $y \in f(f^{-1}(U))$. By Definition 3.4.1, there exists $x \in f^{-1}(U)$ such that $y = f(x)$, and by Definition 3.4.4, it implies $y \in U$. So for every element of $f(f^{-1}(U))$ is an element of U , and $f(f^{-1}(U)) \subseteq U$. Then we prove that $f(f^{-1}(U)) = U$ since f is surjective. Suppose that $y \in U$. Because for every $y \in U \subseteq Y$, there exists $x \in X$ such that $y = f(x)$. Then we can let $y = f(x)$ for some $x \in X$, and by Definition 3.4.4, x is an element of $f^{-1}(U)$ which is a subset of X . So there exists some $x \in f^{-1}(U)$ such that $y = f(x)$, this means y is an element of $f(f^{-1}(U))$. This is our desired and $f(f^{-1}(U))$ is equal to U . \square

Exercise 3.4.3. *Let A, B be two subsets of a set X , and let $f : X \rightarrow Y$ be a function. Show that $f(A \cap B) \subseteq f(A) \cap f(B)$, that $f(A) \setminus f(B) \subseteq f(A \setminus B)$, $f(A \cup B) = f(A) \cup f(B)$. For the first two statements, is it true that the \subseteq relation can be improved to $=$?*

Proof. Prove $f(A \cap B) \subseteq f(A) \cap f(B)$. Suppose that $y \in f(A \cap B)$. By Definition 3.4.1, there exists $y = f(x)$ for some $x \in A \cap B$. It holds both for $x \in A$ and $x \in B$. So by Definition 3.4.1 again, both of $y \in f(A)$ and $y \in f(B)$ are true, and $y \in f(A) \cap f(B)$.

Prove $f(A) \setminus f(B) \subseteq f(A \setminus B)$. Suppose that $y \in f(A) \setminus f(B)$, which means that $y \in f(A)$ but $y \notin f(B)$. By Definition 3.4.1, there exists $y = f(x)$ for some $x \in A$ but $x \notin B$, then $x \in A \setminus B$. So $y \in f(A \setminus B)$.

Prove $f(A \cup B) = f(A) \cup f(B)$. Suppose that $y \in f(A \cup B)$. By Definition 3.4.1, this means that $y = f(x)$ for some $x \in A \cup B$. If $x \in A$, we have $y \in f(A)$. While if $x \in B$, then $y \in f(B)$. So at least one of $y \in f(A)$ or $y \in f(B)$ is true, then $y \in f(A) \cup f(B)$. It is similar to prove that every element of $f(A) \cup f(B)$ is an element of $f(A \cup B)$. So these two sets are equal to each other.

For the first two statements, if f is injective that the \subseteq relation can

be improved to $=$. □

Exercise 3.4.4. Let $f : X \rightarrow Y$ be a function from one set X to another set Y , and let U, V be subsets of Y . Show that $f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$, that $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$, and that $f^{-1}(U \setminus V) = f^{-1}(U) \setminus f^{-1}(V)$.

Proof. Prove the first statement. Suppose that $x \in f^{-1}(U \cup V)$. By Definition 3.4.4, for every x , $x \in X$ and $f(x) \in U \cup V$. Then there at least one of $f(x) \in U$ or $f(x) \in V$ is true for $x \in X$. So by Definition 3.4.4 again, $x \in f^{-1}(U) \cup f^{-1}(V)$, and every element of $f^{-1}(U \cup V)$ is an element of $f^{-1}(U) \cup f^{-1}(V)$. A similar argument shows that every element of $f^{-1}(U) \cup f^{-1}(V)$ is an element of $f^{-1}(U \cup V)$. Thus these two sets are equal to each other.

Prove the second statement. Suppose that $x \in f^{-1}(U \cap V)$. By Definition 3.4.4, for every x , $x \in X$ and $f(x) \in U \cap V$. Then both of $f(x) \in U$ and $f(x) \in V$ are true for $x \in X$. So by Definition 3.4.4 again, $x \in f^{-1}(U) \cap f^{-1}(V)$, and every element of $f^{-1}(U \cap V)$ is an element of $f^{-1}(U) \cap f^{-1}(V)$. A similar argument shows that every element of $f^{-1}(U) \cap f^{-1}(V)$ is an element of $f^{-1}(U \cap V)$. Thus these two sets are equal to each other.

Prove the third statement. Suppose that $x \in f^{-1}(U \setminus V)$. By Definition 3.4.4, for every x , $x \in X$ and $f(x) \in U \setminus V$. Then both of $f(x) \in U$ and $f(x) \notin V$ are true for $x \in X$. So by Definition 3.4.4 again, $x \in f^{-1}(U) \setminus f^{-1}(V)$, and every element of $f^{-1}(U \setminus V)$ is an element of $f^{-1}(U) \setminus f^{-1}(V)$. A similar argument shows that every element of $f^{-1}(U) \setminus f^{-1}(V)$ is an element of $f^{-1}(U \setminus V)$. Thus these two sets are equal to each other. □

Exercise 3.4.5. Let $f : X \rightarrow Y$ be a function from one set X to another set Y . Show that $f(f^{-1}(S)) = S$ for every $S \subseteq Y$ if and only if f is surjective. Show that $f^{-1}(f(S)) = S$ for every $S \subseteq X$ if and only if f is injective.

Proof. See Exercise 3.4.2. □

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 $f^{-1}(\{1\})$.) See also Exercise 3.5.11.

Lemma 3.4.9. Let X be a set. Then the set

$$\{Y : Y \text{ is a subset of } X\}$$

is a set.

Proof. Let $P(f, Y)$ be the following formula:

$$(f \in \{0, 1\}^X \wedge Y = f^{-1}(\{1\})) \vee (f \notin \{0, 1\}^X \wedge Y = \emptyset).$$

Notice that for each $f \in \{0, 1\}^X$ there is at most one Y for $P(f, Y)$ is true. Then we use replacement axiom to conclude that there is a set,

$$\begin{aligned} z \in G &:= \{Y : P(f, Y) \text{ for some } f \in \{0, 1\}^X\} \\ &\iff P(f, z) \text{ is true for some } f \in \{0, 1\}^X, \end{aligned}$$

i.e.,

$$Y \in G \iff Y = f^{-1}(\{1\}) \text{ for some } f \in \{0, 1\}^X.$$

Now, we are going to show that for every element of G , $Y \in G$ is equivalent to $Y \subseteq X$. It is obviously that Y is a subset of X if we suppose that Y is an element of G by Definition 3.4.4. While we suppose that Y is a subset of X . Let $Y = \{1\}$, then by Definition 3.4.4, $f^{-1}(\{1\}) = \{x \in X : f(x) \in \{1\}\}$ for some $f \in \{0, 1\}^X$ is true. Otherwise, we have $Y = \emptyset$, and still have $Y \in G$. This is our desired. \square

Remark 3.4.6. We cannot begin by letting such a set $\{Y : Y \subseteq X\}$ because, by Axiom 3.5 and Axiom 3.6, we don't yet know that this object exists: that's what we're trying to prove. Axiom of Replacement and Specification only say that we chose some elements x from set A but we don't know how can we select some subsets Y from X .

Exercise 3.4.7. Let X, Y be sets. Define a **partial function** from X to Y to be any function $f : X' \rightarrow Y'$ whose domain X' is a subset of X , and whose range Y' is a subset of Y . Show that the collection of all partial functions from X to Y is itself a set. (Hint: use Exercise 3.4.6, the power

set axiom, the replacement axiom, and the union axiom.)³

Proof. We first fix some Y_0 which is a subset of Y . By Lemma 3.4.9, we construct a set which include all subsets X' of X : $2^X = \{X' : X' \subseteq X\}$. Then by replacement axiom, $Y^{X'} = \{f \in Y_0^{X'} : X' \in 2^X\}$, this is a set include every partial function from arbitrary X' to fixed Y_0 , it means that if there exists such a partial function $f : X_0 \rightarrow Y_0$ let X_0 is a subset of X , then we collect all of them; for all such partial function $f : X' \rightarrow Y_0$ we collect them form a set. Now we free the Y_0 and by Lemma 3.4.9, construct a set which include all subset Y' of Y : $2^Y = \{Y' : Y' \subseteq Y\}$, and by union axiom, $(Y')^{X'} = \{Y^{X'} : Y' \in 2^Y\}$. Thus we have a set of all partial function that from X to Y . \square

Exercise 3.4.8. Show that Axiom 3.4 can be deduced from Axiom 3.1, Axiom 3.3 and Axiom 3.11.

Proof. By Axiom 3.1, set A and B are objects. By Axiom 3.4, we have a set $X \in \{A, B\}$ if and only if $X = A$ or $X = B$, which X is a non-empty set with element of x . By Axiom 3.11, $x \in A$ for $A \in \{A, B\}$ or $x \in B$ for $B \in \{A, B\}$. So for any object x , $x \in A$ or $x \in B$, and $\bigcup\{A, B\} = A \cup B$. \square

³Some advices of Terence Tao that 1) You can begin by showing that the set of pairs (X', Y') , with $X' \subset X$ and $Y' \subset Y$, is a set. Then you can use the axiom of replacement with a predicate to capture the assertion that a given set is equal to $(Y')^{X'}$. 2) If one wishes to avoid the use of ordered pairs, one can use the axiom of replacement twice rather than once, first to collect all the partial functions with a fixed domain X' and arbitrary range Y' (or vice versa), and then again to allow X' to vary (the axiom of union will be useful here too).

³We write these formulae (p.60) here which will be used in Exercise 3.4.9, 3.4.10 and 3.4.11:

$$y \in \bigcup_{\alpha \in I} A_\alpha \iff (y \in A_\alpha \text{ for some } \alpha \in I) \quad (3.2)$$

$$\bigcap_{\alpha \in I} A_\alpha := \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} \quad (3.3)$$

$$y \in \bigcap_{\alpha \in I} A_\alpha \iff (y \in A_\alpha \text{ for all } \alpha \in I) \quad (3.4)$$

Exercise 3.4.9. Show that if β and β' are two elements of a set I , and to each $\alpha \in I$ we assign a set A_α , then

$$\{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\} = \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\},$$

and so the definition of $\bigcap_{\alpha \in I} A_\alpha$ defined in (3.3) does not depend on β . Also explain why (3.4) is true.

Proof. Suppose that $y \in \{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\}$. This means that $y \in A_\beta$ and $y \in A_\alpha$ for all $\alpha \in I$ is true. We assign a set $A_{\beta'}$ to β' , then $y \in A_{\beta'}$ and $y \in A_\alpha$ for all $\alpha \in I$ is true. By definition, $y \in \{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\}$. A similar argument shows that for every element of $\{x \in A_{\beta'} : x \in A_\alpha \text{ for all } \alpha \in I\}$ is an element of $\{x \in A_\beta : x \in A_\alpha \text{ for all } \alpha \in I\}$. So these two sets are equal to each other. \square

Exercise 3.4.10. Suppose that I and J are two sets, and for all $\alpha \in I \cup J$ let A_α be a set. Show that $(\bigcup_{\alpha \in I} A_\alpha) \cup (\bigcup_{\alpha \in J} A_\alpha) = \bigcup_{\alpha \in I \cup J} A_\alpha$. If I and J are non-empty, show that $(\bigcap_{\alpha \in I} A_\alpha) \cap (\bigcap_{\alpha \in J} A_\alpha) = \bigcap_{\alpha \in I \cup J} A_\alpha$.

Proof. We prove the first statement. Suppose that $x \in (\bigcup_{\alpha \in I} A_\alpha) \cup (\bigcup_{\alpha \in J} A_\alpha)$. By Axiom 3.4, $x \in \bigcup_{\alpha \in I} A_\alpha$ or $x \in \bigcup_{\alpha \in J} A_\alpha$. If $x \in \bigcup_{\alpha \in J} A_\alpha$, By (3.2), $x \in A_\alpha$ for some $\alpha \in J$. While if $x \in \bigcup_{\alpha \in I} A_\alpha$, By (3.2) again, $x \in A_\alpha$ for some $\alpha \in I$. So we have $x \in A_\alpha$ for some $\alpha \in I \cup J$ by Axiom 3.4. By (3.2) again, $x \in \bigcup_{\alpha \in I \cup J} A_\alpha$, as we desired. A similar argument shows that every element of $\bigcup_{\alpha \in I \cup J} A_\alpha$ is an element of $(\bigcup_{\alpha \in I} A_\alpha) \cup (\bigcup_{\alpha \in J} A_\alpha)$, so these sets are equal to each other.

Then we prove the second statement. Suppose that $x \in (\bigcap_{\alpha \in I} A_\alpha) \cap (\bigcap_{\alpha \in J} A_\alpha)$. This means that $x \in \bigcap_{\alpha \in I} A_\alpha$ and $x \in \bigcap_{\alpha \in J} A_\alpha$. By (3.4), $y \in A_\alpha$ for all $\alpha \in I$ and $y \in A_\alpha$ for all $\alpha \in J$. Because I and J are non-empty set, so we have $y \in A_\alpha$ for all $\alpha \in I \cup J$. Then by (3.4) again, $x \in \bigcap_{\alpha \in I \cup J} A_\alpha$. A similar argument shows that every element of $\bigcap_{\alpha \in I \cup J} A_\alpha$ is an element of $(\bigcap_{\alpha \in I} A_\alpha) \cap (\bigcap_{\alpha \in J} A_\alpha)$, so these sets are equal to each other. \square

Remark 3.4.10. Notice that we require I and J are non-empty sets in the second claim but in the first claim.

Exercise 3.4.11. Let X be a set, let I be a non-empty set, and for all $\alpha \in I$ let A_α be a subset of X . Show that

$$X \setminus \bigcup_{\alpha \in I} A_\alpha = \bigcap_{\alpha \in I} (X \setminus A_\alpha)$$

and

$$X \setminus \bigcap_{\alpha \in I} A_\alpha = \bigcup_{\alpha \in I} (X \setminus A_\alpha).$$

This should be compared with de Morgan's laws in Proposition 3.1.28 (although one cannot derive the above identities directly from de Morgan's laws, as I could be infinite).

Proof. We prove the first statement. Suppose that $x \in X \setminus \bigcup_{\alpha \in I} A_\alpha$. This means that $x \in X$ and $x \notin \bigcup_{\alpha \in I} A_\alpha$. By (3.2), $x \notin A_\alpha$ for all $\alpha \in I$. So $x \in X \setminus A_\alpha$ for all $\alpha \in I$. By (3.4), $x \in \bigcap_{\alpha \in I} (X \setminus A_\alpha)$. A similar argument shows that every element of $\bigcap_{\alpha \in I} (X \setminus A_\alpha)$ is an element of $X \setminus \bigcup_{\alpha \in I} A_\alpha$, so these two sets are equal to each other.

Then we prove the second statement. Suppose that $x \in X \setminus \bigcap_{\alpha \in I} A_\alpha$. This means that $x \in X$ and $x \notin \bigcap_{\alpha \in I} A_\alpha$. By (3.4), $x \notin A_\alpha$ for some $\alpha \in I$. Then $x \in X \setminus A_\alpha$ for some $\alpha \in I$. By (3.2), $x \in \bigcup_{\alpha \in I} (X \setminus A_\alpha)$. A similar argument shows that every element of $\bigcup_{\alpha \in I} (X \setminus A_\alpha)$ is an element of $X \setminus \bigcap_{\alpha \in I} A_\alpha$, so these sets are equal to each other. \square

3.5

Cartesian products

Exercise 3.5.1. Suppose we **define** the ordered pair (x, y) for any objects x and y by the formula $(x, y) := \{\{x\}, \{x, y\}\}$ (thus using several applications of Axiom 3.3). Thus for instance $(1, 2)$ is the set $\{\{1\}, \{1, 2\}\}$, $(2, 1)$ is the set $\{\{2\}, \{2, 1\}\}$, and $(1, 1)$ is the set $\{\{1\}\}$. Show that such a definition

indeed obeys the property (3.5), and also whenever X and Y are sets, the Cartesian product $X \times Y$ is also a set. Thus this definition can be validly used as a definition of an ordered pair. For an additional challenge, show that the alternate definition $(x, y) := \{x, \{x, y\}\}$ also verifies (3.5) and is thus also an acceptable definition of ordered pair. (For this latter task one needs the axiom of regularity, and in particular Exercise 3.2.2.)⁴

Proof. Prove that such a definition obeys the property (3.5). We need to show that

$$\{\{x\}, \{x, y\}\} = \{\{x'\}, \{x', y'\}\} \iff x = x' \text{ and } y = y'.$$

We first suppose the left-hand side and by Axiom 3.3, $\{x\}$ and $\{x, y\}$ are the elements of $\{\{x\}, \{x, y\}\}$; and $\{x'\}$ and $\{x', y'\}$ are the elements of $\{\{x'\}, \{x', y'\}\}$. Because singleton set is not equal to pair set and these two sets have same elements. So $\{x\} = \{x'\}$ and $\{x, y\} = \{x', y'\}$, and therefore $x = x'$ and $y = y'$.

Then we suppose right-hand side and use contradiction. Suppose $x = x'$, $y = y'$ and $\{\{x\}, \{x, y\}\} \neq \{\{x'\}, \{x', y'\}\}$. There is at least one object z such that $z \in (x, y)$ but $z \notin (x', y')$. By Axiom 3.3, ordered pair is defined by pair set, so if $z = \{x\}$, then $\{x\} = \{x'\}$ for $x = x'$, then $z \in (x', y')$. While if $z = \{x, y\}$, then $\{x, y\} = \{x', y'\}$ for $x = x'$ and $y = y'$, then $z \in (x', y')$. So for every elements of (x, y) is an element of (x', y') . This is contradictive.

Now we prove that $X \times Y$ is a set. Suppose that X and Y are non-empty set, then there exists objects $x \in X$ and $y \in Y$. By Axiom 3.4, $x, y \in X \cup Y$ and $\{x, y\} \subseteq X \cup Y$. We can generate a set $2^{\{x, y\}} = \{\emptyset, \{x\}, \{y\}, \{x, y\}\}$, where $\{\{x\}, \{x, y\}\} \subseteq 2^{\{x, y\}}$. Let $a = \{\{x\}, \{x, y\}\}$. For non-empty set X and Y there exists $x \in X$ and $y \in Y$ such that $a = \{\{x\}, \{x, y\}\}$. Thus $X \times Y$ is a set.

Now we consider the alternate definition $(x, y) := \{x, \{x, y\}\}$, prove this definition obeys the property (3.5). Suppose that $(x, y) = (x', y')$. If

⁴property (3.5):

$$(x, y) = (x', y') \iff (x = x' \text{ and } y = y') \quad (3.5)$$

x is a set, By Exercise 3.2.2, $x \notin x$ and $x \notin \{x, y\}$, latter one is obviously false. Thus by Axiom 3.9, x is not a set and $x \neq \{x, y\}$. So $x = x'$ and $\{x, y\} = \{x', y'\}$ implies $y = y'$. It is easy to show that $(x, y) = (x', y')$ when $x = x'$ and $y = y'$. This is our desired. \square

Exercise 3.5.2. Suppose we **define** an ordered n -tuple to be a surjective function $x : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X$ whose range is some arbitrary set X (so different ordered n -tuples are allowed to have different ranges); we then write x_i for $x(i)$, and also write x as $(x_i)_{1 \leq i \leq n}$. Using this definition, verify that we have $(x_i)_{1 \leq i \leq n} = (y_i)_{1 \leq i \leq n}$ if and only if $x_i = y_i$ for all $1 \leq i \leq n$. Also, show that if $(X_i)_{1 \leq i \leq n}$ are an ordered n -tuple of sets, then the Cartesian product, as defined in Definition 3.5.7, is indeed a set. (Hint: use Exercise 3.4.7 and the axiom of specification.)

Proof. Suppose that $(x_i)_{1 \leq i \leq n} = (y_i)_{1 \leq i \leq n}$. Because ordered n -tuple is a surjective function, for every $(x_i)_{1 \leq i \leq n}, (y_i)_{1 \leq i \leq n} \in X$, there exists $1 \leq i \leq n$ such that $(x_i)_{1 \leq i \leq n} = x_i$ and $(y_i)_{1 \leq i \leq n} = y_i$. So that $x_i = y_i$ for all $1 \leq i \leq n$. Now we suppose for sake of contradiction that $(x_i)_{1 \leq i \leq n} \neq (y_i)_{1 \leq i \leq n}$ and $x_i = y_i$ for all $1 \leq i \leq n$. By definition of surjection, for every $(x_i)_{1 \leq i \leq n} \in X$ and $(y_i)_{1 \leq i \leq n} \in X$, we can find an $1 \leq i' \leq n$ such that $x_{i'} \neq y_{i'}$, a contradiction.

Then we turn to show that Cartesian product $\prod_{1 \leq i \leq n} X_i$ is a set. Consider the partial function $x : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X_i$, we have $x_i \in X_i$. And we know that X_i is a subset of $\bigcup_{1 \leq i \leq n} X_i$. By Exercise 3.4.7, the collection of all partial functions from $\{i \in \mathbf{N} : 1 \leq i \leq n\}$ to $\bigcup_{1 \leq i \leq n} X_i$ is a set. Then by Axiom 3.10, we can write such a set by

$$(x_i)_{1 \leq i \leq n} \in \left(\bigcup_{1 \leq i \leq n} X_i \right)^{\{1 \leq i \leq n\}}.$$

Now we can use axiom of specification and obtain

$$\prod_{1 \leq i \leq n} X_i := \left\{ (x_i)_{1 \leq i \leq n} \in \left(\bigcup_{1 \leq i \leq n} X_i \right)^{\{1 \leq i \leq n\}} : x_i \in X_i \text{ for all } 1 \leq i \leq n \right\}$$

as desired. \square

Exercise 3.5.3. *Show that the definitions of equality for ordered pair and ordered n -tuple obey the reflexivity, symmetry, and transitivity axioms.*

Proof. Prove definition of equality for ordered pair. Reflexivity: given any two objects x and y , $x = x$ and $y = y$, then $(x, y) = (x, y)$. Symmetry: given any objects x, y and x', y' , if $(x, y) = (x', y')$, by Definition 3.5.1, $x = x'$ and $y = y'$. So $(x', y') = (x, y)$. Transitivity: given any objects x, y, x', y' and x'', y'' , if $(x, y) = (x', y')$ and $(x', y') = (x'', y'')$, by Definition 3.5.1, $x = x'$ and $y = y'$, $x' = x''$ and $y' = y''$. So $x = x''$ and $y = y''$, and $(x, y) = (x'', y'')$.

Then we consider ordered n -tuple. Reflexivity: Since $x_i = x_i$ for all $1 \leq i \leq n$, then $(x_i)_{1 \leq i \leq n} = (x_i)_{1 \leq i \leq n}$. Symmetry: Given $(x_i)_{1 \leq i \leq n} = (y_i)_{1 \leq i \leq n}$, then $x_i = y_i$ for all $1 \leq i \leq n$. Then $y_i = x_i$ for all $1 \leq i \leq n$, by definition, we have $(y_i)_{1 \leq i \leq n} = (x_i)_{1 \leq i \leq n}$. Transitivity: Given $(x_i)_{1 \leq i \leq n} = (y_i)_{1 \leq i \leq n}$ and $(y_i)_{1 \leq i \leq n} = (z_i)_{1 \leq i \leq n}$. Then $x_i = y_i$ and $y_i = z_i$ for all $1 \leq i \leq n$. Then we have $x_i = z_i$ for all $1 \leq i \leq n$. By definition, $(x_i)_{1 \leq i \leq n} = (z_i)_{1 \leq i \leq n}$. \square

Exercise 3.5.4. *Let A, B, C be sets. Show that $A \times (B \cup C) = (A \times B) \cup (A \times C)$, that $A \times (B \cap C) = (A \times B) \cap (A \times C)$, and that $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$. (One can of course prove similar identities in which the rôles of the left and right factors of the Cartesian product are reversed.)*

Proof. Prove $A \times (B \cup C) = (A \times B) \cup (A \times C)$. Suppose that $a \in A \times (B \cup C)$, by Definition 3.5.4, $a = (x, y)$ for some $x \in X$ and $y \in B \cup C$, then we have $y \in B$ or $y \in C$. If $y \in B$, then $a = (x, y)$ for some $x \in X$ and $y \in B$, by Definition 3.5.4 again, $a \in A \times B$. While if $y \in C$, $a = (x, y)$ for some $x \in X$ and $y \in C$, then $a \in A \times C$. So $a \in (A \times B) \cup (A \times C)$. A similar argument shows that every element of $(A \times B) \cup (A \times C)$ is an element of $A \times (B \cup C)$, so equality holds.

Prove $A \times (B \cap C) = (A \times B) \cap (A \times C)$. Suppose that $a \in A \times (B \cap C)$, by Definition 3.5.4, $a = (x, y)$ for some $x \in X$ and $y \in B \cap C$, then we have $y \in B$ and $y \in C$. Because $y \in B$, then $a = (x, y)$ for some $x \in X$ and $y \in B$, by Definition 3.5.4 again, $a \in A \times B$. Also we have $y \in C$, $a = (x, y)$ for some $x \in X$ and $y \in C$, then $a \in A \times C$. So $a \in (A \times B) \cap (A \times C)$.

A similar argument shows that every element of $(A \times B) \cap (A \times C)$ is an element of $A \times (B \cap C)$, so equality holds.

Prove $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$. Suppose that $a \in A \times (B \setminus C)$, by Definition 3.5.4, $a = (x, y)$ for some $x \in X$ and $y \in B \setminus C$, then we have $y \in B$ and $y \notin C$. Because $y \in B$, then $a = (x, y)$ for some $x \in X$ and $y \in B$, by Definition 3.5.4 again, $a \in A \times B$. But we have $y \notin C$, $a = (x, y)$ for some $x \in X$ and $y \notin C$, then $a \notin A \times C$. So $a \in (A \times B) \setminus (A \times C)$. A similar argument shows that every element of $(A \times B) \setminus (A \times C)$ is an element of $A \times (B \setminus C)$, so equality holds. \square

Exercise 3.5.5. Let A, B, C, D be sets. Show that $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$. Is it true that $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$? Is it true that $(A \times B) \setminus (C \times D) = (A \setminus C) \times (B \setminus D)$?

Proof. The first claim is easy to verify by Definition 3.5.4. We only indicate that the second claim have such relation: $(A \times B) \cup (C \times D) \subseteq (A \cup C) \times (B \cup D) = (A \times C) \cup (A \times D) \cup (B \times C) \cup (B \times D)$. And the third claim is false when $A = C$ or $B = D$. So the second and third claims are not hold. \square

Exercise 3.5.6. Let A, B, C, D be non-empty sets. Show that $A \times B \subseteq C \times D$ if and only if $A \subseteq C$ and $B \subseteq D$, and that $A \times B = C \times D$ if and only if $A = C$ and $B = D$. What happens if the hypotheses that the A, B, C, D are all non-empty are removed?

Proof. It is obvious by definition, we consider the case of empty set. If $A = C = \emptyset$, then there do not exists $x \in A$ such that $a \in A \times B \implies a \in C \times D$. So claims are false when we removed the hypotheses of non-empty. \square

Exercise 3.5.7. Let X, Y be sets, and let $\pi_{X \times Y \rightarrow X} : X \times Y \rightarrow X$ and $\pi_{X \times Y \rightarrow Y} : X \times Y \rightarrow Y$ be the maps $\pi_{X \times Y \rightarrow X}(x, y) := x$ and $\pi_{X \times Y \rightarrow Y}(x, y) := y$; these maps are known as the **co-ordinate functions** on $X \times Y$. Show that for any functions $f : Z \rightarrow X$ and $g : Z \rightarrow Y$, there exists a unique function $h : Z \rightarrow X \times Y$ such that $\pi_{X \times Y \rightarrow X} \circ h = f$ and $\pi_{X \times Y \rightarrow Y} \circ h = g$.

(Compare this to the last part of Exercise 3.3.8, and to Exercise 3.1.7.) This function h is known as the **direct sum** of f and g and is denoted $h = f \oplus g$.

Proof. We first prove the existence. Since $f : Z \rightarrow X$ and $g : Z \rightarrow Y$, we can define $h : Z \rightarrow X \times Y$ by following rule: $h(z) := (f(z), g(z))$. We can see that $f(z) \in X$ and $g(z) \in Y$. Then by definition, for all $z \in Z$

$$\begin{aligned} (\pi_{X \times Y \rightarrow X} \circ h)(z) &= \pi_{X \times Y \rightarrow X}(h(z)) = \pi_{X \times Y \rightarrow X}(f(z), g(z)) = f(z); \\ (\pi_{X \times Y \rightarrow Y} \circ h)(z) &= \pi_{X \times Y \rightarrow Y}(h(z)) = \pi_{X \times Y \rightarrow Y}(f(z), g(z)) = g(z). \end{aligned}$$

Now we prove h is uniqueness. Suppose there exists another function $\tilde{h} : Z \rightarrow X \times Y$ satisfying the equation. Then for all $z \in Z$, we have

$$\begin{aligned} (\pi_{X \times Y \rightarrow X} \circ \tilde{h})(z) &= \pi_{X \times Y \rightarrow X}(f(z), g(z)) = f(z); \\ (\pi_{X \times Y \rightarrow Y} \circ \tilde{h})(z) &= \pi_{X \times Y \rightarrow Y}(f(z), g(z)) = g(z). \end{aligned}$$

So that $\tilde{h}(z) = (f(z), g(z)) = h(z)$, as desired. \square

Exercise 3.5.8. Let X_1, \dots, X_n be sets. Show that the Cartesian product $\prod_{i=1}^n X_i$ is empty if and only if at least one of the X_i is empty.

Proof. Suppose $\prod_{i=1}^n X_i$ is an empty set, use the contrapositive of Lemma 3.5.12, it immediately follows that there exists an X_i is empty set. Now suppose that $X_{i'}$ is an empty set for some $1 \leq i' \leq n$, then by definition, for all

$$(x_{i'})_{1 \leq i' \leq n} \notin \left(\bigcup_{1 \leq i \leq n} X_i \right)^{\{1 \leq i \leq n\}}.$$

such that $x_{i'} \in X_{i'}$ for all $1 \leq i' \leq n$. \square

Exercise 3.5.9. Suppose that I and J are two sets, and for all $\alpha \in I$ let A_α be a set, and for all $\beta \in J$ let B_β be a set. Show that $(\bigcup_{\alpha \in I} A_\alpha) \cap (\bigcup_{\beta \in J} B_\beta) = \bigcup_{(\alpha, \beta) \in I \times J} (A_\alpha \cap B_\beta)$.

Proof. Suppose that $x \in (\bigcup_{\alpha \in I} A_\alpha) \cap (\bigcup_{\beta \in J} B_\beta)$. Then we have $x \in (\bigcup_{\alpha \in I} A_\alpha)$ and $x \in (\bigcup_{\beta \in J} B_\beta)$. By definition, this means that $x \in A_\alpha$ for

some $\alpha \in I$ and $x \in B_\beta$ for some $\beta \in J$. By Definition 3.5.4, we can see that there exists a pair (α, β) for some $\alpha \in I$ and $\beta \in J$. Then we have $x \in A_\alpha \cup B_\beta$ for some $(\alpha, \beta) \in I \times J$. By definition again, $x \in \bigcup_{(\alpha, \beta) \in I \times J} (A_\alpha \cap B_\beta)$. A similar argument show that for every element of $\bigcup_{(\alpha, \beta) \in I \times J} (A_\alpha \cap B_\beta)$ is an element of $(\bigcup_{\alpha \in I} A_\alpha) \cap (\bigcup_{\beta \in J} B_\beta)$. So these two sets are equal to each other, as desired. \square

Exercise 3.5.10. *If $f : X \rightarrow Y$ is a function, define the **graph** of f to be the subset of $X \times Y$ defined by $\{(x, f(x)) : x \in X\}$. Show that two functions $f : X \rightarrow Y$, $\tilde{f} : X \rightarrow Y$ are equal if and only if they have the same graph. Conversely, if G is any subset of $X \times Y$ with the property that for each $x \in X$, the set $\{y \in Y : (x, y) \in G\}$ has exactly one element (or in other words, G obeys the **vertical line test**), show that there is exactly one function $f : X \rightarrow Y$ whose graph is equal to G .*

Proof. Suppose that $f = \tilde{f}$, by definition, for every $x \in X$, we have $f(x) = \tilde{f}(x)$. So $(x, f(x)) = (x, \tilde{f}(x))$ for all $x \in X$. Then suppose that $(x, f(x)) = (x, \tilde{f}(x))$ for all $x \in X$, by definition, $f(x) = \tilde{f}(x)$ for all $x \in X$. Thus $f = \tilde{f}$.

Since G obeys the vertical line test, we let $f(x) = y$. If there exists another function $f' : X \rightarrow Y$ whose graph is equal to G , i.e., $G = \{(x, f'(x)) : x \in X\}$. Then we have $(x, f'(x)) \in G$, and it satisfying that that $\{f'(x) \in Y : (x, f'(x)) \in G\}$. But for such x , the element of the set is uniqueness, so we have $f'(x) = y$. So for every $x \in X$, $f(x) = f'(x)$, $f = f'$. \square

Exercise 3.5.11. *Show that Axiom 3.10 can in fact be deduced from Lemma 3.4.9 and the other axioms of set theory, and thus Lemma 3.4.9 can be used as an alternate formulation of the power set axiom. (Hint: for any two sets X and Y , use Lemma 3.4.9 and the axiom of specification to construct the set of all subsets of $X \times Y$ which obey the vertical line test. Then use Exercise 3.5.10 and the axiom of replacement.)*

Proof. By Lemma 3.4.9, let $\mathcal{P}(X)$ and $\mathcal{P}(Y)$ be the sets that including all

of the subsets of X and Y . Then of course $\mathcal{P}(X) \cup \mathcal{P}(Y)$ is a set, and so does $\mathcal{P}(\mathcal{P}(X) \cup \mathcal{P}(Y))$ by Lemma 3.4.9. Now we use axiom of specification to generate the set $X \times Y$,

$$X \times Y := \{a \in \mathcal{P}(\mathcal{P}(X) \cup \mathcal{P}(Y)) : a = (x, y) \text{ for all } x \in X \text{ and } y \in Y\}.$$

And we have $\mathcal{P}(X \times Y) \subseteq X \times Y$. By Exercise 3.5.10, $\mathcal{P}(X \times Y)$ obeys the vertical line test, then there is exactly one function $f : X \rightarrow Y$ whose graph is equal to $\mathcal{P}(X \times Y)$. By axiom of replacement, we have

$$Y^X = \left\{ f \mid \begin{array}{l} \text{the graph of } f \text{ equal to } \mathcal{P}(X \times Y) \\ \text{is true for some } (x, y) \in \mathcal{P}(X \times Y) \end{array} \right\}$$

as desired. \square

Exercise 3.5.12. *This exercise will establish a rigorous version of Proposition 2.1.16. Let X be an arbitrary set. Let $f : \mathbf{N} \times X \rightarrow X$ be a function, and let c be an element of X . Show that there exists a function $a : \mathbf{N} \rightarrow X$ such that*

$$a(0) = c$$

and

$$a(n++) = f(n, a(n)) \text{ for all } n \in \mathbf{N},$$

and furthermore that this function is unique. (Hint: first show inductively, by a modification of the proof of Lemma 3.5.12, that for every natural number $N \in \mathbf{N}$, there exists a unique function $a_N : \{n \in \mathbf{N} : n \leq N\} \rightarrow X$ such that $a_N(0) = c$ and $a_N(n++) = f(n, a_N(n))$ for all $n \in \mathbf{N}$ such that $n < N$.)⁵

For an additional challenge, prove this result without using any properties of the natural numbers other than the Peano axioms directly (in particular, without using the ordering of the natural numbers, and without appealing to Proposition 2.1.16). (Hint: first show inductively, using

⁵In Exercise 3.5.12, add “Let X be an arbitrary set” after the first sentence, and let f be a function from $\mathbf{N} \times X$ to X rather than from $\mathbf{N} \times \mathbf{N}$ to \mathbf{N} ; also c should be an element of X rather than a natural number. This generalisation will help for instance in establishing Exercise 3.5.13. — Errata from Tao.

only the Peano axioms and basic set theory, that for every natural number $N \in \mathbf{N}$, there exists a unique pair A_N, B_N of subsets of \mathbf{N} which obeys the following properties:

- (a) $A_N \cap B_N = \emptyset$,
- (b) $A_N \cup B_N = \mathbf{N}$,
- (c) $0 \in A_N$,
- (d) $N ++ \in B_N$,
- (e) Whenever $n \in B_N$, we have $n ++ \in B_N$.
- (f) Whenever $n \in A_N$ and $n \neq N$, we have $n ++ \in A_N$.

Once one obtains these sets, use A_N as a substitute for $\{n \in \mathbf{N} : n \leq N\}$ in the previous argument.)⁶

Proof. We want to show that for every natural number $N \in \mathbf{N}$, there exists a unique function $a_N : \{n \in \mathbf{N} : n \leq N\} \rightarrow X$ such that $a_N(0) = c$ and $a_N(n++) = f(n, a_N(n))$ for all $n \in \mathbf{N}$ such that $n < N$.

Existence of a_N . We use induction on N . Consider the base case $N = 0$. Since $c \in X$, we can define the function $a_0 : \{0\} \rightarrow X$ by $a_0(0) = c$ and $a_0(1) = f(0, c)$ is vacuously true for that there is no $n \in \mathbf{N}$ such that $n < 0$. Now we inductively suppose that there exists a function a_N . We prove the existence of a_{N++} . Define function a_{N++} as following rule: $a_{N++}(n) := a_N(n)$ for all $n \in \mathbf{N}$ such that $n < N$, and $a_{N++}(n++) := f(n, a_{N++}(n))$ for all $n \in \mathbf{N}$ such that $n = N$. Therefore, we can see that $a_{N++}(0) = a_N(0) = c$ and $a_{N++}(n++) = f(n, a_{N++}(n))$ for all $n \in \mathbf{N}$ such that $n < N++$. This close the induction.

Uniqueness of a_N . We need to show that there is at most one function a_N satisfying the condition. Suppose there exists some another function $\tilde{a}_N :$

⁶One can use ordinary induction on N to establish the uniqueness of the pair A_N, B_N . (One may first need to establish a preliminary lemma that if A_N, B_N obey the required properties, then $N \in A_N$. In particular, given a pair A_{N++}, B_{N++} obeying the required properties, one can delete $N++$ from A_{N++} and add it to B_{N++} to create a pair A_N, B_N that one can then verify obeys the required properties also.) — Tao.

$\mathbf{N} \rightarrow X$ where $\tilde{a}_N \neq a_N$ such that $\tilde{a}(0) = c$ and $\tilde{a}_N(n++) = f(n, \tilde{a}_N(n))$ for all $n \in \mathbf{N}$ such that $n < N$. Since we have $a_N(0) = c = \tilde{a}_N(0)$, assertion is not hold for $n = 0$. By definition, we have $a_N(m) = \tilde{a}_N(m)$ for all $m \in \mathbf{N}$ such that $m < n < N$. By definition again, we have $a_N(m++) = f(m, a_N(m)) = f(m, \tilde{a}_N(m)) = \tilde{a}_N(m++)$ for all $m++ \in \mathbf{N}$ such that $m++ \leq n < N$. Thus we have $a_N(n++) = f(n, a_N(n)) = f(n, \tilde{a}_N(n)) = \tilde{a}_N(n++)$.

Now we conclude that there exists a unique function $a_N : \{n \in \mathbf{N} : n \leq N\} \rightarrow X$ such that $a_N(0) = c$ and $a_N(n++) = f(n, a_N(n))$ for all $n \in \mathbf{N}$ such that $n < N$.

Let G and G_N be the graph of $a : \mathbf{N} \rightarrow X$ and $a_N : \{n \in \mathbf{N} : n \leq N\} \rightarrow X$. By Exercise 3.4.7, $G = \bigcup_{N \in \mathbf{N}} G_N$, and by Exercise 3.5.10, $G \subseteq \mathbf{N} \times X$, so that $G_N \subseteq \mathbf{N} \times X$. Thus there is exactly one function whose graph is equal to G . This means that there exists a unique function $a : \mathbf{N} \rightarrow X$ such that $a(0) = c$ and $a(n++) = f(n, a(n))$ for all $n \in \mathbf{N}$. \square

Now we turn to consider the additional challenges. Before we prove this assertion, we need a lemma:

Lemma. *If A_N, B_N obey the required properties, then $N \in A_N$. In particular, given a pair A_{N++}, B_{N++} obeying the required properties, one can delete $N++$ from A_{N++} and add it to B_{N++} to create a pair A_N, B_N that one can then verify obeys the required properties also.*

Proof. Suppose A_N, B_N obey the required properties. We need to show that $N \in A_N$. Let $m++ = N$, then $m \neq N$. We induct on m . When $m = 0$, we have $0 \in A_N$. Now inductively suppose that $m \in A_N$ (we know $m \neq N$). By property (f), we have $m++ \in A_N$, which means $N \in A_N$. In particular, If A_{N++}, B_{N++} obey the required properties, then $N++ \in A_{N++}$. By property (a), $N++ \notin B_{N++}$, then by property (e), we have $N \notin B_{N++}$. Now we can rewrite it as $A_{N++} := A_N \cup \{N++\}$ and $B_{N++} := B_N \setminus \{N\}$, which means A_N and B_N also obey the required properties, as desired. \square

Now we prove the assertion.

Proof of assertion. For uniqueness. Suppose that $B_M \neq B_N$ (e.g., $B_M \supsetneq B_N$), and pair A_N, B_N and pair A_N, B_M both obey the required properties.

We use induction on N . When $N = 0$, by Lemma, we can let $A_0 = \{0\}$ and $B_0 = \mathbf{N} \setminus \{0\}$. Since $B_M \supsetneq B_N$, assume that there is some natural number $M \neq 0$ such that $M \in B_M$. By property (e) and (f), since $M \neq 0$, $M++ \in A_N$; and since $M \in B_M$, $M++ \in B_M$. Then we have $A_N \cap B_M \neq \emptyset$, a contradiction. Now we inductively suppose that the assertion is true. We need to prove that pair A_{N++}, B_{N++} and pair A_{N++}, B_{M++} both obey the required properties. By Lemma, pair A_{N++}, B_{N++} obeys the required properties. Consider pair A_{N++}, B_{M++} . By Lemma, we can let $A_{N++} := A_N \cup \{N++\}$ and $B_{M++} := B_M \setminus \{M\}$. Then $A_{N++} \cap B_{M++} = (A_N \cup \{N++\}) \cap (B_M \setminus \{M\}) \neq \emptyset$. This close the induction.

For existence, this immediately follows from Lemma. Therefore, there is exactly one pair A_N, B_N obeys the required properties. \square

Exercise 3.5.13. *The purpose of this exercise is to show that there is essentially only one version of the natural number system in set theory (cf. the discussion in Remark 2.1.12). Suppose we have a set \mathbf{N}' of “alternative natural numbers”, an “alternative zero” $0'$, and an “alternative increment operation” which takes any alternative natural number $n' \in \mathbf{N}'$ and returns another alternative natural number $n'++' \in \mathbf{N}'$ such that the Peano axioms (Axioms 2.1-2.5) all hold with the natural numbers, zero, and increment replaced by their alternative counterparts. Show that there exists a bijection $f : \mathbf{N} \rightarrow \mathbf{N}'$ from the natural numbers to the alternative natural numbers such that $f(0) = 0'$, and such that for any $n \in \mathbf{N}$ and $n' \in \mathbf{N}'$, we have $f(n) = n'$ if and only if $f(n++) = n'++'$ (Hint: use Exercise 3.5.12.)*

3.6

Cardinality of sets

Exercise 3.6.1. *Prove Proposition 3.6.4.*

Proposition 3.6.4. Let X, Y, Z be sets. Then X has equal cardinality with X . If X has equal cardinality with Y , then Y has equal cardinality

with X . If X has equal cardinality with Y and Y has equal cardinality with Z , then X has equal cardinality with Z .

Proof. Let X, Y, Z be sets. We prove that cardinality is reflexive, symmetric and transitive.

First, X has equal cardinality with X , because $f : X \rightarrow X$ defined by $f(x) = x$ is a bijection. So the cardinality relation is reflexive.

Then we prove that if X has equal cardinality with Y , then Y has equal cardinality with X . Suppose that X has equal cardinality with Y , therefore there exists a bijection $f : X \rightarrow Y$. As we have proved in Exercise 3.3.6 that $f^{-1} : Y \rightarrow X$ is a bijection when $f : X \rightarrow Y$ is a bijection. Therefore the cardinality relation is symmetric.

Finally we prove that the cardinality relation is transitive. Suppose that $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be bijections. We prove that $g \circ f : X \rightarrow Z$ is a bijection, which is proved in Exercise 3.3.7. Therefore the cardinality relation is transitive. \square

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

Proof. Suppose that X has cardinality 0, by Definition 3.6.5, it has equal cardinality with $\{i \in \mathbf{N} : \leq i \leq 0\}$, which is empty set. Then suppose that X is an empty set, and $f : X \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq n\}$ should be a bijection. As we show in Exercise 3.3.3, empty function is bijective iff both X and $\{i \in \mathbf{N} : 1 \leq i \leq n\}$ are empty set. Because i is a natural number, therefore i need to equal 0. So X has cardinality 0. \square

Exercise 3.6.3. *Let n be a natural number, and let $f : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow \mathbf{N}$ be a function. Show that there exists a natural number M such that $f(i) \leq M$ for all $1 \leq i \leq n$. (Hint: induct on n . You may also want to peek at Lemma 5.1.14.) Thus finite subsets of the natural numbers are bounded.*

Proof. We prove this by induction on n . When $n = 1$ it's clearly true, for

if we choose $M := 1$ then clearly we have $f(i) \leq 1$ for all $1 \leq i \leq n$. Now suppose that we have already proved the claim for some $n \geq 1$; we now prove it for $n + 1$. By the induction hypothesis we know that $f(n) \leq M$ by some M , and f is a bijection, therefore $f(n + 1) \leq M + 1$. This close the induction. \square

Exercise 3.6.4. *Prove Proposition 3.6.14*

Proof.

- (a) *Let X be a finite set, and let x be an object which is not an element of X . Then $X \cup \{x\}$ is finite and $\#(X \cup \{x\}) = \#(X) + 1$.*

By Definition 3.6.10, X is finite, it has cardinality N for some natural number. So $\#(X) + 1 = N + 1$. Let y be an element of X . Then we have a bijection $f : X \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N\}$, and $f(y)$ is a natural number between 1 and N . Now we define the function $g : X \cup \{x\} \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N + 1\}$ by following rule: for any $z \in X \cup \{x\}$, we define $g(z) := f(y)$ if $z = y$, and $g(z) := N + 1$ if $z = x$. This map is a bijection and $\#(X \cup \{x\}) = N + 1$. Therefore, $\#(X \cup \{x\})$ is finite and it has same cardinality with $\#(X) + 1$.

- (b) *Let X and Y be finite sets. Then $X \cup Y$ is finite and $\#(X \cup Y) \leq \#(X) + \#(Y)$. If in addition X and Y are disjoint (i.e., $X \cap Y = \emptyset$), then $\#(X \cup Y) = \#(X) + \#(Y)$.*

By Definition 3.6.10, suppose that $\#(X) = N, \#(Y) = M$, for some natural numbers N, M . Then $f : X \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N\}$ and $g : Y \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq M\}$ are bijections. Let $Z = Y \setminus (X \cap Y)$, which is a subset of Y . Then by Definition 3.4.1, $g(Z)$ is a subset of $\{i \in \mathbf{N} : 1 \leq i \leq M\}$; $g(Z) := \{i \in \mathbf{N} : 1 \leq i \leq K\}$ for some natural number $K \leq M$. It is clear that $X \cup Y = X \cup (Y \setminus (X \cap Y)) = X \cup Z$. Because every element of X is not an element of Z . Therefore, we define the function h from $X \cup Z$ to $\{i \in \mathbf{N} : 1 \leq i \leq N + K\}$ by following rule: for any $a \in X \cup Z$ we define $h(a) := f(a)$ if $a \in X$; $h(a) := g(a) + N$ if $a \in Z$. We can see that h is bijective. Hence, $\#(X \cup$

$Y) = \#(X \cup Z) = N + K \leq N + M = \#(X) + \#(Y)$. Particularly, when $X \cap Y = \emptyset$, $K = M$, so $\#(X \cup Y) = \#(X) + \#(Y) = N + M$.

- (c) *Let X be a finite set, and let Y be a subset of X . Then Y is finite, $\#(Y) \leq \#(X)$. If in addition $Y \neq X$ (i.e., Y is a proper subset of X), then we have $\#(Y) < \#(X)$.*

Let X be a finite set and $\#(X) = N$ for some natural number N . Then by Definition 3.6.10, there exists a bijective function $f : X \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N\}$. Let Y be a proper subset of X . We need to prove that there exists no bijection from Y to $\{i \in \mathbf{N} : 1 \leq i \leq N\}$, but there exists a bijection from Y to $\{i \in \mathbf{N} : 1 \leq i \leq M\}$ for some $M < N$. Then by Definition 3.6.5, this implies that $\#(Y) < \#(X)$.

We induct on N . When $N = 1$, Y must be empty set, by Definition 3.6.5, $\#(Y) = 0$. Hence $\#(Y) = 0 < 1 = \#(X)$.

Now we inductively suppose that assertion is hold for N . We need to prove the case of $N + 1$. Suppose $f : X \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N + 1\}$ is bijective and Y be a non-empty proper subset of X . Choose an element x_0 of Y and an element x_1 of $X \setminus Y$. Then by Lemma 3.6.9, we have a bijection

$$g : X \setminus \{x_0\} \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N\}.$$

Now $Y \setminus \{x_0\}$ is a proper subset of $X \setminus \{x_0\}$, for that $x_1 \in X \setminus \{x_0\}$ but $x_1 \notin Y \setminus \{x_0\}$. On the one hand, by induction hypothesis, there exists no bijection of $Y \setminus \{x_0\}$ with $\{i \in \mathbf{N} : 1 \leq i \leq N\}$. By Lemma 3.6.9, there exists no bijection of Y with $\{i \in \mathbf{N} : 1 \leq i \leq N + 1\}$.

Other hand, since induction hypothesis means that the assertion is hold for natural number N , we have either $Y \setminus \{x_0\} = \emptyset$, or there exists a bijection of $Y \setminus \{x_0\}$ with $\{i \in \mathbf{N} : 1 \leq i \leq k\}$ for some $k < N$. If $Y \setminus \{x_0\} = \emptyset$, by Lemma 3.6.9, we have $\#(Y) = 1$. While if $Y \setminus \{x_0\} \neq \emptyset$, by Lemma 3.6.9, $\#(Y) = k + 1$. Thus in either case, there is a bijection from Y to $\{i \in \mathbf{N} : 1 \leq i \leq M\}$ for some $M < N + 1$, as desired. This close the induction.

In particular, when $Y = X$, Y still be a subset of X . There is a identity map $\iota_{Y \rightarrow X} : Y \rightarrow X$, then bijection $f \circ \iota : Y \rightarrow \{i \in \mathbf{N} : 0 \leq i \leq N\}$. Hence $\#(Y) = \#(X)$. Together with preceding conclusion, when Y be a subset of X , we have $\#(Y) \leq \#(X)$.

- (d) *If X is a finite set, and $f : X \rightarrow Y$ is a function, then $f(X)$ is a finite set with $\#(f(X)) \leq \#(X)$. If in addition f is one-to-one, then $\#(f(X)) = \#(X)$.*

Let X is a finite set, and $f : X \rightarrow Y$. We suppose that $\#(X) = N$. By Definition 3.6.5, there is a bijection $g : X \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N\}$. Since $f(X) \subseteq Y$, by Definition 3.4.1, $f : X \rightarrow f(X)$ is surjection, so that f^{-1} is injection. By Exercise 3.3.5, $g \circ f^{-1} : f(X) \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq N\}$ is injection. Thus there is an $M \leq N$ such that $g \circ f^{-1} : f(X) \rightarrow \{i \in \mathbf{N} : 1 \leq i \leq M\}$ be a bijection. This means that $\#(f(X)) \leq \#(X)$. In particular, if f is bijection, then we have $f : X \rightarrow f(X)$ is a bijection. This implies that $\#(f(X)) = \#(X)$.

- (e) *Let X and Y be finite sets. Then Cartesian product $X \times Y$ is finite $\#(X \times Y) = \#(X) \times \#(Y)$.*

By Definition 3.6.10, suppose that $\#(X) = N$ and $\#(Y) = M$. If either N or M equals to 0, then by Exercise 3.6.2, $\#(X \times Y) = \#(X) \times \#(Y) = 0$. Now we use induction on M . When $M = 1$, we let $Y = \{y\}$ and define $f : X \rightarrow X \times \{y\}$ by following rule: $f(x) = (x, y)$ for all $x \in X$. f is bijective. We have $\#(X) \times \#(Y) = \#(X) = \#(X \times Y) = N$. Now we inductively suppose that $\#(X \times Y) = \#(X) \times \#(Y)$ is hold, we prove the case of $M + 1$. Let $y' \notin Y$ and $y' \notin X$, then by Proposition 3.6.14(a), we have $Y \cup \{y'\} = M + 1$. By Exercise 3.5.4, we have $X \times (Y \cup \{y'\}) = (X \times Y) \cup (X \times \{y'\})$. By induction hypothesis, $\#(X \times Y) = NM$ and $\#(X \times \{y'\}) = N \times 1$. Hence $\#(X \times (Y \cup \{y'\})) = \#(X) \times \#(Y \cup \{y'\}) = N(M + 1)$. This close the induction.

- (f) *Let X and Y be finite sets. Then the set Y^X (defined in Axiom 3.10) is finite and $\#(Y^X) = \#(Y)^{\#(X)}$*

By Definition 3.6.10, suppose that $\#(X) = N$ and $\#(Y) = M$. We use induction on N . When $N = 0$, we have $\#(Y^X) = \#(Y)^{\#(X)} = 1$, the only function defined in Y^X is empty function. Now we inductively suppose that $\#(Y^X) = \#(Y)^{\#(X)}$ is hold for N . We prove that $N + 1$. Let $x_0 \notin X$, then we have $X \cup \{x_0\}$ with cardinality $N + 1$. Any subsets of $X \cup \{x_0\}$ can be divided into the set containing x_0 or not. Then there is $\#(Y)^{\#X}$ function that map from subsets of $X \cup \{x_0\}$ which not contain $\{x_0\}$ to Y , and $\#(Y)^{\#X}$ function that map from subsets of $X \cup \{x_0\}$ which contain $\{x_0\}$ to Y . Thus we have $\#(Y^{X \cup \{x_0\}}) = \#(Y)^{\#X+1}$. This close the induction. \square

Exercise 3.6.5. *Let A and B be sets. Show that $A \times B$ and $B \times A$ have equal cardinality by constructing an explicit bijection between the two sets. Then use Proposition 3.6.14 to conclude an alternate proof of Lemma 2.3.2*

Proof. Define $f : A \times B \rightarrow B \times A$ by

$$f(a, b) = (b, a),$$

for all $(a, b) \in A \times B$. We can see that f is injection, so that $A \times B$ and $B \times A$ has same cardinality.

Suppose that $\#(A) = n, \#(B) = m$, then by Proposition 3.6.14(e), we have

$$\#(A \times B) = n \times m = m \times n = \#(B \times A).$$

\square

Exercise 3.6.6. *Let A, B, C be sets. Show that the sets $(A^B)^C$ and $A^{B \times C}$ have equal cardinality by constructing an explicit bijection between the two sets. Conclude that $(a^b)^c = a^{bc}$ for any natural numbers a, b, c . Use a similar argument to also conclude $a^b \times a^c = a^{b+c}$.*

Proof.

3.7

***Zermelo-Fraenkel axioms of set theory**

The axioms of set theory that we have introduced (Axioms 3.1-3.11, excluding the dangerous Axiom 3.8) are known as the *Zermelo-Fraenkel axioms of set theory*, after Ernest Zermelo (1871–1953) and Abraham Fraenkel (1891–1965).

We left the representation of ZF axiom in Terence Tao’s *Analysis I* and Ralf Schindler’s *Set Theory: Exploring Independence and Truth* here respectively. It is useful to compare with them.

3.7.1 ZFC axiom in Analysis I of Terence Tao

Axiom 3.0 (Extensionality) *Two sets A and B are equal, $A = B$, iff every element of A is an element of B and vice versa. To put it another way, $A = B$ if and only if every element x of A belongs also to B , and every element y of B belongs also to A .*

Axiom 3.1 (Sets are objects). *If A is a set, then A is also an object. In particular, given two sets A and B , it is meaningful to ask whether A is also an element of B .*

Axiom 3.2 (Empty set). *There exists a set \emptyset , known as the empty set, which contains no elements, i.e., for every object x we have $x \notin \emptyset$.*

Axiom 3.3 (Singleton sets and pair sets). *If a is an object, then there exists a set $\{a\}$ whose only element is a , i.e., for every object y , we have $y \in \{a\}$ if and only if $y = a$; we refer to $\{a\}$ as the singleton set whose element is a . Furthermore, if a and b are objects, then there exists a set $\{a, b\}$ whose only elements are a and b ; i.e., for every object y , we have $y \in \{a, b\}$ if and only if $y = a$ or $y = b$; we refer to this set as the pair set formed by a and b .*

Axiom 3.4 (Pairwise union). *Given any two sets A, B , there exists a set $A \cup B$, called the union $A \cup B$ of A and B , whose elements consist*

of all the elements which belong to A or B or both. In other words, for any object x ,

$$x \in A \cup B \iff (x \in A \text{ or } x \in B).$$

Axiom 3.5 (Axiom of specification). *Let A be a set, and for each $x \in A$, let $P(x)$ be a property pertaining to x (i.e., $P(x)$ is either a true statement or a false statement). Then there exists a set, called $\{x \in A : P(x) \text{ is true}\}$ (or simply $\{x \in A : P(x)\}$ for short), whose elements are precisely the elements x in A for which $P(x)$ is true. In other words, for any object y ,*

$$y \in \{x \in A : P(x) \text{ is true}\} \iff (y \in A \text{ and } P(y) \text{ is true}).$$

Axiom 3.6 (Replacement). *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 $\{y : P(x, y) \text{ is true for some } x \in A\}$, such that for any object z ,*

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

Axiom 3.7 (Infinity). *There exists a set \mathbf{N} , whose elements are called natural numbers, as well as an object 0 in \mathbf{N} , and an object $n++$ assigned to every natural number $n \in \mathbf{N}$, such that the Peano axioms (Axioms 2.1-2.5) hold.*

Axiom 3.9 (Regularity). *If A is a non-empty set, then there is at least one element x of A which is either not a set, or is disjoint from A .*

Axiom 3.10 (Power set axiom). *Let X and Y be sets. Then there exists a set, denoted Y^X , which consists of all the functions from X to Y , thus*

$$f \in Y^X \iff (f \text{ is a function with domain } X \text{ and range } Y).$$

Axiom 3.11 (Union). *Let A be a set, all of whose elements are themselves sets. Then there exists a set $\bigcup A$ whose elements are precisely those objects which are elements of the elements of A , thus for all objects x*

$$x \in \bigcup A \iff (x \in S \text{ for some } S \in A).$$

Axiom 8.1 (Choice). *Let I be a set, and for each $\alpha \in I$, let X_α be a non-empty set. Then $\prod_{\alpha \in I} X_\alpha$ is also non-empty. In other words, there exists a function $(x_\alpha)_{\alpha \in I}$ which assigns to each $\alpha \in I$ an element $x_\alpha \in X_\alpha$.*

3.7.2 ZFC axiom in Set Theory of Ralf Schindler

The language of set theory, which we denote by \mathcal{L}_\in , is the usual language of first order logic (with one type of variables) equipped with just one binary relation symbol, \in . The intended domain of set theoretical discourse (i.e., the range of the variables) is the universe of all sets, and the intended interpretation of \in is “is an element of.” We shall use $x, y, z, \dots, a, b, \dots$, etc. as variables to range over sets.

Axiom of extensionality. *Two sets are equal iff they contain the same elements:*

$$\forall x \forall y (x = y \iff \forall z (z \in x \iff z \in y)). \quad (\text{Ext})$$

A set x is a *subset* of y , abbreviated by $x \subseteq y$, if $\forall z (z \in x \implies z \in y)$. (Ext) is then logically equivalent to

$$\forall x \forall y (x \subseteq y \wedge y \subseteq x \implies x = y).$$

Axiom of foundation. *Each nonempty set has an \in -minimal member*

$$\forall x (\exists y y \in x \implies \exists y (y \in x \wedge \nexists z (z \in y \wedge z \in x))). \quad (\text{Fund})$$

We write $x = \emptyset$ for $\nexists y y \in x$ (and $x \neq \emptyset$ for $\exists y y \in x$), and $x \cap y = \emptyset$ for $\nexists z (z \in y \wedge z \in x)$. (Fund) then says that

$$\forall x (x \neq \emptyset \implies \exists y (y \in x \wedge y \cap x = \emptyset)).$$

Axiom of pairing. Let us write $z = \{x, y\}$ instead of

$$y \in z \wedge x \in z \wedge \forall u(u \in z \implies (u = y \vee u = x)).$$

In particular, when $x = y$, we have $z = \{x\}$. Now we have

$$\forall x \forall y \exists z z = \{x, y\}. \quad (\text{Pair})$$

In the presence of (Pair), (Fund) implies that there cannot be a set x with $x \in x$: if $x \in x$, then x is the only element of $\{x\}$, but $x \cup \{x\}$, as $x \in x \cup \{x\}$.

Axiom of union. Let us write $y = \bigcup x$ for

$$\forall z(z \in y \iff \exists u(u \in x \wedge z \in u)).$$

Then we have

$$\forall x \exists y y = \bigcup x. \quad (\text{Union})$$

Power set axiom. For every set x , the set of all subsets of x exists. We write $y = \mathcal{P}(x)$ for

$$\forall z(z \in y \iff z \subseteq x)$$

and formulate

$$\forall x \exists y y = \mathcal{P}(x). \quad (\text{Pow})$$

Axiom of infinity. There is a set which contains all of the following sets as members:

$$\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \dots$$

To make this precise, we call a set x inductive iff

$$\emptyset \in x \wedge \forall y(y \in x \implies y \cup \{y\} \in x).$$

We then say:

$$\exists x(x \text{ is inductive}). \quad (\text{Inf})$$

Axiom of separation. Let φ be a formula of \mathcal{L}_\in in which exactly the variables x, v_1, \dots, v_p (which all differ from b) occur freely. The axiom of separation corresponding to φ runs as follows.

$$\forall v_1 \dots \forall v_p \forall a \exists b \forall x(x \in b \implies x \in a \wedge \varphi). \quad (\text{Sep}_\varphi)$$

Let us write $b = \{x \in a : \varphi\}$ for $\forall x(x \in b \iff x \in a \wedge \varphi)$. If we suppress v_1, \dots, v_p , (Sep_φ) then says that

$$\forall a \exists b \, b = \{x \in a : \varphi\}.$$

Writing $z = x \cap y$ for

$$\forall u(u \in z \iff u \in x \wedge u \in y),$$

$(\text{Sep}_{x \in c})$ proves that $\forall a \forall c \exists b \, b = a \cap c$. Writing $z = x \setminus y$ for

$$\forall u(u \in z \iff u \in x \wedge u \notin y),$$

$(\text{Sep}_{x \notin c})$ proves that $\forall a \forall c \exists b \, b = a \setminus c$. Also, if we write $x = \bigcap y$ for

$$\forall z(z \in x \iff \forall u(u \in y \implies z \in u)),$$

then $(\text{Sep}_{\forall u(u \in y \implies z \in u)})$, applied to any member of y proves that

$$\forall y(y \neq \emptyset \implies \exists x \, x = \bigcap y).$$

Axiom of replacement. Let φ be a formula of \mathcal{L}_\in in which exactly the variables x, v_1, \dots, v_p (which all differ from b) occur freely. The axiom of separation corresponding to φ runs as follows.

$$\begin{aligned} & \forall v_1 \dots \forall v_p (\forall x \exists y' (\forall y (y = y' \iff \varphi)) \\ & \implies \forall a \exists b \forall y (y \in b \iff \exists (x \in a \wedge \varphi))). \end{aligned} \quad (\text{Rep}_\varphi)$$

Axiom of choice. For each family of pairwise disjoint non-empty sets there is a “choice set”, i.e.

$$\begin{aligned} & \forall x (\forall y (y \in x \implies y \neq \emptyset) \wedge \forall y \forall y' (y \in x \wedge y' \in x \wedge y \neq y' \implies \\ & y \cap y' = \emptyset) \implies \exists z \forall y (y \in x \implies \exists u \forall u' (u' = u \iff u' \in z \cap y))). \end{aligned} \quad (\text{AC})$$

In what follows we shall always abbreviate $\forall y(y \in x \implies \varphi)$ by $\forall y \in x \, \varphi$ and $\exists y(y \in x \wedge \varphi)$ by $\exists y \in x \, \varphi$. We may then also formulate (AC) as

$$\begin{aligned} & \forall x (\forall y \in x \, y \neq \emptyset \wedge \forall y \in x \forall y' \in x (y \neq y' \implies y \cap y' = \emptyset) \\ & \implies \exists z \forall y \in x \exists u \, z \cap y = \{u\}), \end{aligned}$$

i.e., for each member of x , z contains exactly one “representative.”

CHAPTER IV

Integers and rationals

4.1

The integers

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

Proof. Prove reflexivity. $a - b = a - b$ iff $a + b = a + b$, because a and b are natural numbers, so it is obviously true. Then prove symmetry. For natural numbers a, b and c , suppose $a - b = c - d$, by Definition 4.1.1, $a + d = c + b$, so we have $c + b = a + d$, by Definition 4.1.1 again, $c - d = a - b$. This is our desired. \square

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).*

Proof. Suppose $(a - b) = (a' - b')$, by Definition 4.1.1, we have $a + b' = a' + b$. By Definition 4.1.4, $-(a - b) = -(a' - b') \implies (b - a) = (b' - a')$. If they equal to each other, it should satisfy that $b + a' = b' + a$. It is true by

condition. So, the definition of negation on the integers is well-defined. \square

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

Proof. a is a integer, let $a = (a' - b')$, in which a' and b' are arbitrary natural numbers, by Definition 4.1.4, we need to prove that $(0 - 1) \times (a' - b') = (b' - a')$. By Definition 4.1.2, $(0 - 1) \times (a - b) = (0a + 1b) - (0b + 1a) = b - a$. This is our desired. \square

Exercise 4.1.4. Prove the remaining identities in Proposition 4.1.6. (Hint: one can save some work by using some identities to prove others. For instance once you know that $xy = yx$, you get for free that $x1 = 1x$, and once you also prove $x(y + z) = xy + xz$, you automatically get $(y + z)x = yx + zx$ for free.)

Proposition 4.1.6 (Laws of algebra for integers). Let x, y, z be integers. Then we have

$$\begin{aligned}
 x + y &= y + x \\
 (x + y) + z &= x + (y + z) \\
 x + 0 &= 0 + x = x \\
 x + (-x) &= (-x) + x = 0 \\
 xy &= yx \\
 (xy)z &= x(yz) \\
 x1 &= 1x = x \\
 x(y + z) &= xy + xz \\
 (y + z)x &= yx + zx.
 \end{aligned}$$

Proof. Let $x = a - b, y = c - d, z = e - f$ a, b, c, d, e, f are natural numbers.

(a) Prove $x + y = y + x$.

$$\begin{aligned}
 x + y &= (a - b) + (c - d) = (a + c) - (b + d) \\
 y + x &= (c - d) + (a - b) = (c + a) - (d + b) = (a + c) - (b + d)
 \end{aligned}$$

(b) Prove $(x + y) + z = x + (y + z)$.

$$\begin{aligned}
 (x + y) + z &= [(a + c) - (b + d)] + (e - f) \\
 &= [(a + c) + e] - [(b + d) + f] \\
 &= (a + c + e) - (b + d + f) \\
 x + (y + z) &= (a - b) + [(c + e) - (d + f)] \\
 &= [a + (c + e)] - [b + (d + f)] \\
 &= (a + c + e) - (b + d + f)
 \end{aligned}$$

(c) Prove $x + 0 = 0 + x = x$. By Definition 4.1.4, let $n = 0$, then $0 = 0 - 0$.

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

(d) Prove $x + (-x) = (-x) + x = 0$.

$$x + (-x) = (a - b) + (b - a) = (a + b) - (a + b) = (a + b) - (a + b)$$

Because $a + b + 0 = a + b + 0$, by Definition 4.1.1, $x + (-x) = (a + b) - (a + b) = 0 - 0 = 0$.

(e) Prove $xy = yx$.

$$\begin{aligned}
 xy &= (a - b) \times (c - d) = (ac + bd) - (ad + bc) \\
 yx &= (c - d) \times (a - b) \\
 &= (ca + db) - (da + cb) \\
 &= (ac + bd) - (ad + bc)
 \end{aligned}$$

(f) Prove $x1 = 1x = x$. Let $n = 1$, then $1 = 1 - 0$.

$$x1 = (a1 + b0) - (a0 + b1) = a - b = x$$

(g) Prove $x(y + z) = xy + xz$.

$$\begin{aligned}
 x(y + z) &= (a - b) \times [(c - d) + (e - f)] \\
 &= (a - b) \times [(c + e) - (d + f)] \\
 &= [a(c + e) + b(d + f)] - [a(d + f) + b(c + e)] \\
 &= [ac + bd + ae + bf] - [ad + bc + af + be] \\
 xy + xz &= (a - b)(c - d) + (a - b)(e - f) \\
 &= [(ac + bd) - (ad + bc)] + [(ae + bf) - (af + be)] \\
 &= [(ac + bd) + (ae + bf)] - [(ad + bc) + (af + be)] \\
 &= [ac + bd + ae + bf] - [ad + bc + af + be]
 \end{aligned}$$

(h) Prove $(y + z)x = yx + zx$. By conclusions of (e) and (g), we have

$$\begin{aligned}
 (y + z)x &= x(y + z) \\
 yx + zx &= xy + xz
 \end{aligned}$$

□

Exercise 4.1.5. Prove Proposition 4.1.8. (Hint: while this proposition is not quite the same as Lemma 2.3.3, it is certainly legitimate to use Lemma 2.3.3 in the course of proving Proposition 4.1.8.)

Proposition 4.1.8 (Integers have no zero divisors). Let a and b be integers such that $ab = 0$. Then either $a = 0$ or $b = 0$ (or both).

Proof. By Definition 4.1.4, $ab = 0 \iff (a - 0)(b - 0) = 0 - 0$, then we have $(ab + 0 \cdot 0) - (a0 + 0b) = ab - 0 = 0 - 0 \implies ab = 0$. By Lemma 2.3.3, at least one of natural number a and b equal to 0. By Trichotomy of integers, at least one of integer a and b equal to 0. □

Exercise 4.1.6. Prove Corollary 4.1.9. (Hint: there are two ways to do this One is to use Proposition 4.1.8 to conclude that $a - b$ must be zero. Another way is to combine Corollary 2.3.7 with Lemma 4.1.5.)

Corollary 4.1.9 (Cancellation law for integers). If a, b, c are integers such that $ac = bc$ and c is non-zero. then $a = b$.

Proof. By Laws of algebra for integers, $ac = bc \implies (a-b)c = 0$. Then by Proposition 4.1.8 and by $c \neq 0$, $a-b = 0 \implies a = b$. \square

Exercise 4.1.7. Prove Lemma 4.1.11. (Hint: use the first part of this lemma to prove all the others.

Proof. Let a, b, c be integers.

(a) $a > b$ if and only if $a - b$ is a positive natural number.

Suppose that $a > b$, there exists a natural number n such that $a = b+n$ and $a \neq b$. By Definition 4.1.1, Definition 4.1.4, $a - b = n - 0 = 0$. Because n is a positive natural number, then $a - b$ also a positive natural number. Then we assume that $a - b$ is a positive natural number, denoted by n . By Definition 4.1.1, Definition 4.1.4, $a = b + n$. By Lemma 4.1.5, $a \neq b$. Therefore, by Definition 4.1.10, $a > b$.

(b) (Addition preserves order) If $a > b$, then $a + c > b + c$.

By claim (a), $a + c > b + c$ if and only if $(a + c) - (b + c) = a - b$ is a positive natural number. Because $a > b$, statement is true.

(c) (Positive multiplication preserves order) If $a > b$ and c is positive then $ac > bc$.

Because $ac - bc = (a - b)c$ and c is positive, so $a - b$ also is positive. Therefore, $ac - bc$ also positive, $ac > bc$.

(d) (Negation reverses order) If $a > b$, then $-a < -b$.

$(-b) - (-a) = -b + a = a - b$, and $a - b > 0$. So $-a < -b$.

(e) (Order is transitive) If $a > b$ and $b > c$, then $a > c$.

By condition we have $a - b > 0$ and $b - c > 0$. Then $(a - b) + (b - c) = a - c > 0$. So $a > c$.

(f) (*Order trichotomy*) Exactly one of the statements $a > b$, $a < b$, or $a = b$ is true.

By Trichotomy of natural number, at least one of $a - b$ greater than 0, less than 0 or equal 0 is true. If $a - b > 0$, by claim (a), $a > b$. If $a - b < 0 \implies b - a > 0$, then $a < b$. While if $a - b = 0$, $a = b$. This is our desired. \square

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 . Thus induction is not as useful a tool for dealing with the integers as it is with the natural numbers. (The situation becomes even worse with the rational and real numbers, which we shall define shortly.)

Proof. Consider the inequality:

$$n + 1 \geq 0.$$

This is hold for $n = 0$ that $1 \geq 0$. Suppose that $n + 1 \geq 0$ is hold for all integer n . Then for case of $n + 1$, we have

$$(n + 1) + 1 \geq 0 + 1 = 1 \geq 0.$$

Thus $n + 1 \geq 0$ is hold for all integer n . But if we pick $n = -2$, we have $-1 < 0$, a contradiction. \square

4.2

The rationals

Exercise 4.2.1. Show that the definition of equality for the rational numbers is reflexive, symmetric, and transitive. (Hint: for transitivity, use Corollary 4.1.9.)

Proof. Let a, b, c, d, e, f are integers, and b, d, f are not equal 0.

Prove reflexivity. $a // b = a // b$ if and only if $ab = ab$. Because a and b are integers. Then it is obviously true.

Prove symmetry. Assume that $a // b = c // d$. By Definition 4.2.1, $ad = bc$, i.e. $cb = ad$. $c // d = a // b$.

Prove transitivity. Assume that $a // b = c // d$ and $c // d = e // f$. By Definition 4.2.1, $ad = cb$ and $cf = ed$. Then $adcf = cb ed$. Because $d \neq 0$, By Corollary 4.1.9, $caf = ceb$. Now we consider c : if $c \neq 0$, by Corollary 4.1.9 again, $af = eb$, i.e. $a // b = e // f$. While if $c = 0$, let $a = e = 0$, then $af = eb$. Therefore, $a // b = e // f$. \square

Exercise 4.2.2. Prove the remaining components of Lemma 4.2.3.

Lemma 4.2.3. The sum, product, and negation operations on rational numbers are well-defined, in the sense that if one replaces $a // b$ with another rational number $a' // b'$ which is equal to $a // b$, then the output of the above operations remains unchanged, and similarly for $c // d$.

Proof. Suppose $a // b = a' // b'$, so that b and b' are non-zero and $ab' = a'b$. We now show that $(a // b) * (c // d) = (a' // b') * (c // d)$. By definition, the left-hand side is $(ac) // (bd)$ and the right-hand side is $(a'c) // (b'd)$, so we have to show that

$$(ac) // (bd) = (a'c) // (b'd),$$

which expands to

$$ab'cd = a'bcd$$

But since $ab' = a'b$, the claim follows. Similarly if one replaces $c // d$ by $c // d$. \square

Exercise 4.2.3. Prove the remaining components of Proposition 4.2.4. (Hint as with Proposition 4.1.6, you can save some work by using some identities to prove others.)

Proposition 4.2.4 (Laws of algebra for rationals). Let x, y, z be ratio-

nals. Then the following laws of algebra hold:

$$x + y = y + x$$

$$(x + y) + z = x + (y + z)$$

$$x + 0 = 0 + x = x$$

$$x + (-x) = (-x) + x = 0$$

$$xy = yx$$

$$(xy)z = x(yz)$$

$$x1 = 1x = x$$

$$x(y + z) = xy + xz$$

$$(y + z)x = yx + zx.$$

If x is non-zero, we also have

$$xx^{-1} = x^{-1}x = 1.$$

Proof. Let $x = a // b, y = c // d, z = e // f$, for some integers a, c, e and non-zero integers b, d, f .

(a) Prove $x + y = y + x$.

$$x + y = (a // b) + (c // d) = (ad + bc) // (bd);$$

$$y + x = (c // d) + (a // b) = (ad + bc) // (bd).$$

(b) Prove $(x + y) + z = x + (y + z)$.

$$(x + y) + z = ((a // b) + (c // d)) + (e // f)$$

$$= ((ad + bc) // bd) + (e // f)$$

$$= (adf + bcf + bde) // bdf;$$

$$x + (y + z) = (a // b) + ((c // d) + (e // f))$$

$$= (a // b) + ((cf + de) // df)$$

$$= (adf + bcf + bde) // bdf.$$

(c) Prove $x + 0 = 0 + x = x$. Let $y = 0$, then $x + 0 = 0 + x$. Because $0 = 0 // 1$

$$x + 0 = (a // b) + (0 // 1) = (a1 + b0) // b1 = a // b = x.$$

(d) Prove $x + (-x) = (-x) + x = 0$. Let $y = -x$, then $x + (-x) = (-x) + x$.

$$x + (-x) = (a // b) + ((-a) // b) = (ab - ab) // b^2 = 0.$$

(e) Prove $xy = yx$.

$$xy = (a // b) \times (c // d) = (ac // bd);$$

$$yx = (c // d) \times (a // b) = (ac // bd).$$

(f) Prove $x1 = 1x = x$. let $y = 1$, then $x1 = 1x$. Because $1 = 1 // 1$,

$$x1 = (a // b) \times (1 // 1) = a // b = x.$$

(g) Prove $x(y + z) = xy + xz$.

$$x(y + z) = (a // b) \times ((c // d) + (e // f))$$

$$= (a // b) \times ((cf + ed) // df)$$

$$= (acf + aed) // bdf;$$

$$xy + xz = (a // b)(c // d) + (a // b)(e // f)$$

$$= (ac) // (bd) + (ae) // (bf)$$

$$= (b(acf + aed)) // (b(bdf))$$

$$= (acf + aed) // bdf.$$

(h) Prove $(y + z)x = yx + zx$. By conclusions of (e) and (g), we have

$$(y + z)x = x(y + z);$$

$$yx + zx = xy + xz.$$

(i) Prove $xx^{-1} = x^{-1}x = 1$. Let $y = x^{-1}$, then $xx^{-1} = x^{-1}x$.

$$xx^{-1} = (a // b) \times (b // a) = (ab) // (ab) = 1 // 1 = 1.$$

□

Exercise 4.2.4. Prove Lemma 4.2.7. (Note that, as in Proposition 2.2.13, you have to prove two different things: firstly, that at least one of (a), (b), (c) is true; and secondly, that at most one of (a), (b), (c) is true.)

Lemma 4.2.7 (Trichotomy of rationals). Let x be 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. (c) x is a negative rational number.

Proof. First we show that at least one of (a), (b), (c) is true. By Definition 4.2.1, $x = a/b$ for some integers a and non-zero integers b . And we let $y = n/m$ for some positive integers n and positive non-zero integers m . By Lemma 4.1.5, we have $a, b > 0$ or $a, b < 0$ or $a = 0$. If $a, b > 0$, we let $a = n, b = m$, and if $a, b < 0$, let $a = -n, b = -m$. Suppose that $b = m > 0$. If $a = n > 0$ then $x = n/m$ is positive by Definition 4.2.6, and if $a = -n < 0$ then $x = (-n)/m < 0$ by definition. If $a = 0$ then $x = 0$ by definition.¹ Then we suppose that $b = -m < 0$. If $a = n > 0$, then $x = -(n/m) < 0$, and if $a = -n < 0$ then $x = n/m > 0$. If $a = 0$ then $x = 0$. Therefore, at least one of (a), (b), (c) is true.

Now we show that no more than one of (a), (b), (c) can hold at a time. By Lemma 4.1.5, Definition 4.2.6, whatever b , $x = 0$ iff $a = 0$. Thus, (a) and (c) or (b) and (c) cannot simultaneously be true. If (a) and (b) were simultaneously true, fixed b , then it contradicting with Lemma 4.1.5. And it is similarly to fix a . Thus exactly one of (a), (b), (c) is true for any rational x . \square

Exercise 4.2.5. *Prove Proposition 4.2.9.*

Proof. Let x, y , be rational numbers.

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

Let $a = x - y$, then by Lemma 4.2.7, exactly one of $a = 0, a < 0$ or $a > 0$ is true. By Definition 4.2.7, exactly one of $x = y, x < y$, or $x > y$ is true.

(b) (*Order is anti-symmetric*) One has $x < y$ if and only if $y > x$.

¹A rational number $a // b$ is equal to $0 = 0 // 1$ if and only if $a \times 1 = b \times 0$, i.e., if the numerator a is equal to 0. (p.83)

If $x < y$, then by Definition 4.2.7, $x - y$ is a negative rational number. Then $-(x - y) = y - x$ is positive and, by Definition 4.2.7 again, $y > x$.

(c) (*Order is transitive*) If $x < y$ and $y < z$, then $x < z$.

If $x < y$ and $y < z$, by Definition 4.2.7, $y - x$ and $z - y$ are positive. Let $y - x = a/b$ and $z - y = c/d$ for some positive integers a, c and positive non-zero integers b, d . Then we have $(y - x) + (z - y) = z - x = a/b + c/d = (ad + bc)/bd$, which is positive. Thus, $z - x$ is positive and $x < z$.

(d) (*Addition preserves order*) If $x < y$, then $x + z < y + z$.

If $x < y$, then $y - x$ is positive. We need to show that $x + z < y + z$ iff $(y + z) - (x + z) = y - x$ is positive. This is our desired.

(e) (*Positive multiplication preserves order*) If $x < y$ and z is positive, then $xz < yz$.

Let $y - x = a/b, z = c/d$ for some positive integers a, c and non-zero positive integers b, d . We have known that $y - x$ and z are positive. We need to show that $yz - xz$ is positive. Because $yz - xz = (y - x)z = (a/b) \times (c/d) = (ac)/(bd)$, which is a positive. Therefore, $xz < yz$. \square

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

Proof. By Definition 4.2.8, $y - x$ is a positive rational number, we let $y - x = a/b, z = (-c)/d$ for some positive integers a, c and non-zero positive integers b, d . Then $yz - xz = z(y - x) = ((-c)/d) \times (a/b) = (-ac)/bd$, which is negative. Thus, $xz > yz$ by Definition 4.2.8. \square

4.3

Absolute value and exponentiation

Exercise 4.3.1. *Prove Proposition 4.3.3. (Hint: while all of these claims can be proven by dividing into cases, such as when x is positive, negative, or zero several parts of the proposition can be proven without such a tedious division into cases. For instance one can use earlier parts of the proposition to prove later ones.)*

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

- (a) *(non-degeneracy of absolute value)* We have $|x| \geq 0$. Also, $|x| = 0$ if and only if x is 0.

By Lemma 4.2.7, we have exactly one of $x > 0$, $x < 0$ or $x = 0$ is true. If $x > 0$, then $|x| > 0$; and if $x < 0$, then $|x| = -x > 0$. If $x = 0$, then $|x| = 0$. Thus, we have $|x| > 0$ or $|x| = 0$ is true, i.e., $|x| \geq 0$.

- (b) *(Triangle inequality for absolute value)* We have $|x + y| \leq |x| + |y|$.

If $0 \leq x$ and $0 \leq y$, then $0 \leq x + y$, $|x + y| = x + y$, $|x| = x$, and $|y| = y$, so that equality holds in this case.

If $x \leq 0$ and $y \leq 0$, then $x + y < 0$, $|x + y| = -(x + y) = -x - y$, $|x| = -x$, $|y| = -y$, and again we have equality.

Now suppose one of the numbers is negative and the other positive, for example, $x < 0 < y$. Then either $x < x + y < 0$ or $0 < x + y < y$. In the first case $|x + y| < |x|$, and in the second case $|x + y| < |y|$, so that in both cases $|x + y| \leq |x| + |y|$.

- (c) *We have the inequalities $-y \leq x \leq y$ if and only if $y \geq |x|$. In particular, we have $-|x| \leq x \leq |x|$.*

If $y \leq 0$, then $|x| = -x$, $|-y| = y$ and $|y| = -y$, because $y \geq -x$ and $-y \leq x$, so that equality holds in this case.

If $-y \geq 0$, i.e., $y \leq 0$, it is same with former case. Therefore, $y \leq |x|$ when $-y \leq x \leq y$. In particular, we let $y = |x|$, then $-|x| \leq x \leq |x|$.

- (d) (*Multiplicativity of absolute value*) We have $|xy| = |x||y|$. In particular, $|-x| = |x|$.

If $0 \leq x$ and $0 \leq y$, then $0 \leq xy$, $|xy| = xy$, $|x| = x$, and $|y| = y$, so that equality holds in this case.

If $x \leq 0$ and $y \leq 0$, then $xy > 0$, $|xy| = xy$, $|x| = -x$, $|y| = -y$, and $|x||y| = xy$, again we have equality.

Now suppose one of the numbers is negative and the other positive, for example, $x < 0 < y$, then $|xy| = -xy$, and $|x| = -x$, $|y| = y$, so we have equality. There is a similar case that $y < 0 < x$, and we still have equality. so that in both cases $|xy| = |x||y|$.

- (e) (*Non-degeneracy of distance*) We have $d(x, y) \geq 0$. Also, $d(x, y) = 0$ if and only if $x = y$.

Let $z = x - y$, then $d(x, y) = |z|$. Proof is same as statement (a).

- (f) (*Symmetry of distance*) $d(x, y) = d(y, x)$.

If $x < y$, $d(x, y) = |x - y| = -(x - y) = y - x$, and $d(y, x) = |y - x| = y - x$.

If $y < x$, $d(x, y) = |x - y| = x - y$, and $d(y, x) = |y - x| = -(y - x) = x - y$.

If $x = y$, then $d(x, y) = d(y, x) = 0$.

- (g) (*Triangle inequality for distance*) $d(x, z) \leq d(x, y) + d(y, z)$.

Let $a = x - y$, $b = y - z$, because $x - z = (x - y) + (y - z) = a + b$.

So $d(x, z) \leq d(x, y) + d(y, z)$ means $|a + b| \leq |a| + |b|$, which we have proved in statement (b). \square

Exercise 4.3.2. Prove the remaining claims in Proposition 4.3.7.

Proof.

- (a) If $x = y$, then x is ε -close to y for every $\varepsilon > 0$. Conversely, if x is ε -close to y for every $\varepsilon > 0$, then we have $x = y$.

If $x = y$, then $d(x, y) = 0 < \varepsilon$. Use contradiction. Suppose that $x \neq y$. Then $d(x, y) = |x - y| > 0$. Because x is ε -close to y , $d(x, y) = |x - y| \leq \varepsilon$ by definition. Let $\varepsilon = |x - y|/2$, then $|x - y| = 2\varepsilon > \varepsilon$. It's contradictory, $x = y$.

(b) Let $\varepsilon > 0$. If x is ε -close to y , then y is ε -close to x .

Suppose that x is ε -close to y , then $d(x, y) = |x - y| \leq \varepsilon$ by definition. By Proposition 4.3.3 (f), $d(y, x) = d(x, y) \leq \varepsilon$ for every $\varepsilon > 0$. Thus, y is ε -close to x .

(c) Let $\varepsilon, \delta > 0$. If x is ε -close to y , and y is δ -close to z , then x and z are $(\varepsilon + \delta)$ -close.

Suppose that x and y are ε -close, y and z are δ -close, then $d(x, y) = |x - y| \leq \varepsilon$, $d(y, z) = |y - z| \leq \delta$. Then

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

Therefore, x and z are $(\varepsilon + \delta)$ -close.

(d) Let $\varepsilon, \delta > 0$. If x and y are ε -close, and z and w are δ -close, then $x + z$ and $y + w$ are $(\varepsilon + \delta)$ -close, $x - z$ and $y - w$ are also $(\varepsilon + \delta)$ -close.

Suppose x and y are ε -close, and z and w are δ -close, then $d(x, y) = |x - y| \leq \varepsilon$, $d(z, w) = |z - w| \leq \delta$. Because

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

$x + z$ and $y + w$ are $(\varepsilon + \delta)$ -close. And from

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

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

(e) Let $\varepsilon > 0$. If x and y are ε -close, they are also ε' -close for every $\varepsilon' > \varepsilon$.

If x and y are ε -close and $\varepsilon' > \varepsilon$, then $d(x, y) \leq \varepsilon < \varepsilon'$. So $d(x, y) \leq \varepsilon'$ and they are ε' -close.

- (f) Let $\varepsilon > 0$. If y and z are both ε -close to x , and w is between y and z (i.e., $y \leq w \leq z$ or $z \leq w \leq y$), then w is also ε -close to x .

Suppose that y and z are both ε -close to x , $d(y, x) = |y - x| \leq \varepsilon$, $d(z, x) = |z - x| \leq \varepsilon$. If $y \leq w \leq z$, then $y - x \leq w - x \leq z - x$. By hypotheses and Proposition 4.3.3 (c), $-\varepsilon \leq y - x \leq \varepsilon$ and $-\varepsilon \leq z - x \leq \varepsilon$. Therefore, $-\varepsilon \leq w - x \leq \varepsilon$, by Proposition 4.3.3 (c) again, $d(w, x) = |w - x| \leq \varepsilon$. A similar argument shows that $d(w, x) \leq \varepsilon$ when $z \leq w \leq y$. So, w is also ε -close to x .

- (g) Let $\varepsilon > 0$. If x and y are ε -close, and z is non-zero, then xz and yz are $\varepsilon|z|$ -close.

Suppose that x and y are ε -close, then $d(x, y) = |x - y| \leq \varepsilon$. z is non-zero, $\varepsilon|z| > 0$, and $d(xz, yz) = |xz - yz| = |x - y||z| \leq \varepsilon|z|$. \square

Exercise 4.3.3. Prove Proposition 4.3.10. (Hint: use induction.)

Proof. Let x, y be rational numbers. and let n, m be natural numbers.

- (a) We have $x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, and $(xy)^n = x^n y^n$.

Prove $x^n x^m = x^{n+m}$. We fix m and induct on n . Consider the base case $n = 0$, $x^0 \times x^m = 1 \times x^m = x^m$ is hold by Definition 4.3.9. Now assume inductively that $x^n \times x^m = x^{n+m}$. We have to show that $x^{n+1} \times x^m = x^{n+m+1}$. By Definition 4.3.9, $x^{n+1} = x^n \times x$, then by induction hypothesis, $x^{n+1} \times x^m = x^n \times x \times x^m = x^{m+n} \times x = x^{m+n+1}$. This close the induction.

Prove $(x^n)^m = x^{nm}$. We fix n and induct on m . Since $(x^n)^0 = 1$, and $x^{n \times 0} = x^0 = 1$, equality is hold when $m = 0$. Now assume inductively that $(x^n)^m = x^{nm}$. We have to show that $(x^n)^{m+1} = x^{n(m+1)}$. By Definition 4.3.9, $(x^n)^{m+1} = (x^n)^m \times x^n$, then $x^{nm} \times x^n = x^{nm+n} = x^{n(m+1)}$. This close the induction.

Prove $(xy)^n = x^n y^n$. Use induction on n . When $n = 0$, $(xy)^0 = 1$, and $x^0 y^0 = 1 \times 1 = 1$. So, equality is hold. Now assume inductively that $(xy)^n = x^n y^n$. We have to show that $(xy)^{n+1} = x^{n+1} y^{n+1}$. By

Definition 4.3.9, $(xy)^{n+1} = x^n \times y^n \times x \times y = x^{n+1}y^{n+1}$. This close the induction.

(b) Suppose $n > 0$. Then we have $x^n = 0$ if and only if $x = 0$.

$x = 0 \implies x^n = 0$. Suppose that $x = 0$, we need to show that $x^n = 0$. Use induction on n . When $n = 1$, $0^1 = 0$. Now we inductively assume that $0^n = 0$, we have to show that $x^{n+1} = 0$. By Definition 4.3.9, $0^{n+1} = 0^n \times 0 = 0 \times 0 = 0$. This close the induction.

$x^n = 0 \implies x = 0$. Suppose for contradiction that $x^n = 0$ for $n > 0$ and $x \neq 0$. Then there exists $n = 1$ such that $x^n \neq 0$, a contradiction.

(c) If $x \geq y \geq 0$, then $x^n \geq y^n \geq 0$. If $x > y \geq 0$ and $n > 0$, then $x^n > y^n \geq 0$.

Suppose $x \geq y \geq 0$, we use induction on n . When $n = 0$, inequality is hold. Inductively assume that $x^n \geq y^n \geq 0$, we have to show that $x^{n+1} \geq y^{n+1} \geq 0$. Since

$$x^{n+1} - y^{n+1} = x^n x - y^n y \geq x^n y - y^n y = (x^n - y^n)y \geq 0,$$

this is hold for $y \geq 0$ and $x^n \geq y^n \geq 0$. Hence $x^n x \geq y^n y$. And we also have $x^{n+1} \geq 0, y^{n+1} \geq 0$. Thus, $x^{n+1} \geq y^{n+1} \geq 0$ is hold. This close the induction.

Then we suppose that $x > y \geq 0$ and $n > 0$. Induct on n . inequality is hold when $n = 1$, and in the case of $n = 0$ is vacuously true. Now inductively assume that $x^n > y^n \geq 0$. We have to show that $x^{n+1} > y^{n+1} \geq 0$. Since $y \geq 0$ and $y^n \geq 0$, we have $y^{n+1} \geq 0$. We have

$$x^{n+1} - y^{n+1} = x^n x - y^n y > x^n y - y^n y = (x^n - y^n)y.$$

When $y = 0$, $x^{n+1} > y^{n+1}$; otherwise, $x^n > y^n$ implies that $x^{n+1} > y^{n+1}$. Thus we have $x^{n+1} > y^{n+1} \geq 0$. This close the induction.

(d) We have $|x^n| = |x|^n$.

We use induction on n . Equality is hold when $n = 0$. Now we inductively assume that $|x^n| = |x|^n$. We need to show that $|x^{n+1}| = |x|^{n+1}$. Consider left-hand side, $|x^{n+1}| = |x^n \times x| = |x^n||x| = |x|^n|x|$. Then

consider right-hand side, $|x|^{n+1} = |x|^n|x|$. So that equality is hold, and this close the induction. \square

Exercise 4.3.4. *Prove Proposition 4.3.12. (Hint: induction is not suitable here. Instead, use Proposition 4.3.10.)*

Proof. Let x, y be non-zero rational numbers. and let n, m be integers.

(a) We have $x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, and $(xy)^n = x^n y^n$.

Prove $x^n x^m = x^{n+m}$. If $n < 0, m < 0$,

$$x^n x^m = 1/x^{-n} \times 1/x^{-m} = 1/(x^{-n}x^{-m}) = 1/x^{-(n+m)} = x^{n+m}.$$

If $n < 0 \leq m$,

$$x^n x^m = 1/x^{-n} \times x^m = x^m/x^{-n} = 1/x^{-(n+m)} = x^{n+m}.$$

If $m < 0 \leq n$,

$$x^n x^m = x^n \times 1/x^{-m} = x^n/x^{-m} = 1/x^{-(n+m)} = x^{n+m}.$$

Prove $(x^n)^m = x^{nm}$. If $n < 0, m < 0$,

$$(x^n)^m = (1/x^{-n})^m = 1/x^{-nm} = x^{nm}.$$

If $n < 0 \leq m$,

$$(x^n)^m = (1/x^{-n})^m = 1/x^{-nm} = x^{nm} = x^{nm}.$$

If $m < 0 \leq n$,

$$(x^n)^m = (x^n)^m = 1/(x^{-n})^m = 1/x^{-nm} = x^{nm}.$$

Prove $(xy)^n = x^n y^n$. If $n < 0$,

$$(xy)^n = 1/(xy)^{-n} = 1/(x^{-n}y^{-n}) = 1/x^{-n} \times 1/y^{-n} = x^n y^n.$$

(b) If $x \geq y > 0$, then $x^n \geq y^n > 0$ if n is positive, and $0 < x^n \leq y^n$ if n is negative.

By Proposition 4.3.10(c), if $x > y$ and $n > 0$, then $x^n > y^n$. We replace x by y and replace y by 0. Hence we have $y > 0$ and $n > 0$, then $y^n > 0$. By Proposition 4.3.10(c) again, we have $x \geq y > 0$, then $x^n \geq y^n > 0$ if $n > 0$.

Let $n < 0$. If $x \geq y > 0$, by Proposition 4.3.10(c), $x^{-n} \geq y^{-n} > 0$. Then

$$x^{-n} \geq y^{-n} > 0 \implies 1/x^n \geq 1/y^n > 0 \implies y^n \geq x^n > 0.$$

(c) If $x, y > 0, n \neq 0$, and $x^n = y^n$, then $x = y$.

If $n > 0$. Suppose that $x \neq y$. Then we can divide it into two cases: $x > y > 0$ or $y > x > 0$. Consider $x > y > 0$, by Proposition 4.3.10(c), $x^n > y^n$. This is contradictive with hypothesis. A similar argument shows that $y^n > x^n$ when $y > x > 0$. Thus if $n > 0$ and $x^n = y^n$, then $x = y$.

If $n < 0$. Suppose $x \neq y$. Then we can divide it into two cases: $x > y > 0$ or $y > x > 0$. Consider $x > y > 0$, by Proposition 4.3.12(b), $x^n < y^n$. This is contradictive with hypothesis. A similar argument shows that $y^n < x^n$ when $y > x > 0$. Thus $n < 0$ and $x^n = y^n$, then $x = y$.

Thus if $x, y > 0, n \neq 0$, and $x^n = y^n$, then $x = y$.

(d) We have $|x^n| = |x|^n$.

If $n < 0$. We have

$$|x^n| = |(1/x)^{-n}| = |1/x|^{-n} = |x|^n.$$

If $n \geq 0$, we following from Proposition 4.3.10(d). □

Exercise 4.3.5. Prove that $2^N \geq N$ for all positive integers N . (Hint: use induction.)

Proof. Use induction on N . When $N = 1$, $2^1 = 2 \geq 1$. Now we inductively assume that $2^N \geq N$, we have to show that $2^{N+1} \geq N + 1$:

$$2^{N+1} = 2^N \times 2 \geq N + N \geq N + 1, \text{ for } N \geq 1.$$

This close the induction. □

4.4

Gaps in the rational numbers

Exercise 4.4.1. *Prove Proposition 4.4.1. (Hint: use Proposition 2.3.9.)*

Proposition 4.4.1 (Interspersing of integers by rationals). Let x be a rational number. Then there exists an integer such that $n \leq x < n + 1$. In fact, this integer is unique (i.e., for each x there is only one n for which $n \leq x < n + 1$). In particular, there exists a natural number N such that $N > x$ (i.e., there is no such thing as a rational number which is larger than all the natural numbers).

Proof. By Lemma 4.2.7, one of $x > 0$, $x < 0$ or $x = 0$ is true. First of all consider $x > 0$. Let $x = a/b$ for some positive integers a and b . By Proposition 2.3.9, there exist natural numbers n, r such that $0 \leq r < b$ and

$$a = nb + r,$$

so that

$$a/b = n + r/b.$$

Because $0 \leq r < b$, then $0 \leq r/b < 1$. Therefore, $n \leq a/b < n + 1$.

If $x < 0$, let $x = -a/b$ for some positive integers a and b . Again, by Proposition 2.3.9, there exist natural numbers m, r such that $0 \leq r < b$ and

$$-a = -mb - r \implies -a/b = -m - r/b = (1 - m) + (1 - r/b)$$

in which, $0 \leq 1 - r/b < 1$. Let $n = 1 - m$, then $n \leq -a/b < n + 1$.

If $x = 0$, this follows immediately from Proposition 4.1.1.

Now we turn to prove that such integer n is unique. Suppose that there exist two integer numbers n_1 and n_2 satisfy the equality, which means:

$$n_1 \leq x < n_1 + 1, \quad \text{and} \quad n_2 \leq x < n_2 + 1,$$

then we have $n_1 < n_2 + 1$ and $n_2 < n_1 + 1$, so that $n_1 \leq n_2$ and $n_2 \leq n_1$. Therefore, $n_1 = n_2$. \square

Exercise 4.4.2. *A definition: a sequence a_0, a_1, a_2, \dots 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., $a_0 > a_1 > a_2 > \cdots$).

- (a) Prove the **principle of infinite descent**: that it is not possible to have a sequence of **natural numbers** which is in infinite descent. (Hint: assume for sake of contradiction that you can find a sequence of natural numbers which is in infinite descent. Since all the a_n are natural numbers, you know that $a_n \geq 0$ for all n . Now use induction to show in fact that $a_n \geq k$ for all $k \in \mathbf{N}$ and all $n \in \mathbf{N}$, and obtain a contradiction.)
- (b) Does the principle of infinite descent work if the sequence a_1, a_2, a_3, \dots 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.

Proof. Suppose we can find a sequence of natural numbers a_0, a_1, a_2, \dots which is in infinite descent. Since all the a_n are natural numbers, by definition, for all n , $a_n > a_{n+1} \geq 0$. We use induction on k to show that $a_N > k$ for all $k \in \mathbf{N}$ and all $N < n \in \mathbf{N}$. inequality is hold when $k = 0$. Now we inductively assume that $a_N > k$, we have to show that $a_N > k + 1$. Use contradiction, assume that $a_N < k + 1$, then $a_N \leq k$. This is a contradiction with $a_N > k$. Therefore, $a_n \geq k$ for all $k \in \mathbf{N}$ and all $n \in \mathbf{N}$. Let $k = a_0$, then $a_n \geq a_0$, which contradicts the principle of infinite descent. \square

Exercise 4.4.3. Fill in the gaps marked (why?) in the proof of Proposition 4.4.4.

Proposition 4.4.4. There does not exist any rational number x for which $x^2 = 2$.

Proof.

Every natural number is either even or odd, but not both (why?).

For every natural number p , it is even if $p = 2k$ and odd if $p = 2k + 1$ for some natural numbers k . Suppose p is both odd and even, let $p = 2n$, then $2n = 2k + 1 \implies n = k + 1/2$, which is not a natural number.

If p is odd, then p^2 is also odd (why?), which contradicts $p^2 = 2q^2$.

Let $p = 2k + 1$, Then $p^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2a + 1$ in which $2k = 4k^2 + 4k$.

Since $p^2 = 2q^2$, we have $q < p$ (why?).

CHAPTER V

The real numbers

5.1

Cauchy sequences

Exercise 5.1.1. *Prove Lemma 5.1.15. (Hint: use the fact that a_n is eventually 1-steady, and thus can be split into a finite sequence and a 1-steady sequence. Then use Lemma 5.1.14 for the finite part. Note there is nothing special about the number 1 used here; any other positive number would have sufficed.)*

Lemma 5.1.15 (Cauchy sequences are bounded). Every Cauchy sequence $(a_n)_{n=1}^{\infty}$ is bounded.

Proof. Suppose that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence, then we have $|a_n - a_k| \leq 1$ for all $n, k \geq N$. Therefore, there exists N such that $|a_n - a_N| \leq 1$. By Lemma 5.1.14, $(a_n)_{n=1}^N$ is a finite sequence, then $|a_N| < M$ for some $M \geq 0$. Because

$$|a_n| = |a_n - a_N + a_N| \leq |a_n - a_N| + |a_N| \leq 1 + M.$$

By Definition 5.1.12, $(a_n)_{n=1}^{\infty}$ is bounded. □

5.2

Equivalent Cauchy sequences

Exercise 5.2.1. Show that if $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent sequences of rationals, then $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence if and only if $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence.

Proof. Since $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent sequences of rationals, by Definition 5.2.6, for every rational $\varepsilon_1 > 0$, there exists an $N_1 \geq 1$ such that $|a_i - b_i| \leq \varepsilon_1$ for all $i \geq N_1$. Suppose that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence, then for every $\varepsilon_2 > 0$, there exists an $N_2 \geq 1$ such that $|a_j - a_k| \leq \varepsilon_2$ for all $j, k \geq N_2$. Let $N = N_1 + N_2$, $\varepsilon = 2\varepsilon_1 + \varepsilon_2$. If $N_1 > N_2$, we have

$$\begin{aligned} |b_n - b_m| &= |b_n - a_i + a_i - a_j + a_j - b_m| \\ &\leq |b_n - a_i| + |a_i - a_j| + |a_j - b_m| \\ &\leq \varepsilon_1 + \varepsilon_2 + \varepsilon_1 = \varepsilon, \end{aligned}$$

for all $n, m \geq N$. Otherwise, we have

$$\begin{aligned} |b_n - b_m| &= |b_n - a_j + a_j - a_k + a_k - b_m| \\ &\leq |b_n - a_j| + |a_j - a_k| + |a_k - b_m| \\ &\leq \varepsilon_1 + \varepsilon_2 + \varepsilon_1 = \varepsilon, \end{aligned}$$

for all $n, m \geq N$. By Definition 5.1.8, $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence. A similar argument shows that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence when we suppose that $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence. This is our desired. \square

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.

Proof. Since $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are eventually ε -close, there exists an $N \geq 1$ such that $|a_n - b_n| \leq \varepsilon$ for all $n \geq N$. We suppose that $(a_n)_{n=1}^{\infty}$ is bounded by $M_1 \geq 0$, by Definition 5.1.12, we have $|a_i| \leq M_1$ for all $i \geq 1$. Since we have

$$|b_n| = |b_n - a_n + a_n| \leq |a_n - b_n| + |a_n| \leq \varepsilon + M_1$$

Let $M = \varepsilon + M_1$, then for all $i > 1$, we have $|b_i| \leq M$. By Definition 5.1.12, $(b_n)_{n=1}^\infty$ is bounded. A similar argument shows that $(a_n)_{n=1}^\infty$ is bounded when we assume that $(b_n)_{n=1}^\infty$ is bounded. \square

5.3**The construction of the real numbers**

Exercise 5.3.1. *Prove Proposition 5.3.3. (Hint: you may find Proposition 4.3.7 to be useful.)*

Proposition 5.3.3 (Formal limits are well-defined). Let $x = \text{LIM}_{n \rightarrow \infty} a_n$, $y = \text{LIM}_{n \rightarrow \infty} b_n$, and $z = \text{LIM}_{n \rightarrow \infty} c_n$ be real numbers. Then, with the above definition of equality for real numbers, we have $x = x$. Also, if $x = y$, then $y = x$. Finally, if $x = y$ and $y = z$, then $x = z$.

Proof. For every $\varepsilon > 0$, By Definition 5.2.6 and Definition 5.3.1, we obviously have an $N \geq 0$ such that $|a_n - a_n| = 0 \leq \varepsilon$ for all $n \geq N$. So that $x = x$.

Suppose that $x = y$. By Definition 5.2.6 and Definition 5.3.1, there exists an $N \geq 0$ such that $|a_n - b_n| \leq \varepsilon$, then $|b_n - a_n| \leq \varepsilon$ for all $n \geq N$. Hence, $(b_n)_{n=1}^\infty$ and $(a_n)_{n=1}^\infty$ are equivalent Cauchy sequences and $y = x$.

Suppose that $x = y$ and $y = z$. By Definition 5.2.6 and Definition 5.3.1, there exists $N_1, N_2 \geq 0$ such that $|a_j - b_j| \leq \varepsilon_1$ and $|b_k - c_k| \leq \varepsilon_1$ for all $j \geq N_1, k \geq N_2$. Let $\varepsilon = \varepsilon_1 + \varepsilon_2 > 0$ and $N = N_1 + N_2 \geq 0$. If $N_1 > N_2$, then we have

$$\begin{aligned} |a_n - c_n| &= |a_n - b_j + b_j - c_n| \\ &\leq |a_n - b_j| + |b_j - c_n| \\ &\leq \varepsilon_1 + \varepsilon_2 = \varepsilon. \end{aligned}$$

for all $n \geq N$. Otherwise, we have

$$\begin{aligned} |a_n - c_n| &= |a_n - b_k + b_k - c_n| \\ &\leq |a_n - b_k| + |b_k - c_n| \\ &\leq \varepsilon_1 + \varepsilon_2 = \varepsilon. \end{aligned}$$

Therefore, $(a_n)_{n=1}^\infty$ and $(c_n)_{n=1}^\infty$ are equivalent Cauchy sequences and $x = z$. \square

Exercise 5.3.2. Prove Proposition 5.3.10. (Hint: again, Proposition 4.3.7 may be useful.)

Proposition 5.3.10 (Multiplication is well defined). Let $x = \text{LIM}_{n \rightarrow \infty} a_n$, $y = \text{LIM}_{n \rightarrow \infty} b_n$, and $x' = \text{LIM}_{n \rightarrow \infty} a'_n$ be real numbers. Then xy is also a real number. Furthermore, if $x = x'$, then $xy = x'y$.

Proof. We need to show that for every $\varepsilon > 0$, the sequence $(a_n b_n)_{n=1}^{\infty}$ is eventually ε -steady. Since $(a_n)_{n=1}^{\infty}$ and $(c_n)_{n=1}^{\infty}$ are Cauchy sequences, by Lemma 5.1.15, there exist some rational numbers $A \geq 0$ and $B \geq 0$ such that $|a_n| \leq A$ and $|b_n| \leq B$ for all $i \geq 1$.

Let $\delta', \delta'' > 0$. Since $(a_n)_{n=1}^{\infty}$ is eventually δ' -steady, there exists an $N' \geq 1$ such that a_n and a_m are δ' -close for every $n, m \geq N'$. Similarly, there exists an $N'' \geq 1$ such that b_n and b_m are δ'' -close for every $n, m \geq N''$. By Proposition 4.3.7(h), we know that

$$|a_n b_n - a_m b_m| \leq \delta'' |a_n| + \delta' |b_n| + \delta' \delta''.$$

Let $\varepsilon > 0$, and let

$$\delta' = \min \left(1, \frac{\varepsilon}{3(B+1)} \right), \quad \delta'' = \min \left(1, \frac{\varepsilon}{3(A+1)} \right).$$

Since

$$|a_n| \leq A < A+1, \quad |b_n| \leq B < B+1,$$

it's easy to verify that ¹

$$\delta'' |a_n| \leq \min \left(1, \frac{\varepsilon}{3(A+1)} \right) \cdot (A+1) \leq \frac{\varepsilon}{3},$$

$$\delta' |b_n| \leq \min \left(1, \frac{\varepsilon}{3(B+1)} \right) \cdot (B+1) \leq \frac{\varepsilon}{3},$$

$$\delta' \delta'' = \min \left(1, \frac{\varepsilon}{3(B+1)} \right) \cdot \min \left(1, \frac{\varepsilon}{3(A+1)} \right) \leq \frac{\varepsilon}{3}.$$

Hence for every $n, m \geq N = \max(N', N'')$, we have

$$|a_n b_n - a_m b_m| \leq \varepsilon,$$

i.e., $(a_n b_n)_{n=1}^{\infty}$ and $(a_m b_m)_{n=1}^{\infty}$ are ε -close. This implies that the sequence $(a_n b_n)_{n=1}^{\infty}$ is eventually ε -close, as desired.

Now we prove that the multiplication of equivalent Cauchy sequences are equivalent. Since x and x' are equal, we know that the Cauchy sequences $(a_n)_{n=1}^{\infty}$ and $(a'_n)_{n=1}^{\infty}$ are equivalent. In other words, they are eventually ε -close, so that ε/M -close for each $\varepsilon > 0$. We need to show that the sequences

¹If $1 < \frac{\varepsilon}{3(A+1)}$, then $A+1 < \frac{\varepsilon}{3}$. We have $\delta'' |a_n| \leq 1 \cdot (A+1) < \frac{\varepsilon}{3}$. Conversely, if $\frac{\varepsilon}{3(A+1)} < 1$, then we have $\delta'' |a_n| \leq \frac{\varepsilon}{3(A+1)} \cdot (A+1) = \frac{\varepsilon}{3}$. This is similar to show that $\delta' |b_n| \leq \frac{\varepsilon}{3}$ and $\delta' \delta'' \leq \frac{\varepsilon}{3}$.

$(a_n b_n)_{n=1}^{\infty}$ and $(a'_n b_n)_{n=1}^{\infty}$ are eventually ε -close for each $\varepsilon > 0$. But we already know that there is an $N \geq 1$ such that $(a_n)_{n=1}^{\infty}$ and $(a'_n)_{n=1}^{\infty}$ are ε/M -close, i.e., a_n and a'_n are ε/M -close for each $n \geq N$. Since $(b_n)_{n=1}^{\infty}$ is a Cauchy sequences, there exist a rational number $M \geq 0$ such that $|b_n| \leq M$. By Proposition 4.3.7(g), $a_n b_n$ and $a'_n b_n$ are ε -close for every $n \geq N$. This implies that $(a_n b_n)_{n=1}^{\infty}$ and $(a'_n b_n)_{n=1}^{\infty}$ are eventually ε -close for every $\varepsilon > 0$. \square

Exercise 5.3.3. *Let a, b be rational numbers. Show that $a = b$ if and only if $\text{LIM}_{n \rightarrow \infty} a = \text{LIM}_{n \rightarrow \infty} b$ (i.e., the Cauchy sequences a, a, a, a, \dots and b, b, b, b, \dots equivalent if and only if $a = b$). This allows us to embed the rational numbers inside the real numbers in a well-defined manner.*

Proof. Suppose that $a = b$, then for every $\varepsilon > 0$, there exists an $N \geq 0$ such that $|a - b| = 0 \leq \varepsilon$ for all $n \geq N$. So $(a)_{n=1}^{\infty}$ and $(b)_{n=1}^{\infty}$ are equivalent Cauchy sequences, then $\text{LIM}_{n \rightarrow \infty} a = \text{LIM}_{n \rightarrow \infty} b$.

We use contradiction. Suppose that $\text{LIM}_{n \rightarrow \infty} a = \text{LIM}_{n \rightarrow \infty} b$ and $a \neq b$. By definition, $(a)_{n=1}^{\infty}$ and $(b)_{n=1}^{\infty}$ are equivalent Cauchy sequences, then $|a - b| \leq \varepsilon$ for all $n \geq N$. Because $a \neq b$, let $\delta = |a - b| > 0$ and let $\varepsilon = \delta/3$. So that $\varepsilon = |a - b|/3 < |a - b|$, a contradiction. \square

Exercise 5.3.4. *Let $(a_n)_{n=0}^{\infty}$ be a sequence of rational numbers which is bounded. Let $(b_n)_{n=0}^{\infty}$ be another sequence of rational numbers which is equivalent to $(a_n)_{n=0}^{\infty}$. Show that $(b_n)_{n=0}^{\infty}$ is also bounded. (Hint: use Exercise 5.2.2.)*

Proof. Since $(b_n)_{n=0}^{\infty}$ is equivalent to $(a_n)_{n=0}^{\infty}$, they are eventually ε -close for $\varepsilon > 0$. Then $(b_n)_{n=0}^{\infty}$ is bounded by Exercise 5.2.2. \square

Exercise 5.3.5. *Show that $\text{LIM}_{n \rightarrow \infty} 1/n = 0$.*

Proof. Let $\varepsilon > 0$. There exists an $N = 1/\varepsilon \geq 0$ such that

$$\left| \frac{1}{n} - 0 \right| = \left| \frac{1}{n} \right| \leq \frac{1}{N} = \varepsilon, \quad \text{for all } n \geq N$$

Thus, $(1/n)_{n=1}^{\infty}$ is equivalent to $0 = (0)_{n=1}^{\infty}$. \square

5.4

Ordering the reals

Exercise 5.4.1. *Prove Proposition 5.4.4. (Hint: if x is not zero, and x is the formal limit of some sequence $(a_n)_{n=1}^{\infty}$, then this sequence cannot be eventually ε -close to the zero sequence $(0)_{n=1}^{\infty}$ for every single $\varepsilon > 0$. Use this to show that the sequence $(a_n)_{n=1}^{\infty}$ is eventually either positively bounded away from zero or negatively bounded away from zero.)*

Proposition 5.4.4 (Basic properties of positive reals). For every real number x , exactly one of the following three statements is true: (a) x is zero; (b) x is positive; (c) x is negative. A real number x is negative if and only if $-x$ is positive. If x and y are positive, then so are $x + y$ and xy .

Proof. If $x \neq 0$, by Lemma 5.3.14, $x = \text{LIM}_{n \rightarrow \infty} a_n$ for some Cauchy sequences $(a_n)_{n=1}^{\infty}$ which is bounded away from zero, i.e., there exists a rational number $c > 0$ such that $|a_n| \geq c$, so that $a_n \geq c$ or $a_n \leq -c$ for all $n \geq 1$. (b) and (c) would not hold at the time.

If $x = 0$, then $\text{LIM}_{n \rightarrow \infty} a_n = \text{LIM}_{n \rightarrow \infty} 0$. But it's not hold for that $a_n = 0 \geq c$ when $c > 0$ or $a_n = 0 \leq -c$ when $-c < 0$. Therefore (a) and (b), or (a) and (c) would not hold at the time.

Since x is negative, $x = \text{LIM}_{n \rightarrow \infty} a_n$ for some Cauchy sequences $(a_n)_{n=1}^{\infty}$ is negatively bounded away from zero, i.e., we have a negative rational $-c < 0$ such that $a_n \leq -c$ for all $n \geq 1$. Let $b_n = -a_n$. Then $b_n \geq c$ for all $n \geq 1$ and positive rational $c > 0$. Because $\text{LIM}_{n \rightarrow \infty} -a_n = -\text{LIM}_{n \rightarrow \infty} a_n = -x$. Thus $-x$ is positive. The claim that x is negative is proven similarly when we assume that $-x$ is positive.

Since real numbers x and y are positive, by definition, we know that $x = \text{LIM}_{n \rightarrow \infty} a_n$ and $y = \text{LIM}_{n \rightarrow \infty} b_n$ for some Cauchy sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ which are positively bounded away from zero, i.e., there exists a positive rational $c > 0$ such that $a_n \geq c/2$ and $b_n \geq c/2$ for all $n \geq 1$.

Then we have $a_n + b_n \geq 2c \geq c$ for all $n \geq 1$. Hence, $x + y = (\text{LIM}_{n \rightarrow \infty} a_n) + (\text{LIM}_{n \rightarrow \infty} b_n) = \text{LIM}_{n \rightarrow \infty} (a_n + b_n)$ is positive. Again, let $c = \max(1, c)$, then

$$a_n b_n \geq c^2 = \max(1, c) \cdot \max(1, c) \geq c, \quad \text{for all } n \geq 1$$

Therefore, $xy = (\text{LIM}_{n \rightarrow \infty} a_n)(\text{LIM}_{n \rightarrow \infty} b_n) = \text{LIM}_{n \rightarrow \infty} (a_n b_n)$ is positive. \square

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

Proposition 5.4.7. All the claims in Proposition 4.2.9 which held for rationals, continue to hold for real numbers

Proof.

- (a) (*Order trichotomy*) Exactly one of the three statements $x = y$, $x < y$, or $x > y$ is true.

Let $a = x - y$, then by Proposition 5.4.4, exactly one of $a = 0$, $a < 0$ or $a > 0$ is true. Then, exactly one of $x = y$, $x < y$, or $x > y$ is true.

- (b) (*Order is anti-symmetric*) One has $x < y$ if and only if $y > x$.

If $x < y$, then $x - y$ is a negative real number. Then $-(x - y) = y - x$ is positive, so that $y > x$.

- (c) (*Order is transitive*) If $x < y$ and $y < z$, then $x < z$.

If $x < y$ and $y < z$, then $y - x$ and $z - y$ are positive. We have $(y - x) + (z - y) = z - x$ is positive. Thus $x < z$.

- (d) (*Addition preserves order*) If $x < y$, then $x + z < y + z$.

If $x < y$, then $y - x$ is positive. We need to show that $x + z < y + z$ iff $(y + z) - (x + z) = y - x$ is positive. \square

Exercise 5.4.3. *Show that for every real number x there is exactly one integer N such that $N \leq x < N + 1$. (This integer N is called the integer part of x , and is sometimes denoted $N = \lfloor x \rfloor$.)*

Proof. There exists a positive integer such that $(M - 1)\varepsilon \leq x < M\varepsilon$. This follows immediately from the Proposition 5.4.12 and Archimedean

property. We have to show that such an M is unique. Suppose that there exist two integer numbers M_1 and M_2 satisfying the inequality. Then we have $(M_1 - 1)\varepsilon \leq x < M_2\varepsilon$ and $(M_2 - 1)\varepsilon \leq x < M_2\varepsilon$. If $M_1 \neq M_2$, we can assume that $M_1 < M_2$, then

$$(M_1 - 1)\varepsilon \leq x < M_1\varepsilon \leq (M_2 - 1)\varepsilon \leq x < M_2\varepsilon,$$

so that

$$x < x.$$

This is a contradiction. Now we rewrite the *Archimedean property* following:

Archimedean property. Let x and ε be any positive real numbers. Then there exists a unique positive integer M such that $(M - 1)\varepsilon \leq x < M\varepsilon$.

Then $N \leq x < N + 1$ follows immediately from the Archimedean property. \square

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

Proof. Replace x by 1, ε by x and M by N , then following by Corollary 5.4.13,

$$Nx > 1 > 0 \implies x > 1/N > 0.$$

Conclusion follows. \square

Exercise 5.4.5. Prove Proposition 5.4.14. (Hint: use Exercise 5.4.4. You may also need to argue by contradiction.)

Proposition 5.4.14. Given any two real numbers $x < y$, we can find a rational number q such that $x < q < y$.

Proof. Since $x < y$, we have $y - x > 0$, and by Corollary 5.4.13, there is a integer n such that

$$n(y - x) > 1.$$

Apply Archimedean property again, to obtain positive integers m_1 and m_2

such that $m_1 > nx, m_2 > -nx$. Then

$$-m_2 < nx < m_1.$$

Hence there is an integer m (with $-m_2 \leq m \leq m_1$) such that

$$m - 1 \leq nx < m.$$

If we combine these inequalities, we obtain

$$nx < m \leq 1 + nx < ny.$$

Since $n > 0$, it follows that

$$x < \frac{m}{n} < y.$$

This proves claim, with $q = m/n$. □

Remark 5.4.5. This proposition may be stated by saying that \mathbf{Q} is dense in \mathbf{R} : Between any two real numbers there is a rational one.

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 $|x - y| \leq \varepsilon$ if and only if $y - \varepsilon \leq x \leq y + \varepsilon$.

Proof. Suppose that $|x - y| < \varepsilon$. If $x = y$, we have $0 < \varepsilon$. While if $x > y$, then

$$|x - y| = x - y < \varepsilon \implies x < y + \varepsilon.$$

If $x < y$, then

$$|x - y| = y - x < \varepsilon \implies y - \varepsilon < x.$$

Combine the inequalities we have $y - \varepsilon < x < y + \varepsilon$.

Then we suppose that $y - \varepsilon < x < y + \varepsilon$. Then we have

$$y - x < \varepsilon, \quad \text{and} \quad x - y < \varepsilon.$$

So, $|x - y| < \varepsilon$.

We only prove the equality in the second claim. Suppose that $|x - y| = \varepsilon$. We see that $|x - y| = x - y = \varepsilon$, then $y + \varepsilon = x$ when $x > y$, and $|x - y| = y - x = \varepsilon$, then $y - \varepsilon = x$ when $x < y$, and $x \neq y$ for that $\varepsilon > 0$. Combine the first statement we have $y - \varepsilon \leq x \leq y + \varepsilon$. Then suppose that $y - \varepsilon = x = y + \varepsilon$, it implies that $\varepsilon = -(x - y) = x - y$. So that $|x - y| = \varepsilon$. Therefore, $|x - y| \leq \varepsilon$. □

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

Proof. Since $x \leq y$, $x \leq y + \varepsilon$ for all real numbers $\varepsilon > 0$. Now we use contradiction, suppose that $x > y$ and $x \leq y + \varepsilon$. Because $x - y > 0$, we can let $\varepsilon = (x - y)/2$. Then

$$x \leq y + \varepsilon = y + \frac{x - y}{2} = \frac{x + y}{2},$$

it implies that $x \leq y$, which contradicts our hypothesis.

$|x - y| \leq \varepsilon$ is obviously hold when $x = y$. For sake of contradiction, suppose that $x \neq y$. It splits into two cases: $x < y$ or $x > y$. If $x < y$, then

$$|x - y| = y - x \leq \varepsilon \iff y \leq x$$

So that $y < x$, a contradiction. A similar argument shows that if we assume that $x > y$, then we obtain $x < y$. Thus $x = y$. \square

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 $a_n \leq x$ for all $n \geq 1$, then $\text{LIM}_{n \rightarrow \infty} a_n \leq x$. Similarly, show that if $a_n \geq x$ for all $n \geq 1$, then $\text{LIM}_{n \rightarrow \infty} a_n \geq x$. (Hint: prove by contradiction. Use Proposition 5.4.14 to find a rational between $\text{LIM}_{n \rightarrow \infty} a_n$ and x , and then use Proposition 5.4.9 or Corollary 5.4.10.)

Proof. Suppose for sake of contradiction that $\text{LIM}_{n \rightarrow \infty} a_n > x$ and $a_n \leq x$. By proposition 5.4.14, there exists a rational number q such that $x < q < \text{LIM}_{n \rightarrow \infty} a_n$. Since $a_n \leq x$ for all $n \geq 1$, by Corollary 5.4.10, $\text{LIM}_{n \rightarrow \infty} a_n \leq x$. So that

$$\text{LIM}_{n \rightarrow \infty} a_n \leq x < q < \text{LIM}_{n \rightarrow \infty} a_n$$

This is a contradiction. It's similarly to show the second claim. \square

5.5

The least upper bound property

Exercise 5.5.1. Let E be a subset of the real numbers \mathbf{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., $-M = \inf(E)$.

Proof. Since $M = \sup(E)$, choose some $x_0 \in E$, then we have $x_0 \leq M$ and any other upper bound M' for E must satisfy that $M \leq M'$. By definition, for every $x_0 \in E$ we have $-x_0 \in -E$, and there exists an $-M \leq -x_0$ for each $-x_0 \in -E$. Thus $-M$ is a lower bound for $-E$. We have to show that $-M$ is a greatest lower bound for $-E$. Suppose that there exists some L such that $-M < L \leq -x_0$, this implies that $M > L \geq x_0$, so that L is an upper bound for E , a contradiction. So $-M$ is a greatest lower bound for $-E$. \square

Exercise 5.5.2. Let E be a non-empty subset of \mathbf{R} , let $n > 1$ be an integer, and let $L < K$ be integers. Suppose that 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 $L < m \leq K$ 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. It may also help to draw a picture of the situation.)

Proof. Suppose for sake of contradiction that for all integers $L < m \leq K$ such that m/n and $(m-1)/n$ are or not an upper bound for E at the time, or m/n is not an upper bound for E but that $(m-1)/n$ is an upper bound for E . But we can exclude latter case for that $m < K \leq (m-1)$ is not hold.

Since m/n is not an upper bound for E , $(m-1)/n$ is not an upper

bound for E . We use strong principle of induction (Proposition 2.2.14) to show that for each $K > L$, we have following inequality: L/n is not an upper bound for E by hypothesis, and for each $L < m - 1 < K$ is hold for inequality, then K/n also is not an upper bound for E , but it contradicts hypothesis. \square

Exercise 5.5.3. Let E be a non-empty subset of \mathbf{R} , let $n \geq 1$ 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 (Hint: again, drawing a picture will be helpful).

Proof. We use contradiction. Suppose that $m \neq m'$, it splits into two cases: $m < m'$ or $m > m'$. Consider that if $m < m'$, we have $m \leq m' - 1$. When $(m' - 1)/n$ is not an upper bound for E , by definition, so that m/n is not, a contradiction. $m > m'$ case following same process. \square

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

Proof. For every $\varepsilon > 0$, we can find an $M > 1/\varepsilon$ by the Archimedean property such that

$$|q_n - q_{n'}| \leq \frac{1}{M} < \varepsilon, \quad \text{for all } n, n' \geq M.$$

Therefore $(q_n)_{n=1}^{\infty}$ is a Cauchy sequence. Furthermore, we have

$$|q_M - q_n| \leq \frac{1}{M} \implies q_M - \frac{1}{M} \leq q_n \leq q_M + \frac{1}{M},$$

for all $n \geq M \geq 1$. By Exercise 5.4.8, we have $|q_M - S| \leq 1/M$. \square

Exercise 5.5.5. Establish an analogue of Proposition 5.4.14, in which “rational” is replaced by “irrational”.

Proof.

Proposition. Given any two real numbers $x < y$, we can find a irrational number p such that $x < p < y$.

Since $x < y$, we have $y - x > 0$, and by Corollary 5.4.13, there is an integer n such that

$$n(y - x) > 1$$

Apply Archimedean property again, to obtain positive integers m_1 and m_2 such that $m_1 > nx, m_2 > -nx$. Then

$$-m_2 < nx < m_1.$$

Hence there is a real $m - \sqrt{2}$ (with $-m_2 \leq m + \sqrt{2} \leq m_1$) such that

$$(m + \sqrt{2}) - 1 \leq nx < m + \sqrt{2}.$$

If we combine these inequalities, we obtain

$$nx < m + \sqrt{2} \leq 1 + nx < ny.$$

Since $n > 0$, it follows that

$$x < \frac{m + \sqrt{2}}{n} < y.$$

We can see that $\frac{m+\sqrt{2}}{n}$ is not a rational: Suppose that $\frac{m+\sqrt{2}}{n} = \frac{a}{b}$, this implies $\sqrt{2} = \frac{an}{b} - m$. Left-hand side of equality is a rational, but right-hand side is not a rational. So that $\frac{m+\sqrt{2}}{n}$ is an irrational. This proves claim, with $p = \frac{m+\sqrt{2}}{n}$. \square

5.6

Real exponentiation, part I

Exercise 5.6.1. Prove Lemma 5.6.6. (Hints: review the proof of Proposition 5.5.12. Also, you will find proof by contradiction a useful tool, especially when combined with the trichotomy of order in Proposition 5.4.7 and Proposition 5.4.12. The earlier parts of the lemma can be used to prove later parts of the lemma. With part (e), first show that if $x > 1$ then $x^{1/n} > 1$, and if $x < 1$ then $x^{1/n} < 1$.)

Proof. Let $x, y \geq 0$ be non-negative reals, and let $n, m \geq 1$ be positive integers. Let set $E := \{k \in \mathbf{R} : k \geq 0 \text{ and } k^n \leq x\}$.

To prove assertion (a), we first prove a lemma:

Lemma. Let $0 < a < b$ and $n \leq m$, we have

$$b^n - a^n \leq m(b - a)b^{n-1}$$

for all $n \in \mathbf{Z}$.

Proof. We use induction on n . Consider the base case $n = 1$, we have $b - a \leq (b - a)m$ where $m \geq 1$. Now we inductively suppose that inequality is hold for n . We need to show that the case of $n + 1$. Since we have

$$\begin{aligned} m(b - a)b^n &\geq (n + 1)(b - a)b^{n-1} \\ &\geq b(n(b - a)b^{n-1} + (b - a)b^{n-1}) \\ &\geq b(b^n - a^n + b^n - ab^{n-1}) \\ &\geq 2b^{n+1} + ab(b^{n-1} - a^{n-1}) \\ &\geq b^{n+1} - a^{n+1}. \end{aligned}$$

This close the induction. □

(a) If $y = x^{1/n}$, then $y^n = x$.²

If $y = x^{1/n}$, by definition, we have $y = \sup(E)$ which is a least upper bound of E . Then for every $k \geq 0$, we have $y \geq k \geq 0$, by Proposition 5.6.3 and Proposition 4.3.10(c), $y^n \geq k^n \geq 0$. Now we show that $y^n = x$.

We argue this by contradiction. First suppose that $y^n < x$. By Corollary 5.4.13, there exists an integer m such that

$$0 < x - y^n < m(y + 1)^n \implies 0 < \frac{x - y^n}{m(y + 1)^{n-1}} < 1.$$

Let $0 < \varepsilon < \frac{x - y^n}{m(y + 1)^{n-1}}$ be some sufficiently small number. By Lemma, we have

$$(y + \varepsilon)^n - y^n \leq \varepsilon m(y + \varepsilon)^{n-1} < \varepsilon m(y + 1)^{n-1} < x - y^n.$$

²First prove that $(y - \varepsilon)^n \leq x$ and $(y + \varepsilon)^n > x$ for all sufficiently small $\varepsilon > 0$. — Tao.

There should be $(y + \varepsilon)^n < x$ and $(y - \varepsilon)^n \geq x$.

Hence $(y + \varepsilon)^n < x$, this means that $y + \varepsilon \in E$, but this contradicts the fact that y is a least upper bound of E .

While if we suppose that $y^n > x$, then $y^n - x > 0$. By the Archimedean property, there exists an integer m such that

$$0 < y^n - x < my^n \implies 0 < \frac{y^n - x}{my^{n-1}} < y.$$

Let $0 < \varepsilon < \frac{y^n - x}{my^{n-1}}$ be some sufficiently small number. By Lemma, we have

$$y^n - (y - \varepsilon)^n \leq \varepsilon my^{n-1} < y^n - x.$$

Hence $(y - \varepsilon)^n > x$, this means that $y - \varepsilon \notin E$. Since $y - \varepsilon < y$, this contradicts the fact that y is a least upper bound of E .

(b) *Conversely, if $y^n = x$, then $y = x^{1/n}$.*

Suppose that $y^n = x$, we can see that $y \in E$ and it is an upper bound for E . Since $k^n \leq x$, we have $k^n \leq y^n$, this implies that $k \leq y$ for every element $k \in \mathbf{R}$. Because $x^{1/n}$ is the least upper bound for E , and for every real number, if it is greater than y , then it is not the least upper bound. So suppose that $x^{1/n} < y$. This following a contradiction immediately. So we only have $y = x^{1/n}$

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

Since $0 \in E$, we must have $0 \leq x^{1/n}$. By Lemma 5.6.6(a) and (b), we have $x^{1/n} \neq 0$ iff $x \neq 0^n = 0$. Together our conclusion, $x^{1/n}$ is positive if and only if x is positive.

(d) *We have $x > y$ if and only if $x^{1/n} > y^{1/n}$.*

Let $a = x^{1/n}$ and $b = y^{1/n}$, then we have $a^n = x$ and $b^n = y$. Suppose that $x > y \geq 0$, which means $a^n > b^n \geq 0$. Then, by Proposition 5.6.3 and Proposition 4.3.10(c), we have $a > b > 0$, i.e., $x^{1/n} > y^{1/n} > 0$. While if $x^{1/n} > y^{1/n}$, this implies, via similar process we have $x > y$.

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

then $x^{1/k}$ is an increasing function of k . If $x = 1$, then $x^k = 1$ for all k .³

It's easy to check that for every $k \geq 1$, $x^{1/k} > 1$ when $x > 1$, and $x^{1/k} < 1$ when $x < 1$. Let $a = x^{1/k}$ and $b = x^{1/(k+1)}$, then $x = a^k = b^{k+1}$.

We suppose that $x > 1$, we can see that a and b are greater than 1, we have

$$x = b^{k+1} = b^k b > b^k \implies a^k > b^k.$$

By Proposition 5.6.3, we have $a > b$, so that $x^{1/k} > x^{1/(k+1)}$, function is decreasing.

If we suppose that $x < 1$, we can see that a and b are less than 1, we have

$$x = b^{k+1} = b^k b < b^k \implies a^k < b^k.$$

By Proposition 5.6.3, we have $a < b$, so that $x^{1/k} < x^{1/(k+1)}$, function is increasing.

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

By definition, $(xy)^{1/n} = \sup\{k \in \mathbf{R} : k \geq 0 \text{ and } k^n \leq xy\}$. We have to show that $x^{1/n}y^{1/n}$ is the least upper bound for such a set. Let $a = x^{1/n}, b = y^{1/n}$, then $xy = a^n b^n = (ab)^n$. We have $k^n \leq (ab)^n$, so that $k \leq ab$, which means that $k \leq x^{1/n}y^{1/n}$. Then $x^{1/n}y^{1/n}$ is an upper bound.

Now we want to show that $x^{1/n}y^{1/n}$ is the least upper bound. Suppose for sake of contradiction that there exists an $h \in E$ and $x^{1/n}y^{1/m} < h$. Let $q = h^n$, then $x^{1/n}y^{1/m} < h \leq q$. Since $h \in E$, we have $q \leq xy$. Since $x^{1/n}y^{1/m} < q$, we have $(ab)^n < q^n$, this implies that $xy < q$, a contradiction.

³Page 122: Before Lemma 5.6.6: “ n^{th} root” should be “ n^{th} roots”. In (e), add “Here k ranges over the positive integers”, and after “decreasing”, add “(i.e., $x^{1/k} < x^{1/l}$ whenever $l > k$)”. One can also replace $x < 1$ by $0 < x < 1$ for clarity. — Errata from Tao.

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

Similarly, by definition, $(x^{1/n})^{1/m} = \sup\{k \in \mathbf{R} : k \geq 0 \text{ and } k^m \leq x^{1/n}\}$. We have to show that $x^{1/nm}$ is the least upper bound for such a set. Let $a = (x^{1/n})^{1/m}$, then $x^{1/n} = a^m \implies x = a^{nm}$, so that $a = x^{1/nm}$. Since $k^m \leq a^m$, we have $k \leq a$, so that $k \leq x^{1/nm}$, then $x^{1/nm}$ is an upper bound.

Now we want to show that $x^{1/nm}$ is the least upper bound. Suppose for sake of contradiction that there exists an $h \in E$ and $x^{1/nm} < h$. Let $q = h^m$, then $x^{1/nm} < h \leq q$. Since $q \in E$, we have $q^m \leq x^{1/n}$. Since $x^{1/nm} < q$, we have $x^{1/n} = a^m < q^m$, a contradiction. \square

Exercise 5.6.2. Prove Lemma 5.6.9. (Hint: you should rely mainly on Lemma 5.6.6 and on algebra.)

Proof. Let $x, y > 0$ be positive reals, and let q, r be rationals. Let $q = a/b, r = c/d$.

(a) x^q is a positive real.

By definition, $x^q = (x^{1/b})^a$. By Lemma 5.6.6(c), $x^{1/b}$ is a positive real number, so that x^q is a positive real number.

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

By Lemma 5.6.6, we have

$$\begin{aligned} x^{q+r} &= x^{a/b+c/d} \\ &= x^{(ad+bc)/bd} \\ &= (x^{1/bd})^{ad+bc} \\ &= (x^{1/bd})^{ad} (x^{1/bd})^{bc} \\ &= (x^{1/b})^a (x^{1/d})^c = x^q x^r, \end{aligned}$$

and

$$(x^q)^r = (x^{a/b})^{c/d} = (x^{a/bd})^c = x^{ac/bd} = x^{(a/b)(c/d)} = x^{qr}.$$

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

By Lemma 5.6.6, we have

$$x^{-q} = (x^{1/b})^{-a} = \frac{1}{(x^{1/b})^a} = 1/x^q.$$

(d) If $q > 0$, then $x > y$ if and only if $x^q > y^q$.

Since $q > 0$, $a > 0$. By Lemma 5.6.6 and Proposition 5.6.3, we have

$$\begin{aligned} x > y &\iff x^{1/b} > y^{1/b} \\ &\iff (x^{1/b})^a > (y^{1/b})^a \\ &\iff x^q > y^q. \end{aligned}$$

(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 $q < r$.

Since $x > 1$, and $x^q > x^r$, then $x^{q/r} > x > 1$. By Lemma 5.6.6(e)

$$\frac{q}{r} = \frac{ad}{bc} > 1 \implies ad > bc \implies \frac{a}{b} > \frac{c}{d}.$$

So that $q > r$, vice versa.

Since $x < 1$, and $x^q > x^r$. By Lemma 5.6.6(e)

$$\frac{q}{r} = \frac{ad}{bc} < 1 \implies ad < bc \implies \frac{a}{b} < \frac{c}{d}.$$

So that $q < r$, vice versa. □

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

Proof. We divide into three cases: $x = 0$, $x > 0$ and $x < 0$. If $x = 0$, then $|x| = 0$ and $(x^2)^{1/2} = (0^2)^{1/2} = 0^{1/2} = 0$. If $x > 0$, then $|x| = x$ and $(x^2)^{1/2} = x$. If $x < 0$, then $|x| = -x$ and $(x^2)^{1/2} = ((-x)^2)^{1/2} = -x$. □

CHAPTER VI

Limits of sequences

6.1

Convergence and limit laws

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.)*

Proof. Let $m = n + k$ for some natural number $k \geq 1$, then m satisfy that $m > n$. We keeping n fixed and induct on k . $k = 0$ is vacuously true and meaningless, and $k = 1$ is hold for condition. Now we inductively suppose that $a_{n+k} > a_n$ is hold for all k . We have to show that $a_{n+k+1} > a_n$ is true. Since $a_{n+k+1} > a_{n+k}$ by condition, so that $a_{n+k+1} > a_{n+k} > a_n$. This close the induction and $a_m > a_n$ is hold for all $m > n$. \square

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 if and only if, given any real $\varepsilon > 0$, one can find an $N \geq m$ such that $|a_n - L| \leq \varepsilon$ for all $n \geq N$.*

Proof. Suppose that $(a_n)_{n=m}^{\infty}$ converges to L . By Definition 6.1.5, there exists an $N \geq m$ such that $(a_n)_{n=N}^{\infty}$ is ε -close to L . Then we have $|a_n - L| \leq \varepsilon$ for every $n \geq N$.

Suppose that given any real $\varepsilon > 0$, one can find an $N \geq m$ such that $|a_n - L| \leq \varepsilon$ for all $n \geq N$. Then we can find an $N \geq m$ such that $(a_n)_{n=N}^{\infty}$ is ε -close to L . Hence, $(a_n)_{n=m}^{\infty}$ converges to L . \square

Exercise 6.1.3. Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let c be a real number, and let $m' > m$ 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 .

Proof. Let $\varepsilon > 0$ be a real number. Suppose that $(a_n)_{n=m}^{\infty}$ converges to c . Then there exists an $N' \geq m$ such that $|a_n - c| \leq \varepsilon$ for all $n \geq N'$. Let $N = \min(N', m')$, we can find an $N \geq m$ such that $|a_n - c| \leq \varepsilon$ for all $n \geq N$. So that $(a_n)_{n=m'}^{\infty}$ converges to c .

Suppose that $(a_n)_{n=m'}^{\infty}$ converges to c . Then there exists an $N \geq m'$, so that $N \geq m$ such that $|a_n - c| \leq \varepsilon$ for all $n \geq N$. Therefore, $(a_n)_{n=m}^{\infty}$ converges to c . \square

Exercise 6.1.4. Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let c be a real number, and let $k \geq 0$ 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 .

Proof. Let $\varepsilon > 0$ be a real number. Suppose that $(a_n)_{n=m}^{\infty}$ converges to c . Then there exists an $N \geq m$ such that $|a_n - c| \leq \varepsilon$ for all $n \geq N$. Since $k \geq 0$, we have $|a_{n+k} - c| \leq \varepsilon$ for all $n + k \geq N$. Hence $(a_{n+k})_{n=m}^{\infty}$ converges to c .

Suppose $(a_{n+k})_{n=m}^{\infty}$ converges to c . By Exercise 6.1.3, $(a_{n+k})_{n=m+k}^{\infty}$ also converges to c . Then there exists an $N \geq m + k$ such that $|a_{n+k} - c| \leq \varepsilon$ for all $n + k \geq N$. We can find an $N' = N - k > m$ such that $|a_n - c| \leq \varepsilon$ for all $n \geq N'$. So $(a_n)_{n=m}^{\infty}$ converges to c . \square

Exercise 6.1.5. Prove Proposition 6.1.12. (Hint: use the triangle inequality, or Proposition 4.3.7.)

Proposition 6.1.12 (Convergent sequences are Cauchy). Suppose that $(a_n)_{n=m}^{\infty}$ is a convergent sequence of real numbers. Then $(a_n)_{n=m}^{\infty}$ is also a Cauchy sequence.

Proof. Suppose that $(a_n)_{n=m}^{\infty}$ is a convergent sequence of real numbers. Then for any $\varepsilon > 0$, there exist some $N', N'' \geq m$ such that $|a_i - c| \leq \varepsilon/2$ and $|a_j - c| \leq \varepsilon/2$ for all $i \geq N'$ and $j \geq N''$. Let $N = \max(N', N'')$, then for $N \geq m$ such that

$$|a_i - a_j| = |(a_i - c) - (a_j - c)| \leq |a_i - c| + |a_j - c| \leq \varepsilon$$

for all $i, j \geq N$. So that $(a_n)_{n=m}^{\infty}$ is also a Cauchy sequence. \square

Exercise 6.1.6. Prove Proposition 6.1.15, using the following outline. Let $(a_n)_{n=m}^{\infty}$ be a Cauchy sequence of rationals, and write $L := \text{LIM}_{n \rightarrow \infty} a_n$. We have to show that $(a_n)_{n=m}^{\infty}$ converges to L . Let $\varepsilon > 0$. Assume for sake of contradiction that sequence a_n is not eventually ε -close to L . Use this, and the fact that $(a_n)_{n=m}^{\infty}$ is Cauchy, to show that there is an $N \geq m$ such that either $a_n > L + \varepsilon/2$ for all $n \geq N$, or $a_n \leq L - \varepsilon/2$ for all $n \geq N$. Then use Exercise 5.4.8.

Proposition 6.1.15 (Formal limits are genuine limits). Suppose that $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence of rational numbers. Then $(a_n)_{n=1}^{\infty}$ converges to $\text{LIM}_{n \rightarrow \infty} a_n$, i.e.

$$\text{LIM}_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} a_n.$$

Proof. Let $(a_n)_{n=m}^{\infty}$ be a Cauchy sequence of rationals, and write $L := \text{LIM}_{n \rightarrow \infty} a_n$. We have to show that $(a_n)_{n=m}^{\infty}$ converges to L . Let $\varepsilon > 0$. Assume for sake of contradiction that sequence a_n is not eventually ε -close to L . Then for all $N \geq m$ there exists an $n \geq N$ such that $|a_n - L| > \varepsilon$. This means that $a_n > L + \varepsilon$ or $a_n < L - \varepsilon$.

Other hand, $(a_n)_{n=m}^{\infty}$ is Cauchy, there is an $N' \geq 0$ such that $|a_i - a_k| \leq \varepsilon/2$ for all $i, k \geq N'$. We can find an $N > N'$ such that

$$\begin{aligned} |a_n - L| &= |(a_n - L) + (a_k - a_n)| \\ &\geq |a_n - L| - |a_k - a_n| \geq \varepsilon/2, \end{aligned}$$

for all $n \geq N$. Then we have $a_n > L + \varepsilon/2$ for all $n \geq N$, or $a_n \leq L - \varepsilon/2$ for all $n \geq N$. By Exercise 5.4.8, we have $L > L + \varepsilon/2$ for all $n \geq N$ or $L \leq L - \varepsilon/2$ for all $n \geq N$, a contradiction. \square

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.)*

Proof. Suppose $(a_n)_{n=m}^{\infty}$ is bounded in the sense of Definition 6.1.15, we have $|a_n| \leq M$ for all $n \geq m$. Let $m = 1$, we are in the sense of Definition 5.1.12.

Now suppose that $(a_n)_{n=m}^{\infty}$ is bounded in the sense of Definition 5.1.12. Then $|a_i| \leq M$ for all $i \geq 1$. Since $n \geq i \geq 1$, we have $|a_n| \leq M$. So that we are in the sense of Definition 6.1.15. \square

Exercise 6.1.8. *Prove Theorem 6.1.19. (Hint: you can use some parts of the theorem to prove others, e.g., (b) can be used to prove (c); (a), (c) can be used to prove (d); and (b), (e) can be used to prove (f). The proofs are similar to those of Lemma 5.3.6, Proposition 5.3.10, and Lemma 5.3.15. For (e), you may need to first prove the auxiliary result that any sequence whose elements are non-zero, and which converges to a non-zero limit, is bounded away from zero.)*

Proof. Let $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ be convergent sequences of real numbers, and let x, y be the real numbers $x := \lim_{n \rightarrow \infty} a_n$ and $y := \lim_{n \rightarrow \infty} b_n$.

(a) *The sequence $(a_n + b_n)_{n=m}^{\infty}$ converges to $x + y$; in other words,*

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

Prove (a). Since $x := \lim_{n \rightarrow \infty} a_n$ and $y := \lim_{n \rightarrow \infty} b_n$. For every $\varepsilon > 0$, there exist $N', N'' \geq m$ such that $|a_n - x| \leq \varepsilon/2$ and $|b_n - y| \leq \varepsilon/2$ for all $n \geq N'$ and $n \geq N''$. Then for $n \geq N = \max(N', N'')$

$$|(a_n + b_n) - (x + y)| \leq |a_n - x| + |b_n - y| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which means $\lim_{n \rightarrow \infty} (a_n + b_n) = x + y$.

(b) The sequence $(a_n b_n)_{n=m}^{\infty}$ converges to xy ; in other words,

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

Prove (b). Let

$$\delta' = \min \left(1, \frac{\varepsilon}{3(|y| + 1)} \right), \quad \text{and} \quad \delta'' = \min \left(1, \frac{\varepsilon}{3(|x| + 1)} \right).$$

For every $\varepsilon > 0$, there exist $N', N'' \geq m$ such that $|a_n - x| \leq \delta'$ for all $n \geq N'$ and $|b_n - y| \leq \delta''$ for all $n \geq N''$. We have inequality

$$\begin{aligned} |a_n| &\leq |x| + \delta' \leq |x| + 1, \\ |b_n| &\leq |y| + \delta'' \leq |y| + 1, \end{aligned}$$

and

$$\begin{aligned} \delta'' |a_n| &\leq \min \left(1, \frac{\varepsilon}{3(|x| + 1)} \right) \cdot (|x| + 1) \leq \frac{\varepsilon}{3}, \\ \delta' |b_n| &\leq \min \left(1, \frac{\varepsilon}{3(|y| + 1)} \right) \cdot (|y| + 1) \leq \frac{\varepsilon}{3}, \\ \delta' \delta'' &< \min \left(1, \frac{\varepsilon}{3(|y| + 1)} \right) \cdot \min \left(1, \frac{\varepsilon}{3(|x| + 1)} \right) \leq \frac{\varepsilon}{3}. \end{aligned}$$

Then for $n \geq N = \max(N', N'')$ we have

$$|a_n b_n - xy| \leq \delta'' |a_n| + \delta' |b_n| + \delta' \delta'' \leq \varepsilon.$$

So that $\lim_{n \rightarrow \infty} (a_n b_n) = xy$.

(c) For any real number c , the sequence $(ca_n)_{n=m}^{\infty}$ converges to cx ; in other words,

$$\lim_{n \rightarrow \infty} (ca_n) = c \lim_{n \rightarrow \infty} a_n.$$

Prove (c). Let $(b_n)_{n=m}^{\infty} = (c)_{n=m}^{\infty}$, then $\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} c = c$. By Theorem 6.1.19(b), we have

$$\lim_{n \rightarrow \infty} (ca_n) = \left(\lim_{n \rightarrow \infty} c \right) \left(\lim_{n \rightarrow \infty} a_n \right) = c \lim_{n \rightarrow \infty} a_n.$$

This means that $\lim_{n \rightarrow \infty} (ca_n) = cx$.

(d) The sequence $(a_n - b_n)_{n=m}^{\infty}$ converges to $x - y$; in other words,

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

Prove (d). This follows immediately from (a), (c).

- (e) Suppose that $y \neq 0$, and that $b_n \neq 0$ for all $n \geq m$. Then the sequence $(b_n^{-1})_{n=m}^{\infty}$ converges to y^{-1} ; in other words,

$$\lim_{n \rightarrow \infty} (b_n^{-1}) = \left(\lim_{n \rightarrow \infty} b_n \right)^{-1}.$$

Prove (e). For every $\varepsilon > 0$, there exists $N \geq m$ such that $|b_n - y| \leq y^2 \varepsilon$ for all $n \geq N$. Then for $n \geq N$, we have

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

This means $\lim_{n \rightarrow \infty} (b_n^{-1}) = y^{-1}$.

- (f) Suppose that $y \neq 0$, and that $b_n \neq 0$ for all $n \geq m$. Then the sequence $(a_n/b_n)_{n=m}^{\infty}$ converges to x/y ; in other words,

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

Prove (f). This follows immediately from (b), (e).

- (g) The sequence $(\max(a_n, b_n))_{n=m}^{\infty}$ converges to $\max(x, y)$; in other words,

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

Prove (g). For every $\varepsilon', \varepsilon'' > 0$, there exist $N', N'' \geq m$ such that $|a_n - x| \leq \varepsilon'$ and $|b_n - y| \leq \varepsilon''$ for all $n \geq N'$ and $n \geq N''$. Then for $n \geq N = \max(N', N'')$, we can verify that

$$\begin{aligned} |a_n - y| &= |(a_n - x) + (x - y)| < \varepsilon' + x, \\ |b_n - x| &= |(b_n - y) + (y - x)| < \varepsilon'' + y. \end{aligned}$$

Hence, let $\varepsilon \geq \max(\varepsilon' + x, \varepsilon'' + y)$, we have

$$|\max(a_n, b_n) - \max(x, y)| < \max(\varepsilon' + x, \varepsilon'' + y) \leq \varepsilon,$$

which means $\lim_{n \rightarrow \infty} \max(a_n, b_n) = \max(x, y)$.

- (h) The sequence $(\min(a_n, b_n))_{n=m}^{\infty}$ converges to $\min(x, y)$; in other words,

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

Prove (h). This is similar to prove as assertion (g). □

Exercise 6.1.9. *Explain why Theorem 6.1.19(f) fails when the limit of the denominator is 0. (To repair that problem requires **L'Hôpital's rule**, see Section 10.5.)*

Proof. Suppose that $b_n = 1/n$ and $a_n = 1$, then we have

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

There is no real number such that $(n)_{n=m}^{\infty}$ to be convergent. This is easy to verify and we leave it out. \square

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 $\varepsilon > 0$. (Hint: modify the proof of Proposition 6.1.4.)*

Proof. Suppose first that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are equivalent where ε is a positive real, they are eventually ε -close for every $\varepsilon > 0$. In particular, $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are eventually ε -close for every rational $\varepsilon > 0$, which is in the sense of Definition 5.2.6.

Now suppose that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are equivalent in the sense of Definition 5.2.6, then they are eventually ε -close for every rational $\varepsilon > 0$. If ε is a real number, then there exists a rational $\varepsilon' > 0$ which is smaller than ε , by Proposition 5.4.12. Since ε' is rational, we know that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are eventually ε' -steady; since $\varepsilon' < \varepsilon$, this implies that they are eventually ε -steady. Since ε is an arbitrary positive real number, we thus see that $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ are equivalent for every ε is a positive real. \square

6.2

The extended real number system

Exercise 6.2.1. *Prove Proposition 6.2.5. (Hint: you may need Proposition 5.4.7.)*

Proof. Let x, y, z be extended real numbers.

(a) (*Reflexivity*) We have $x \leq x$.

Prove (a). If $x \in \mathbf{R}$, then $x \leq x$. By Definition 6.2.3, if $x = +\infty$, then $+\infty \leq +\infty$. While if $x = -\infty$, then $-\infty \leq -\infty$.

(b) (*Trichotomy*) Exactly one of the statements $x < y$, $x = y$, or $x > y$ is true.

Prove (b). Since $x, y \in \mathbf{R}$, trichotomy is hold. Hence there at least one of $x < y$, $x = y$ or $x > y$ is hold. By Definition 6.2.3, $x < y$ and $x = y$ can't hold at the time, and $x = y$ and $x > y$ also can't hold at the time. Suppose that $x < y$ and $x > y$ hold at the time, and let $x = +\infty, y = -\infty$, we have $+\infty < -\infty$, a contradiction.

(c) (*Transitivity*) if $x \leq y$ and $y \leq z$, then $x \leq z$.

Prove (c). Since $x \leq y$ and $y \leq z$. If $x, y \in \mathbf{R}$, claim is hold. If $x = -\infty$, then $x < z$ for all $z \in \mathbf{R}^*$. If $z = +\infty$, then $x < z$ for all $x \in \mathbf{R}^*$. If $y = +\infty$, then we must have $z = +\infty$; while if $y = -\infty$, then we must have $x = -\infty$, so that $x < z$.

(d) (*Negation reverses order*) If $x \leq y$, then $-y \leq -x$.

Prove (d). Since $x \leq y$. Claim is hold for all $x, y \in \mathbf{R}$. Suppose that $x = -\infty$, then $-x = +\infty$, we have $-x \geq -y$ for all $-y \in \mathbf{R}^*$. If $y = +\infty$, then $-y = -\infty$, we have $-y \leq -x$ for all $-x \in \mathbf{R}^*$. \square

Exercise 6.2.2. *Prove Theorem 6.2.11. (Hint: you may need to break into cases depending on whether $+\infty$ or $-\infty$ belongs to E . You can of course*

use Definition 5.5.10, provided that E consists only of real numbers.)

Theorem 6.2.11. Let E be a subset of \mathbf{R}^* . Then the following statements are true.

- (a) For every $x \in E$ we have $x \leq \sup(E)$ and $x \geq \inf(E)$.
- (b) Suppose that $M \in \mathbf{R}^*$ is an upper bound for E , i.e., $x \leq M$ for all $x \in E$. Then we have $\sup(E) \leq M$.
- (c) Suppose that $M \in \mathbf{R}^*$ is a lower bound for E , i.e., $x \leq M$ for all $x \in E$. Then we have $\inf(E) \geq M$.

Proof.

- (a) If E is contained in \mathbf{R} . By Definition 6.2.6, $\sup(E)$ be as defined in Definition 5.5.10. Hence for every $x \in E$ we have $x \leq \sup(E)$ and $x \geq \inf(E)$.

If E contains $+\infty$. By Definition 6.2.6, $\sup(E) := +\infty$. By Definition 6.2.3, for all $x \in \mathbf{R}^*$, $x \leq \sup(E) = +\infty$, and by Definition 5.5.10, $x \geq \inf(E)$ for all $x \in E$.

If E does not contain $+\infty$ but does contain $-\infty$. By Definition 6.2.6, $\sup(E) := \sup(E \setminus \{-\infty\})$. Since $-\infty \leq x$ for all $x \in \mathbf{R}^*$, by Definition 6.2.6, we have $x \geq \inf(E) = -\infty$, and $x \leq \sup(E) = \sup(E \setminus \{-\infty\})$.

- (b) If E is contained in \mathbf{R} . $\sup(E) \leq M$ is hold by Definition 5.5.5 and Definition 5.5.10.

If E contains $+\infty$. Suppose for contradiction that $\sup(E) > M$ for all $M \in \mathbf{R}^*$. Let $M = +\infty$, by Definition 6.2.3 and Definition 6.2.3, we have $+\infty > +\infty$, a contradiction.

If E does not contain $+\infty$ but does contain $-\infty$. Since M is an upper bound of E , by Definition 5.5.10, $\sup(E) \leq M$ is hold for all $x \in E$.

- (c) If E is contained in \mathbf{R} . $\inf(E) \geq M$ is hold Definition 5.5.10.

If E contains $+\infty$. Since M is an lower bound of E , by Definition 5.5.10, $\inf(E) \geq M$ is hold for all $x \in E$.

If E does not contain $+\infty$ but does contain $-\infty$. Suppose for contradiction that $\inf(E) < M$ for all $M \in \mathbf{R}^*$. Let $M = -\infty$, by Definition 6.2.3 and Definition 6.2.3, we have $-\infty < -\infty$, a contradiction. \square

6.3

Suprema and Infima of sequences

Exercise 6.3.1. *Verify the claim in Example 6.3.4.*

Example 6.3.4. Let $a_n := 1/n$; thus $(a_n)_{n=1}^\infty$ is the sequence $1, 1/2, 1/3, \dots$. Then the set $\{a_n : n \geq 1\}$ is the countable set $\{1, 1/2, 1/3, 1/4, \dots\}$. Thus $\sup(a_n)_{n=1}^\infty = 1$ and $\inf(a_n)_{n=1}^\infty = 0$.

Proof. Let $a_n = 1/n$. We need to verify $\sup(a_n)_{n=1}^\infty = 1$ and $\inf(a_n)_{n=1}^\infty = 0$. Since $n \geq 1$, we know that for every integer n we have $1/n \leq 1$. So 1 is an upper bound of $(a_n)_{n=1}^\infty$. We have to show that 1 is least. Because for every real number $\delta < 1$ and which is an upper bound would implies that $\delta < a_1$. So that 1 is the least upper bound of $(a_n)_{n=1}^\infty$.

Now we verify that $\inf(a_n)_{n=1}^\infty = 0$. For every integer $n \geq 1$, we have $0 < 1/n$. So that 0 is a lower bound of $(a_n)_{n=1}^\infty$. We need to show that 0 is greatest. Suppose there exists a lower bound $\delta > 0$ such that $a_n \geq \delta$. We can find an $N = 1/\delta$ such that $1/n < \delta$ for all $n > N$. So that 0 is the greatest lower bound of $(a_n)_{n=1}^\infty$. \square

Exercise 6.3.2. *Prove Proposition 6.3.6. (Hint: use Theorem 6.2.11.)*

Proposition 6.3.6 (Least upper bound property). Let $(a_n)_{n=m}^\infty$ be a sequence of real numbers, and let x be the extended real number $x := \sup(a_n)_{n=m}^\infty$. Then we have $a_n \leq x$ for all $n \geq m$. Also, whenever $M \in \mathbf{R}^*$ is an upper bound for an (i.e., $a_n \leq M$ for all $n \geq m$), we have $x \leq M$. Finally, for every extended real number y for which $y < x$, there exists at least one $n \geq m$ for which $y < a_n \leq x$.

Proof. Since $(a_n)_{n=m}^\infty$ is a sequence of real numbers, $E := \{a_n : n \geq m\}$

is a subset of \mathbf{R}^* . Since $n \geq m$, by Theorem 6.2.11, for every $a_n \in E$ we have $a_n \leq x$. Also, since $M \in \mathbf{R}^*$ is an upper bound for E , so that $x \leq M$. Finally, suppose for sake of contradiction that $y \leq a_n$ for all $n \geq m$. Since $y \in \mathbf{R}^*$ and $y < x$, we have $y = -\infty$ or $y \in \mathbf{R}$. If $y = -\infty$, we must have $a_n > y$, a contradiction. While if $y \in \mathbf{R}$, then y is an upper bound of E . But $x > y$ is the least upper bound of E , a contradiction. Therefore, there exists at least one $n \geq m$ for $y < a_n \leq x$. \square

Exercise 6.3.3. Prove Proposition 6.3.8. (Hint: use Proposition 6.3.6, together with the assumption that a_n is increasing, to show that a_n converges to $\sup(a_n)_{n=m}^\infty$.)

Proposition 6.3.8 (Monotone bounded sequences converge). Let $(a_n)_{n=m}^\infty$ be a sequence of real numbers which has some finite upper bound $M \in \mathbf{R}$, and which is also increasing (i.e., $a_{n+1} \geq a_n$ for all $n \geq m$). Then $(a_n)_{n=m}^\infty$ is convergent, and in fact

$$\lim_{n \rightarrow \infty} a_n = \sup(a_n)_{n=m}^\infty \leq M.$$

Proof. Let $x = \sup(a_n)_{n=m}^\infty$, and suppose that a_n is increasing. We need to show that $\lim_{n \rightarrow \infty} a_n = x$, i.e., let $\varepsilon > 0$, there exists an $N \geq m$ such that $|a_n - x| \leq \varepsilon$ for all $n \geq N$. By definition of least upper bound, for every $\varepsilon > 0$ there exists an $a_m \in E$ such that $x - \varepsilon < a_N \leq x$. Since a_n is increasing, we have $x - \varepsilon < a_N \leq x_n \leq x$ for all $n > N$. That is, $|a_n - x| \leq x - a_n \leq \varepsilon$. Hence we proved that $\lim_{n \rightarrow \infty} a_n = x$. By Proposition 6.3.6, we have $\lim_{n \rightarrow \infty} a_n = \sup(a_n)_{n=m}^\infty \leq M$. \square

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 $x > 1$. (Hint: prove by contradiction and use the identity $(1/x)^n x^n = 1$ and the limit laws in Theorem 6.1.19.) Compare this with the argument in Example 1.2.3; can you now explain the flaws in the reasoning in that example?

Proof. Since $x > 1$, we can see that $0 < 1/x < 1$. Suppose for sake of contradiction that $(x^n)_{n=1}^\infty$ converges to some limit L . Following Proposition

6.3.10, $\lim_{n \rightarrow \infty} (1/x)^n = 0$. Since identity, $(1/x)^n x^n = 1$ and Theorem 6.1.19,

$$\lim_{n \rightarrow \infty} \left(\frac{1}{x}\right)^n x^n = \lim_{n \rightarrow \infty} \left(\frac{1}{x}\right)^n \cdot \lim_{n \rightarrow \infty} x^n = 0 \cdot \lim_{n \rightarrow \infty} x^n = 1.$$

This is a contradiction. □

6.4

Limsup Liminf, and limit points

Exercise 6.4.1. *Prove Proposition 6.4.5.*

Proposition 6.4.5 (Limits are limit points). Let $(a_n)_{n=m}^{\infty}$ be a sequence which converges to a real number c . Then c is a limit point of $(a_n)_{n=m}^{\infty}$, and in fact it is the only limit point of $(a_n)_{n=m}^{\infty}$.

Proof. Since $(a_n)_{n=m}^{\infty}$ converges to a real number c , for every $\varepsilon > 0$, there exists an $N \geq m$ such that $|a_n - c| \leq \varepsilon$ for all $n \geq N$. Then for every $N_n \geq m$ we can find an n_n to correspond which satisfied the inequality. We choose $N = \max(N_n)$, then for every $N \geq m$, there exists an $n \geq N$ such that $|a_n - c| \leq \varepsilon$. Hence, c is a limit point of $(a_n)_{n=m}^{\infty}$. Now we want to show that limit point c is unique. Suppose for contradiction that there exists another limit point $c' \neq c$. By definition, for every $M \geq m$, we have $|a_n - c'| \leq \varepsilon$ for all $n \geq M$. Let $\varepsilon = |c - c'|/3$. For $n \geq \max(N, M)$, we have

$$|c - c'| \leq |a_n - c| + |a_n - c'| \leq 2\varepsilon,$$

so we have $|c - c'| \leq 2|c - c'|/3$ for $c \neq c'$, a contradiction. □

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.*

Proposition I. Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let c be a real number, and let $m' \geq m$ be an integer. Then c is a limit point of $(a_n)_{n=m}^{\infty}$ if and only if c is a limit point of $(a_n)_{n=m'}^{\infty}$.

Proof. If c is a limit point of $(a_n)_{n=m}^\infty$, for every $\varepsilon > 0$, it is ε -adherent to $(a_n)_{n=m}^\infty$ for every $N \geq m$. Because $m' \geq m$, x is ε -adherent to $(a_n)_{n=m}^\infty$ for every $N \geq m'$. Thus c is a limit point of $(a_n)_{n=m'}^\infty$. While if c is a limit point of $(a_n)_{n=m'}^\infty$, for every $\varepsilon > 0$, it is ε -adherent to $(a_n)_{n=m}^\infty$ for every $N \geq m'$. Since $m' \geq m$, for every $m' \geq M \geq m$ such that a_n is ε' -close to c . So that c is continually δ -adherent to $(a_n)_{n=m}^\infty$ for every $\delta = \max(\varepsilon, \varepsilon') > 0$. \square

Proposition II. We define $(a_N^+)_{N=m}^\infty$ and $(a_N^+)_{N=m'}^\infty$ by the formula

$$a_N^+ := \sup(a_n)_{n=N}^\infty.$$

Then $(a_n)_{n=m}^\infty$ is a limit superior if and only if $(a_n)_{n=m'}^\infty$ is a limit superior, i.e.,

$$\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m}^\infty = \inf(a_N^+)_{N=m'}^\infty.$$

Proof. We first want to prove that $(a_N^+)_{N=m}^\infty$ is decreasing. Suppose for sake of contradiction that $a_{N+1} > a_N$ for all $N \geq m$. By definition, we have $\sup(a_n)_{n=N+1}^\infty > \sup(a_n)_{n=N}^\infty$. Since $a_{N+1} \in (a_n)_{n=N}^\infty$ and a_N is an upper bound of the set, we have $a_{N+1} = \sup(a_n)_{n=N+1}^\infty = \sup(a_n)_{n=N}^\infty$, a contradiction.

Suppose $(a_n)_{n=m}^\infty$ is a limit superior, then $\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m}^\infty$ for all $N \geq m$. Since $m' \geq m$ and $(a_N^+)_{N=m}^\infty$ is decreasing. This is hold for $N \geq m'$, i.e., $\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m'}^\infty$, $(a_n)_{n=m'}^\infty$ is a limit superior.

Now suppose that $(a_n)_{n=m'}^\infty$ is a limit superior, then $\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m'}^\infty$ for all $N \geq m'$. Since $\inf(a_N^+)_{N=m'}^\infty$ is the greatest lower bound of sequence which is decreasing, then we also have $\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m}^\infty$. \square

Proposition III. We define $(a_N^-)_{N=m}^\infty$ and $(a_N^-)_{N=m'}^\infty$ by the formula

$$a_N^- := \inf(a_n)_{n=N}^\infty.$$

Then $(a_n)_{n=m}^\infty$ is a limit inferior if and only if $(a_n)_{n=m'}^\infty$ is a limit inferior, i.e.,

$$\liminf_{n \rightarrow \infty} a_n = \sup(a_N^-)_{N=m}^\infty = \sup(a_N^-)_{N=m'}^\infty.$$

Proof. This is similarly to prove as last assertion. \square

Proposition IV. Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let c be a real number, and let $k \geq 0$ be a non-negative integer. Then c is a limit point of $(a_n)_{n=m}^{\infty}$ if and only if c is a limit point of $(a_{n+k})_{n=m}^{\infty}$.

Proof. If c is a limit point of $(a_n)_{n=m}^{\infty}$, for every $\varepsilon > 0$ and every $N \geq m$, there exists an $n \geq N$, so that $n + k \geq N$, such that $|a_{n+k} - c| \leq \varepsilon$. Thus c is a limit point of $(a_{n+k})_{n=m}^{\infty}$. While if c is a limit point of $(a_{n+k})_{n=m}^{\infty}$, for every $\varepsilon > 0$ and every $N \geq m$, there exists an $n + k \geq N$ such that $|a_{n+k} - c| \leq \varepsilon$. Then for every $N + k \geq N \geq m$, there exists an $n \geq N$ such that $|a_n - c| \leq \varepsilon$. c is a limit point of $(a_n)_{n=m}^{\infty}$. \square

Proposition V: We define $(a_N^+)_{N=m}^{\infty}$ and $(a_{N+k}^+)_{N=m}^{\infty}$ by the formula

$$a_N^+ := \sup(a_n)_{n=N}^{\infty} \quad \text{and} \quad a_{N+k}^+ := \sup(a_{n+k})_{n=N}^{\infty}.$$

Then $(a_n)_{n=m}^{\infty}$ is a limit superior if and only if $(a_{n+k})_{n=m}^{\infty}$ is a limit superior, i.e.,

$$\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m}^{\infty} = \inf(a_{N+k}^+)_{N=m}^{\infty}.$$

Proof. Suppose that $(a_n)_{n=m}^{\infty}$ is a limit superior, then $\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m}^{\infty}$ for all $N \geq m$. Since $n + k \geq n$, we can see that $a_N^+ \geq a_{N+k}^+$ for all $N \geq m$. So that $\limsup_{n \rightarrow \infty} a_n = \inf(a_{N+k}^+)_{N=m}^{\infty}$. Now suppose that $(a_{n+k})_{n=m}^{\infty}$ is a limit superior, then $\limsup_{n \rightarrow \infty} a_n = \inf(a_{N+k}^+)_{N=m}^{\infty}$ for all $N + k \geq m$. This is hold also for all $N \geq m$. So $\limsup_{n \rightarrow \infty} a_n = \inf(a_N^+)_{N=m}^{\infty}$. \square

Proposition VI: We define $(a_N^-)_{N=m}^{\infty}$ and $(a_{N+k}^-)_{N=m}^{\infty}$ by the formula

$$a_N^- := \inf(a_n)_{n=N}^{\infty} \quad \text{and} \quad a_{N+k}^- := \inf(a_{n+k})_{n=N}^{\infty}.$$

Then $(a_n)_{n=m}^{\infty}$ is a limit inferior if and only if $(a_{n+k})_{n=m}^{\infty}$ is a limit inferior, i.e.,

$$\liminf_{n \rightarrow \infty} a_n = \sup(a_N^-)_{N=m}^{\infty} = \sup(a_{N+k}^-)_{N=m}^{\infty}.$$

Proof. This is similarly to prove as last assertion. \square

Exercise 6.4.3. Prove parts (c), (d), (e), (f) of Proposition 6.4.12. (Hint:

you can use earlier parts of the proposition to prove later ones.)

Proof.

(c) We have $\inf(a_n)_{n=m}^{\infty} \leq L^- \leq L^+ \leq \sup(a_n)_{n=m}^{\infty}$.

Prove (c). Suppose for sake of contradiction that $L^+ > \sup(a_n)_{n=m}^{\infty}$, by definition, we have $\inf(a_N^+)_{N=m}^{\infty} > \sup(a_n)_{n=m}^{\infty}$. By Proposition 6.3.6, we have $a_N^+ > \sup(a_n)_{n=m}^{\infty}$ for all $N \geq m$. By definition of a_N^+ , this means that $\sup(a_n)_{n=N}^{\infty} > \sup(a_n)_{n=m}^{\infty}$. Since $N \geq m$, we have $(a_n)_{n=N}^{\infty} \subseteq (a_n)_{n=m}^{\infty}$, by Theorem 6.2.11 (a), $\sup(a_n)_{n=N}^{\infty} \leq \sup(a_n)_{n=m}^{\infty}$, a contradiction. Inequality $\inf(a_n)_{n=m}^{\infty} \leq L^-$ is proven similarly. Now we need to prove that $L^- \leq L^+$. Suppose for contradiction that $L^- > L^+$, then we have $\inf(a_n)_{n=m}^{\infty} > \sup(a_n)_{n=m}^{\infty}$. But for every element of $a_n \in (a_n)_{n=m}^{\infty}$ we must have $\inf(a_n)_{n=m}^{\infty} \leq a_n \leq \sup(a_n)_{n=m}^{\infty}$ for all $n \geq m$, a contraction. Therefore, $\inf(a_n)_{n=m}^{\infty} \leq L^- \leq L^+ \leq \sup(a_n)_{n=m}^{\infty}$.

(d) If c is any limit point of $(a_n)_{n=m}^{\infty}$, then we have $L^- \leq c \leq L^+$.

Prove (d). Since c is a limit point of $(a_n)_{n=m}^{\infty}$, by definition, for every $\varepsilon > 0$ and every $N \geq m$, there exists an $n \geq N$ such that $|a_n - c| \leq \varepsilon$. Suppose for sake of contradiction that $c > L^+$, i.e., $c > \sup(a_n)_{n=m}^{\infty}$ for all $n \geq N$. Since $a_n \leq L^+ < c$, we have $0 < |c - L^+| \leq |a_n - c| \leq \varepsilon$. Let $\varepsilon = |c - L^+|/4$, there exists an $n \geq N$ such that

$$\begin{aligned} |c - L^+| &\leq |a_n - c| + |a_n - L^+| = |a_n - c| + |(a_n - c) + (c - L^+)| \\ &\leq |a_n - c| + |a_n - c| + |L^+ - c| \leq 3\varepsilon = \frac{3}{4}|c - L^+|. \end{aligned}$$

a contradiction. $L^- \leq c$ is proven similarly.

(e) If L^+ is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$. Similarly, if L^- is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$.¹

Prove (e). By Proposition 6.4.12 (b), for every $\varepsilon > 0$, we have $L^+ - \varepsilon < L^+$, and every $N \geq m$, there exists an $n \geq N$ such that $a_n > L^+ - \varepsilon$, so that $|a_n - L^+| < \varepsilon$. This proves that L^+ is a limit point, case of L^- is proven similarly.

¹By Definition 6.2.1, extended real number x is finite iff it is a real number ($x \neq \pm\infty$).

- (f) Let c be a real number. If $(a_n)_{n=m}^{\infty}$ converges to c , then we must have $L^+ = L^- = c$. Conversely, if $L^+ = L^- = c$, then $(a_n)_{n=m}^{\infty}$ converges to c .

Prove (f). Since $(a_n)_{n=m}^{\infty}$ converges to c , by Proposition 6.4.5, c is a only limit point. By Proposition 6.4.12. (e), $L^+ = L^- = c$. If $L^+ = L^- = c$, we suppose for sake of contradiction that $(a_n)_{n=m}^{\infty}$ don't converges to c . This means that there exist $\varepsilon > 0$ and $n \geq N$ such that $|a_n - c| > \varepsilon$ for all $n \geq m$. Since $L^+ = L^-$, there exists an $N \geq m$ such that $\sup(a_n)_{n=N}^{\infty} = \inf(a_n)_{n=N}^{\infty}$. This implies that for all $n \geq N$, we have $\inf(a_n)_{n=N}^{\infty} = a_n = \sup(a_n)_{n=N}^{\infty}$. Therefore, there exists an $N \geq m$ such that $|a_n - c| \leq \varepsilon$ for all $\varepsilon > 0$. \square

Exercise 6.4.4. Prove Lemma 6.4.13.

Lemma 6.4.13 (Comparison principle). Suppose that $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ are two sequences of real numbers such that $a_n \leq b_n$ for all $n \geq m$. Then we have the inequalities

$$\begin{aligned}\sup(a_n)_{n=m}^{\infty} &\leq \sup(b_n)_{n=m}^{\infty} \\ \inf(a_n)_{n=m}^{\infty} &\leq \inf(b_n)_{n=m}^{\infty} \\ \limsup_{n \rightarrow \infty} a_n &\leq \limsup_{n \rightarrow \infty} b_n \\ \liminf_{n \rightarrow \infty} a_n &\leq \liminf_{n \rightarrow \infty} b_n\end{aligned}$$

Proof. Prove $\sup(a_n)_{n=m}^{\infty} \leq \sup(b_n)_{n=m}^{\infty}$. Suppose for sake of contradiction that $\sup(a_n)_{n=m}^{\infty} > \sup(b_n)_{n=m}^{\infty}$. By definition, we have $a_n \leq \sup(a_n)_{n=m}^{\infty}$ and $b_n \leq \sup(b_n)_{n=m}^{\infty}$ for some $n \geq m$. Consider that if $\sup(a_n)_{n=m}^{\infty} \leq b_n$, then inequality holds. Again, consider that if $\sup(a_n)_{n=m}^{\infty} > b_n$, since $a_n \leq b_n$ for all $n \geq m$, then we see that b_n is the least upper bound of $(a_n)_{n=m}^{\infty}$. So we only have $\sup(a_n)_{n=m}^{\infty} \leq b_n$ and in this case, inequality is hold.

Prove $\inf(a_n)_{n=m}^{\infty} \leq \inf(b_n)_{n=m}^{\infty}$. This is proven similarly.

Prove $\limsup_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} b_n$. Since $a_n \leq b_n$, we have $a_N^+ \leq b_N^+$. This implies that $\limsup_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} b_n$.

Prove $\liminf_{n \rightarrow \infty} a_n \leq \liminf_{n \rightarrow \infty} b_n$. This is proven similarly. \square

Exercise 6.4.5. Use Lemma 6.4.13 to prove Corollary 6.4.14.

Corollary 6.4.14 (Squeeze test). Let $(a_n)_{n=m}^{\infty}$, $(b_n)_{n=m}^{\infty}$, and $(c_n)_{n=m}^{\infty}$ be sequences of real numbers such that

$$a_n \leq b_n \leq c_n$$

for all $n \geq m$. Suppose also that $(a_n)_{n=m}^{\infty}$ and $(c_n)_{n=m}^{\infty}$ both converge to the same limit L . Then $(b_n)_{n=m}^{\infty}$ is also convergent to L .

Proof. Suppose $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = L$. Given $\varepsilon > 0$ choose N' and N'' such that $L - \varepsilon < a_n$ and $c_n < L + \varepsilon$ for all $n \geq N''$. Then for $n > N = \max(N', N'')$ we shall have $L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon$, which means $|b_n - L| \leq \varepsilon$, that is $\lim_{n \rightarrow \infty} b_n = L$. \square

Exercise 6.4.6. Given 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 $n \geq 1$, but that $\sup(a_n)_{n=1}^{\infty} \not\leq \sup(b_n)_{n=1}^{\infty}$. Explain why this does not contradict Lemma 6.4.13.

Proof. Let $a_n = 0$ and $b_n = 1/n$. Then we have $a_n < b_n$ for all $n \geq 1$. But $\lim_{n \rightarrow \infty} 1/n = 0$. \square

Exercise 6.4.7. Prove Corollary 6.4.17. Is the corollary still true if we replace zero in the statement of this Corollary by some other number?

Corollary 6.4.17 (Zero test for sequences). Let $(a_n)_{n=M}^{\infty}$ be a sequence of real numbers. Then the limit $\lim_{n \rightarrow \infty} a_n$ exists and is equal to zero if and only if the limit $\lim_{n \rightarrow \infty} |a_n|$ exists and is equal to zero.

Proof. Since $\lim_{n \rightarrow \infty} a_n = 0$, for every ε there exists an $N \geq M$ such that $||a_n| - 0| \leq \varepsilon$ for all $n \geq N$. So that $\lim_{n \rightarrow \infty} |a_n| = 0$. Now suppose that $\lim_{n \rightarrow \infty} |a_n| = 0$, then $\lim_{n \rightarrow \infty} -|a_n| = 0$. Because we have $-|a_n| \leq a_n \leq |a_n|$, by squeeze test, $\lim_{n \rightarrow \infty} a_n = 0$. \square

Exercise 6.4.8. Let us say that a sequence $(a_n)_{n=M}^{\infty}$ of real numbers has $+\infty$ as a limit point iff it has no finite upper bound, and that it has $-\infty$ as a limit point iff it has no finite lower bound. With this definition, show that

$\limsup_{n \rightarrow \infty} a_n$ is a limit point of $(a_n)_{n=M}^{\infty}$, and furthermore that it is larger than all the other limit points of $(a_n)_{n=M}^{\infty}$; in other words, the limit superior is the largest limit point of a sequence. Similarly, show that the limit inferior is the smallest limit point of a sequence. (One can use Proposition 6.4.12 in the course of the proof.)

Proof. Let $L^+ = \sup(a_n)_{n=M}^{\infty}$. We know that L^+ is finite or infinite. Since it is finite, by Proposition (e), L^+ is a limit point of $(a_n)_{n=M}^{\infty}$. And by (d), it's the largest limit point of the sequence. Since it's infinite, this means that $a_n \leq +\infty$ for all $n \geq M$, and there haven't other upper bound less than $+\infty$. By definition, $+\infty$ is a limit point of $(a_n)_{n=M}^{\infty}$. Suppose another limit point of the sequence as $x \in \mathbf{R}$, then we must have $x < +\infty$. Otherwise, $x = +\infty$. This proves the limit superior case, limit inferior case is proven similarly. \square

Exercise 6.4.9. Using the definition in Exercise 6.4.8, construct a sequence $(a_n)_{n=N}^{\infty}$ which has exactly three limit points, at $-\infty$, 0, and $+\infty$.

Proof. Let $(a_n)_{n=1}^{\infty} = n \sin(n\pi/2)$ denote the sequence

$$1, 0, -3, -0, 5, 0, -7, -0, 9, \dots$$

Then it has exactly three limit points, at $-\infty$, 0, and $+\infty$. \square

Exercise 6.4.10. Let $(a_n)_{n=N}^{\infty}$ be a sequence of real numbers, and let $(b_m)_{m=M}^{\infty}$ be another sequence of real numbers such that each b_m is a limit point of $(a_n)_{n=N}^{\infty}$. Let c be a limit point of $(b_m)_{m=M}^{\infty}$. Prove that c is also a limit point of $(a_n)_{n=N}^{\infty}$. (In other words, limit points of limit points are themselves limit points of the original sequence.)

Proof. Since c is a limit point of $(b_m)_{m=M}^{\infty}$, by definition, for every $\varepsilon > 0$ and every $K' \geq m$, there exists an $m \geq K'$ such that $|b_m - c| \leq \varepsilon/2$. And since b_m is a limit point of $(a_n)_{n=N}^{\infty}$, for every $\varepsilon > 0$ and every $K'' \geq n$, there exists an $n \geq K''$ such that $|a_n - b_m| \leq \varepsilon/2$. Then for every $\varepsilon > 0$ and every $K \geq \max(m, n)$, there exists an $n \geq \max(K', K'')$ such that

$$|a_n - c| \leq |a_n - b_m| + |b_m - c| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Thus, c is a limit point of $(a_n)_{n=N}^{\infty}$. \square

6.5

Some standard limits

Exercise 6.5.1. Show that $\lim_{n \rightarrow \infty} 1/n^q = 0$ for any rational $q > 0$. (Hint: use Corollary 6.5.1 and the limit laws, Theorem 6.1.19.) Conclude that the limit $\lim_{n \rightarrow \infty} n^q$ does not exist. (Hint: argue by contradiction using Theorem 6.1.19(e).)

Proof. Let $q = a/b$ for some non-negative integer b and integer a . Then we have

$$\lim_{n \rightarrow \infty} \frac{1}{n^q} = \lim_{n \rightarrow \infty} \left(\frac{1}{n^{1/b}} \right)^a.$$

By Corollary 6.5.1, $\lim_{n \rightarrow \infty} 1/n^{1/b} = 0$, so that $\lim_{n \rightarrow \infty} (1/n^{1/b})^a = 0$.

Now suppose that $\lim_{n \rightarrow \infty} n^q$ exist. Let $L = \lim_{n \rightarrow \infty} n^q$. Since $a > 0$, we have $n^q \neq 0$. Then $\lim_{n \rightarrow \infty} 1/n^q = 1/L = 0$, a contradiction. \square

Exercise 6.5.2. Prove Lemma 6.5.2. (Hint: use Proposition 6.3.10, Exercise 6.3.4, and the squeeze test.)

Lemma 6.5.2. Let x be a real number. Then the limit $\lim_{n \rightarrow \infty} x^n$ 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 $x > 1$.

Proof. Since $|x| < 1$, by Proposition 6.3.10, $\lim_{n \rightarrow \infty} |x| = \lim_{n \rightarrow \infty} -|x| = 0$. Because $-|x|^n \leq x^n \leq |x|^n$, by squeeze test, we have $\lim_{n \rightarrow \infty} x^n = 0$. When $x = 1$, then $x^n \equiv 1$, so that $\lim_{n \rightarrow \infty} x^n = 1$. When $x = -1$, we can see that $L^+ = 1$ and $L^- = -1$, in which $L^+ \neq L^-$. So the sequence is divergent. When $|x| > 1$, assume that L is the limit of the sequence, and let $\varepsilon = 1/2 + L$. Then there exists an N such that

$$|x^n - L| \leq |x|^n + L \leq \varepsilon = 1/2 + L$$

for all $n \geq N$. Hence, $|x| \leq 1/2^{1/n}$. When $n \geq 0$, we have $|x| \leq 1/2^{1/n} < 1$, a contradiction. \square

Exercise 6.5.3. Prove Lemma 6.5.3. (Hint: you may need to treat the cases $x \geq 1$ 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 $M^{1/n} \leq 1 + \varepsilon$.)

Lemma 6.5.3. For any $\varepsilon > 0$, we have $\lim_{n \rightarrow \infty} x^{1/n} = 1$.

Proof. Assume first that $x \geq 1$. For any $\varepsilon > 0$ there exists $N \in \mathbf{N}$ such that $1 \leq x < (1 + \varepsilon)^n$ for all $n \geq N$, and we then have $1 \leq x^{1/n} < 1 + \varepsilon$ for all $n \geq N$, which says $\lim_{n \rightarrow \infty} x^{1/n} = 1$.

For $0 < x < 1$, we have $1 < 1/x$, and then

$$\lim_{n \rightarrow \infty} x^{1/n} = \lim_{n \rightarrow \infty} \frac{1}{1/x^{1/n}} = \frac{1}{\lim_{n \rightarrow \infty} (1/x)^{1/n}} = 1.$$

This is desired. □

6.6

Subsequences

Exercise 6.6.1. Prove Lemma 6.6.4.

Lemma 6.6.4. Let $(a_n)_{n=0}^\infty$, $(b_n)_{n=0}^\infty$ and $(c_n)_{n=0}^\infty$ be sequences of real numbers. Then $(a_n)_{n=0}^\infty$ is a subsequence of $(a_n)_{n=0}^\infty$. Furthermore, if $(b_n)_{n=0}^\infty$ is a subsequence of $(a_n)_{n=0}^\infty$, and $(c_n)_{n=0}^\infty$ is a subsequence of $(b_n)_{n=0}^\infty$, then $(c_n)_{n=0}^\infty$ is a subsequence of $(a_n)_{n=0}^\infty$.

Proof. Prove the first part of statement. Let the function $f : \mathbf{N} \rightarrow \mathbf{N}$ defined by $f(n) := n$, then $a_n = a_{f(n)}$ for all $n \in \mathbf{N}$.

For the second part. Since $(b_n)_{n=0}^\infty$ is a subsequence of $(a_n)_{n=0}^\infty$, and $(c_n)_{n=0}^\infty$ is a subsequence of $(b_n)_{n=0}^\infty$, we have $b_n = a_{f(n)}$ and $c_n = b_{g(n)}$ for all $n \in \mathbf{N}$. Let $g \circ f : \mathbf{N} \rightarrow \mathbf{N}$ which is a strictly increasing function. We have $c_n = a_{(g \circ f)(n)}$ for all $n \in \mathbf{N}$. □

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?

Proof. Let $a_n := \sin(n\pi)$ and $b_n := \cos(n\pi)$. Then we have $b_n = a_{n+1/2}$ and $a_n = b_{n+3/2}$ for all $n \in \mathbf{N}$. \square

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 $\lim_{n \rightarrow \infty} 1/b_n$ exists and is equal to zero. (Hint: for each natural number j , recursively introduce the quantity $n_j := \min\{n \in \mathbf{N} : |a_n| \geq j; n > n_{j-1}\}$ (omitting the condition $n > n_{j-1}$ when $j = 0$), first explaining why the set $\{n \in \mathbf{N} : |a_n| \geq j; n > n_{j-1}\}$ is non-empty. Then set $b_j := a_{n_j}$.)²

Proof. Let $n_0 = 0$. Then $\{n \in \mathbf{N} : |a_n| \geq 1; n > n_0\}$ is non-empty because $(a_n)_{n=0}^\infty$ is not bounded, there exists an $n \geq 1$ such that $|a_n| > 1$. Now we inductively suppose that

$$\{n \in \mathbf{N} : |a_n| \geq j; n > n_{j-1}\}$$

is non-empty. We need to show that the set

$$\{n \in \mathbf{N} : |a_n| \geq j + 1; n > n_j\}$$

is non-empty. For if it were empty, we have $|a_n| < j + 1$ for all $n > n_j$. By induction hypothesis, this means that $|a_n|$ is bounded by

$$\max(|a_0|, |a_{n_1}|, \dots, |a_{n_j}|, j),$$

a contradiction. By Proposition 8.1.4 (well ordering principle, footnote#2 explained why can we use it), n_j exists and is unique. Let $b_j = a_{n_j}$. Then we have $|b_j| \geq j$ for all $j \in \mathbf{N}$. Thus $|1/b_j| \leq 1/j$, and $\lim_{j \rightarrow \infty} 1/b_j = 0$. \square

Exercise 6.6.4. Prove Proposition 6.6.5. (Note that one of the two implications has a very short proof.)

Proposition 6.6.5 (Subsequences related to limits). Let $(a_n)_{n=0}^\infty$ be a sequence of real numbers, and let L be a real number. Then the following two statements are logically equivalent (each one implies the other):

²“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 5.5.9).” Similarly for Exercise 6.6.5.

- (a) The sequence $(a_n)_{n=0}^{\infty}$ converges to L .
- (b) Every subsequence of $(a_n)_{n=0}^{\infty}$ converges to L .

Proof. (a) \implies (b). Suppose there is a function $f : \mathbf{N} \rightarrow \mathbf{N}$ such that $(a_{f(n)})_{n=0}^{\infty}$ be a subsequence of $(a_n)_{n=0}^{\infty}$. Since $(a_n)_{n=0}^{\infty}$ converges to L . Let $\varepsilon > 0$, there exists an $N \geq 0$ such that $|a_n - L| \leq \varepsilon$ for all $n \geq N$. Because f is strictly increasing, we can find an m such that $f(m) \geq N$. Then there is $f(m) \geq N$ such that $|a_{f(n)} - L| \leq \varepsilon$ for all $n \geq m$. Hence $(a_{f(n)})_{n=0}^{\infty}$ converges to L .

(b) \implies (a). Let $f(n) = n$, then we have $a_{f(n)} = a_n$, and $(a_n)_{n=0}^{\infty}$ converges to L . \square

Exercise 6.6.5. *Prove Proposition 6.6.6. (Hint: to show that (a) implies (b), define the numbers n_j for each natural numbers j by the formula $n_j := \min\{n > n_{j-1} : |a_n - L| \leq 1/j\}$, with the convention $n_0 := 0$, explaining why the set $\{n > n_{j-1} : |a_n - L| \leq 1/j\}$ is non-empty. Then consider the sequence a_{n_j} .)*

Proposition 6.6.6 (Subsequences related to limit points). Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let L be a real number. Then the following two statements are logically equivalent.

- (a) L is a limit point of $(a_n)_{n=0}^{\infty}$.
- (b) There exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges to L .

Proof. (a) \implies (b). Let $n_0 = 0$. Define n_j recursively, for every $j \geq 1$,

$$n_j := \min \{n > n_{j-1} : |a_n - L| \leq 1/j\}.$$

Since L is a limit point of $(a_n)_{n=0}^{\infty}$, by Definition 6.4.1, for every $\varepsilon = 1/j > 0$ and every $N \geq n_0 = 0$, there exists an $n \geq N$ such that $|a_n - L| \leq 1/j$. Thus $\{n > n_{j-1} : |a_n - L| \leq 1/j\}$ is non-empty. By Proposition 8.1.4, n_j exists and is unique. Let a_{n_j} be a subsequence of a_n . By definition, there exists $N \geq 0$ such that $|a_{n_j} - L| \leq 1/j$ for all $n_j \geq N$. Hence $(a_{n_j})_{j=1}^{\infty}$ converges to L .

For the converse, let $(b_n)_{n=0}^\infty$ be a subsequence of $(a_n)_{n=0}^\infty$ which converges to L . By Proposition 6.6.6, $(a_n)_{n=0}^\infty$ converges to L . By Proposition 6.4.5, L is a limit point of $(a_n)_{n=0}^\infty$. \square

6.7**Real exponentiation, part II**

Exercise 6.7.1. *Prove the remaining components of Proposition 6.7.3.*

Proposition 6.7.3. All the results of Lemma 5.6.9, which held for rational numbers q and r , continue to hold for real numbers q and r .

Proof. Let q and r be real numbers. Then we can write $q = \lim_{n \rightarrow \infty} q_n$ and $r = \lim_{n \rightarrow \infty} r_n$ for some sequences $(q_n)_{n=1}^\infty$ and $(r_n)_{n=1}^\infty$ of rationals by the definition of real numbers (and Proposition 6.1.15).

(a) x^q is a positive real.

By Lemma 5.6.9, x^{q_n} is positive for every $n \geq 1$. Then by Definition 5.4.3, $x^q = \lim_{n \rightarrow \infty} x^{q_n}$ is positive.

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

By the limit laws, $q + r$ is the limit of $(q_n + r_n)_{n=1}^\infty$. By definition of real exponentiation, we have

$$x^{q+r} = \lim_{n \rightarrow \infty} x^{q_n + r_n}; \quad x^q = \lim_{n \rightarrow \infty} x^{q_n}; \quad x^r = \lim_{n \rightarrow \infty} x^{r_n}.$$

But by Lemma 5.6.9(b) (applied to rational exponents) we have $x^{q_n + r_n} = x^{q_n} x^{r_n}$. Thus by limit laws we have $x^{q+r} = x^q x^r$, as desired.

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

By the limit laws, $-q$ is the limit of $(-q_n)_{n=1}^\infty$. By definition of real exponentiation, we have

$$x^{-q} = \lim_{n \rightarrow \infty} x^{-q_n}.$$

But by Lemma 5.6.9(c) we have $x^{-q_n} = 1/x^{q_n}$. Thus by limit laws we have $x^{-q} = 1/x^q$, as desired.

(d) If $q > 0$, then $x > y$ if and only if $x^q > y^q$.

Since $x > y$, by Lemma 5.6.9(d), we have $x^{q_n} > y^{q_n}$ for all $n \geq 1$. By Corollary 5.4.10,

$$\lim_{n \rightarrow \infty} x^{q_n} > \lim_{n \rightarrow \infty} y^{q_n}.$$

Hence by Definition 6.7.2, $x^q > y^q$.

(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 $q < r$.

If $x > 1$. We can see that

$$\begin{aligned} x^q > x^r &\iff x^q - x^r > 0 \\ &\iff x^r(x^{q-r} - 1) > 0 \\ &\iff x^{q-r} > x^0 \\ &\iff q - r > 0 \\ &\iff q > r. \end{aligned}$$

While if $x < 1$, we have

$$\begin{aligned} x^q > x^r &\iff x^q - x^r < 0 \\ &\iff x^r(x^{q-r} - 1) < 0 \\ &\iff x^{q-r} < x^0 \\ &\iff q - r < 0 \\ &\iff q < r. \end{aligned}$$

□

CHAPTER VII

Series

7.1

Finite 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.)*

Proof.

- (a) *Let $m \leq n < p$ be integers, and let a_i be a real number assigned to each integer $m \leq i \leq p$. Then we have*

$$\sum_{i=m}^n a_i + \sum_{i=n+1}^p a_i = \sum_{i=m}^p a_i.$$

Prove (a). We induct on p . When $p = m + 1$, then $n = m$, and we have

$$\sum_{i=m}^m a_i + \sum_{i=m+1}^{m+1} a_i = a_m + a_{m+1} = \sum_{i=m}^{m+1} a_i.$$

Now suppose inductively that equality is hold in the case of p . We

have

$$\sum_{i=m}^n a_i + \sum_{i=n+1}^{p+1} a_i = \sum_{i=m}^n a_i + \sum_{i=n+1}^p a_i + a_{p+1} = \sum_{i=m}^p a_i + a_{p+1} = \sum_{i=m}^{p+1} a_i.$$

This close the induction.

- (b) Let $m \leq n$ be integers, k be another integer, and let a_i be a real number assigned to each integer $m \leq i \leq n$. Then we have

$$\sum_{i=m}^n a_i = \sum_{j=m+k}^{n+k} a_{j-k}.$$

Prove (b). We use induction on n . When $n = m$, we have

$$\sum_{i=m}^m a_i = a_m = \sum_{j=m+k}^{m+k} a_{j-k}.$$

Then inductively suppose that n is hold for equality, then we have

$$\sum_{i=m}^{n+1} a_i = \sum_{i=m}^n a_i + a_{n+1} = \sum_{j=m+k}^{n+k} a_{j-k} + a_{n+1} = \sum_{j=m+k}^{n+k+1} a_{j-k}.$$

This close the induction.

- (c) Let $m \leq n$ be integers, and let a_i, b_i be real numbers assigned to each integer $m \leq i \leq n$. Then we have

$$\sum_{i=m}^n (a_i + b_i) = \left(\sum_{i=m}^n a_i \right) + \left(\sum_{i=m}^n b_i \right).$$

Prove (c). Use induction on n . When $n = m$, we obviously have

$$\sum_{i=m}^m (a_i + b_i) = a_m + b_m = \left(\sum_{i=m}^m a_i \right) + \left(\sum_{i=m}^m b_i \right).$$

Suppose inductively that n is hold for the equality. We have

$$\begin{aligned} \sum_{i=m}^{n+1} (a_i + b_i) &= \sum_{i=m}^n (a_i + b_i) + a_{n+1} + b_{n+1} \\ &= \left(\sum_{i=m}^n a_i \right) + \left(\sum_{i=m}^n b_i \right) + a_{n+1} + b_{n+1} \\ &= \left(\sum_{i=m}^{n+1} a_i \right) + \left(\sum_{i=m}^{n+1} b_i \right). \end{aligned}$$

This close the induction.

- (d) Let $m \leq n$ be integers, and let a_i be a real number assigned to each integer $m \leq i \leq n$, and let c be another real number. Then we have

$$\sum_{i=m}^n (ca_i) = c \left(\sum_{i=m}^n a_i \right).$$

Prove (d). We induct on n . When $n = m$, we have

$$\sum_{i=m}^m (ca_i) = ca_m = c \left(\sum_{i=m}^m a_i \right).$$

Now inductively suppose that n is hold for the equality. We have

$$\sum_{i=m}^{n+1} (ca_i) = \sum_{i=m}^n (ca_i) + ca_{n+1} = c \left(\sum_{i=m}^n a_i + a_{n+1} \right) = c \left(\sum_{i=m}^{n+1} a_i \right).$$

This close the induction.

- (e) (Triangle inequality for finite series) Let $m \leq n$ be integers, and let a_i be a real number assigned to each integer $m \leq i \leq n$. Then we have

$$\left| \sum_{i=m}^n a_i \right| \leq \sum_{i=m}^n |a_i|.$$

Prove (e). We induct on n . When $n = m$, we have $|a_m| \leq |a_m|$ is true.

Now inductively suppose that n is hold for inequality. We have

$$\left| \sum_{i=m}^{n+1} a_i \right| = \left| \sum_{i=m}^n a_i + a_{n+1} \right| \leq \sum_{i=m}^n |a_i| + |a_{n+1}| = \sum_{i=m}^{n+1} |a_i|.$$

This close the induction.

- (f) (Comparison test for finite series) Let $m \leq n$ be integers, and let a_i, b_i be real numbers assigned to each integer $m \leq i \leq n$. Suppose that $a_i \leq b_i$ for all $m \leq i \leq n$. Then we have

$$\sum_{i=m}^n a_i \leq \sum_{i=m}^n b_i.$$

Prove (f). We induct on n . This is obviously when we suppose that $n = m$. Then we inductively suppose that $n - 1$ is hold for inequality.

We have

$$\sum_{i=m}^n a_i = \sum_{i=m}^{n-1} a_i + a_n \leq \sum_{i=m}^{n-1} b_i + b_n = \sum_{i=m}^n b_i.$$

This close the induction. \square

Exercise 7.1.2. *Prove Proposition 7.1.11. (Hint: this is not as lengthy as it may first appear. It is largely a matter of choosing the right bijections to turn these sums over sets into finite series, and then applying Lemma 7.1.4.)*

Proof.

- (a) *If X is empty, and $f : X \rightarrow \mathbf{R}$ is a function (i.e., f is the empty function), we have*

$$\sum_{x \in X} f(x) = 0.$$

Prove (a). Since X is an empty. By Exercise 3.3.3, there is a bijection $g : \{i \in \mathbf{N} : 1 \leq i \leq 0\} \rightarrow X$. Then we have

$$\sum_{x \in X} f(x) = \sum_{i=1}^0 f(g(i)) = 0.$$

- (b) *If X consists of a single element, $X = \{x_0\}$, and $f : X \rightarrow \mathbf{R}$ is a function, we have*

$$\sum_{x \in X} f(x) = f(x_0).$$

Prove (b). Let $g : \{i \in \mathbf{N} : 1 \leq i \leq 1\} \rightarrow X$ be a bijection defined by $x_0 = g(1)$. Then we have

$$\sum_{x \in X} f(x) = \sum_{i=1}^1 f(g(i)) = f(x_0).$$

- (c) *(Substitution, part I) If X is a finite set, $f : X \rightarrow \mathbf{R}$ is a function, and $g : Y \rightarrow X$ is a bijection, then*

$$\sum_{x \in X} f(x) = \sum_{y \in Y} f(g(y)).$$

Prove (c). Suppose that X are finite sets with n elements, then Y also has n elements. Let $h : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow Y$ be a bijection. Then $g \circ h : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X$ also is a bijection. We have

$$\begin{aligned} \sum_{y \in Y} f(g(y)) &= \sum_{y \in Y} (f \circ g)(y) \\ &= \sum_{i=1}^n (f \circ g)(h(i)) \\ &= \sum_{i=1}^n f(g \circ h(i)) = \sum_{x \in X} f(x). \end{aligned}$$

(d) (*Substitution, part II*) Let $n \leq m$ be integers, and let X be the set $X := \{i \in \mathbf{Z} : n \leq i \leq m\}$. If a_i is a real number assigned to each integer $i \in X$, then we have

$$\sum_{i=n}^m a_i = \sum_{i \in X} a_i.$$

Prove (d). Let $a : X \rightarrow \mathbf{R}$ and define $g : \{i \in \mathbf{Z} : n \leq i \leq m\} \rightarrow X$ by $g(i) = i$, which is a bijection. Then we have

$$\sum_{i \in X} a_i = \sum_{i=n}^m a_{g(i)} = \sum_{i=n}^m a_i.$$

(e) Let X, Y be disjoint finite sets (so $X \cap Y = \emptyset$), and $f : X \cup Y \rightarrow \mathbf{R}$ is a function. Then we have

$$\sum_{z \in X \cup Y} f(z) = \left(\sum_{x \in X} f(x) \right) + \left(\sum_{y \in Y} f(y) \right).$$

Prove (e). Suppose that X, Y are finite sets with n and m elements respectively. Let $h : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X$ and $\tilde{h} : \{i \in \mathbf{N} : n+1 \leq i \leq n+m\} \rightarrow Y$ be bijections. Then we have

$$\left(\sum_{x \in X} f(x) \right) + \left(\sum_{y \in Y} f(y) \right) = \left(\sum_{i=1}^n f(h(i)) \right) + \left(\sum_{i=n+1}^{n+m} f(\tilde{h}(i)) \right).$$

Let define bijection $g : \{i \in \mathbf{N} : 1 \leq i \leq n+m\} \rightarrow X \cup Y$ by following

rule:

$$g(i) = \begin{cases} h(i), & 1 \leq i \leq n; \\ \tilde{h}(i), & n+1 \leq i \leq n+m. \end{cases}$$

Hence

$$\begin{aligned} \sum_{z \in X \cup Y} f(z) &= \sum_{i=1}^{n+m} f(g(i)) \\ &= \left(\sum_{i=1}^n f(g(i)) \right) + \left(\sum_{i=n+1}^{n+m} f(g(i)) \right) \\ &= \left(\sum_{i=1}^n f(h(i)) \right) + \left(\sum_{i=n+1}^{n+m} f(\tilde{h}(i)) \right). \end{aligned}$$

Thus we have

$$\sum_{z \in X \cup Y} f(z) = \left(\sum_{x \in X} f(x) \right) + \left(\sum_{y \in Y} f(y) \right).$$

(f) (*Linearity, part I*) Let X be a finite set, and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions. Then

$$\sum_{x \in X} (f(x) + g(x)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

Prove (f). Suppose that X is a finite set with n elements. Let $h : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X$ be a bijection. Then we have

$$\begin{aligned} \sum_{x \in X} (f(x) + g(x)) &= \sum_{i=1}^n (f(h(i)) + g(h(i))) \\ &= \sum_{i=1}^n f(h(i)) + \sum_{i=1}^n g(h(i)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x). \end{aligned}$$

(g) (*Linearity, part II*) Let X be a finite set, let $f : X \rightarrow \mathbf{R}$ be a function, and let c be a real number. Then

$$\sum_{x \in X} cf(x) = c \sum_{x \in X} f(x).$$

Prove (g). Suppose that X is a finite set with n elements. Let $g : \{i \in$

$\mathbf{N} : 1 \leq i \leq n\} \rightarrow X$ be a bijection. Then we have

$$\sum_{x \in X} cf(x) = \sum_{i=1}^n cf(g(i)) = c \sum_{i=1}^n f(g(i)) = c \sum_{x \in X} f(x).$$

(h) (*Monotonicity*) Let X be a finite set, and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$. be functions such that $f(x) \leq g(x)$ for all $x \in X$. Then we have

$$\sum_{x \in X} f(x) \leq \sum_{x \in X} g(x).$$

Prove (h). Suppose that X is a finite set with n elements. Let $h : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X$. Since $f(x) \leq g(x)$ for all $x \in X$, we have $f(h(i)) \leq g(h(i))$, for all $1 \leq i \in \mathbf{N}$. Therefore,

$$\sum_{i=1}^n f(h(i)) \leq \sum_{i=1}^n g(h(i)) \implies \sum_{x \in X} f(x) \leq \sum_{x \in X} g(x).$$

(i) (*Triangle inequality*) Let X be a finite set, and let $f : X \rightarrow \mathbf{R}$ be a function, then

$$\left| \sum_{x \in X} f(x) \right| \leq \sum_{x \in X} |f(x)|.$$

Prove (i). Suppose that X is a finite set with n element. Let $h : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow X$ be a bijection. By Lemma 7.1.4(e), we have

$$\left| \sum_{i=1}^n f(h(i)) \right| \leq \sum_{i=1}^n |f(h(i))|.$$

Therefore,

$$\left| \sum_{x \in X} f(x) \right| \leq \sum_{x \in X} |f(x)|.$$

□

Exercise 7.1.3. Form a definition for the finite products $\prod_{i=1}^n a_i$ and $\prod_{x \in X} f(x)$. Which of the above results for finite series have analogues for products? (Note that it is dangerous to apply logarithms because some of the a_i or $f(x)$ could be zero or negative. Besides, we haven't defined logarithms yet.)

Proof.

Definition 1 (Finite products). Let m, n be integers, and let $(a_i)_{i=m}^n$ be a finite sequence of real numbers, assigning a real number a_i to each integer i between m and n inclusive (i.e., $m \leq i \leq n$). Then we define the finite product by the recursive formula

$$\prod_{i=m}^n a_i := 0 \text{ whenever } n < m;$$

$$\prod_{i=m}^{n+1} a_i := \left(\prod_{i=m}^n a_i \right) \times a_{n+1} \text{ whenever } n \geq m - 1.$$

Thus for instance we have the identities

$$\prod_{i=m}^{m-2} a_i = 0; \quad \prod_{i=m}^{m-1} a_i = 0; \quad \prod_{i=m}^m a_i = a_m;$$

$$\prod_{i=m}^{m+1} a_i = a_m \times a_{m+1}; \quad \prod_{i=m}^{m+2} a_i = a_m \times a_{m+1} \times a_{m+2}.$$

Because of this, we sometimes express $\prod_{i=m}^n a_i$ less formally as

$$\prod_{i=m}^n a_i = a_m \times a_{m+1} \times \cdots \times a_n.$$

Definition 2 (Productions over finite sets). Let X be a finite set with n elements (where $n \in \mathbf{N}$), and let $f : X \rightarrow \mathbf{R}$ be a function from X to the real numbers (i.e., f assigns a real number $f(x)$ to each element of X). Then we can define the finite product $\prod_{x \in X} f(x)$ as follows. We first select any bijection g from $\{i \in \mathbf{N} : 1 \leq i \leq n\}$ to X ; such a bijection exists since X is assumed to have n elements. We then define

$$\prod_{x \in X} f(x) := \prod_{i=1}^n f(g(i)).$$

Proposition (Basic properties of production over finite sets).

- (a) If X is empty, and $f : X \rightarrow \mathbf{R}$ is a function (i.e., f is the empty function), we have

$$\prod_{x \in X} f(x) = 0.$$

- (b) If X consists of a single element, $X = \{x_o\}$, and $f : X \rightarrow \mathbf{R}$ is a function, we have

$$\prod_{x \in X} f(x) = f(x_o).$$

- (c) (Substitution, part I) If X is a finite set, $f : X \rightarrow \mathbf{R}$ is a function, and $g : Y \rightarrow X$ is a bijection, then

$$\prod_{x \in X} f(x) = \prod_{y \in Y} f(g(y)).$$

- (d) (Substitution, part II) Let $n \leq m$ be integers, and let X be the set $X := \{i \in \mathbf{Z} : n \leq i \leq m\}$. If a_i is a real number assigned to each integer $i \in X$, then we have

$$\prod_{i=n}^m a_i = \prod_{i \in X} a_i.$$

- (e) Let X, Y be disjoint finite sets (so $X \cap Y = \emptyset$), and $f : X \cup Y \rightarrow \mathbf{R}$ is a function. Then we have

$$\prod_{z \in X \cup Y} f(z) = \left(\prod_{x \in X} f(x) \right) \times \left(\prod_{y \in Y} f(y) \right).$$

- (f) (Linearity, part I) Let X be a finite set, and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions. Then

$$\prod_{x \in X} (f(x) \times g(x)) = \left(\prod_{x \in X} f(x) \right) \times \left(\prod_{x \in X} g(x) \right).$$

- (g) (Linearity, part II) Let X be a finite set with n elements, let $f : X \rightarrow \mathbf{R}$ be a function, and let c be a real number. Then

$$\prod_{x \in X} cf(x) = c^n \prod_{x \in X} f(x).$$

- (h) (Monotonicity) Let X be a finite set, and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions such that $f(x) \leq g(x)$ for all $x \in X$. Then we have

$$\prod_{x \in X} f(x) \leq \prod_{x \in X} g(x).$$

(i) (Triangle inequality) Let X be a finite set, and let $f : X \rightarrow \mathbf{R}$ be a function, then

$$\left| \prod_{x \in X} f(x) \right| \leq \prod_{x \in X} |f(x)|.$$

□

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 .)

Proof. We use induction on n . When $n = 0$, we have $(x+y)^0 = 1$ and

$$\sum_{j=0}^0 \frac{n!}{j!(n-j)!} x^j y^{n-j} = 1.$$

Now we inductively assume that the equality is hold. Then we have

$$\begin{aligned} (x+y)^{n+1} &= (x+y)^n \times (x+y) \\ &= \left(\sum_{j=0}^n \frac{n!}{j!(n-j)!} x^j y^{n-j} \right) \times (x+y) \\ &= x \left(\sum_{j=0}^n \frac{n!}{j!(n-j)!} x^j y^{n-j} \right) + y \left(\sum_{j=0}^n \frac{n!}{j!(n-j)!} x^j y^{n-j} \right) \\ &= \left(\sum_{j=0}^n \frac{n!}{j!(n-j)!} x^{j+1} y^{n-j} \right) + \left(\sum_{j=0}^n \frac{n!}{j!(n-j)!} x^j y^{n-j+1} \right). \end{aligned} \tag{7.1}$$

We pick the term that $j = 1$ out of the right-hand side of the second term of (7.1). Then we can thus write the right-hand side of (7.1) as

$$= \left(\sum_{j=0}^n \frac{n!}{j!(n-j)!} x^{j+1} y^{n-j} \right) + \left(\sum_{j=1}^n \frac{n!}{j!(n-j)!} x^j y^{n-j+1} \right) + y^{n+1}. \tag{7.2}$$

Again, we pick the $(j = n)^{th}$ term out of the first term of (7.2). Then we write (7.2) as

$$= \left(\sum_{j=0}^{n-1} \frac{n!}{j!(n-j)!} x^{j+1} y^{n-j} \right) + \left(\sum_{j=1}^n \frac{n!}{j!(n-j)!} x^j y^{n-j+1} \right) + x^{n+1} + y^{n+1}. \quad (7.3)$$

Let $j = k - 1$ in the first term of (7.3), and let $j = k$ in the second term of (7.3). Then

$$\begin{aligned} &= \left(\sum_{k=1}^n \frac{n!}{(k-1)!(n-k+1)!} x^k y^{n-k+1} \right) \\ &\quad + \left(\sum_{k=1}^n \frac{n!}{k!(n-k)!} x^k y^{n-k+1} \right) + x^{n+1} + y^{n+1} \\ &= \left(\sum_{k=1}^n \left[\frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \right] x^k y^{n-k+1} \right) + x^{n+1} + y^{n+1}. \end{aligned} \quad (7.4)$$

Because we have the identity, which is called the *Pascal's formula*

$$\frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} = \frac{(n+1)!}{k!(n-k+1)!}.$$

Therefore, (7.4) can write as

$$= \left(\sum_{k=1}^n \frac{(n+1)!}{k!(n-k+1)!} x^k y^{n-k+1} \right) + x^{n+1} + y^{n+1}. \quad (7.5)$$

Since we have known that x^{n+1} is the $(k = n + 1)^{th}$ term and y^{n+1} is the $(k = 0)^{th}$ term of (7.5). So we obtain

$$= \left(\sum_{k=0}^{n+1} \frac{(n+1)!}{k!(n-k+1)!} x^k y^{n-k+1} \right).$$

from (7.5). Finally, we can safely replace k by j , and obtain that

$$(x + y)^{n+1} = \left(\sum_{j=0}^{n+1} \frac{(n+1)!}{j!(n-j+1)!} x^j y^{n-j+1} \right),$$

as desired. This close our induction. \square

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

$$\lim_{n \rightarrow \infty} \sum_{x \in X} a_n(x) = \sum_{x \in X} \lim_{n \rightarrow \infty} a_n(x).$$

(Hint: induct on the cardinality of X , and use Theorem 6.1.19(a).) Thus we may always interchange finite sums with convergent limits. Things however get trickier with infinite sums: see Exercise 11.47.11.

Proof. We induct on the cardinality of X . Suppose $\#(X) = 0$, then $\sum_{x \in X} a_n(x) = 0$, and we have $\lim_{n \rightarrow \infty} 0 = 0$. Now we inductively suppose that the equality is hold for $\#(X) = n$, we need to prove that the case of $\#(X) = n + 1$. Let X be a set with $n + 1$ elements. Then $X - \{x_0\}$ has n elements. By Proposition 7.1.11(e) and Theorem 6.1.19(a), we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{x \in X} a_n(x) &= \lim_{n \rightarrow \infty} \left(\sum_{x \in X - \{x_0\}} a_n(x) + a_n(x_0) \right) \\ &= \lim_{n \rightarrow \infty} \sum_{x \in X - \{x_0\}} a_n(x) + \lim_{n \rightarrow \infty} a_n(x_0) \\ &= \sum_{x \in X - \{x_0\}} \lim_{n \rightarrow \infty} a_n(x) + \lim_{n \rightarrow \infty} a_n(x_0) \\ &= \sum_{x \in X} \left(\lim_{n \rightarrow \infty} a_n(x) + \lim_{n \rightarrow \infty} a_n(x_0) \right) = \sum_{x \in X} \lim_{n \rightarrow \infty} a_n(x). \end{aligned}$$

This close the induction. □

7.2

Infinite series

Exercise 7.2.1. Is the series $\sum_{n=1}^{\infty} (-1)^n$ convergent or divergent? Justify your answer. Can you now resolve the difficulty in Example 1.2.2?

Proof. Consider the formal infinite series

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

For partial sums we have

$$S_{2N} = \sum_{n=1}^{2N} (-1)^n = (-1)^{2N};$$

$$S_{2N+1} = \sum_{n=1}^{2N+1} (-1)^n = 1 + (-1)^{2N+1}.$$

The sequence S_{2N} converges to 1, but S_{2N+1} converges to 0. Hence $\sum_{n=1}^{\infty} (-1)^n$ is divergent. \square

Exercise 7.2.2. Prove Proposition 7.2.5. (Hint: use Proposition 6.1.12 and Theorem 6.4.18.)

Proposition 7.2.5. Let $\sum_{n=m}^{\infty} a_n$ be a formal series of real numbers. Then $\sum_{n=m}^{\infty} a_n$ converges if and only if, for every real number $\varepsilon > 0$, there exists an integer $N \geq m$ such that

$$\sum_{n=p}^q a_n \leq \varepsilon \text{ for all } p, q \geq N.$$

Proof. Let S_p and S_q be the partial sums $S_p := \sum_{n=m}^p a_n$ and $S_q := \sum_{n=m}^q a_n$ for any $p, q \geq m$. Suppose that $\sum_{n=m}^{\infty} a_n$ converges, by Proposition 6.1.12, $\sum_{n=m}^{\infty} a_n$ is a Cauchy sequence. Then for every $\varepsilon > 0$, there exists an $N \geq m$ such that

$$|S_q - S_p| = \left| \sum_{n=p}^q a_n \right| \leq \varepsilon \text{ for all } p, q \geq N.$$

Now suppose that for every $\varepsilon > 0$, there exists an $N \geq m$ such that $|\sum_{n=p}^q a_n| \leq \varepsilon$ for all $p, q \geq N$. This means $(a_n)_{n=m}^{\infty}$ is a Cauchy sequence. By Theorem 6.4.18, $(a_n)_{n=m}^{\infty}$ is convergent. \square

Exercise 7.2.3. Use Proposition 7.2.5 to prove Corollary 7.2.6.

Corollary 7.2.6 (Zero test). Let $\sum_{n=m}^{\infty} a_n$ be a convergent series of real numbers. Then we must have $\lim_{n \rightarrow \infty} a_n = 0$. To put this another way, if $\lim_{n \rightarrow \infty} a_n$ is non-zero or divergent, then the series $\sum_{n=m}^{\infty} a_n$ divergent.

Proof. Since $\sum_{n=m}^{\infty} a_n$ be a convergent series of real numbers, for every $\varepsilon > 0$ there is an $N \geq m$ such that

$$\left| \sum_{n=p}^q a_n \right| \leq \varepsilon \text{ for all } p, q \geq N.$$

Let $p = q = n \geq N$. We have

$$\left| \sum_{n=n}^n a_n \right| = |a_n| \leq \varepsilon \text{ for all } n \geq N.$$

By definition of the limit of a sequence, $\lim_{n \rightarrow \infty} a_n = 0$. □

Exercise 7.2.4. Prove Proposition 7.2.9. (Hint: use Proposition 7.2.5 and Proposition 7.1.4(e).)

Proposition 7.2.9 (Absolute convergence test). Let $\sum_{n=m}^{\infty} a_n$ be a formal series of real numbers. If this series is absolutely convergent, then it is also conditionally convergent. Furthermore, in this case we have the triangle inequality

$$\left| \sum_{n=m}^{\infty} a_n \right| \leq \sum_{n=m}^{\infty} |a_n|.$$

Proof. Since $\sum_{n=m}^{\infty} a_n$ is absolutely convergent, by Lemma 7.1.4(e) and Proposition 7.2.5, we have

$$\left| \sum_{n=p}^q a_n \right| \leq \sum_{n=p}^q |a_n| = \left| \sum_{n=p}^q |a_n| \right| \leq \varepsilon.$$

Hence $\sum_{n=m}^{\infty} a_n$ is convergent. □

Exercise 7.2.5. Prove Proposition 7.2.14. (Hint: use Theorem 6.1.19.)

Proof.

- (a) If $\sum_{n=m}^{\infty} a_n$ is a series of real numbers converging to x , and $\sum_{n=m}^{\infty} b_n$ is a series of real numbers converging to y , then $\sum_{n=m}^{\infty} (a_n + b_n)$ is also a convergent series, and converges to $x + y$. In particular, we have

$$\sum_{n=m}^{\infty} (a_n + b_n) = \sum_{n=m}^{\infty} a_n + \sum_{n=m}^{\infty} b_n.$$

Prove (a). Let A_N and B_N be the partial sums $A_N := \sum_{n=m}^N a_n$ and $B_N := \sum_{n=m}^N b_n$. Then we have $\lim_{N \rightarrow \infty} A_N = x$ and $\lim_{N \rightarrow \infty} B_N = y$. By Theorem 6.1.19(a)

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=m}^N (a_n + b_n) &= \lim_{N \rightarrow \infty} \left(\sum_{n=m}^N a_n + \sum_{n=m}^N b_n \right) \\ &= \lim_{N \rightarrow \infty} \sum_{n=m}^N a_n + \lim_{N \rightarrow \infty} \sum_{n=m}^N b_n \\ &= \lim_{N \rightarrow \infty} A_N + \lim_{N \rightarrow \infty} B_N = x + y. \end{aligned}$$

$\sum_{n=m}^N (a_n + b_n)$ is the partial sum of $\sum_{n=m}^{\infty} (a_n + b_n)$, so it converges to $x + y$.

- (b) If $\sum_{n=m}^{\infty} a_n$ is a series of real numbers converging to x , and c is a real number, then $\sum_{n=m}^{\infty} (ca_n)$ is also a convergent series, and converges to cx . In particular, we have

$$\sum_{n=m}^{\infty} (ca_n) = c \sum_{n=m}^{\infty} a_n.$$

Prove (b). Let A_N be the partial sum $A_N := \sum_{n=m}^N a_n$. Then we have $\lim_{N \rightarrow \infty} A_N = x$. By Theorem 6.1.19(c)

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=m}^N (ca_n) &= \lim_{N \rightarrow \infty} \left(c \sum_{n=m}^N a_n \right) \\ &= \lim_{N \rightarrow \infty} cA_N \\ &= c \lim_{N \rightarrow \infty} A_N = cx. \end{aligned}$$

$\sum_{n=m}^N (ca_n)$ is the partial sum of $\sum_{n=m}^{\infty} (ca_n)$, so it converges to cx .

- (c) Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers, and let $k \geq 0$ be an integer. If one of the two series $\sum_{n=m}^{\infty} a_n$ and $\sum_{n=m+k}^{\infty} a_n$ are convergent, then the other one is also, and we have the identity

$$\sum_{n=m}^{\infty} a_n = \sum_{n=m}^{m+k-1} a_n + \sum_{n=m+k}^{\infty} a_n.$$

Prove (c). Let S_N and K_N be the partial sums $S_N := \sum_{n=m}^N a_n$ and $K_N := \sum_{n=m+k}^N a_n$. We use induction on k , when $k = 0$, we have

$$\sum_{n=m}^{m-1} a_n + \sum_{n=m}^{\infty} a_n = \sum_{n=m}^{m-1} a_n + \lim_{N \rightarrow \infty} S_N = 0 + \lim_{N \rightarrow \infty} S_N = \sum_{n=m}^{\infty} a_n.$$

Now inductively suppose that k is hold for the equality. Consider the case of $k + 1$

$$\begin{aligned} & \sum_{n=m}^{m+k} a_n + \sum_{n=m+k+1}^{\infty} a_n \\ &= \lim_{N \rightarrow \infty} \sum_{n=m}^{m+k} a_n + \lim_{N \rightarrow \infty} \sum_{n=m+k+1}^N a_n \\ &= \lim_{N \rightarrow \infty} \left(\sum_{n=m}^{m+k-1} a_n + \sum_{n=m+k}^N a_n + a_{m+k} - a_{m+k} \right) \\ &= \lim_{N \rightarrow \infty} \sum_{n=m}^N a_n = \sum_{n=m}^{\infty} a_n. \end{aligned}$$

This close the induction.

- (d) Let $\sum_{n=m}^{\infty} a_n$ be a series of real numbers converging to x , and let k be an integer. Then $\sum_{n=m+k}^{\infty} a_{n-k}$ also converges to x .

Prove (d). Let A_N and B_N be the partial sums $A_N := \sum_{n=m}^N a_n$ and $K_N := \sum_{n=m+k}^N a_{n-k}$. Then we have $\lim_{N \rightarrow \infty} A_N = x$. Let $s = n - k$

$$\lim_{N \rightarrow \infty} \sum_{n=m+k}^N a_{n-k} = \lim_{N \rightarrow \infty} \sum_{n-k=m}^N a_{n-k} = \lim_{N \rightarrow \infty} \sum_{s=m}^N a_s = x.$$

□

Exercise 7.2.6. Prove Lemma 7.2.15. (Hint: First work out what the partial sums $\sum_{n=0}^N (a_n - a_{n+1})$ should be, and prove your assertion using induction. How does the proposition change if we assume that a_n does not converge to zero, but instead converges to some other real number L ?

Lemma 7.2.15 (Telescoping series). Let $(a_n)_{n=0}^\infty$ be a sequence of real numbers which converge to 0, i.e., $\lim_{n \rightarrow \infty} a_n = 0$. Then the series $\sum_{n=0}^\infty (a_n - a_{n+1})$ converges to a_0 .

Proof. Let S_N be the partial sum $S_N := \sum_{n=0}^N (a_n - a_{n+1})$. We induct on n . When $n = 0$, we have

$$S_0 = \sum_{n=0}^0 (a_n - a_{n+1}) = a_0 - a_1$$

Now we inductively suppose that $S_N = a_0 - a_{N+1}$ in the case N . We show that

$$\begin{aligned} S_{N+1} &= \sum_{n=0}^{N+1} (a_n - a_{n+1}) \\ &= \sum_{n=0}^N (a_n - a_{n+1}) + a_{N+1} - a_{N+2} \\ &= a_0 - a_{N+1} + a_{N+1} - a_{N+2} = a_0 - a_{N+2} \end{aligned}$$

This close the induction, and we have $S_N = a_0 - a_{N+1}$. When $N \rightarrow \infty$, by hypothesis, $S_N = a_0$. Furthermore, when a_n converges to some other real number L , then $\sum_{n=0}^\infty (a_n - a_{n+1})$ converges to $a_0 + L$. \square

7.3

Sums of non-negative numbers

Exercise 7.3.1. Use Proposition 7.3.1 to prove Corollary 7.3.2.

Corollary 7.3.2 (Comparison test). Let $\sum_{n=m}^\infty a_n$ and $\sum_{n=m}^\infty b_n$ be two formal series of real numbers, and suppose that $|a_n| \leq b_n$ for all $n \geq m$. Then if $\sum_{n=m}^\infty b_n$ is convergent, then $\sum_{n=m}^\infty a_n$ is absolutely convergent, and in fact

$$\left| \sum_{n=m}^\infty a_n \right| \leq \sum_{n=m}^\infty |a_n| \leq \sum_{n=m}^\infty b_n.$$

Proof. Since $\sum_{n=m}^{\infty} b_n$ is convergent, by Proposition 7.3.1, there is a real number M such that

$$\sum_{n=m}^N b_n \leq M$$

for all integers $N \geq m$. Since $|a_n| \leq b_n$ for all $n \geq m$, by Lemma 7.1.4(f)

$$\sum_{n=m}^N |a_n| \leq \sum_{n=m}^N b_n \leq M$$

for all integers $N \geq m$. This means that $\sum_{n=m}^{\infty} a_n$ is absolutely convergent. By Proposition 7.2.9, it also is conditionally convergent. Then we have

$$\left| \sum_{n=m}^{\infty} a_n \right| \leq \sum_{n=m}^{\infty} |a_n| \leq \sum_{n=m}^{\infty} b_n.$$

□

Exercise 7.3.2. Prove Lemma 7.3.3. (Hint: for the first part, use the zero test. For the second part, first use induction to establish the **geometric series formula**

$$\sum_{n=0}^N x^n = (1 - x^{N+1})/(1 - x)$$

and then apply Lemma 6.5.2.)

Lemma 7.3.3 (Geometric series). Let x be a real number. If $|x| \geq 1$, then the series $\sum_{n=0}^{\infty} x^n$ is divergent. If however $|x| < 1$, then the series is absolutely convergent and

$$\sum_{n=0}^{\infty} x^n = 1/(1 - x).$$

Proof. If $|x| \geq 1$, we have $\lim_{n \rightarrow \infty} x^n \neq 0$, so that $\sum_{n=0}^{\infty} x^n$ is divergent. While if $|x| < 1$, let S_N be the partial sum $S_N := \sum_{n=0}^N x^n$. We use induction on N . When $N = 0$,

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

Now assume inductively that geometric series formula is hold for N . We need to show that case $N + 1$. Because

$$\sum_{n=0}^{N+1} x^n = \sum_{n=0}^N x^n + x_{N+1} = \frac{1 - x^{N+1}}{1 - x} + x^{N+1} = \frac{1 - x^{N+2}}{1 - x}.$$

This close the induction, and we have the identity

$$\sum_{n=0}^N x^n = (1 - x^{N+1})/(1 - x).$$

$\sum_{n=0}^N x^n = 1/(1 - x)$ when $N \rightarrow \infty$. In particular we can see that

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=0}^N x^n &= \lim_{N \rightarrow \infty} \frac{1 - x^{N+1}}{1 - x} = \frac{1}{1 - x}; \\ \lim_{N \rightarrow \infty} \sum_{n=0}^N |x^n| &= \lim_{N \rightarrow \infty} \sum_{n=0}^N |x|^n = \lim_{N \rightarrow \infty} \frac{1 - |x|^{N+1}}{1 - |x|} = \frac{1}{1 - |x|}, \end{aligned}$$

when $|x| < 1$. So $\sum_{n=0}^{\infty} x^n$ is absolutely convergent. \square

Exercise 7.3.3. Let $\sum_{n=0}^{\infty} a_n$, be an absolutely convergent series of real numbers such that $\sum_{n=0}^{\infty} |a_n| = 0$. Show that $a_n = 0$ for every natural number n .

Proof. Suppose for sake of contradiction that there is a natural number m such that $a_m \neq 0$, so that $|a_m| > 0$. Let S_N be the partial sum $S_N := \sum_{n=0}^N |a_n|$, in which $N \geq m$. Then we have $\lim_{N \rightarrow \infty} S_N = 0$. Then

$$\begin{aligned} \lim_{N \rightarrow \infty} \sum_{n=0}^N |a_n| &= \lim_{N \rightarrow \infty} \left(\sum_{n=0}^{m-1} |a_n| + \sum_{n=m+1}^N |a_n| + |a_m| \right) \\ &= \lim_{N \rightarrow \infty} \left(\sum_{n=1}^{m-1} |a_{n+1}| + \sum_{n=m}^{N-1} |a_{n+1}| \right) + \lim_{N \rightarrow \infty} |a_m| \\ &= \lim_{N \rightarrow \infty} \left(\sum_{n=1}^{N-1} |a_{n+1}| \right) + |a_m| = |a_m|, \end{aligned}$$

a contradiction. \square

7.4**Rearrangement of series**

Exercise 7.4.1. Let $\sum_{n=0}^{\infty} a_n$ be an absolutely convergent series of real numbers. Let $f : \mathbf{N} \rightarrow \mathbf{N}$ be an increasing function (i.e., $f(n+1) > f(n)$ for all $n \in \mathbf{N}$). Show that $\sum_{n=0}^{\infty} a_{f(n)}$ is also an absolutely convergent series. (Hint: try to compare each partial sum of $\sum_{n=0}^{\infty} a_{f(n)}$ with a (slightly different) partial sum of $\sum_{n=0}^{\infty} a_n$.)

Proof. Let $L := \sum_{n=0}^{\infty} |a_n|$.

7.5**The root and ratio tests**

Exercise 7.5.1. Prove the first inequality in Lemma 7.5.2.

Lemma 7.5.2. Let $(c_n)_{n=m}^{\infty}$ be a sequence of positive numbers. Then we have

$$\liminf_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} \leq \liminf_{n \rightarrow \infty} c_n^{1/n} \leq \limsup_{n \rightarrow \infty} c_n^{1/n} \leq \limsup_{n \rightarrow \infty} \frac{c_{n+1}}{c_n}.$$

Proof. Prove the inequality

$$\liminf_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} \leq \liminf_{n \rightarrow \infty} c_n^{1/n}.$$

Write $G := \liminf_{n \rightarrow \infty} c_{n+1}/c_n$, and assume that $L \in \mathbf{R}$. Since c_{n+1}/c_n is always positive, we know that $G \geq 0$.

Let $\varepsilon > 0$. By Proposition 6.4.12(a), we know that there exists an $N \geq m$ such that $G - \varepsilon \leq c_{n+1}/c_n$ for all $n \geq N$. This implies that $c_n(G - \varepsilon) \leq c_{n+1}$ for all $n \geq N$. By induction, we show that this implies that

$$c_N(G - \varepsilon)^{n-N} \leq c_n \text{ for all } n \geq N.$$

We induct on n . When $n = N$, we have $c_N \leq c_N$. Now inductively

suppose that inequality is hold for n . Consider case of $n + 1$

$$\begin{aligned} c_N(G - \varepsilon)^{n-N+1} &= c_N(G - \varepsilon)^{n-N}(G - \varepsilon) \\ &\leq c_n(G - \varepsilon) \leq c_{n+1}. \end{aligned}$$

This close the induction. If we write $A := c_N(G - \varepsilon)^{-N}$, then we have

$$A(G - \varepsilon)^n \leq c_n$$

and thus

$$A^{1/n}(G - \varepsilon) \leq c$$

for all $n \geq N$. But we have

$$\lim_{n \rightarrow \infty} A^{1/n}(G - \varepsilon) = G - \varepsilon.$$

Thus by the comparison principle (Lemma 6.4.13) we have

$$G - \varepsilon \leq \liminf_{n \rightarrow \infty} c_n^{1/n}.$$

For show that $G \leq \liminf_{n \rightarrow \infty} c_n^{1/n}$, we suppose for sake of contradiction that $G > \liminf_{n \rightarrow \infty} c_n^{1/n}$. Let $\varepsilon = (G + \liminf_{n \rightarrow \infty} c_n^{1/n})/2 > 0$, then we have

$$G - \varepsilon = G/2 + (\liminf_{n \rightarrow \infty} c_n^{1/n})/2 \leq \liminf_{n \rightarrow \infty} c_n^{1/n}.$$

This implies that

$$G/2 \leq \liminf_{n \rightarrow \infty} c_n^{1/n}/2 \implies G \leq \liminf_{n \rightarrow \infty} c_n^{1/n}$$

a contradiction. So we conclude that

$$G \leq \liminf_{n \rightarrow \infty} c_n^{1/n}$$

as desired. □

Exercise 7.5.2. Let x be a real number with $|x| < 1$, and q be a real number. Show that the series $\sum_{n=1}^{\infty} n^q x^n$ is absolutely convergent, and that $\lim_{n \rightarrow \infty} n^q x^n = 0$.

Proof. Use ratio test. Let $a_n = n^q x^n$, then

$$\frac{|a_{n+1}|}{|a_n|} = \frac{|(n+1)^q x^{n+1}|}{|n^q x^n|} = |x| \left(\frac{n+1}{n} \right)^q.$$

Since the limit of such sequence is exist, and

$$\limsup_{n \rightarrow \infty} |x| \left(\frac{n+1}{n} \right)^q = |x| < 1.$$

This means that $\sum_{n=1}^{\infty} n^q x^n$ is absolutely convergent. By Corollary 7.2.6, we have $\lim_{n \rightarrow \infty} n^q x^n = 0$. \square

Exercise 7.5.3. Give an example of a divergent series $\sum_{n=m}^{\infty} a_n$ of positive numbers a_n such that $\lim_{n \rightarrow \infty} a_{n+1}/a_n = \lim_{n \rightarrow \infty} a_n^{1/n} = 1$, and give an example of a convergent series $\sum_{n=m}^{\infty} b_n$ of positive numbers b_n such that $\lim_{n \rightarrow \infty} b_{n+1}/b_n = \lim_{n \rightarrow \infty} b_n^{1/n} = 1$. (Hint: use Corollary 7.3.7.) This shows that the ratio and root tests can be inconclusive even when the summands are positive and all the limits converge.

Proof. Let $a_n = 1/n$ and $b_n = 1/n^2$, by Corollary 7.3.7, $\sum_{n=m}^{\infty} 1/n$ is divergent and $\sum_{n=m}^{\infty} 1/n^2$ is convergent. But

$$\begin{aligned}\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} &= \lim_{n \rightarrow \infty} \frac{n}{n+1} = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{1+n}\right) = 1; \\ \lim_{n \rightarrow \infty} \frac{b_{n+1}}{b_n} &= \lim_{n \rightarrow \infty} \left(\frac{n}{n+1}\right)^2 = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{1+n}\right)^2 = 1.\end{aligned}$$

This means that ratio test is inconclusive. Then for root test, by Lemma 6.5.3, we still have

$$\begin{aligned}\lim_{n \rightarrow \infty} \left(\frac{1}{n}\right)^{1/n} &= \lim_{n \rightarrow \infty} \frac{1}{n^{1/n}} = 1; \\ \lim_{n \rightarrow \infty} \left(\frac{1}{n^2}\right)^{1/n} &= \lim_{n \rightarrow \infty} \left(\frac{1}{n^{1/n}}\right)^2 = 1.\end{aligned}$$

This means that root test is inconclusive. \square

CHAPTER VIII

Infinite sets

8.1

Countability

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.)*

Proof. Suppose that X is a finite set with cardinality n . Since $Y \subsetneq X$, by Proposition 3.6.14(c), we have $\#(Y) < \#(X)$. But Y has the same cardinality with X , a contradiction. Thus X is infinite by Definition 3.6.10.

Then we suppose that X is infinite. First of all we prove that X must contain a countable subset. We need to prove that there exists some sets $A_n := \{a_k \in X : 0 \leq k \leq n\}$ for all $n \in \mathbf{N}$ inductively. By single choice, we have $A_0 = \{a_0\}$ for $a_0 \in X$, this is base case. Now we inductively suppose that there is $A_n := \{a_k \in X : 0 \leq k \leq n\}$ for some $k \geq 0$. By Proposition 3.6.14(a), we can show that $X - A_n$ is infinite set (suppose for sake of contradiction that $X - A_n$ is finite). Then choose an element $a_{n+1} \in X$ but $a_{n+1} \notin A_n$, and define $A_{n+1} := A_n \cup \{a_{n+1}\}$, by Proposition 3.6.14(a) again, A_{n+1} is finite, this close the induction.

Define

$$A := \bigcup_{n \in \mathbf{N}} A_n.$$

Then A is a countable set and $A \subseteq X$ (we can define function $f : A \rightarrow \mathbf{N}$ as $f(a_n) = n$, and every element of A is selected from X).

Now we define a bijection $f : X \rightarrow X \setminus \{a_1\}$ onto a proper subset of X by following rule:

$$\begin{aligned} f(a_n) &= a_{n+1} & \text{for } a_n \in A, \\ f(x) &= x & \text{for } x \notin A. \end{aligned}$$

Let $Y = X \setminus \{a_1\}$, then we have X and the proper subset Y has same cardinality. \square

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 integer? What if we replace the natural numbers by the positive rationals? Explain.*

Proposition 8.1.4 (Well ordering principle). Let X be a non-empty subset of the natural numbers \mathbf{N} . Then there exists exactly one element $n \in X$ such that $n \leq m$ for all $m \in X$. In other words, every non-empty set of natural numbers has a minimum element.

Proof. We prove that for every $n \in \mathbf{N}$, every non-empty subset of $A := \{m \in \mathbf{N} : 0 \leq m \leq n\}$ has a smallest element.

We prove this assertion inductively. When $n = 0$, $\{0\}$ is only non-empty subset of A , and 0 is the smallest element. Now inductively suppose that claim is true for n . We need to prove that for every non-empty subset, denoted by B , of $\{m \in \mathbf{N} : 0 \leq m \leq n + 1\}$ has a smallest element. If B include $n + 1$ only, then $n + 1$ is the smallest element of B . Otherwise, consider the set $B \cap \{m \in \mathbf{N} : 0 \leq m \leq n\}$, since for every $n \in A$ has a smallest element, then it also the smallest element of B . This close the induction.

Now we suppose that X is a non-empty subset of \mathbf{N} . We pick an n from X , then we have $A = X \cap \{m \in \mathbf{N} : 0 \leq m \leq n\}$, which consist a smallest element. Such element is also the smallest element of X , as desired. \square

Exercise 8.1.3. Fill in the gaps marked (?) in Proposition 8.1.5.

Proposition 8.1.5. Let X be an infinite subset of the natural numbers \mathbf{N} . Then there exists a unique bijection $f : \mathbf{N} \rightarrow X$ which is increasing, in the sense that $f(n+1) > f(n)$ for all $n \in \mathbf{N}$. In particular, X has equal cardinality with \mathbf{N} and is hence countable.

Proof. We now define a sequence a_0, a_1, a_2, \dots of natural numbers recursively by the formula

$$a_n := \min\{x \in X : x \neq a_m \text{ for all } m < n\}.$$

Since X is infinite, the set $\{x \in X : x \neq a_m \text{ for all } m < n\}$ is infinite, hence non-empty. We show this by induction principle. First of all we define the set

$$A := \{x \in X : x = a_m \text{ for all } m < n\}$$

inductively. Such a set is finite for all $n \in \mathbf{N}$. When $n = 0$, A is empty, and $\#(A) = 0$. If we consider $n = 1$, then $A = \{a_0\}$ and $\#(A) = 1$. This is base case. Now we suppose inductively that A has the cardinality with n . Since

$$\{x \in X : x = a_m \text{ for all } m < n+1\} = A \cup \{a_n\}.$$

By induction hypothesis, it has cardinality with $n+1$. This close the induction. Thus for all $n \in \mathbf{N}$, A has the cardinality with n , which means it is finite.

Now we can safely define the set

$$\{x \in X : x \neq a_m \text{ for all } m < n\} = X \setminus A.$$

By Proposition 3.6.14(a), it is infinite, hence non-empty.

Thus by the well-ordering principle, the minimum,

$$\min\{x \in X : x \neq a_m \text{ for all } m < n\}$$

is always well-defined.

One can show that a_n is an increasing sequence, i.e.

$$a_0 < a_1 < a_2 < \cdots$$

Let

$$X_n := \{x \in X : x \neq a_m \text{ for all } m < n\},$$

then $a_m = \min(X_n)$. We can see that

$$X_0 \supsetneq X_1 \supsetneq X_2 \supsetneq \cdots$$

By well-ordering principle, since a_0 is the only smallest element of X_0 , for every element of X_1, X_2, X_3, \dots , we have $a_0 < \min(X_1), a_0 < \min(X_2), a_0 < \min(X_3), \dots$. This is easy to prove with induction principle. In particular, if $a_n = a_m$ for some $m \neq n$. Since a_n is increasing sequence, for $a_n < a_m$ or $a_m < a_n$ we have $a_n < a_n$, a contradiction. Thus, there has $a_n \neq a_m$ for all $n \neq m$. Also, we have $a_n \in X$ for each natural number n by definition.

Now define the function $f : \mathbf{N} \rightarrow X$ by $f(n) := a_n$. From the previous paragraph we know that f is one-to-one. Now we show that f is onto. In other words, we claim that for every $x \in X$, there exists an n such that $a_n = x$.

Let $x \in X$. Suppose for sake of contradiction that $a_n \neq x$ for every natural number n . Then this implies that x is an element of the set $\{x \in X : x \neq a_m \text{ for all } m < n\}$ for all n . To show this, we induct on n . When $n = 1$ ($n = 0$ is vacuously true), we have $x \neq a_0$. By definition

$$x \in \{x \in X : x \neq a_0 \text{ for all } m < 1\}.$$

Now inductively suppose that the assertion is hold for n . Consider case $n+1$. We know that $x \neq a_n$, and the set $\{x \in X : x \neq a_m \text{ for all } m < n+1\}$ is equal to the union of $\{x \in X : x \neq a_m \text{ for all } m < n\}$ and $\{x \in X : x \neq a_n\}$. By induction hypothesis

$$x \in \{x \in X : x \neq a_m \text{ for all } m < n+1\}.$$

This close the induction. By definition of a_n , this implies that $x \geq a_n$ for every natural number. However, since a_n is an increasing sequence, we have $a_n > n$, and hence $x \geq n$ for every natural number n . Otherwise, we have $a_n \geq n$. Since $a_n \in X$ for all natural number n , and a_n is increasing, we

only have $a_n < a_{n+1} \leq n < n + 1$. When $n = 0$, we have $a_0 < 0$, which means $a_0 \notin X$, a contradiction. In particular, we have $x \geq x + 1$, which is a contradiction. Thus we must have $a_n = x$ for some natural number n , and hence f is onto.

Since $f : \mathbf{N} \rightarrow X$ is both one-to-one and onto, it is a bijection. We have thus found at least one increasing bijection f from \mathbf{N} to X . Now suppose for sake of contradiction that there was at least one other increasing bijection g from \mathbf{N} to X which was not equal to f . Then the set $\{x \in \mathbf{N} : g(n) \neq f(n)\}$ is non-empty, and define $m := \{x \in \mathbf{N} : g(n) \neq f(n)\}$, thus in particular $g(m) \neq f(m) = a_m$ for that m is the smallest element satisfy the condition, and $g(n) = f(n) = a_n$ for all $n < m$. Since g is increasing, for some $x = g(m)$

$$g(m) = \min\{x \in X : x \neq g(t) \text{ for all } t < m\}.$$

Replace $g(t)$ by a_t for all $t < m$, we then must have

$$g(m) = \{x \in X : x \neq a_t \text{ for all } t < m\} = a_m,$$

a contradiction. Thus there is no other increasing bijection from \mathbf{N} to X other than f . \square

Exercise 8.1.4. *Prove Proposition 8.1.8. (Hint: the basic problem here is that f is not assumed to be one-to-one. Define A to be the set*

$$A := \{n \in \mathbf{N} : f(m) \neq f(n) \text{ for all } 0 \leq m < n\};$$

informally speaking, A is the set of natural numbers n for which $f(n)$ does not appear in the sequence $f(0), f(1), \dots, f(n-1)$. Prove that when f is restricted to A , it becomes a bijection from A to $f(\mathbf{N})$. Then use Corollary 8.1.6)

Proposition 8.1.8. Let Y be a set, and let $f : \mathbf{N} \rightarrow Y$ be a function. Then $f(\mathbf{N})$ is at most countable.

Proof. If Y is finite, then it is at most countable. By Corollary 8.1.6, the result follows for that $f(\mathbf{N}) \subseteq Y$.

We consider that Y is infinite. Define A to be the set

$$A := \{n \in \mathbf{N} : f(m) \neq f(n) \text{ for all } 0 \leq m < n\}.$$

And denote f which be restricted to A by $f|_A : A \rightarrow f(\mathbf{N})$.

We first show $f|_A$ is one-to-one. Suppose that $m \neq n$ for $m, n \in A$, and $f|_A(m) = f|_A(n)$. If $m < n$, by definition of A , we have $n \notin A$; while if $m > n$, we have $m \notin A$, a contradiction. Thus $f|_A$ is one-to-one.

Now we show that $f|_A$ is onto. Let $y \in f(\mathbf{N})$. We claim that for every $n \in A$, there exists an $f|_A(n)$ such that $y = f|_A(n)$. Suppose for sake of contradiction that $f|_A(n) \neq y$ for every $n \in A$. By definition of image, for all $m \in \mathbf{N}$ and $f(m) = y$. If $m = n$, then $f|_A(n) = f(n)$, a contradiction. So $f|_A$ is onto.

Since $f|_A : A \rightarrow \mathbf{N}$ is both one-to-one and onto, it is a bijection. Thus \mathbf{N} has equal cardinality with A . But A is at most countable by Corollary 8.1.6. Hence $f(\mathbf{N})$ is also at most countable. \square

Remark 8.1.4. If f is one-to-one, there is a bijection $g : \mathbf{N} \rightarrow f(\mathbf{N})$ where $f(\mathbf{N}) \subseteq Y$. Then $f(\mathbf{N})$ is countable or finite. Since we don't know whether f is one-to-one or not, we construct a set A by the following rule: $0 \in A$ for that there is no m such that $f(m) = f(0)$. Consider 1, if $f(0) = f(1)$, then we except it; while if $f(0) \neq f(1)$, we have $1 \in A$, and so on. Then there is a injective function from A to $f(\mathbf{N})$ which is a partial function.

Exercise 8.1.5. Use Proposition 8.1.8 to prove Corollary 8.1.9.

Corollary 8.1.9. Let X be a countable set, and let $f : X \rightarrow Y$ be a function. Then $f(X)$ is at most countable.

Proof. Since X is countable, by definition, there is a bijective function $g : \mathbf{N} \rightarrow X$. Then we have $f \circ g : \mathbf{N} \rightarrow Y$. By Proposition 8.1.8, $f(X) = (f \circ g)(\mathbf{N})$ is at most countable. \square

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 $f : A \rightarrow \mathbf{N}$ from A to \mathbf{N} .

Proof. Since A is at most countable, it is either countable or finite. If A is countable, there is a bijective function from A to \mathbf{N} , assertion follows.

While if A is finite, by definition, it has cardinality n for some natural

numbers, i.e., there is a bijective function $f : \{i \in \mathbf{N} : 0 \leq i \leq n\} \rightarrow A$. Then for every $a \in A$ there exists a natural number n such that $f^{-1}(a) = n$, so $f^{-1}(\{a\})$ is non-empty. By well-ordering principle, we can now define $g : A \rightarrow \{i \in \mathbf{N} : 0 \leq i \leq n\}$ by following rule:

$$g(a) := \min (f^{-1}(\{a\})) .$$

We see that $g(a)$ is unique. For $a, a' \in A$, if $a \neq a'$, we have $f^{-1}(a) \neq f^{-1}(a')$, this means $g(a) \neq g(a')$. Thus g is injective.

Now suppose that $f : A \rightarrow \mathbf{N}$ is injective. There must exist a subset B of \mathbf{N} such that $f : A \rightarrow B$ is injection. By Corollary 8.1.6, B is countable, so that A is countable, then of course is at most countable. \square

Exercise 8.1.7. *Prove Proposition 8.1.10. (Hint: by hypothesis, we have a bijection $f : \mathbf{N} \rightarrow X$, and a bijection $g : \mathbf{N} \rightarrow Y$. Now define $h : \mathbf{N} \rightarrow X \cup Y$ by setting $h(2n) := f(n)$ and $h(2n + 1) := g(n)$ for every natural number n , and show that $h(\mathbf{N}) = X \cup Y$. Then use Corollary 8.1.9, and show that $X \cup Y$ cannot possibly be finite.)*

Proposition 8.1.10. Let X be a countable set, and let Y be a countable set. Then $X \cup Y$ is a countable set.

Proof. We show that h is onto. Suppose for sake of contradiction, there exists $a \in X \cup Y$ such that $h(n) \neq a$ for all $n \in \mathbf{N}$. If $a \in X$, then $a = f(n)$. For all $2n \in \mathbf{N}$ we have $h(2n) = f(n)$. While if $a \in Y$, then $a = g(n)$ and for all $2n + 1 \in \mathbf{N}$, we have $h(2n + 1) = g(n)$. A contradiction. Thus h is onto. Then we have $h(\mathbf{N}) = X \cup Y$. By Corollary 8.1.9, $X \cup Y$ is at most countable.

If $X \cup Y$ is finite, by Proposition 3.6.14, we must have $\#(X) < \#(X \cup Y)$. But X has same cardinality with \mathbf{N} which is infinite set, a contradiction. Thus $X \cup Y$ cannot be finite. \square

Exercise 8.1.8. *Use Corollary 8.1.13 to prove Corollary 8.1.14.*

Corollary 8.1.14. If X and Y are countable, then $X \times Y$ is countable.

Proof. Since X and Y are countable, there are bijective functions $f : \mathbf{N} \rightarrow$

X and $g : \mathbf{N} \rightarrow Y$. We need to show that $h : \mathbf{N} \times \mathbf{N} \rightarrow X \times Y$ is bijective. We define f by following equation, $h(n, m) = (f(n), g(m))$. Since f and g are bijective, h also bijective. By Corollary 8.1.13, $X \times Y$ is countable. \square

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.)*

Proof. Since I is at most countable, it is either finite or countable. Suppose that I is finite, and for each $\alpha \in I$ let A_α be a finite set. By definition, there is a bijective function $\tilde{f} : \{0, 1, \dots, n\} \rightarrow I$. We show that $\bigcup_{\alpha \in \tilde{f}(\mathbf{N})} A_{\tilde{f}(n)}$ is finite. Use induction on n . When $n = 0$, this is obviously true. Now we inductively suppose that $n - 1$ is true for assertion. Then

$$\bigcup_{\alpha \in \tilde{f}(\mathbf{N})} A_{\tilde{f}(n)} = \left(\bigcup_{\alpha \in \tilde{f}(\mathbf{N})} A_{\tilde{f}(n-1)} \right) \cup \{A_{\tilde{f}(n)}\}.$$

By induction hypothesis, it is finite.

Then we suppose that I is countable, and for each $\alpha \in I$ let A_α be a countable set. By definition, we can find a bijective function $f : \mathbf{N} \rightarrow A_\alpha$ and a bijective function $g : \mathbf{N} \rightarrow I$. Now we define

$$h : \mathbf{N} \times \mathbf{N} \rightarrow \bigcup_{\alpha \in g(\mathbf{N})} A_{g(n)}.$$

by the equation

$$h(n, m) = f_{g(n)}(m).$$

Since this is bijective, and by Corollary 8.1.13, $\bigcup_{\alpha \in g(\mathbf{N})} A_{g(n)}$ is countable. Together our results, $\bigcup_{\alpha \in I} A_\alpha$ is at most countable. In particular, countable unions of countable sets are countable. \square

Exercise 8.1.10. *Find a bijection $f : \mathbf{N} \rightarrow \mathbf{Q}$ from the natural numbers to the rationals. (Warning this is actually rather tricky to do explicitly; it is difficult to get f to be simultaneously injective and surjective.)*

Proof. We just give an example without proof. Every rational number q

can be represented as a finite continued fraction

$$q = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n}}}}$$

where $a_n \in \mathbf{N}$. □

8.2

Summation on infinite sets

Exercise 8.2.1. *Prove Lemma 8.2.3. (Hint: you may find Exercise 3.6.3 to be useful.)*¹

Lemma 8.2.3. Let X be a countable set, and let $f : X \rightarrow \mathbf{R}$ be a function. Then the series $\sum_{x \in X} f(x)$ is absolutely convergent if and only if

$$\sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite} \right\} < \infty.$$

Proof. Let

$$E = \left\{ \sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite} \right\}.$$

First of all we suppose the series $\sum_{x \in X} f(x)$ is absolutely convergent. Since X is countable and $f : X \rightarrow \mathbf{R}$ be a function, by Definition 8.2.1, for some bijection $g : \mathbf{N} \rightarrow X$, the sum $\sum_{n=0}^{\infty} f(g(n))$ is absolutely convergent. Let $\sum_{n=0}^{\infty} |f(g(n))|$ converges to L .

¹Let n be a natural number, and let $f : \{i \in \mathbf{N} : 1 \leq i \leq n\} \rightarrow \mathbf{N}$ be a function. Then there exists a natural number M such that $f(i) \leq M$ for all $1 \leq i \leq n$. Thus the finite subsets of the natural numbers are bounded. See p.36.

“In Lemma 8.2.3, X should be assumed to be *countable*, rather than *at most countable*.”
— Errata from Tao.

For every finite set $A \subsetneq X$, there is a finite set $N = \{n \in \mathbf{N} : 0 \leq n \leq m\} \subsetneq \mathbf{N}$ such that $g^{-1}(A) = N$. By Exercise 3.6.3, $g^{-1}(x)$ is bounded above by some real number M . Thus

$$\begin{aligned} \sum_{x \in A} |f(x)| &= \sum_{n \in g^{-1}(A)} |f(g(n))| \\ &= \sum_{n=0}^M |f(g(n))| \\ &\leq \sum_{n=0}^{\infty} |f(g(n))| \\ &\leq \sum_{n \in \mathbf{N}} |f(g(n))| = L. \end{aligned}$$

Hence

$$\sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite} \right\} < \infty.$$

For the converse, we suppose that

$$L := \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite} \right\} < \infty.$$

There is some finite subsets $N \subsetneq \mathbf{N}$ such that $g(N) = A$. Hence $\sum_{x \in A} |f(x)| \leq L$. But

$$\sum_{x \in A} |f(x)| = \sum_{x \in g(N)} |f(x)| = \sum_{n \in N} |f(g(n))| = \sum_{n=0}^n |f(g(n))|.$$

So we have $\sum_{n=0}^m |f(g(n))| \leq L$ for all natural number m . By Proposition 7.3.1, $\sum_{n=0}^{\infty} |f(g(n))|$ is convergent, so that $\sum_{x \in X} f(x)$ is absolutely convergent. \square

Exercise 8.2.2. Prove Lemma 8.2.5. (Hint: first show if M is the quantity

$$M := \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite} \right\}$$

then the sets $\{x \in X : |f(x)| > 1/n\}$ are finite with cardinality at most Mn for every positive integer n . Then use Exercise 8.1.9 (which uses the axiom of choice, see Section 8.4).)

Lemma 8.2.5. Let X be a set (which could be uncountable), and let $f : X \rightarrow \mathbf{R}$ be a function such that the series $\sum_{x \in X} f(x)$ is absolutely convergent. Then the set $\{x \in X : f(x) \neq 0\}$ is at most countable. (This result requires the axiom of choice, see Section 8.4.)

Proof. Since $\sum_{x \in X} f(x)$ is absolutely convergent, by Definition 8.2.4, we can let

$$M := \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X, A \text{ finite} \right\} < \infty.$$

This means for every subsets A of X , we have $\sum_{x \in A} |f(x)| \leq M$. Suppose that A has cardinality with natural number m , and let $f(x_m)$ satisfies that $f(x_m) \geq f(x)$ for every $x \in X$. We induct on n . When $n = 1$, we have

$$\sum_{x \in A} |f(x_m)| \leq (m+1)|f(x_m)| \leq M \implies |f(x_m)| \leq \frac{M}{m+1}.$$

If $|f(x)| > 1$, we must have $m \leq M$. Now we suppose inductively that there is at most Mn 's x such that $|f(x)| > 1/n$. Since

$$\frac{1}{n+1} < |f(x_m)| \leq \frac{M}{m+1},$$

It requires that $m+1 < M(n+1)$, which means $m \leq M(n+1)$. This close the induction. The set $\{x \in X : |f(x)| > 1/n\}$ are finite with cardinality at most Mn for every positive integer n .

Then by Exercise 8.1.9,

$$\{x \in X : f(x) \neq 0\} = \bigcup_{n=0}^{\infty} \left\{ x \in X : |f(x)| > \frac{1}{n} \right\}$$

is at most countable. □

Exercise 8.2.3. Prove Proposition 8.2.6. (Hint: you may of course use all the results from Chapter 7 to do this.)

Proof. Let X be arbitrary set (possibly uncountable), and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions such that the series $\sum_{x \in X} f(x)$ and $\sum_{x \in X} g(x)$ are both absolutely convergent.

(a) The series $\sum_{x \in X} (f(x) + g(x))$ is absolutely convergent, and

$$\sum_{x \in X} (f(x) + g(x)) = \sum_{x \in X} f(x) + \sum_{x \in X} g(x).$$

Prove (a). Let A be a finite subset of X . Since $\sum_{x \in X} f(x)$ and $\sum_{x \in X} g(x)$ are both absolutely convergent, by Definition 8.2.4,

$$\sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X \right\} + \sup \left\{ \sum_{x \in A} |g(x)| : A \subseteq X \right\} < \infty.$$

By Proposition 7.1.11, we know that

$$\sum_{x \in A} |f(x) + g(x)| \leq \sum_{x \in A} |f(x)| + \sum_{x \in A} |g(x)|,$$

which means

$$\begin{aligned} \sum_{x \in A} |f(x) + g(x)| &\leq \sup \left\{ \sum_{x \in A} |f(x) + g(x)| : A \subseteq X \right\} \\ &\leq \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X \right\} + \sup \left\{ \sum_{x \in A} |g(x)| : A \subseteq X \right\} < \infty. \end{aligned}$$

Hence, $\sum_{x \in X} (f(x) + g(x))$ is absolutely convergent.

Since X is either at most countable or uncountable. If X is finite, by Proposition 7.1.11, equation is obviously true. If X is countable, by Definition 8.2.1 and Proposition 7.2.14, there is a bijection $h : \mathbf{N} \rightarrow X$ such that

$$\sum_{n=0}^{\infty} (f(h(n)) + g(h(n))) = \sum_{n=0}^{\infty} f(h(n)) + \sum_{n=0}^{\infty} g(h(n)).$$

So the equation also hold for countable set X . While if X is uncountable, then we have

$$\begin{aligned} \sum_{x \in X} f(x) + \sum_{x \in X} g(x) &= \sum_{x \in X: f(x) \neq 0} f(x) + \sum_{x \in X: g(x) \neq 0} g(x) \\ &= \sum_{x \in X: f(x) + g(x) \neq 0} (f(x) + g(x)) \\ &= \sum_{x \in X} (f(x) + g(x)), \end{aligned}$$

where $\{x \in X : f(x) \neq 0\}$, $\{x \in X : g(x) \neq 0\}$ and $\{x \in X : f(x) + g(x) \neq 0\}$ are countable sets.

Thus for arbitrary set X , we have $\sum_{x \in A} (f(x) + g(x)) = \sum_{x \in A} f(x) + \sum_{x \in A} g(x)$.

(b) If c is a real number, then $\sum_{x \in X} cf(x)$ is absolutely convergent, and

$$\sum_{x \in X} cf(x) = c \sum_{x \in X} f(x).$$

Prove (b). Let A be a finite subset of X . By Definition 8.2.4, we have

$$\sum_{x \in A} |f(x)| \leq \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X \right\} < \infty.$$

Thus by Proposition 7.1.11,

$$\sum_{x \in A} |cf(x)| = |c| \sum_{x \in A} |f(x)| \leq |c| \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X \right\}.$$

Then we have

$$\sum_{x \in A} |cf(x)| \leq \sup \left\{ \sum_{x \in A} |cf(x)| : A \subseteq X \right\} < \infty.$$

Thus $\sum_{x \in X} cf(x)$ is absolutely convergent.

When X is at most countable, equation immediately follows from Proposition 7.1.11(g) and Proposition 7.2.13(b). If X is uncountable, then

$$c \sum_{x \in X} f(x) = c \sum_{x \in X: f(x) \neq 0} f(x) = \sum_{x \in X: cf(x) \neq 0} cf(x) = \sum_{x \in X} cf(x).$$

Together these results, we have $\sum_{x \in X} cf(x) = c \sum_{x \in X} f(x)$.

(c) If $X = X_1 \cup X_2$ for some disjoint sets X_1 and X_2 , then $\sum_{x \in X_1} f(x)$ and $\sum_{x \in X_2} f(x)$ are absolutely convergent, and

$$\sum_{x \in X_1 \cup X_2} f(x) = \sum_{x \in X_1} f(x) + \sum_{x \in X_2} f(x).$$

Conversely, if $h : X \rightarrow \mathbf{R}$ is such that $\sum_{x \in X_1} h(x)$ and $\sum_{x \in X_2} h(x)$ are absolutely convergent, then $\sum_{x \in X_1 \cup X_2} h(x)$ is also absolutely convergent, and

$$\sum_{x \in X_1 \cup X_2} h(x) = \sum_{x \in X_1} h(x) + \sum_{x \in X_2} h(x).$$

Prove the first part of (c). Let A_1, A_2 be a finite subset of X_1 and X_2 , and let $A = A_1 \cup A_2$ for some disjoint sets A_1 and A_2 . By Proposition 7.1.11 and Definition 8.2.4,

$$\sum_{x \in A_1} f(x) \leq \sum_{x \in A} f(x) \leq \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X \right\} < \infty,$$

$$\sum_{x \in A_2} f(x) \leq \sum_{x \in A} f(x) \leq \sup \left\{ \sum_{x \in A} |f(x)| : A \subseteq X \right\} < \infty.$$

Then,

$$\sum_{x \in A_1} f(x) \leq \sup \left\{ \sum_{x \in A_1} |f(x)| : A_1 \subseteq X_1 \right\} < \infty,$$

$$\sum_{x \in A_2} f(x) \leq \sup \left\{ \sum_{x \in A_2} |f(x)| : A_2 \subseteq X_1 \right\} < \infty.$$

Thus $\sum_{x \in X_1} f(x)$ and $\sum_{x \in X_2} f(x)$ are absolutely convergent.

When X is at most countable, equation immediately follows from Proposition 7.1.11(e) and Proposition 7.2.13. If X is uncountable, then

$$\begin{aligned} \sum_{x \in X_1} f(x) + \sum_{x \in X_2} f(x) &= \sum_{x \in X_1: f(x) \neq 0} f(x) + \sum_{x \in X_1: f(x) \neq 0} f(x) \\ &= \sum_{x \in X_1 \cup X_2: f(x) \neq 0} f(x) = \sum_{x \in X_1 \cup X_2} f(x). \end{aligned}$$

Now we prove the second part of (c). By Definition 8.2.4, we have

$$\sum_{x \in X_1} h(x) \leq \sup \left\{ \sum_{x \in A_1} |f(x)| : A_1 \subseteq X_1 \right\} < \infty,$$

$$\sum_{x \in X_2} h(x) \leq \sup \left\{ \sum_{x \in A_2} |f(x)| : A_2 \subseteq X_2 \right\} < \infty.$$

By Proposition 7.1.11(e), we have

$$\sum_{x \in X_1 \cup X_2} |h(x)| = \sum_{x \in X_1} |h(x)| + \sum_{x \in X_2} |h(x)|.$$

This means that

$$\begin{aligned} \sum_{x \in X_1 \cup X_2} |h(x)| &\leq \sup \left\{ \sum_{A_1 \cup A_2} |h(x)| : A_1 \cup A_2 \subseteq X_1 \cup X_2 \right\} \\ &= \sup \left\{ \sum_{x \in A_1} |f(x)| : A_1 \subseteq X_1 \right\} + \sup \left\{ \sum_{x \in A_2} |f(x)| : A_2 \subseteq X_2 \right\} < \infty. \end{aligned}$$

Thus $\sum_{x \in X_1 \cup X_2} h(x)$ is also absolutely convergent.

(d) If Y is another set, and $\phi : Y \rightarrow X$ is a bijection, then $\sum_{y \in Y} f(\phi(y))$ is absolutely convergent, and

$$\sum_{y \in Y} f(\phi(y)) = \sum_{x \in X} f(x).$$

(This result requires the axiom of choice when X is uncountable, see Section 8.4.)

Exercise 8.2.4. Prove Lemma 8.2.7. (Hint: prove by contradiction, and use limit laws.)

Lemma 8.2.7. Let $\sum_{n=0}^{\infty} a_n$ be a series of real numbers which is conditionally convergent, but not absolutely convergent. Define the sets $A_+ := \{n \in \mathbf{N} : a_n \geq 0\}$ and $A_- := \{n \in \mathbf{N} : a_n < 0\}$, thus $A_+ \cup A_- = \mathbf{N}$ and $A_+ \cap A_- = \emptyset$. Then both of the series $\sum_{n \in A_+} a_n$ and $\sum_{n \in A_-} a_n$ are not conditionally convergent (and thus not absolutely convergent).

Proof.

Exercise 8.2.5. Explain the gaps marked (why?) in the proof of Theorem 8.2.8.

Proof.

Exercise 8.2.6. Let $\sum_{n=0}^{\infty} a_n$ be a series which is conditionally convergent, but not absolutely convergent. Show that there exists a bijection $f : \mathbf{N} \rightarrow \mathbf{N}$

such that $\sum_{m=0}^{\infty} a_{f(m)}$ diverges to $+\infty$, or more precisely that

$$\lim_{N \rightarrow \infty} \inf \sum_{m=N}^{\infty} a_{f(m)} = \lim_{N \rightarrow \infty} \sup \sum_{m=N}^{\infty} a_{f(m)} = +\infty.$$

(Of course, a similar statement holds with $+\infty$ replaced by $-\infty$.)

Proof.

8.3

Uncountable sets

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 . (Hint: use induction on n .)

Proof. Let

$$2^X := \{A : A \subseteq X\}.$$

We use induction on n . When $n = 0$, we have $X = \emptyset$. There is only $A = \emptyset$ lies in 2^X ; hence $2^0 = 1$. Now we inductively suppose that X has cardinality with n and 2^X has cardinality with 2^n . Let $x_0 \notin X$, then $X \cup \{x_0\}$ has cardinality with $n + 1$. We can see that there is 2^n subsets that not contain $\{x_0\}$, and 2^n subsets contain $\{x_0\}$. Thus there is totally 2^{n+1} subsets of $X \cup \{x_0\}$. This close the induction. \square

Exercise 8.3.2. Let A, B, C be sets such that $A \subseteq B \subseteq C$, and suppose that there is a injection $f : C \rightarrow A$. Define the sets D_0, D_1, D_2, \dots recursively by setting $D_0 := B \setminus A$, and then $D_{n+1} := f(D_n)$ for all natural numbers n . Prove that the sets D_0, D_1, \dots are all disjoint from each other (i.e., $D_n \cap D_m = \emptyset$ whenever $n \neq m$). Also show that if $g : A \rightarrow B$ is the function defined by setting $g(x) := f^{-1}(x)$ when $x \in \bigcup_{n=1}^{\infty} D_n$, and $g(x) := x$ when $x \notin \bigcup_{n=1}^{\infty} D_n$, then g does indeed map A to B and is a bijection between the two. In particular, A and B have the same cardinality.

Proof. First of all, we show that for all n , $f(D_n) \subseteq A$. Since f is an injection, we have $f(C) \subseteq A$, and for every subset S of C , $f(S) \subseteq A$. For

the base case we have $D_0 \subseteq C$, so that $D_1 = f(D_0) \subseteq A$. Since $A \subseteq C$, we have $D_2 = f(f(D_0)) \subseteq A$. It is easy to see that for all $n \in \mathbf{N}$, we have $D_{n+1} = f(D_n) \subseteq A$. Thus for all n , $D_0 \cap D_{n+1} = \emptyset$. By Exercise 3.4.3, since f is injective, $S \cap U = \emptyset$ implies that $f(S) \cap f(U) = \emptyset$. Then we iterate the process and obtain that

$$\begin{aligned} f(f(S) \cap f(U)) &= f(f(S)) \cap f(f(U)) = \emptyset, \\ f(f(f(S) \cap f(U))) &= f(f(f(S))) \cap f(f(f(U))) = \emptyset, \\ &\dots \\ \underbrace{(f \circ \dots \circ f)}_n(f(S) \cap f(U)) &= \underbrace{(f \circ \dots \circ f)}_{n+1}(S) \cap \underbrace{(f \circ \dots \circ f)}_{n+1}(U) = \emptyset. \end{aligned}$$

Base on this, we have $D_n \cap D_m = \emptyset$ for all $n \neq m$ where we can let $m = n+k$ for some $k \in \mathbf{N}$.

We restrict function f from $\bigcup_{n=0}^{\infty} D_n$ to $f(\bigcup_{n=0}^{\infty} D_n)$, we denote it by $f|_{\bigcup_{n=0}^{\infty} D_n}$, which is obviously an bijection. And by definition, $\bigcup_{n=1}^{\infty} D_n = f(\bigcup_{n=0}^{\infty} D_n)$. Thus there is a bijection $f^{-1} : \bigcup_{n=1}^{\infty} D_n \rightarrow \bigcup_{n=0}^{\infty} D_n$ for all $x \in \bigcup_{n=1}^{\infty} D_n$. So that g is bijection for all $x \in \bigcup_{n=1}^{\infty} D_n$. Consider the case of $x \notin \bigcup_{n=1}^{\infty} D_n$. We must have $x \in A \setminus \bigcup_{n=1}^{\infty} D_n$, such a set is non-empty for that $\bigcup_{n=1}^{\infty} D_n \subseteq A$. Since

$$B \setminus \bigcup_{n=0}^{\infty} D_n = \left((B \setminus A) \cup A \right) \setminus \left((B \setminus A) \cup \bigcup_{n=1}^{\infty} D_n \right) = A \setminus \bigcup_{n=1}^{\infty} D_n.$$

Then $A \setminus \bigcup_{n=1}^{\infty} D_n \subseteq B \setminus f(\bigcup_{n=1}^{\infty} D_n)$ for that $\bigcup_{n=2}^{\infty} D_n = f(\bigcup_{n=1}^{\infty} D_n) \subseteq \bigcup_{n=1}^{\infty} D_n$. By Exercise 3.3.8, $g : A \setminus \bigcup_{n=1}^{\infty} D_n \rightarrow B \setminus f(\bigcup_{n=1}^{\infty} D_n)$ is identity. Hence g also is bijective for all $x \notin \bigcup_{n=1}^{\infty} D_n$. Thus g is a bijection. In particular, A and B have the same cardinality. \square

Remark 8.3.2. We rewrite Exercise 8.3.2 more straightforward, this exercise constructed a lemma which we can be used to prove the Schröder-Bernstein theorem (Exercise 8.3.3). Lemma is following below:

Lemma. Let there be injection $f : B \rightarrow A$ from a set B into its subset $A \subseteq B$. Then there exists a bijection between B and A . This means that A and B has equal cardinality.

Exercise 8.3.3. Recall from Exercise 3.6.7 that a set A is said to have lesser or equal cardinality than a set B iff there is an injective map $f : A \rightarrow B$ from A to B . Using Exercise 8.3.2, show that if A, B are sets such that A has lesser or equal cardinality to B and B has lesser or equal cardinality to A , then A and B have equal cardinality. (This is known as the **Schröder-Bernstein theorem** after Ernst Schröder (1841-1902) and Felix Bernstein (1878-1956).)

Proof. Let $f : A \rightarrow B$ and $g : B \rightarrow A$ to be injections. We have bijection $g : B \rightarrow g(B)$, so that injection $g \circ f : A \rightarrow g(B)$. Since $g(B) \subseteq A$, by Exercise 8.3.2, there is a bijection $g \circ f : A \rightarrow g(B)$. Then we have bijection $g^{-1} \circ (g \circ f) : A \rightarrow B$. Hence A and B has equal cardinality. \square

Exercise 8.3.4. Let us say that a set A has strictly lesser cardinality than a set B if A has lesser than or equal cardinality to B (in the sense of Exercise 3.6.7) but A does not have equal cardinality to B . Show that for any set X , that X has strictly lesser cardinality than 2^X . Also, show that if A has strictly lesser cardinality than B , and B has strictly lesser cardinality than C , then A has strictly lesser cardinality than C .

Proof. Consider that X to be a singleton set, where $x \in X$. Then $f(x) = \{x\}$ is a bijection from X to the subset of 2^X and X has lesser or equal cardinality to 2^X . By Theorem 8.3.1, equality is not hold for X . Thus X has strictly lesser cardinality than 2^X .

For the second assertion. Let $f : A \rightarrow B$ and $g : B \rightarrow C$ to be two injections, and $\#(A) < \#(B)$ and $\#(B) < \#(C)$. By Exercise 3.3.2, $g \circ f : A \rightarrow C$ is also an injection, and we have $\#(A) \leq \#(C)$. Suppose that $\#(A) = \#(C)$, then we have $\#(B) < \#(A)$, which contradict $\#(A) < \#(B)$. Thus $\#(A) < \#(C)$ and A has strictly lesser cardinality than C . \square

Exercise 8.3.5. Show that no power set (i.e., a set of the form 2^X for some set X) can be countably infinite.

Proof. If X is finite, by Exercise 8.3.1, 2^X also finite. While if X is count-

able. By definition, X has same cardinality with \mathbf{N} , then $\#(2^X) = \#(2^{\mathbf{N}})$. But by Corollary 8.3.3, $2^{\mathbf{N}}$ is uncountable. \square

8.4

The axiom of choice

Exercise 8.4.1. *Show that the axiom of choice implies Proposition 8.4.7. (Hint: consider the sets $Y_x := \{y \in Y : P(x, y) \text{ is true}\}$ for each $x \in X$.) Conversely, show that if Proposition 8.4.7 is true. then the axiom of choice is also true.*

Proposition 8.4.7. Let X and Y be sets, and let $P(x, y)$ be a property pertaining to an object $x \in X$ and an object $y \in Y$ such that for every $x \in X$ there is at least one $y \in Y$ such that $P(x, y)$ as true. Then there exists a function $f : X \rightarrow Y$ such that $P(x, f(x))$ is true for all $x \in X$.

Proof. Let X be a set, and $Y_x := \{y \in Y : P(x, y) \text{ is true}\}$ is non-empty for each $x \in X$. In other words, $P(x, y)$ is a property pertaining to an object $x \in X$ and an object $y \in Y$ such that for every $x \in X$ there is at least one $y \in Y$ such that $P(x, y)$ as true. By axiom of choice, $\prod_{x \in X} Y_x$ is non-empty, i.e., there is a function $f : X \rightarrow \bigcup_{x \in X} Y_x$ such that $P(x, f(x))$ is true for all $x \in X$. Since $\bigcup_{x \in X} Y_x \subseteq Y$, $f : X \rightarrow Y$ such that $P(x, f(x))$ is also true for all $x \in X$.

For the converse, let I be a set, and for each $\alpha \in I$, let X_α be a non-empty set. Then for every $\alpha \in I$ there is at least one $x_\alpha \in \bigcup_{\beta \in I} X_\beta$ such that $x_\alpha \in X_\alpha$ for all $\alpha \in I$. By Proposition 8.4.7, there exists a function $(x_\alpha)_{\alpha \in I} \in (\prod_{\beta \in I} X_\beta)^I$ such that $x_\alpha \in X_\alpha$ for all $\alpha \in I$. \square

Exercise 8.4.2. *Let I be a set, and for each $\alpha \in I$ let X_α be a non-empty set. Suppose that all the sets X_α are disjoint from each other, i.e., $X_\alpha \cap X_\beta = \emptyset$ for all distinct $\alpha, \beta \in I$. Using the axiom of choice, show that there exists a set Y such that $\#(Y \cap X_\alpha) = 1$ for all $\alpha \in I$ (i.e., Y intersects each X_α in exactly one element). Conversely, show that if the*

above statement was true for an arbitrary choice of sets I and non-empty disjoint sets X_α , then the axiom of choice is true. (Hint: the problem is that in Axiom 8.1 the sets X_α are not assumed to be disjoint. But this can be fixed by the trick by looking at the sets $\{\alpha\} \times X_\alpha = \{(\alpha, x) : x \in X_\alpha\}$ instead.)

Proof. Suppose that $X_\alpha \cap X_\beta = \emptyset$ for all distinct $\alpha, \beta \in I$. And by axiom of choice, there is a function $(x_\alpha)_{\alpha \in I}$ such that $x_\alpha \in X_\alpha$ for all $\alpha \in I$. Let $Y := \{x_\alpha : \alpha \in I\}$. Then $Y \cap X_\alpha = x_\alpha$, so that $\#(Y \cap X_\alpha) = 1$, for all $\alpha \in I$.

Conversely, let $\{\alpha\} \times X_\alpha = \{(\alpha, x) : x \in X_\alpha\}$. Suppose above statement is hold. Recall Exercise 3.5.10, we need to show that there exists a function $(x_\alpha)_{\alpha \in I}$ whose graph is equal to $\{\alpha\} \times X_\alpha$. This means that $(x_\alpha)_{\alpha \in I}$ assigns to each $\alpha \in I$ an element $x_\alpha \in X_\alpha$. We can see that $\{\alpha\} \times X_\alpha \subseteq I \times \bigcup_{\beta \in I} X_\beta$ such that for each $\alpha \in I$, there is one $x \in X_\alpha$ such that $(\alpha, x_\alpha) \in \{\alpha\} \times X_\alpha$. Otherwise, we can find another $x' \in X_\alpha$ such that $\#(Y \cap (\{\alpha\} \times X_\alpha)) = 2$ where $Y := \{x, x'\}$, a contradiction. We denote such unique x as x_α . Then by Exercise 3.5.10, there is exactly one function, denoted by $(x_\alpha)_{\alpha \in I}$, whose graph is $\{\alpha\} \times X_\alpha$ for each $\alpha \in I$. This proved the existence of function $(x_\alpha)_{\alpha \in I} \in (\bigcup_{\beta \in I} X_\beta)^I$. \square

Exercise 8.4.3. Let A and B be sets such that there exists a surjection $g : B \rightarrow A$. Using the axiom of choice, show that there then exists an injection $f : A \rightarrow B$; in other words A has lesser or equal cardinality to B in the sense of Exercise 3.6.7. (Hint: consider the inverse images $g^{-1}(\{a\})$ for each $a \in A$.) Compare this with Exercise 3.6.8. Conversely, show that if the above statement is true for arbitrary sets A, B and surjections $g : B \rightarrow A$, then the axiom of choice is true. (Hint: use Exercise 8.4.2.)

Proof.

8.5

Ordered sets

Exercise 8.5.1. Consider the empty set \emptyset with the empty order relation \leq_\emptyset this relation is vacuous because the empty set has no elements). Is this set partially ordered? totally ordered? well-ordered? Explain.

Proof. The empty set is partially ordered, totally ordered and well-ordered, for that all of definitions are vacuously true. \square

Exercise 8.5.2. Give examples of a set X and a relation \leq such that

- (a) The relation \leq is reflexive and anti-symmetric, but not transitive;
- (b) The relation \leq is reflexive and transitive, but not anti-symmetric;
- (c) The relation \leq is anti-symmetric and transitive, but not reflexive.

Proof. Example of (a). Suppose that

$$\leq_X := \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 3)\},$$

¹For the purposes of my book, the only relevant thing about a relation is what it does, which is described in the parenthetical comment in Definition 8.5.1: if x and y are in X , then $x \leq_X y$ is either a true or a false statement. One can model a relation, as you say, as a function from $X \times X$ to a two-valued set such as $\{0, 1\}$, or $\{\text{true}, \text{false}\}$, or one can also model a relation as a subset of $X \times X$ as is given in the Wikipedia definition. (The connection between a relation \leq_X , and a model of that relation, is analogous to the relationship between a function f , and the graph of that function.)

However, as this text does not study the theory of relations per se, we do not need to fix precisely what set-theoretic model one would use to model the relation concept, and can instead proceed on an ad hoc basis, so that whenever one actually needs to apply, say, Zorn's lemma, one verifies that the relation supplied by the application (e.g. subset inclusion) obeys the relevant properties required of a relation (in this case, that one can attempt to compare any two elements in the domain). The situation is similar to that of ordered pairs, discussed in Section 3.5, where again the exact set-theoretic model of an ordered pair is not particularly relevant for applications.

and

$$X := \{1, 2, 3\}.$$

Then by definition, \leq_X is reflexive and anti-symmetric on X , but not transitive for that $(1, 2) \in \leq_X$ and $(2, 3) \in \leq_X$ but $(1, 3) \notin \leq_X$.

Example of (b). Suppose that

$$\leq_X := \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 3), (1, 3), (2, 1)\},$$

and

$$X := \{1, 2, 3\}.$$

Then by definition, \leq_X is reflexive and transitive on X , but not anti-symmetric for that $(1, 2) \in \leq_X$ and $(2, 1) \in \leq_X$ but $1 \neq 2$.

Example of (c). Suppose that

$$\leq_X := \{(1, 2), (2, 3), (1, 3), (2, 1), (3, 2), (3, 1)\},$$

and

$$X := \{1, 2, 3\}.$$

Then by definition, \leq_X is anti-symmetric and transitive on X , but not reflexive for that $(1, 1) \notin \leq_X$. \square

Exercise 8.5.3. Given two positive integers $n, m \in \mathbf{N} \setminus \{0\}$, 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 $\mathbf{N} \setminus \{0\}$ with the ordering relation $|$ is a partially ordered set but not a totally ordered one. Note that this is a different ordering relation from the usual \leq ordering of $\mathbf{N} \setminus \{0\}$.

Proof. For any $n \in \mathbf{N} \setminus \{0\}$, we have $n = n \times 1$, hence $|$ is reflexive. If $n, m \in \mathbf{N} \setminus \{0\}$ are such that $n|m$ and $m|n$, i.e., there exists positive integers a, b such that $m = na$ and $n = mb$, then $m = mab$. Thus $ab = 1$. Since a, b are integers, we have $a = b = 1$ and $m = n$. Thus $|$ is anti-symmetric. If $n, m, k \in \mathbf{N} \setminus \{0\}$ are such that $n|m$ and $m|k$, then there are integers a, b such that $m = na$ and $k = mb$. Then we have $k = nab$, where ab is an integer, hence $n|k$. Thus $|$ is transitive. Together our conclusion, $|$ is a partially ordered set. But $|$ is not a totally ordered set for that there don't

exist a integer such that $3 = 5a$ or $5 = 3a$, which means both $3|5$ and $5|3$ is not true. \square

Exercise 8.5.4. *Show that the set of positive reals $\mathbf{R}^+ := \{x \in \mathbf{R} : x > 0\}$ have no minimal element.*

Proof. Suppose that y is a minimal element of \mathbf{R}^+ . In particular, we can let $y < 1$. By definition, there is no $y' \in \mathbf{R}^+$ such that $y' < y$. Since we can find a rational number $q > 1$ such that $0 < y^q < y$ which also is a real, let $y' = y^q$, a contradiction. \square

Exercise 8.5.5. *Let $f : X \rightarrow Y$ be a function from one set X to another set Y . Suppose that Y is partially ordered with some ordering relation \leq_Y . Define a relation \leq_X on X by defining $x \leq_X x'$ if and only if $f(x) <_Y f(x')$ or $x = x'$. Show that this relation \leq_X turns X into a partially ordered set. If we know in addition that the relation \leq_Y makes Y totally ordered, does this mean that the relation \leq_X makes X totally ordered also? If not, what additional assumption needs to be made on f in order to ensure that \leq_X makes X totally ordered?*

Proof. Since $x = x$, we have $x \leq_X x$ for all $x \in X$, hence \leq_X is reflexive.

If $x, y \in X$ are such that $x \leq_X y$ and $y \leq_X x$, then we have $f(x) <_Y f(y)$ or $x = y$, and $f(y) <_Y f(x)$ or $y = x$. We can see that 1) $f(x) <_Y f(y)$ and $y = x$; 2) $x = y$ and $f(y) <_Y f(x)$; 3) $x = y$ and $y = x$ can imply that $x = y$. Consider $f(x) <_Y f(y)$ and $f(y) <_Y f(x)$ hold at the time, this implies that $f(x) \leq_Y f(y)$, $f(y) \leq_Y f(x)$ and $f(x) \neq f(y)$ hold at the time, by anti-symmetry of \leq_Y , a contradiction (Notice that this contradiction means that $x \leq_X y$ and $y \leq_X x$ would not implies that $f(x) <_Y f(y)$ and $f(y) <_Y f(x)$ hold at the time, otherwise it would contradict $x \leq_X y$ and $y \leq_X x$). Thus, \leq_X is anti-symmetric.

If $x, y, z \in X$ are such that $x \leq_X y$ and $y \leq_X z$, this means that $f(x) <_Y f(y)$ or $x = y$, and $f(y) <_Y f(z)$ or $y = z$. Because \leq_Y is partially ordered, we have $f(x) <_Y f(z)$ or $x = z$, which means $x \leq_X z$. Hence \leq_X is transitive.

If \leq_Y is totally ordered set can not implies that \leq_X is totally ordered set. While if f is injection, for any $x, x' \in X$, $x \neq x'$ implies that $f(x) \neq f(x')$. Thus at least one of $f(x) <_Y f(x')$ and $f(x') <_Y f(x)$, so that $x \leq_X x'$ or $x' \leq_X x$ is true. While if $f(x) \leq_Y f(x')$ and $f(x') \leq_Y f(x)$ is true at the time, then $f(x) = f(x')$ implies that $x = x'$, so that both $x \leq_X x'$ and $x' \leq_X x$ is true at the time. \square

Exercise 8.5.6. Let X be a partially ordered set. For any x in X , define the **order ideal** $(x) \subset X$ to be the set $(x) := \{y \in X : y \leq x\}$. Let $(X) := \{(x) : x \in X\}$ be the set of all order ideals, and let $f : X \rightarrow (X)$ be the map $f(x) := (x)$ that sends every element of x to its order ideal. Show that f is a bijection, and that given any $x, y \in X$, that $x \leq_X y$ if and only if $f(x) \subseteq f(y)$. This exercise shows that any partially ordered set can be represented by a collection of sets whose ordering relation is given by set inclusion.

Proof. First of all, we show that f is injective. Suppose that for all $(x), (x') \in (X)$ we have $(x) = (x')$. Because $x \in (x), x' \in (x')$ and $(x) = (x')$, we have $x \leq x'$ and $x' \leq x$, so that $x = x'$. Thus f is injective.

For the surjection, for every $(x) \in (X)$, there exists $f(x) \in (X)$ such that $f(x) = \{y \in X : y \leq x\} = (x)$. Thus f is surjective. So that f is an injection.

Suppose that $x \leq_X y$ and $z \in X$. Since \leq_X is partially ordered set, for every $z \leq_X x$ which means $x \in f(x)$, we have $z \leq_X y$, i.e., $z \in f(y)$. Hence $f(x) \subseteq f(y)$. Conversely, suppose that $f(x) \subseteq f(y)$, i.e., for every $z \in f(x)$ implies that $z \in f(y)$. By definition, we have $z \leq_X x$ implies $z \leq_X y$. Let $z = x$, we must have $x \leq_X y$, then $x \leq_X y$. \square

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.

Proof. Let y be a minimal element of Y , suppose there is another minimal

element y' of Y where $y \neq y'$. By Definition 8.5.5, there is no element $x \in Y$ such that $x <_X y$ and $x <_X y'$. Since Y is totally ordered set, we either have $y <_X y'$ or $y' <_X y$. In former case, it contradict $x <_X y'$, and in latter case, it contradict $x <_X y$; a contradiction.

Other hand, let y be a maximal element of Y , suppose there is another maximal element y' of Y where $y \neq y'$. By Definition 8.5.5, there is no element $x \in Y$ such that $y <_X x$ and $y' <_X x$. Since Y is totally ordered set, we either have $y <_X y'$ or $y' <_X y$. In former case, it contradict $x <_X y'$, and in latter case, it contradict $x <_X y$; a contradiction. \square

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.*

Proof. Let X to be a finite non-empty subset of a totally ordered set Y . We use induction to show that every finite non-empty subset of a totally ordered set has a minimum. Consider that $X = \{x\}$, i.e., $\#(X) = 1$. Then x is a minimal element of X . Now we inductively suppose that $X \subseteq Y$ has cardinality with n and $x \in X$ to be the minimal element of X . We pick $x_0 \in Y$ but $x_0 \notin X$. And we have $X \cup \{x_0\}$ has cardinality $n + 1$. Since x is the minimal element of X , we compare x and x_0 , then there is no element $x' \in X \cup \{x_0\}$ such that $x' < \min\{x, x_0\}$. This close the induction. This is similar to show the case of maximum. In particular, every finite totally ordered set is well-defined for which is the subset of itself. \square

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. (Hint: assume for sake of contradiction that X is infinite. Start with the minimal element x_0 of X and then construct an increasing sequence $x_0 < x_1 < \dots$ in X .)*

Proof. Suppose for sake of contradiction that X is infinite, then there is a bijection $f : \mathbb{N} \rightarrow X$. Since X has both a minimum and a maximum, we

define f as following rule:

$$\begin{aligned} f(0) &:= \min(X), \\ f(n) &:= \min(X \setminus f(\{1, \dots, n-1\})). \end{aligned}$$

for all $n \in \mathbf{N}$. We use x_n to represent $f(n)$. Now we have $\{x_n \in X : n \in \mathbf{N}\}$ which is a infinite subset of X but hasn't a maximum. \square

Exercise 8.5.10. *Prove Proposition 8.5.10, without using the axiom of choice (Hint: consider the set*

$$Y := \{n \in X : P(m) \text{ is false for some } n \in X \text{ with } m \leq_X n\},$$

and show that Y being non-empty would lead to a contradiction.

Proposition 8.5.10 (Principle of strong induction). Let X be a well-ordered set with an ordering relation \leq , and let $P(n)$ be a property pertaining to an element $n \in X$ (i.e., for each $n \in X$, $P(n)$ is either a true statement or a false statement). Suppose that for every $n \in X$, we have the following implication: if $P(m)$ is true for all $m \in X$ with $m <_X n$, then $P(n)$ is also true. Prove that $P(n)$ is true for all $n \in X$.

CHAPTER IX

Continuous functions on \mathbb{R}

9.1

Subsets of the real line

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 $\overline{Y} = \overline{X}$.*

Proof. First let us show that every adherent point x of X is adherent to Y , i.e., $\overline{X} \subseteq \overline{Y}$. Let $x \in \overline{X}$ and let $\varepsilon > 0$. By Definition 9.1.5, for every $x \in \overline{X}$, there exists an $x' \in X$ such that $|x - x'| \leq \varepsilon$. Since $X \subseteq Y$, there is an $x' \in Y$ such that $|x - x'| \leq \varepsilon$ for every $x \in \overline{X}$. Thus every adherent point x of X lies in \overline{Y} .

Then we show that every adherent point y of Y is adherent to X , i.e., $\overline{Y} \subseteq \overline{X}$. By Definition 9.1.5, for every $y \in \overline{Y}$, there is a $y' \in Y$ such that $|y - y'| \leq \varepsilon$. Since $Y \subseteq \overline{X}$, there is a $y' \in \overline{X}$ such that $|y - y'| \leq \varepsilon$ for all $y \in \overline{Y}$. Thus every adherent point y of Y is adherent to \overline{X} . By Definition 9.1.10, every adherent point of \overline{X} is adherent point of X , i.e., $\overline{\overline{X}} = \overline{X}$. Therefore, every adherent point y of Y is adherent to X . \square

Exercise 9.1.2. *Prove Lemma 9.1.11.*

Lemma 9.1.11 (Elementary properties of closures). Let X and Y be arbitrary subsets of \mathbf{R} . Then $X \subseteq \overline{X}$, $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$, and $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$. If $X \subseteq Y$, then $\overline{X} \subseteq \overline{Y}$.

Proof. Prove $X \subseteq \overline{X}$. By Definition 9.1.10, every element of X is definitely adherent to itself. Thus $X \subseteq \overline{X}$ is hold.

Prove $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$. Use the conclusion above, we have $X \subseteq \overline{X}$ and $Y \subseteq \overline{Y}$, then $X \cup Y \subseteq \overline{X} \cup \overline{Y}$. Also, we have $X \cup Y \subseteq \overline{X \cup Y}$. First let us show that every element of $\overline{X} \cup \overline{Y}$ is adherent to $X \cup Y$. Let $x \in \overline{X} \cup \overline{Y}$. If $x \in \overline{X}$, then it is adherent to X , so that is adherent to $X \cup Y$. While if $x \in \overline{Y}$, then it is adherent to Y , so that is adherent to $X \cup Y$. Thus every $x \in \overline{X} \cup \overline{Y}$ is adherent to $X \cup Y$, and we have $\overline{X} \cup \overline{Y} \subseteq \overline{X \cup Y}$. Conversely, we need to show that every adherent point of $X \cup Y$ lies in $\overline{X} \cup \overline{Y}$. Suppose for sake of contradiction that $x \notin \overline{X} \cup \overline{Y}$. Then $x \notin \overline{X}$ and $x \notin \overline{Y}$, this means that x neither an adherent point of X nor an adherent point of Y , a contradiction. Thus $\overline{X \cup Y} \subseteq \overline{X} \cup \overline{Y}$.

Prove $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$. By Definition 9.1.10, for every $x \in \overline{X \cap Y}$, we have $\overline{X \cap Y} \subseteq \overline{X}$ and $\overline{X \cap Y} \subseteq \overline{Y}$. This implies that $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$.

Prove that if $X \subseteq Y$, then $\overline{X} \subseteq \overline{Y}$. By Definition 9.1.10, for every $x \in \overline{X}$, there is an $x' \in X$ such that $|x - x'| \leq \varepsilon$. Since $X \subseteq Y$, there is $x' \in Y$ such that $|x - x'| \leq \varepsilon$ for every $x \in \overline{X}$. This means that every adherent point of X is adherent to Y , i.e., $\overline{X} \subseteq \overline{Y}$. \square

Remark 9.1.2. Let $X = (0, 1)$, $Y = (1, 2)$. Then $X \cap Y = \emptyset \implies \overline{X \cap Y} = \emptyset$. But $\overline{X} \cap \overline{Y} = \{1\}$. We must have $\overline{X \cap Y} \subsetneq \overline{X} \cap \overline{Y}$.

Exercise 9.1.3. Prove Lemma 9.1.13. (Hint: for computing the closure of \mathbf{Q} you will need Proposition 5.4.14.)

Lemma 9.1.13. The closure of \mathbf{N} is \mathbf{N} . The closure of \mathbf{Z} is \mathbf{Z} . The closure of \mathbf{Q} is \mathbf{R} , and the closure of \mathbf{R} is \mathbf{R} . The closure of the empty set \emptyset is \emptyset .

Proof. The closure of \mathbf{N} is \mathbf{N} . Let $\overline{\mathbf{N}}$ be the closure of \mathbf{N} . $\mathbf{N} \subseteq \overline{\mathbf{N}}$ follows immediately from Lemma 9.1.11. We prove the converse, suppose for sake

of contradiction that there is an $x \in \overline{\mathbf{N}}$ but $x \notin \mathbf{N}$. Since $x \notin \mathbf{N}$, There is an $m \in \mathbf{N}$ such that $m < x < m + 1$. Let $\varepsilon = \min(x - m, m - x + 1)/2$. Then for every natural number $n \in \mathbf{N}$, we have $|x - n| > \varepsilon$ (because for every natural number n , we have either $n \leq m$ or $n \geq m + 1$, and this is easy to verify that in both case, the assertion is hold). This means that $x \notin \overline{\mathbf{N}}$, a contradiction.

The closure of \mathbf{Z} is \mathbf{Z} . Let $n \in \mathbf{N}$, and define $-\mathbf{N}$ be the set of all negative natural number $-n$. Similarly, we can find an $x \in -\overline{\mathbf{N}}$ which is the closure of $-\mathbf{N}$, but $x \notin -\mathbf{N}$. Then there exists an $M \in \mathbf{N}$ such that $-m - 1 < x < -m$. Let $\varepsilon = \min(-x - m, x + m + 1)/2$. Then for every natural number $n \in \mathbf{N}$, we have $|x - (-n)| > \varepsilon$. This means that $x \notin -\overline{\mathbf{N}}$, a contradiction. So $-\mathbf{N} = -\overline{\mathbf{N}}$. By Lemma 9.1.11, we have $\mathbf{Z} = \mathbf{N} \cup (-\mathbf{N}) = \overline{\mathbf{N} \cup (-\mathbf{N})} = \overline{\mathbf{Z}}$.

The closure of \mathbf{Q} is \mathbf{R} . Let $\overline{\mathbf{Q}}$ be the closure of \mathbf{Q} . Suppose that there exists an $x \in \mathbf{R}$ but $x \notin \overline{\mathbf{Q}}$. By Proposition 5.4.14, for every real number $x \in \mathbf{R}$, there exists a $q \in \mathbf{Q}$ such that $x < q < x + \varepsilon$, then $0 < q - x < \varepsilon$. Therefore, for every $x \in \mathbf{R}$ there exists a $q \in \mathbf{Q}$ such that $|q - x| < \varepsilon$, which means $x \in \overline{\mathbf{Q}}$, a contradiction. Since $\mathbf{Q} \subseteq \mathbf{R}$, every adherent point of \mathbf{Q} is adherent to \mathbf{R} . Thus $\overline{\mathbf{Q}} \subseteq \mathbf{R}$. Since $\overline{\mathbf{Q}} = \mathbf{R}$, closure of \mathbf{Q} is \mathbf{R} .

The closure of \mathbf{R} is \mathbf{R} . We have known that $\mathbf{Q} \subseteq \mathbf{R} = \overline{\mathbf{Q}}$, by Exercise 9.1.1, we have $\overline{\mathbf{R}} = \overline{\mathbf{Q}}$. So that $\overline{\mathbf{R}} = \mathbf{R}$.

The closure of \emptyset is \emptyset . Let $\overline{\emptyset}$ be the closure of \emptyset . Suppose $x \notin \emptyset$ but $x \in \overline{\emptyset}$. Since $x \in \overline{\emptyset}$, there is a $y \in \emptyset$ such that $|x - y| \leq \varepsilon$. But this is impossible, there is no such an $x \in \overline{\emptyset}$. And since we have $\emptyset \subseteq \overline{\emptyset}$. So $\overline{\emptyset} = \emptyset$. \square

Exercise 9.1.4. Give an example of two subsets X, Y of the real line such that $\overline{X \cap Y} \neq \overline{X} \cap \overline{Y}$.

Proof. Let $X = (0, 1)$, $Y = (1, 2)$. Then $X \cap Y = \emptyset \implies \overline{X \cap Y} = \emptyset$. But $\overline{X} \cap \overline{Y} = \{1\}$. We must have $\overline{X \cap Y} \subseteq \overline{X} \cap \overline{Y}$ and $\overline{X \cap Y} \neq \overline{X} \cap \overline{Y}$. \square

Exercise 9.1.5. Prove Lemma 9.1.14. (Hint: in order to prove one of the

two implications here you will need axiom of choice, as in Lemma 8.4.5.)

Lemma 9.1.14. Let X be a subset of \mathbf{R} , and let $x \in \mathbf{R}$. Then x is an adherent point of X if and only if there exists a sequence $(a_n)_{n=0}^\infty$, consisting entirely of elements in X , which converges to x .

Proof. Suppose that x is an adherent point of X . Then for every $n \in \mathbf{N}$, let $\varepsilon = 1/n$, the set

$$X_n := \left\{ y \in X : |y - x| \leq \frac{1}{n} \right\}$$

is non-empty for that the set at least contains x . By choice axiom, we can find $a_n \in X_n$ for every $n \in \mathbf{N}$. Then the sequence $(a_n)_{n=0}^\infty$ in X satisfies that $|a_n - x| \leq 1/n$. We can see that $\lim_{n \rightarrow \infty} a_n = x$. Conversely, if there is a sequence $(a_n)_{n=0}^\infty$ in X and converges to x . Then there is an $N \geq n$ such that $|a_N - x| \leq \varepsilon$ for all $n \in \mathbf{N}$. By definition, x is an adherent point. \square

Exercise 9.1.6. Let X be a subset of \mathbf{R} . Show that \overline{X} is closed (i.e., $\overline{\overline{X}} = \overline{X}$). Furthermore, show that if Y is any closed set that contains X , then Y also contains \overline{X} . Thus the closure \overline{X} of X is the smallest closed set which contains X .

Proof. By Lemma 9.1.11, we have $X \subseteq \overline{X} = \overline{\overline{X}}$. By Exercise 9.1.1, we have $\overline{\overline{X}} = \overline{X}$. If $X \subseteq Y = \overline{Y}$. By Lemma 9.1.1, we have $\overline{X} \subseteq \overline{Y}$, so that $\overline{X} \subseteq Y$. \square

Exercise 9.1.7. Let $n \geq 1$ be a positive integer, and let X_1, \dots, X_n be closed subsets of \mathbf{R} . Show that $X_1 \cup X_2 \cup \dots \cup X_n$ is also closed.

Proof. We use induction on n . When $n = 1$, We have $X_1 = \overline{X_1}$. Now we inductively suppose that $X_1 \cup X_2 \cup \dots \cup X_n$ is closed for n . Let X_{n+1} be closed, then we have

$$(X_1 \cup X_2 \cup \dots \cup X_n) \cup X_{n+1} = \overline{(X_1 \cup X_2 \cup \dots \cup X_n)} \cup \overline{X_{n+1}}.$$

By Lemma 9.1.11,

$$X_1 \cup X_2 \cup \dots \cup X_n \cup X_{n+1} = \overline{X_1 \cup X_2 \cup \dots \cup X_n \cup X_{n+1}}.$$

This close the induction. \square

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 \mathbf{R} . Show that the intersection $\bigcap_{\alpha \in I} X_\alpha$ (defined in (3.3)) is also closed.

Proof. Since X_α is closed, we have $X_\alpha = \overline{X_\alpha}$. By Lemma 9.1.11, we have $\bigcap_{\alpha \in I} X_\alpha \subseteq \overline{\bigcap_{\alpha \in I} X_\alpha}$. We only need to show that every element of $\overline{\bigcap_{\alpha \in I} X_\alpha}$ is an element of $\bigcap_{\alpha \in I} X_\alpha$. Let $x \in \overline{\bigcap_{\alpha \in I} X_\alpha}$ and $\varepsilon > 0$. Therefore, there is an $x' \in \bigcap_{\alpha \in I} X_\alpha$ such that $|x - x'| \leq \varepsilon$ for all $x \in \overline{\bigcap_{\alpha \in I} X_\alpha}$. Then for all $\alpha \in I$, there is an $x \in \overline{X_\alpha}$ such that $|x - x'| \leq \varepsilon$. Since $\overline{X_\alpha} = X_\alpha$, we have $x \in \bigcap_{\alpha \in I} X_\alpha$. \square

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 .

Proof. Let x be an adherent point of X . By Definition 9.1.18, x is either limit point or isolated point of X . Let $\varepsilon > 0$. Suppose that x is limit point of X , then it is an adherent point of $X \setminus \{x\}$. This means there exists a $y \in X \setminus \{x\}$ such that $|x - y| \leq \varepsilon$. Thus x is not an isolated point. While if x is an isolated point, then for $x \in X$ there is some $\varepsilon > 0$ such that $|x - y| > \varepsilon$ for all $y \in X \setminus \{x\}$. So x is not a limit point.

Conversely, if x is a limit point, then it is an adherent point of $X \setminus \{x\} \subsetneq X$. So it also an adherent point of X . While if x is an isolated point, then it of course an adherent point of X by definition. \square

Exercise 9.1.10. If X is a non-empty subset of \mathbf{R} , show that X is bounded if and only if $\inf(X)$ and $\sup(X)$ are finite.

Proof. Suppose that $\inf(X)$ and $\sup(X)$ are finite. Let

$$M := \max(|\inf(X)|, |\sup(X)|),$$

we have $X \subseteq [-M, M]$ is bounded. Conversely, assume that X is bounded. There exists an M such that $X \subseteq [-M, M]$, which means for every $x \in X$ we have $-M \leq x \leq M$. Thus $-M \leq \inf(X) \leq \sup(X) \leq M$, and $\inf(X)$

and $\sup(X)$ are finite. \square

Exercise 9.1.11. *Show that if X is a bounded subset of \mathbf{R} , then the closure \bar{X} is also bounded.*

Proof. Since X is bounded, for every $y \in X$ there is an M such that $|y| \leq M$. Let $\varepsilon > 0$. For all adherent point x of X , there is a $y \in X$ such that $|x - y| \leq \varepsilon$. Since

$$|x| - |y| \leq |x - y| \leq \varepsilon \implies |x| \leq M + \varepsilon.$$

Thus every limit point of X is bounded by $M + \varepsilon$, i.e., $\bar{X} \subseteq [-(M + \varepsilon), M + \varepsilon]$. \square

Exercise 9.1.12. *Show that the union of any finite collection of bounded subsets of \mathbf{R} is still a bounded set. Is this conclusion still true if one takes an infinite collection of bounded subsets of \mathbf{R} ?*

Proof. Let X_1, X_2, \dots, X_n be finite and bounded subsets of \mathbf{R} , where $X_n \subseteq [-M_n, M_n]$. We induct on n . Assertion is obviously true for $n = 1$. Now suppose inductively that

$$\bigcup_{i=1}^n X_i \subseteq [-M', M'].$$

By induction hypothesis,

$$\bigcup_{i=1}^{n+1} X_i = \left(\bigcup_{i=1}^n X_i \right) \cup X_{n+1} \subseteq [-M, M],$$

where $M = M' + M_{n+1}$. This close the induction.

The assertion is not hold for infinite case. Let $X_n := \{n\}$ for all $n \in \mathbf{N}$, then $\bigcup_{n \in \mathbf{N}} X_n = \mathbf{N}$ which is unbounded. \square

Exercise 9.1.13. *Prove Theorem 9.1.24. (Hint: to show (a) implies (b), use the Bolzano-Weierstrass theorem (Theorem 6.6.8) and Corollary 9.1.17. To show (b) implies (a), argue by contradiction, using Corollary 9.1.17 to establish that X is closed. You will need the axiom of choice to show that X is bounded, as in Lemma 8.4.5.)*

Theorem 9.1.24 (Heine-Borel theorem for the line). Let X be a subset of \mathbf{R} . Then the following two statements are equivalent:

- (a) X is closed and bounded.
- (b) Given any sequence $(a_n)_{n=0}^{\infty}$ of real numbers which takes values in X (i.e., $a_n \in X$ for all n), there exists a subsequence $(a_{n_j})_{j=0}^{\infty}$ of the original sequence, which converges to some number L in X .

Proof. (a) \Rightarrow (b). Since X is closed, by Corollary 9.1.17, there is a convergent sequence $(a_n)_{n=0}^{\infty} \in X$, and $\lim_{n \rightarrow \infty} a_n = L \in X$. Since X is bounded, there is an M such that $(a_n)_{n=0}^{\infty} \leq M$, by Bolzano-Weierstrass theorem, there is at least one subsequence of $(a_n)_{n=0}^{\infty}$ which is convergent to L .

(b) \Rightarrow (a). Suppose for sake of contradiction that X is not closed or not bounded. Let $\varepsilon > 0$. If X is not closed. There is an adherent point x of X that $x \notin X$, by Lemma 9.1.14, there is a sequence $(a_n)_{n=0}^{\infty}$ that $\lim_{n \rightarrow \infty} a_n = x$. By hypothesis, there is a subsequence $(a_{n_j})_{j=0}^{\infty}$ of $(a_n)_{n=0}^{\infty}$ which converges to $L \in X$. Since all of subsequence of convergent sequence have same limit, we have $L = x \in X$, a contradiction.

While if X is unbounded. There is an M such that $|x| > M$ for all $x \in X$. Define

$$X_n := \{x \in X : |x| > n \text{ for all } n \in \mathbf{N}\},$$

which is non-empty. By Choice Axiom, we can find a $x_n \in X_n$ for all $n \in \mathbf{N}$ and there is a sequence $(x_n)_{n=1}^{\infty}$ lies in X . By hypothesis, there is a subsequence $(x_{n_j})_{j=0}^{\infty}$ of $(a_n)_{n=0}^{\infty}$ that $\lim_{j \rightarrow \infty} x_{n_j} = L \in X$. Then we have $|x_{n_j} - L| \leq \varepsilon$. Let $\varepsilon = 1/n_j$, then

$$n_j - |L| = 1 < |x_{n_j}| - |L| \leq |x_{n_j} - L| \leq \varepsilon,$$

for all $n_j \geq |L|+1$, a contradiction. Therefore, X is bounded and closed. \square

Exercise 9.1.14. Show that any finite subset of \mathbf{R} is closed and bounded.

Proof. Let X to be a finite subset of \mathbf{R} . Let $X_n = \{x_n\}$. Then X can be represented by the union of X_n for arbitrary $n \leq N$. And this is easy to

see that X_n is closed and bounded for all $n \leq N$. Thus by Exercise 9.1.7 and Exercise 9.1.12, X is closed and bounded. \square

Exercise 9.1.15. Let E be a bounded subset of \mathbf{R} , and let $S := \sup(E)$ be the least upper bound of E . (Note from the least upper bound principle, Theorem 5.5.9, that S is a real number.) Show that S is an adherent point of E , and is also an adherent point of $\mathbf{R} \setminus E$.

Proof. Let $\varepsilon > 0$. By definition, for every $x \in E$, we have $x \leq S$; then there exists an $x \in E$ such that $|x - S| \leq \varepsilon$. This means that S is adherent to E , thus an adherent point of E . For the second statement, define

$$D := \left\{ x \in \mathbf{R} : S + \frac{1}{n} \leq x \text{ for all } n \in \mathbf{N} \right\}.$$

Where D is a subset of $\mathbf{R} \setminus E$. We can see that $(S + \frac{1}{n})_{n=m}^{\infty}$ lies in D and converges to S . Thus S is a limit point, so that an adherent point of D . Since $D \subseteq \mathbf{R} \setminus E$, by Lemma 9.1.11, S is an adherent point of $\mathbf{R} \setminus E$. \square

9.2

The algebra of real-valued functions

Exercise 9.2.1. Let $f : \mathbf{R} \rightarrow \mathbf{R}, g : \mathbf{R} \rightarrow \mathbf{R}, h : \mathbf{R} \rightarrow \mathbf{R}$. Which of the following identities are true, and which ones are false? In the former case, give a proof; in the latter case, give a counterexample.

$$(f + g) \circ h = (f \circ h) + (g \circ h)$$

$$f \circ (g + h) = (f \circ g) + (f \circ h)$$

$$(f + g) \cdot h = (f \cdot h) + (g \cdot h)$$

$$f \cdot (g + h) = (f \cdot g) + (f \cdot h)$$

Proof. Let $f(x) = x^2, g = 2x, h = 2^x$.

$(f + g) \circ h = (f \circ h) + (g \circ h)$ is true for that

$$\begin{aligned} ((f + g) \circ h)(x) &= (f + g)(h(x)) \\ &= f(h(x)) + g(h(x)) \\ &= (f \circ h)(x) + (g \circ h)(x). \end{aligned}$$

$f \circ (g + h) \neq (f \circ g) + (f \circ h)$ for that

$$\begin{aligned} (f \circ (g + h))(x) &= f(g(x) + h(x)) \\ &= f(2x + 2^x) \\ &= (2x + 2^x)^2 = 4x^2 + 2^{2x} + 2^{x+2}x, \\ (f \circ g) + (f \circ h) &= f(g(x)) + f(h(x)) \\ &= f(2x) + f(2^x) \\ &= 4x^2 + 2^{2x}. \end{aligned}$$

$(f + g) \cdot h = (f \cdot h) + (g \cdot h)$ is true for that

$$\begin{aligned} ((f + g) \cdot h)(x) &= (f + g)(x) \cdot h(x) \\ &= (f(x) + g(x)) \cdot h(x) \\ &= f(x) \cdot h(x) + g(x) \cdot h(x) \\ &= (f \cdot g)(x) + (g \cdot h)(x) \\ &= (f \cdot g + g \cdot h)(x). \end{aligned}$$

$f \cdot (g + h) = (f \cdot g) + (f \cdot h)$ is true for that

$$\begin{aligned} (f \cdot (g + h))(x) &= f(x) \cdot (g + h)(x) \\ &= f(x) \cdot (g(x) + h(x)) \\ &= f(x) \cdot g(x) + f(x) \cdot h(x) \\ &= (f \cdot g)(x) + (f \cdot h)(x) \\ &= (f \cdot g + f \cdot h)(x). \end{aligned}$$

□

9.3**Limiting values of functions**

Exercise 9.3.1. *Prove Proposition 9.3.9.*

Proposition 9.3.9. Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, let E be a subset of X , let x_0 be an adherent point of E , and let L be a real number. Then the following two statements are logically 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 .

Proof. (a) \Rightarrow (b). Suppose that f converges to L at x_0 in E . By Definition 9.3.6, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - L| \leq \varepsilon$ for all $x \in E$ such that $|x - x_0| < \delta$. Since $a_n \in X$ for all $n \geq 0$ and $\lim_{n \rightarrow \infty} a_n = x_0$, we can find an $N > 0$ such that $|f(a_n) - L| \leq \varepsilon$ for all $a_n \in E$ and $n \geq N$ such that $|a_n - x_0| < \delta$. Thus $(f(a_n))_{n=0}^\infty$ converges to L .

(b) \Rightarrow (a). Suppose for sake of contradiction that f doesn't converge to L at x_0 in E . By Definition 9.3.6, for all $\delta > 0$, there exists an $x \in E$ and $\varepsilon > 0$ such that $|x - x_0| < \delta$ such that $|f(x) - L| > \varepsilon$. Since $(a_n)_{n=0}^\infty$ converges to x_0 in E , we can find an $N > 0$ such that $|a_n - x_0| < \delta$ for all $n \geq N$. Then implies that $|f(a_n) - L| > \varepsilon$ for all $n \geq N$. This contradict that $(f(a_n))_{n=0}^\infty$ converges to L , as desired. \square

Exercise 9.3.2. *Prove the remaining claims in Proposition 9.3.14.*

Proposition 9.3.14 (Limit laws for functions). Let X be a subset of \mathbf{R} , let E be a subset of X , let x_0 be an adherent point of E , and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions. Suppose that f has a limit L at x_0 in E , and g has a limit M at x_0 in E . Then $f + g$ has a limit

$L + M$ at x_0 in E , $f - g$ has a limit $L - M$ at x_0 in E , $\max(f, g)$ has a limit $\max(L, M)$ at x_0 in E , $\min(f, g)$ has a limit $\min(L, M)$ at x_0 in E and fg has a limit LM at x_0 in E . If c is a real number, then cf has a limit cL at x_0 in E . Finally, if g is non-zero on E (i.e., $g(x) \neq 0$ for all $x \in E$) and M is non-zero, then f/g has a limit L/M at x_0 in E .

Proof. We prove remaining claims. Since x_0 is an adherent point of E , by Lemma 9.1.14, there is a sequence $(a_n)_{n=0}^\infty$ consisting of elements in E , which converges to x_0 . Since f has a limit L at x_0 in E , we thus see by Proposition 9.3.9, 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(d), we conclude that $((f - g)(a_n))_{n=0}^\infty$ converges to $L - M$; by Theorem 6.1.19(g), we have $(\max(f(a_n), g(a_n)))_{n=0}^\infty$ converges to $\max(L, M)$; by Theorem 6.1.19(h), we conclude that $(\min(f(a_n), g(a_n)))_{n=0}^\infty$ converges to $\min(L, M)$; by Theorem 6.1.19(b), we have $((fg)(a_n))_{n=0}^\infty$ converges to LM ; if c is a real number, then by Theorem 6.1.19(c), we conclude that $((cf)(a_n))_{n=0}^\infty$ converges to cL ; if $g(x) \neq 0$ for all $x \in E$ and $M \neq 0$, then by Theorem 6.1.19(f), we conclude that $((f/g)(a_n))_{n=0}^\infty$ converges to L/M . By Proposition 9.3.9 again, these implies that $f - g$ has a limit $L - M$, $\max(f, g)$ has a limit $\max(L, M)$, $\min(f, g)$ has a limit $\min(L, M)$, fg has a limit LM , cf has a limit cL and f/g has a limit L/M , at x_0 in E . \square

Exercise 9.3.3. Prove Lemma 9.3.18.

Proposition 9.3.18 (Limits are local). Let X be a subset of \mathbf{R} , let E be a subset of X , let x_0 be an adherent point of E , let $f : X \rightarrow \mathbf{R}$ be a function, and let L be a real number. Let $\delta > 0$. Then we have

$$\lim_{x \rightarrow x_0; x \in E} f(x) = L$$

if and only if

$$\lim_{x \rightarrow x_0; x \in E \cap (x_0 - \delta, x_0 + \delta)} f(x) = L.$$

Proof. Suppose that

$$\lim_{x \rightarrow x_0; x \in E} f(x) = L.$$

By Definition 9.3.6, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - L| \leq \varepsilon$ for all $x \in E$ such that $|x - x_0| < \delta$. This implies that $x \in E$ and $x \in (x_0 - \delta, x_0 + \delta)$, so that $x \in E \cap (x_0 - \delta, x_0 + \delta)$. Thus by Definition 9.3.6 again,

$$\lim_{x \rightarrow x_0; x \in E \cap (x_0 - \delta, x_0 + \delta)} f(x) = L.$$

For the converse, suppose that

$$\lim_{x \rightarrow x_0; x \in E \cap (x_0 - \delta, x_0 + \delta)} f(x) = L.$$

By Definition 9.3.6, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - L| \leq \varepsilon$ for all $x \in E \cap (x_0 - \delta, x_0 + \delta)$ such that $|x - x_0| < \delta$. This implies that $x \in E$ such that $|x - x_0| < \delta$. Thus by Definition 9.3.6 again,

$$\lim_{x \rightarrow x_0; x \in E} f(x) = L.$$

□

Exercise 9.3.4. *Propose a definition for limit superior $\limsup_{x \rightarrow x_0; x \in E} f(x)$ and limit inferior $\liminf_{x \rightarrow x_0; x \in E} f(x)$, and then propose an analogue of Proposition 9.3.9 for your definition. (For an additional challenge: prove that analogue.)*

Proof.

Definition 1 (Limit superior of function). Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, let E be a subset of X , let x_0 be an adherent point of E , and let L be a real number. We define $f^+(x)$ as the supremum of $f(x)$ for all $x \in E$ such that $|x - x_0| < \delta$. We say that L is the *superior limit* of f , and write $\limsup_{x \rightarrow x_0; x \in E} f(x) = L$, iff for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f^+(x) - L| \leq \varepsilon$ for all $x \in X$ such that $|x - x_0| < \delta$.

Proposition 1. Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, let E be a subset of X , let x_0 be an adherent point of E , and let L be a

real number. Let $f^+(x)$ be a supremum of $f(x)$ for all $x \in E$ such that $|x - x_0| < \delta$. Then the following two statements are logically equivalent:

- (a) L is a superior limit of f 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 .

Proof of Proposition 1. (a) \Rightarrow (b). Suppose that L is a superior limit of f at x_0 in E . By Definition 1, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f^+(x) - L| \leq \varepsilon$ for all $x \in E$ such that $|x - x_0| < \delta$. Since $a_n \in X$ for all $n \geq 0$ and $\lim_{n \rightarrow \infty} a_n = x_0$, we can find an $N > 0$ such that $|f^+(a_n) - L| \leq \varepsilon$ for all $a_n \in E$ and $n \geq N$ such that $|a_n - x_0| < \delta$. Thus $(f(a_n))_{n=0}^\infty$ converges to L .

(b) \Rightarrow (a). Suppose for sake of contradiction that L is not a superior limit of f at x_0 in E . By Definition 1, for all $\delta > 0$, there exists an $x \in E$ and $\varepsilon > 0$ such that $|x - x_0| < \delta$ such that $|f^+(x) - L| > \varepsilon$. Since $(a_n)_{n=0}^\infty$ converges to x_0 in E , we can find an $N > 0$ such that $|a_n - x_0| < \delta$ for all $n \geq N$. Then implies that $|f^+(a_n) - L| > \varepsilon$ for all $n \geq N$. This contradicts that $(f^+(a_n))_{n=0}^\infty$ converges to L , as desired.

Definition 2 (Limit inferior of function). Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, let E be a subset of X , let x_0 be an adherent point of E , and let L be a real number. We define $f^-(x)$ as the infimum of $f(x)$ for all $x \in E$ such that $|x - x_0| < \delta$. We say that L is the *inferior limit* of f , and write $\liminf_{x \rightarrow x_0; x \in E} f(x)$, iff for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f^-(x) - L| \leq \varepsilon$ for all $x \in X$ such that $|x - x_0| < \delta$.

Proposition 2. Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, let E be a subset of X , let x_0 be an adherent point of E , and let L be a real number. Let $f^+(x)$ be a supremum of $f(x)$ for all $x \in E$ such that $|x - x_0| < \delta$. Then the following two statements are logically equivalent:

- (a) L is a superior limit of f 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 .

Proof of Proposition 2. The proof is similar to limit inferior. \square

Exercise 9.3.5. (*Continuous version of squeeze test*) Let X be a subset of \mathbf{R} , let E be a subset of X , let x_0 be an adherent point of E , and let $f : X \rightarrow \mathbf{R}, g : X \rightarrow \mathbf{R}, h : X \rightarrow \mathbf{R}$ be functions such that $f(x) \leq g(x) \leq h(x)$ for all $x \in E$. If we have $\lim_{x \rightarrow x_0; x \in E} f(x) = \lim_{x \rightarrow x_0; x \in E} h(x) = L$ for some real number L , show that $\lim_{x \rightarrow x_0; x \in E} g(x) = L$.

Proof. For every sequence $(a_n)_{n=0}^\infty$, we have $f(a_n) \leq g(a_n) \leq h(a_n)$ for all $n \in \mathbf{N}$. Since f, g converges to L at x_0 in E , by Proposition 9.3.9, we have $(f(a_n))_{n=0}^\infty$ and $(h(a_n))_{n=0}^\infty$ converges to L . By Corollary 6.4.14, $(g(a_n))_{n=0}^\infty$ also converges to L , by Proposition 9.3.9 again, $\lim_{x \rightarrow x_0; x \in E} g(x) = L$. \square

9.4

Continuous functions

Exercise 9.4.1. Prove Proposition 9.4.7. (Hint: this can largely be done by applying the previous propositions and lemmas. Note that to prove (a), (b), and (c) are equivalent, you do not have to prove all six equivalences, but you do have to prove at least three; for instance, showing that (a) implies (b), (b) implies (c), and (c) implies (a) will suffice, although this is not necessarily the shortest or simplest way to do this question.)

Proposition 9.4.7 (Equivalent formulations of continuity). Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, and let x_0 be an element of X . Then the following four statements are logically equivalent:

- (a) f is continuous at x_0 .
- (b) For every sequence $(a_n)_{n=0}^\infty$ consisting of elements of X with $\lim_{n \rightarrow \infty} a_n = x_0$, we have $\lim_{n \rightarrow \infty} f(a_n) = f(x_0)$.

- (c) For every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all $x \in X$ with $|x - x_0| < \delta$.
- (d) For every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for all $x \in X$ with $|x - x_0| \leq \delta$.

Proof. (a) \Rightarrow (b). Suppose that f is continuous at x_0 . By Definition 9.4.1, we have $\lim_{x \rightarrow x_0; x \in X} f(x) = f(x_0)$. Since $\lim_{n \rightarrow \infty} a_n = x_0$, by Proposition 9.3.9, we have $\lim_{n \rightarrow \infty} f(a_n) = f(x_0)$.

(b) \Rightarrow (c). By Proposition 9.3.9, f converges to $f(x_0)$ at $x_0 \in X$. Thus by definition, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all $x \in X$ with $|x - x_0| < \delta$.

(c) \Rightarrow (a). This immediately comes from Definition 9.3.6 and Definition 9.4.1.

Since (c) and (d) obviously are logically equivalent, we proved the equivalency of these four statements. \square

Exercise 9.4.2. Let X be a subset of \mathbf{R} . and let $c \in \mathbf{R}$. Show that the constant function $f : X \rightarrow \mathbf{R}$ defined by $f(x) := c$ is continuous, and show that the identity function $g : X \rightarrow \mathbf{R}$ defined by $g(x) := x$ is also continuous.

Proof. For every real number $x_0 \in X$, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| = |c - c| = 0 < \varepsilon$ for all $x \in X$ such that $|x - x_0| < \delta$. In this case δ can be any positive real number. By Proposition 9.4.7, f is continuous at x_0 .

For every real number $x_0 \in X$, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|g(x) - g(x_0)| = |x - x_0| \leq \varepsilon$ for all $x \in X$ such that $|x - x_0| < \delta$. Where $\delta = \varepsilon$. By Proposition 9.4.7, g is continuous at x_0 . \square

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.)

Proposition 9.4.10 (Exponentiation is continuous, I). Let $a > 0$ be a positive real number. Then the function $f : \mathbf{R} \rightarrow \mathbf{R}$ defined by $f(x) := a^x$ is continuous.

Proof. First of all we show that for all $x \in \mathbf{R}$,

$$\lim_{x \rightarrow 0; x \in \mathbf{R}} a^x = 1.$$

Suppose that $a > 1$. For every natural number $n \in \mathbf{N}$ such that $|x| \leq 1/n$, by Proposition 6.7.3, we have $a^{-1/n} \leq a^x \leq a^{1/n}$. By Lemma 6.5.3, we have $\lim_{n \rightarrow \infty} a^{-1/n} = \lim_{n \rightarrow \infty} a^{1/n} = 1$. Thus by the squeeze test, $\lim_{x \rightarrow 0} a^x = 1$. This is similar to show the case of $0 < a \leq 1$. Then for every $\varepsilon > 0$, there exists a $\delta = 1/n > 0$ such that $|a^x - 1| \leq \varepsilon$ for all $x \in \mathbf{R}$ such that $|x| \leq 1/n$. Therefore, for all real number $a > 0$, we have $\lim_{x \rightarrow 0; x \in \mathbf{R}} a^x = 1$.

Since $x \rightarrow x_0$ implies that $x - x_0 \rightarrow 0$, and a^{x_0} is positive real number for all $x_0 \in \mathbf{R}$. We have for every $\varepsilon/a^{x_0} > 0$, there exists a $\delta > 0$ such that $|a^x - a^{x_0}| = a^{x_0}|a^{x-x_0} - 1| \leq \varepsilon$ for all $x \in \mathbf{R}$ with $|x - x_0| \leq \delta$. By Proposition 9.4.7, f is continuous. \square

Exercise 9.4.4. Prove Proposition 9.4.11. (Hint: from limit laws (Proposition 9.3.14) one can show that $\lim_{x \rightarrow 1} x^n = 1$ for all integers n . From this and the squeeze test (Corollary 6.4.14) deduce that $\lim_{x \rightarrow 1} x^p = 1$ for all real numbers p . Finally, apply Proposition 6.7.3.)

Proposition 9.4.11 (Exponentiation is continuous, II). Let p be a real number. Then the function $f : (0, \infty) \rightarrow \mathbf{R}$ defined by $f(x) := x^p$ is continuous.

Proof. By Proposition 9.3.14, we have

$$\lim_{x \rightarrow 1} x^n = \left(\lim_{x \rightarrow 1} x \right)^n = 1^n = 1.$$

By Exercise 5.4.3, for every real number p , there exists a natural number N such that $N \leq p < N + 1$. Then we have

$$x^N \leq x^p \leq x^{N+1}.$$

By Corollary 6.4.14, this implies that $\lim_{x \rightarrow 1} x^p = 1$.

Since $x \in (0, \infty)$, we know that $x^p > 0$. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|x^p - x_0^p| = x_0^p \left| \left(\frac{x}{x_0} \right)^p - 1 \right| \leq \varepsilon$$

for all $x \in (0, \infty)$ such that $|x - x_0| < \delta$. Thus f is continuous. \square

Exercise 9.4.5. *Prove Proposition 9.4.13.*

Proposition 9.4.13 (Composition preserves continuity). Let X and Y be subsets of \mathbf{R} , and let $f : X \rightarrow Y$ and $g : Y \rightarrow \mathbf{R}$ be functions. Let x_0 be a point in X . If f is continuous at x_0 , and g is continuous at $f(x_0)$, then the composition $g \circ f : X \rightarrow \mathbf{R}$ is continuous at x_0 .

Proof. Since f is continuous, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for all $x \in X$ such that $|x - x_0| < \delta$. Since g is also continuous, for every $\zeta > 0$, there exists $\varepsilon > 0$ such that $|g(f(x)) - g(f(x_0))| \leq \zeta$ for all $f(x) \in Y$ such that $|f(x) - f(x_0)| < \varepsilon$. Thus for every $\zeta > 0$, there exists a $\delta > 0$ such that $|(g \circ f)(x) - (g \circ f)(x_0)| \leq \zeta$ for all $x \in X$ such that $|x - x_0| < \delta$. As desired. \square

Exercise 9.4.6. *Let X be a subset of \mathbf{R} , and let $f : X \rightarrow \mathbf{R}$ be a continuous function. If Y is a subset of X , show that the restriction $f|_Y : Y \rightarrow \mathbf{R}$ of f to Y is also a continuous function. (Hint: this is a simple result, but it requires you to follow the definitions carefully.)*

Proof. Let x_0 be an element of $Y \subseteq X$. Since f is continuous at x_0 on X , for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for all $x \in X$ such that $|x - x_0| < \delta$. Since $Y \subseteq X$, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for all $x \in Y$ such that $|x - x_0| < \delta$. By definition, $f|_Y$ is continuous at x_0 on Y . \square

Exercise 9.4.7. *Let $n \geq 0$ be an integer, and for each $0 \leq i \leq n$ let c_i be a real number. Let $P : \mathbf{R} \rightarrow \mathbf{R}$ be the function*

$$P(x) := \sum_{i=0}^n c_i x^i;$$

such a function is known as a **polynomial of one variable**; a typical example is $P(x) = 6x - 3x^2 + 4$. Show that P is continuous.

Proof. By Proposition 9.4.11, x^i is continuous for all $i \in \mathbf{N}$. Thus by Proposition 9.4.9, $P(x)$ is continuous. \square

9.5

Left and right limits

Exercise 9.5.1. Let E be a subset of \mathbf{R} , let $f : E \rightarrow \mathbf{R}$ 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 $\lim_{x \rightarrow x_0; x \in E} f(x)$ to exist and equal $+\infty$ or $-\infty$. If $f : \mathbf{R} \setminus \{0\} \rightarrow \mathbf{R}$ is the function $f(x) := 1/x$, use your definition to conclude $f(0+) = +\infty$ and $f(0-) = -\infty$. Also, state and prove some analogue of Proposition 9.3.9 when $L = +\infty$ or $L = -\infty$.

Proof.

Definition 1 (Infinite limit of functions at a point). Let E be a subset of \mathbf{R} , let $f : E \rightarrow \mathbf{R}$ be a function, let x_0 be an adherent point of E . We say that f converges to positive infinity at x_0 in E , and write $\lim_{x \rightarrow x_0; x \in E} f(x) = +\infty$, iff for every $\varepsilon > 0$ and $M > 0$, there exists a $\delta > 0$ such that $f(x) > M$ for all $x \in E$ such that $|x - x_0| < \delta$. And f converges to negative infinity at x_0 in E , and write $\lim_{x \rightarrow x_0; x \in E} f(x) = -\infty$, iff for every $\varepsilon > 0$ and $M > 0$, there exists a $\delta > 0$ such that $f(x) < -M$ for all $x \in E$ such that $|x - x_0| < \delta$.

If $f : \mathbf{R} \setminus \{0\} \rightarrow \mathbf{R}$ is the function $f(x) := 1/x$, then $f(0+) = +\infty$ and $f(0-) = -\infty$. We use our definition to prove this assertion: By Definition 9.5.1, we have

$$f(0+) = \lim_{x \rightarrow 0; x \in \mathbf{R} \setminus \{0\} \cap (0, \infty)} \frac{1}{x} = \lim_{x \rightarrow 0; x \in (0, \infty)} \frac{1}{x}.$$

For every $\varepsilon > 0$ and $M > 0$, there exists a $\delta = 1/M$ such that $1/x > M$ for all $x \in (0, \infty)$ such that $|x - 0| = x < 1/M$. By Definition 1, we have

$f(0+) = +\infty$. Other hand, we have

$$f(0-) = \lim_{x \rightarrow 0; x \in \mathbf{R} \setminus \{0\} \cap (-\infty, 0)} \frac{1}{x} = \lim_{x \rightarrow 0; x \in (-\infty, 0)} \frac{1}{x}.$$

Then for every $\varepsilon > 0$ and $M > 0$, there exists a $\delta = 1/M$ such that $1/x < -M$ for all $x \in (-\infty, 0)$ such that $|x - 0| = -x < 1/M$. By Definition 1, we have $f(0-) = -\infty$.

Proposition 1. Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a function, let E be a subset of X , let x_0 be an adherent point of E . Then the following two statements are logically equivalent:

- (a) f converges to $+\infty$ or $-\infty$ 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 $+\infty$ or $-\infty$.

Proof of Proposition 1. (a) \Rightarrow (b). Suppose that f converges to $+\infty$ at x_0 in E . By Definition 1, for every $\varepsilon > 0$ and $M > 0$, there exists a $\delta > 0$ such that $f(x) > M$ for all $x \in E$ such that $|x - x_0| < \delta$. Since $a_n \in X$ for all $n \geq 0$ and $\lim_{n \rightarrow \infty} a_n = x_0$, we can find an $N > 0$ such that $f(a_n) > M$ for all $a_n \in E$ and $n \geq N$ such that $|a_n - x_0| < \delta$. Thus $(f(a_n))_{n=0}^{\infty}$ converges to $+\infty$.

(b) \Rightarrow (a). Suppose for sake of contradiction that $+\infty$ is not a limit of f at x_0 in E . By Definition 1, for all $\delta > 0$, there exists an $x \in E$, $M > 0$ and $\varepsilon > 0$ such that $|x - x_0| < \delta$ such that $f(x) \leq M$. Since $(a_n)_{n=0}^{\infty}$ converges to x_0 in E , we can find an $N > 0$ such that $|a_n - x_0| < \delta$ for all $n \geq N$. Then implies that $f(a_n) > M$ for all $n \geq N$. This contradict that $(f(a_n))_{n=0}^{\infty}$ converges to ∞ , as desired.

This is similar to show that case of $-\infty$. □

9.6**The maximum principle**

Exercise 9.6.1. *Give examples of*

- (a) *a function $f : (1, 2) \rightarrow \mathbf{R}$ which is continuous and bounded, attains its minimum somewhere, but does not attain its maximum anywhere;*
- (b) *a function $f : [0, \infty) \rightarrow \mathbf{R}$ which is continuous, bounded, attains its maximum somewhere, but does not attain its minimum anywhere;*
- (c) *a function $f : [-1, 1] \rightarrow \mathbf{R}$ which is bounded but does not attain its minimum anywhere or its maximum anywhere;*
- (d) *a function $f : [-1, 1] \rightarrow \mathbf{R}$ which has no upper bound and no lower bound.*

Explain why none of the examples you construct violate the maximum principle. (Note: read the assumptions carefully!)

Proof. (a) Let $f(x) = (x - 3/2)^2$. f is continuous and bounded by $1/4$, it attains its minimum in $x = 3/2$ but no maximum. (b) Let $f(x) = e^{-x}$. f is continuous and bounded by 1 , it attains its maximum in $x = 0$ but no minimum. (c) Let $f(x) = x$ for $x \in (-1, 1)$, $f(x) = 0$ for $x = -1$ or $x = 1$. f is bounded by 1 but no maximum and minimum. (d) Let $f(x) = 1/x$. f is unbounded. In (a) and (b), f is not define in closed set; In (c) and (d), f is not continuous. \square

9.7**The intermediate value theorem**

Exercise 9.7.1. *Prove Corollary 9.7.4. (Hint: you may need Exercise 9.4.6 addition to the intermediate value theorem.)*

Corollary 9.7.4 (Images of continuous functions). Let $a < b$, and let $f : [a, b] \rightarrow \mathbf{R}$ be a continuous function on $[a, b]$. Let $M := \sup_{x \in [a, b]} f(x)$ be the maximum value of f , and let $m := \inf_{x \in [a, b]} f(x)$ be the minimum value. Let y be a real number between m and M (i.e., $m \leq y \leq M$). Then there exists a $c \in [a, b]$ such that $f(c) = y$. Furthermore, we have $f([a, b]) = [m, M]$.

Proof. Since f is continuous, there exists $x_m, x_M \in [a, b]$ such that $f(x_m) = m$ and $f(x_M) = M$. Thus we have $[x_m, x_M] \in [a, b]$ or $[x_M, x_m] \in [a, b]$. We assume that $x_m < x_M$, then $[x_m, x_M] \in [a, b]$, this is similar to prove that $x_M < x_m$. By Exercise 9.4.6, $f|_{[x_m, x_M]} : [x_m, x_M] \rightarrow \mathbf{R}$ is continuous. Then by Theorem 9.7.1, there exists a $c \in [x_m, x_M]$ such that $f(c) = y$. \square

Exercise 9.7.2. Let $f : [0, 1] \rightarrow [0, 1]$ be a continuous function. Show that there exists a real number x in $[0, 1]$ such that $f(x) = x$. (Hint: apply the intermediate value theorem to the function $f(x) - x$.) This point x is known as a **fixed point** of f , and this result is a basic example of a **fixed point theorem** which play an important rôle in certain types of analysis.

Proof. Since f is continuous, $g : [0, 1] \rightarrow [-1, 1]$ which is defined as $g(x) := f(x) - x$ also is continuous. Let $0 \in [-1, 1]$. By the intermediate value theorem, there exists a $c \in [0, 1]$ such that $g(c) = 0$. Thus there is an $x \in [0, 1]$ such that $f(x) = x$. \square

9.8

Monotonic functions

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.)

Proof. Let $a < b$ be real numbers and let $f : [a, b] \rightarrow \mathbf{R}$ be a function monotone on $[a, b]$. Then we have either $f(a) \geq f(x)$ or $f(b) \geq f(x)$ for all

$x \in [a, b]$. Thus f attains its maximum at a or b . This obviously hold for strictly monotone. \square

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.)

Proof. Let $f : [-1, 1] \rightarrow \mathbf{R}$ to be a monotonic function

$$f(x) := \begin{cases} x + 1 & \text{if } x \geq 0, \\ x - 1 & \text{if } x < 0. \end{cases}$$

We can see that f is strictly monotone increasing, and $f(-1) = -2, f(1) = 2$. But there is no $x \in [-1, 1]$ such that $f(x) = 0$. \square

Exercise 9.8.3. Let $a < b$ be real numbers, and let $f : [a, b] \rightarrow \mathbf{R}$ be a function which is both continuous and one-to-one. Show that f is strictly monotone (Hint: divide into the three cases $f(a) < f(b)$, $f(a) = f(b)$, $f(a) > f(b)$. The second case leads directly to a contradiction. In the first case, use contradiction and the intermediate value theorem to show that f is strictly monotone increasing; in the third case, argue similarly to show f is strictly monotone decreasing.)

Proof. We divide into three cases: $f(a) < f(b)$, $f(a) = f(b)$, $f(a) > f(b)$. In the first case, suppose that f is not monotone, then there is $c \in [a, b]$ such that $f(a) < f(c)$ and $f(b) < f(c)$, or $f(c) < f(a)$ and $f(c) < f(b)$. In both case, since f is continuous, by the intermediate value theorem, we can find an $x_1 \in [a, c]$ and $x_2 \in [c, b]$ such that $f(x_1) = f(x_2)$, a contradiction. Thus f is strictly monotone increasing. In the second case, Since $a < b$, f is not a one-to-one, a contradiction. In the third case, this is similar to show that f is strictly monotone decreasing. Thus f is strictly monotone. \square

Exercise 9.8.4. Prove Proposition 9.8.3. (Hint: to show that f^{-1} is continuous, it is easiest to use the “epsilon-delta” definition of continuity, Proposition 9.4.7(c).) Is the proposition still true if the continuity assump-

tion is dropped or if strict monotonicity is replaced just by monotonicity? How should one modify the proposition to deal with strictly monotone decreasing functions instead of strictly monotone increasing functions?

Proposition 9.8.3. Let $a < b$ be real numbers, and let $f : [a, b] \rightarrow \mathbf{R}$ be a function which is both continuous and strictly monotone increasing. Then f is a bijection from $[a, b]$ to $[f(a), f(b)]$ and the inverse $f^{-1} : [f(a), f(b)] \rightarrow [a, b]$ is also continuous and strictly monotone increasing.

Proof. First of all, we need to show that f is injective and surjective. Since f is strictly monotone increasing, for every $x \neq x'$, it either $x < x'$ or $x > x'$, thus we have $f(x) < f(x')$ or $f(x) > f(x')$. Thus f is injective. By the intermediate value theorem, for every $y \in [f(a), f(b)]$, there exists a $c \in [a, b]$ such that $f(c) = y$. Thus f is surjective. Therefore, f is bijective and its inverse exists. Let $f^{-1} : [f(a), f(b)] \rightarrow [a, b]$. This is easy to see that f^{-1} is strictly monotone increasing. We show that f^{-1} is continuous. Since f is continuous, by Proposition 9.4.7(c), for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all $x \in [a, b]$ such that $|x - x_0| < \delta$. Since f^{-1} is bijective, there is $x = f^{-1}(y)$ and $x_0 = f^{-1}(y_0)$. Thus for every $\delta > 0$, there exists an $\varepsilon > 0$ such that $|f^{-1}(x) - f^{-1}(x_0)| < \delta$ for all $y \in [f(a), f(b)]$ such that $|y - y_0| < \varepsilon$. By Proposition 9.4.7 again, f^{-1} is continuous at y_0 .

If the continuity assumption is dropped or if strict monotonicity is replaced just by monotonicity, then statement would not hold for that f^{-1} doesn't exist.

When strictly monotone increasing is replaced by strictly monotone decreasing, we have:

Proposition 9.8.3. Let $a < b$ be real numbers, and let $f : [a, b] \rightarrow \mathbf{R}$ be a function which is both continuous and strictly monotone decreasing. Then f is a bijection from $[a, b]$ to $[f(b), f(a)]$ and the inverse $f^{-1} : [f(b), f(a)] \rightarrow [a, b]$ is also continuous and strictly monotone decreasing.

□

Exercise 9.8.5. In this exercise we give an example of a function which has a discontinuity at every rational point, but is continuous at every irrational. Since the rationals are countable, we can write them as $\mathbf{Q} = \{q(0), q(1), q(2), \dots\}$, where $q : \mathbf{N} \rightarrow \mathbf{Q}$ is a bijection from \mathbf{N} to \mathbf{Q} . Now define a function $g : \mathbf{Q} \rightarrow \mathbf{R}$ by setting $g(q(n)) := 2^{-n}$ for each natural number n ; thus g maps $q(0)$ to 1, $q(1)$ to $1/2$, etc. Since $\sum_{n=0}^{\infty} 2^{-n}$ is absolutely convergent, we see that $\sum_{r \in \mathbf{Q}} g(r)$ is also absolutely convergent. Now define the function $f : \mathbf{R} \rightarrow \mathbf{R}$ by

$$f(x) := \sum_{r \in \mathbf{Q}; r < x} g(r).$$

Since $\sum_{r \in \mathbf{Q}} g(r)$ is absolutely convergent, we know that $f(x)$ is well-defined for every real number x

- (a) Show that f is strictly monotone increasing. (Hint: you will need Proposition 5.4.14.)
- (b) Show that for every rational number r , f is discontinuous at r . (Hint: since r is rational, $r = q(n)$ for some natural number n . Show that $f(x) \geq f(r) + 2^{-n}$ for all $x > r$.)
- (c) Show that for every irrational number x , f is continuous at x . (Hint first demonstrate that the functions

$$f_n(x) := \sum_{r \in \mathbf{Q}; r < x, g(r) \geq 2^{-n}} g(r)$$

are continuous at x , and that $|f(x) - f_n(x)| \leq 2^{-n}$.)

Proof.

9.9

Uniform continuity

Exercise 9.9.1. Prove Lemma 9.9.7.

Lemma 9.9.7. Let $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ be sequences of real numbers (not necessarily bounded or convergent). Then $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent if and only if $\lim_{n \rightarrow \infty} (a_n - b_n) = 0$.

Proof. Suppose that $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are equivalent. By Definition 9.9.5, for every $\varepsilon > 0$, there exists an $N \geq 1$ such that they are ε -closed. Thus we have $|a_n - b_n| \leq \varepsilon$ for all $n \geq N$. This means $\lim_{n \rightarrow \infty} (a_n - b_n) = 0$. Conversely, suppose that $\lim_{n \rightarrow \infty} (a_n - b_n) = 0$. By definition, for every $\varepsilon > 0$, there exists an $N > 1$ such that $|a_n - b_n| \leq \varepsilon$ for all $n \geq N$. Thus $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ are eventually ε -closed, so that are equivalent. \square

Exercise 9.9.2. Prove Proposition 9.9.8. (Hint: you should avoid Lemma 9.9.7 and instead go back to the definition of equivalent sequences in Definition 9.9.5.)

Proposition 9.9.8. Let X be a subset of \mathbf{R} , and let $f : X \rightarrow \mathbf{R}$ be a function. Then the following two statements are logically equivalent:

- (a) f is uniformly continuous on X .
- (b) Whenever $(x_n)_{n=0}^{\infty}$ and $(y_n)_{n=0}^{\infty}$ are two equivalent sequences consisting of elements of X , the sequences $(f(x_n))_{n=0}^{\infty}$ and $(f(y_n))_{n=0}^{\infty}$ are also equivalent.

Proof. (a) \Rightarrow (b). Suppose that f is uniformly continuous on X . By Definition 9.9.2, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $f(x)$ and $f(x_0)$ are ε -closed for all $x, x_0 \in X$ such that $|x - x_0| \leq \delta$. Since $(x_n)_{n=0}^{\infty}$ and $(y_n)_{n=0}^{\infty}$ are two equivalent and $x_n, y_n \in X$ for all $n \geq 0$, they are ε -closed in X . Thus for every $\varepsilon > 0$, there exists a $\delta > 0$ and $N \geq 0$ such that $f(x_n)$ and $f(y_n)$ are ε -closed for all $n \geq N$ such that $x_n, y_n \in X$. By Definition 9.9.5, $(f(x_n))_{n=0}^{\infty}$ and $(f(y_n))_{n=0}^{\infty}$ are also equivalent.

(b) \Rightarrow (a). Suppose for sake of contradiction that f is not uniformly continuous on X . Then for every $\delta \geq 0$, there exists an $\varepsilon > 0$ such that $|f(x) - f(x_0)| > \varepsilon$ for some two points are δ -closed. Since $(f(x_n))_{n=0}^{\infty}$ and $(f(y_n))_{n=0}^{\infty}$ are equivalent in X , for every $\varepsilon > 0$, there exists an $N \geq 0$ and

$\delta > 0$ such that $|f(x_n) - f(y_n)| \leq \varepsilon$ for all $n \geq N$ such that $(x_n)_{n=0}^\infty$ and $(y_n)_{n=0}^\infty$ are δ -closed. This means for any $N \geq n$ we cannot find any two points which are ε -closed such that $f(x)$ and $f(x_0)$ be ε -closed, a contradiction. \square

Exercise 9.9.3. *Prove Proposition 9.9.12. (Hint: use Definition 9.9.2 directly.)*

Proposition 9.9.12. Let X be a subset of \mathbf{R} , and let $f : X \rightarrow \mathbf{R}$ be a uniformly continuous function. Let $(x_n)_{n=0}^\infty$ be a Cauchy sequence consisting entirely of elements in X . Then $(f(x_n))_{n=0}^\infty$ is also a Cauchy sequence.

Proof. By Definition 9.9.2, f is a uniformly continuous implies that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \varepsilon$ for any $x, x_0 \in X$ which are δ -closed. Since $(x_n)_{n=0}^\infty$ is a Cauchy sequence, for every $\delta > 0$ such that x_j and x_k are δ -closed for all $j, k \geq N$. Thus for every $\varepsilon > 0$, for every $N > 0$ we have $|f(x_j) - f(x_k)| \leq \varepsilon$ for all $j, k \geq N$. This means that $(f(x_n))_{n=0}^\infty$ is a Cauchy sequence. \square

Exercise 9.9.4. *Use Proposition 9.9.12 to prove Corollary 9.9.14. Use this corollary to give an alternate demonstration of the results in Example 9.9.10.*

Corollary 9.9.14. Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ be a uniformly continuous function, and let x_0 be an adherent point of X . Then the limit $\lim_{x \rightarrow x_0; x \in X} f(x)$ exists (in particular, it is a real number).

Proof. Since x_0 is an adherent point of X , by Lemma 9.1.14, there exists a sequence $(a_n)_{n=1}^\infty$ which converges to x_0 . By Theorem 6.4.18, $(a_n)_{n=1}^\infty$ is a Cauchy sequence. Thus by Proposition 9.9.12, $(f(a_n))_{n=1}^\infty$ is also a Cauchy sequence, by Theorem 6.4.18 again, it is convergent. In particular, if $(a_n)_{n=0}^\infty$ is a sequence of reals, then $f(a_n)_{n=0}^\infty$ also a sequence of reals.

Consider the function $f : (0, 2) \rightarrow \mathbf{R}$ defined by $f(x) := 1/x$. Since $\lim_{x \rightarrow 0; x \in X} 1/x = \infty$, this implies that f is not uniformly continuous by the

contrapositive of Corollary 9.9.14. □

Exercise 9.9.5. *Prove Proposition 9.9.15. (Hint: mimic the proof of Lemma 9.6.3. At some point you will need either Proposition 9.9.12 or Corollary 9.9.14.)*

Proposition 9.9.15. Let X be a subset of \mathbf{R} , and let $f : X \rightarrow \mathbf{R}$ be a uniformly continuous function. Suppose that E is a bounded subset of X . Then $f(E)$ is also bounded.

Proof. Suppose that $f(E)$ is unbounded. Thus for every $M > 0$ there exists an $x \in E$ such that $|f(x)| \geq M$.

In particular, for every natural number n , the set $\{x \in E : |f(x)| \geq n\}$ is non-empty. Let $x_n := \sup\{x \in E : |f(x)| > n\}$, we thus can choose a sequence $(x_n)_{n=0}^\infty$ in E such that $|f(x_n)| \geq n$ for all n . This sequence lies in E and so by Theorem 9.1.24 there exists a subsequence $(a_{n_j})_{j=0}^\infty$ which converges to some limit $L \in E$, where $n_0 < n_1 < n_2 \cdots$ is an increasing sequence of natural numbers. Thus $(a_{n_j})_{j=0}^\infty$ is a Cauchy sequence in X . Since f is uniformly continuous, by Proposition 9.9.12, $(f(a_{n_j}))_{j=0}^\infty$ also a Cauchy sequence, so that bounded, a contradiction. □

Exercise 9.9.6. *Let X, Y, Z be subsets of \mathbf{R} . Let $f : X \rightarrow Y$ be a function which is uniformly continuous on X , and let $g : Y \rightarrow Z$ be a function which is uniformly continuous on Y . Show that the function $g \circ f : X \rightarrow Z$ is uniformly continuous on X .*

Proof. Since f is uniformly continuous, by Definition 9.9.2, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $f(x)$ and $f(x_0)$ are ε -closed for any $x, x_0 \in X$ which are δ -closed. Since g is uniformly continuous, by Definition 9.9.2, for every $\zeta > 0$ such that $g(f(x))$ and $g(f(x_0))$ for any $f(x), f(x_0) \in Y$ which are ζ -closed. Thus for every $\zeta > 0$, there exists a $\delta > 0$ such that $(g \circ f)(x)$ and $(g \circ f)(x_0)$ are ζ -closed for any $x, x_0 \in X$ which are δ -closed. This means that $g \circ f$ is uniformly continuous. □

9.10**Limits at infinity**

Exercise 9.10.1. Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, then a_n can also be thought of as a function from \mathbf{N} to \mathbf{R} , which takes each natural number n to a real number a_n . Show that

$$\lim_{n \rightarrow +\infty; n \in \mathbf{N}} a_n = \lim_{n \rightarrow \infty} a_n$$

where the left-hand limit is defined by Definition 9.10.3 and the right-hand limit is defined by Definition 6.1.8. More precisely, show that if one of the above two limits exists then so does the other, and then they both have the same value. Thus the two notions of limit here are compatible.

CHAPTER X

Differentiation of functions

10.1

Basic definitions

Exercise 10.1.1. *Suppose that X is a subset of \mathbf{R} , x_0 is a limit point of X , and $f : X \rightarrow \mathbf{R}$ is a function which is differentiable at x_0 . Let $Y \subseteq X$ be such that $x_0 \in Y$, and x_0 is also a limit point of Y . Prove that the restricted function $f|_Y : Y \rightarrow \mathbf{R}$ is also differentiable at x_0 , and has the same derivative as f at x_0 . Explain why this does not contradict the discussion in Remark 10.1.2.*

Proof. Suppose that L is the derivative of f at x_0 on X . Since f is differentiable, by Definition 10.1.1, for every $\varepsilon > 0$, there exists an $\delta > 0$ such that

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - L \right| \leq \varepsilon$$

for all $x \in X \setminus \{x_0\}$ such that $|x - x_0| < \delta$. Then it is also hold for $x \in Y \setminus \{x_0\}$ such that $|x - x_0| < \delta$ for that $Y \subseteq X$. Thus by Definition 10.1.1, $f|_Y$ is differentiable at x_0 on Y . Since x_0 is arbitrary, there is no isolated point of $Y \subseteq X$. \square

Exercise 10.1.2. *Prove Proposition 10.1.7. (Hint: the cases $x = x_0$ and $x \neq x_0$ have to be treated separately.)*

Proposition 10.1.7 (Newton's approximation). Let X be a subset of \mathbf{R} , let $x_0 \in X$ be a limit point of X , let $f : X \rightarrow \mathbf{R}$ be a function, and let L be a real number. Then the following statements are logically equivalent:

- (a) f is differentiable at x_0 on X with derivative L .
- (b) For every $\varepsilon > 0$, there exists a $\delta > 0$ such that $f(x)$ is $\varepsilon|x-x_0|$ -close to $f(x_0) + L(x-x_0)$ whenever $x \in X$ is δ -close to x_0 , i.e., we have

$$|f(x) - (f(x_0) + L(x-x_0))| \leq \varepsilon|x-x_0|$$

whenever $x \in X$ and $|x - x_0| \leq \delta$.

Proof. (a) \Rightarrow (b). Suppose (a). By Definition 10.1.1, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that's

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - L \right| \leq \varepsilon$$

for all $x \in X \setminus \{x_0\}$ such that $|x - x_0| < \delta$. If $x \neq x_0$, we can implies that

$$|f(x) - (f(x_0) + L(x - x_0))| \leq \varepsilon|x - x_0|.$$

While if $x = x_0$,

$$|f(x) - (f(x_0) + L(x - x_0))| \leq \varepsilon|x - x_0|$$

is obviously hold. Thus $f(x)$ is $\varepsilon|x-x_0|$ -close to $f(x_0) + L(x-x_0)$ whenever $x \in X$ is δ -close to x_0 .

(b) \Rightarrow (a). Suppose (b), for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - (f(x_0) + L(x - x_0))| \leq \varepsilon|x - x_0|.$$

Since $x - x_0 \neq 0$, we can implies that

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - L \right| \leq \varepsilon$$

for all $x \in X \setminus \{x_0\}$ such that x and x_0 are δ -closed. Thus by Definition 10.1.1, f is differentiable at x_0 on X with derivative L . \square

Exercise 10.1.3. *Prove Proposition 10.1.10. (Hint: either use the limit laws (Proposition 9.3.14), or use Proposition 10.1.7.)*

Proposition 10.1.10 (Differentiability implies continuity). Let X be a subset of \mathbf{R} , let $x_0 \in X$ be a limit point of X , and let $f : X \rightarrow \mathbf{R}$ be a function. If f is differentiable at x_0 , then f is also continuous at x_0 .

Proof. Since f is differentiable at x_0 , by Proposition 10.1.7, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - (f(x_0) + L(x - x_0))| \leq \varepsilon|x - x_0|,$$

so that

$$|f(x) - f(x_0)| \leq (\varepsilon + |L|)|x - x_0|$$

for all $x \in X$ such that $|x - x_0| \leq \delta$. Let $\zeta = (\varepsilon + |L|)|x - x_0| > 0$. Thus for every $\zeta > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| \leq \zeta$ for all $x \in X$ such that $|x - x_0| \leq \delta$. Thus f is continuous at x_0 . \square

Exercise 10.1.4. *Prove Theorem 10.1.13. (Hint: use the limit laws in Proposition 9.3.14. Use earlier parts of this theorem to prove the latter. For the product rule, use the identity*

$$\begin{aligned} f(x)g(x) - f(x_0)g(x_0) &= f(x)g(x) - f(x)g(x_0) + f(x)g(x_0) - f(x_0)g(x_0) \\ &= f(x)(g(x) - g(x_0)) + (f(x) - f(x_0))g(x_0); \end{aligned}$$

this trick of adding and subtracting an intermediate term is sometimes known as the “middle-man trick” and is very useful in analysis.)

Theorem 10.1.13 (Differential calculus). Let X be a subset of \mathbf{R} , let $x_0 \in X$ be a limit point of X , and let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions.

- (a) If f is a constant function, i.e., there exists a real number c such that $f(x) = c$ for all $x \in X$, then f is differentiable at x_0 and $f'(x_0) = 0$.
- (b) If f is the identity function, i.e., $f(x) = x$ for all $x \in X$, then f is differentiable at x_0 and $f'(x_0) = 1$.

- (c) (Sum rule) If f and g are differentiable at x_0 , then $f + g$ is also differentiable at x_0 , and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.
- (d) (Product rule) If f and g are differentiable at x_0 , then fg is also differentiable at x_0 , and $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.
- (e) If f is differentiable at x_0 and c is a real number, then cf is also differentiable at x_0 , and $(cf)'(x_0) = cf'(x_0)$.
- (f) (Difference rule) If f and g are differentiable at x_0 , then $f - g$ is also differentiable at x_0 , and $(f - g)'(x_0) = f'(x_0) - g'(x_0)$.
- (g) If g is differentiable at x_0 , and g is non-zero on X (i.e., $g(x) \neq 0$ for all $x \in X$), then $1/g$ is also differentiable at x_0 , and $(\frac{1}{g})'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}$.
- (h) (Quotient rule) If f and g are differentiable at x_0 , and g is non-zero on X , then f/g is also differentiable at x_0 , and

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$$

Proof.

- (a) Since $f(x) = c$ for all $x \in X$, then for $x_0 \in X$, we have

$$\lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0.$$

Thus f is differentiable at x_0 and $f'(x_0) = 0$.

- (b) Since $f(x) = x$ for all $x \in X$, then for $x_0 \in X$, we have

$$\lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1.$$

Thus f is differentiable at x_0 and $f'(x_0) = 1$.

- (c) By limit laws, we have

$$\begin{aligned} & \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \left(\frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \right) \\ &= \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{g(x) - g(x_0)}{x - x_0} \end{aligned}$$

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

Thus $f + g$ is differentiable at x_0 and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.

(d) By limit laws, we have

$$\begin{aligned} & \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x)(g(x) - g(x_0)) + (f(x) - f(x_0))g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \left(\frac{g(x) - g(x_0)}{x - x_0} f(x) + \frac{f(x) - f(x_0)}{x - x_0} g(x_0) \right) \\ &= g'(x_0)f(x_0) + f'(x_0)g(x_0). \end{aligned}$$

Thus fg is differentiable at x_0 and

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

(e) By limit laws, we have

$$\lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{cf(x) - cf(x_0)}{x - x_0} = c \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = cf'(x_0).$$

Thus cf is differentiable at x_0 , and $(cf)'(x_0) = cf'(x_0)$.

(f) By Theorem 10.1.13(c) and (e), let $c = -1$, then $f - g$ is differentiable at x_0 , and $(f - g)'(x_0) = f'(x_0) - g'(x_0)$.

(g) By limit laws, we have

$$\begin{aligned} & \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{1/g(x) - 1/g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \left(-\frac{g(x) - g(x_0)}{g(x)g(x_0)(x - x_0)} \right) \\ &= -\frac{g'(x_0)}{g(x_0)^2}. \end{aligned}$$

Thus $1/g$ is also differentiable at x_0 , and $(\frac{1}{g})'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}$.

(h) By Theorem 10.1.13(d) and (g), we have

$$\begin{aligned} \left(\frac{f}{g}\right)'(x_0) &= f'(x_0) \left(\frac{1}{g}\right)'(x_0) + f(x_0) \left(\frac{1}{g}\right)'(x_0) \\ &= \frac{f'(x_0)}{g(x_0)} - \frac{f(x_0)g'(x_0)}{g(x_0)^2} \\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}, \end{aligned}$$

as desired. □

Exercise 10.1.5. Let n be a natural number, and let $f : \mathbf{R} \rightarrow \mathbf{R}$ be the function $f(x) := x^n$. Show that f is differentiable on \mathbf{R} and $f'(x) = nx^{n-1}$ for all $x \in \mathbf{R}$. (Hint: use Theorem 10.1.13 and induction.)

Proof. We use induction on n . Base case $n = 0$ is trivial and $n = 1$ is true by Theorem 10.1.13. Suppose inductively that $f(x) = x^n$ and f is differentiable on \mathbf{R} and $f'(x) = nx^{n-1}$. We need to show the case of $n + 1$. Since

$$\begin{aligned} \frac{x^{n+1} - x_0^{n+1}}{x - x_0} &= \frac{x}{x - x_0} \left(x^n - \frac{x_0^{n+1}}{x} \right) \\ &= \frac{x}{x - x_0} \left(x^n - x_0^n + x_0^n - \frac{x_0^{n+1}}{x} \right) \\ &= \frac{x^n - x_0^n}{x - x_0} x + x_0^n. \end{aligned}$$

By induction hypothesis

$$\begin{aligned} \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{x^{n+1} - x_0^{n+1}}{x - x_0} &= \lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \left(\frac{x^n - x_0^n}{x - x_0} x + x_0^n \right) \\ &= nx_0^{n-1} \cdot x_0 + x_0^n \\ &= (n + 1)x_0^n \end{aligned}$$

for all $x_0 \in \mathbf{R}$. This closed the induction. □

Exercise 10.1.6. Let n be a **negative** integer, and let $f : \mathbf{R} \setminus \{0\} \rightarrow \mathbf{R}$ be the function $f(x) := x^n$. Show that f is differentiable on $\mathbf{R} \setminus \{0\}$ and $f'(x) = nx^{n-1}$ for all $x \in \mathbf{R} \setminus \{0\}$. (Hint: use Theorem 10.1.13 and Exercise 10.1.5.)

Proof. Since for every negative integer n , we have $f(x) = x^n = 1/x^{-n}$, where $-n$ is natural number. Let $m = -n$ and let $g : \mathbf{R} \setminus \{0\} \rightarrow \mathbf{R}$ be the function $g(x) := x^m$. By Exercise 10.1.5 and Theorem 10.1.13,

$$f'(x) = \left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2} = -\frac{mx^{m-1}}{x^{2m}} = -\frac{m}{x^{m+1}} = nx^{n-1},$$

as desired. □

Exercise 10.1.7. *Prove Theorem 10.1.15. (Hint one way to do this is via Newton's approximation, Proposition 10.1.7. Another way is to use Proposition 9.3.9 and Proposition 10.1.10 to convert this problem into one involving limits of sequences, however with the latter strategy one has to treat the case $f(x_0) = 0$ separately, as some division-by-zero subtleties can occur in that case.)*

Theorem 10.1.15 (Chain rule). Let X, Y be subsets of \mathbf{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 $f : X \rightarrow Y$ be a function such that $f(x_0) = y_0$, and such that f is differentiable at x_0 . Suppose that $g : Y \rightarrow \mathbf{R}$ is a function which is differentiable at y_0 . Then the function $g \circ f : X \rightarrow \mathbf{R}$ is differentiable at x_0 , and

$$(g \circ f)'(x_0) = g'(y_0)f'(x_0).$$

Proof. Since f is differentiable at x_0 , by Proposition 10.1.7, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - (f(x_0) + f'(x_0)(x - x_0))| \leq \varepsilon|x - x_0|$$

whenever $x \in X$ and $|x - x_0| \leq \delta$. Since g is differentiable at y_0 , by Proposition 10.1.7, for every $\zeta > 0$, there exists a $\varepsilon > 0$ such that

$$|g(y) - (g(y_0) + g'(y_0)(y - y_0))| \leq \zeta|y - y_0|$$

whenever $y \in Y$ and $|y - y_0| \leq \varepsilon$. Thus for every $\zeta > 0$, there exists a $\delta > 0$ such that

$$|g(f(x)) - (g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)))| \leq \zeta|f(x) - f(x_0)|$$

whenever $x \in X$ and $|x - x_0| \leq \delta$. Since

$$\begin{aligned} & |g(f(x)) - (g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)))| \\ &= |g(f(x)) - (g(f(x_0)) + g'(y_0)f'(x_0)(x - x_0)) \\ &\quad - g'(y_0)(f(x) - (f(x_0) + f'(x_0)(x - x_0)))| \\ &\geq |g(f(x)) - (g(f(x_0)) + g'(y_0)f'(x_0)(x - x_0))| \\ &\quad - |g'(y_0)(f(x) - (f(x_0) + f'(x_0)(x - x_0)))|. \end{aligned}$$

Thus we have

$$\begin{aligned} & |g(f(x)) - (g(f(x_0)) + g'(y_0)f'(x_0)(x - x_0))| \\ &\leq |g(f(x)) - (g(f(x_0)) + g'(y_0)(f(x) - f(x_0)))| \\ &\quad + |g'(y_0)(f(x) - (f(x_0) + f'(x_0)(x - x_0)))| \\ &\leq \zeta|f(x) - f(x_0)| + \varepsilon|g'(y_0)(x - x_0)| \\ &= (\zeta|f'(x_0)| + \varepsilon|g'(y_0)|)|x - x_0| \end{aligned}$$

whenever $|x - x_0| \leq \delta$ for $\delta = \zeta|f'(x_0)| + \varepsilon|g'(y_0)| > 0$. By Proposition 10.1.7, $g \circ f$ is differentiable at x_0 on X with derivative $g'(y_0)f'(x_0)$. \square

10.2

Local maxima, local minima, and derivatives

Exercise 10.2.1. Prove Proposition 10.2.6.

Proposition 10.2.6 (Local extrema are stationary). Let $a < b$ be real numbers, and let $f : (a, b) \rightarrow \mathbf{R}$ be a function. If $x_0 \in (a, b)$, f is differentiable at x_0 , and f attains either a local maximum or local minimum at x_0 , then $f'(x_0) = 0$.

Proof. Suppose that f attains local maximum at x_0 . Then there exists a $\delta > 0$ such that $f(x) \leq f(x_0)$ for every $x \in (a, b) \cap (x_0 - \delta, x_0 + \delta)$. Since f is differentiable at x_0 , we have

$$\lim_{x \rightarrow x_0; x \in (a, b) \cap (x_0 - \delta, x_0) \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

and

$$\lim_{x \rightarrow x_0; x \in (a,b) \cap (x_0, x_0 + \delta) \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

Differentiability of f requires above two equations equal to each other, thus $f'(x) = 0$. A similar argument shows the case of minimum. \square

Exercise 10.2.2. Give an example of a function $f : (-1, 1) \rightarrow \mathbf{R}$ 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.

Proof. Consider the function $f : (-1, 1) \rightarrow \mathbf{R}$ which is defined as

$$f(x) := \begin{cases} x & \text{for } x < 0, \\ -x & \text{for } x \geq 0. \end{cases}$$

This function does not contradict Proposition 10.2.6, because differentiability of function is a necessary condition for that f attains its local maximum or local minimum x_0 . \square

Exercise 10.2.3. Give an example of a function $f : (-1, 1) \rightarrow \mathbf{R}$ 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.

Proof. Consider the function $f : (-1, 1) \rightarrow \mathbf{R}$ which is defined as $f(x) := x^3$. Since $f'(x_0) = 0$ is a necessary condition for that f attains its local maximum or local minimum at x_0 , the function does not contradict Proposition 10.2.6. \square

Exercise 10.2.4. Prove Theorem 10.2.7. (Hint: use Corollary 10.1.12 and the maximum principle, Proposition 9.6.7, followed by Proposition 10.2.6. Note that the maximum principle does not tell you whether the maximum or minimum is in the open interval (a, b) or is one of the boundary points a, b , so you have to divide into cases and use the hypothesis $g(a) = g(b)$ somehow.)

Theorem 10.2.7 (Rolle's theorem). Let $a < b$ be real numbers, and let $g : [a, b] \rightarrow \mathbf{R}$ be a continuous function which is differentiable on (a, b) . Suppose also that $g(a) = g(b)$. Then there exists an $x \in (a, b)$ such that $g'(x) = 0$.

Proof. Since f is continuous on $[a, b]$, by maximum principle (Proposition 9.6.7), f attains its maximum and minimum at some points $x_{\max} \in [a, b]$ and $x_{\min} \in [a, b]$ respectively. If x_{\max}, x_{\min} are boundary points of $[a, b]$, then $g(x_{\max}) = g(x_{\min})$, this means that g takes constant on (a, b) . Thus for all $x \in (a, b)$, $g'(x) = 0$. Otherwise, at least one of x_{\max} or x_{\min} lies in (a, b) which such that g attains its maximum or minimum, denote it by $x \in (a, b)$. By Proposition 10.2.6, $g'(x) = 0$. \square

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 .)

Corollary 10.2.9 (Mean value theorem). Let $a < b$ be real numbers, and let $f : [a, b] \rightarrow \mathbf{R}$ be a function which is continuous on $[a, b]$ and differentiable on (a, b) . Then there exists an $x \in (a, b)$ such that $f'(x) = \frac{f(b)-f(a)}{b-a}$.

Proof. Consider the function $g : [a, b] \rightarrow \mathbf{R}$, which is defined as

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}x.$$

By Proposition 9.4.9 and Theorem 10.1.13, g is continuous on $[a, b]$ and differentiable on (a, b) . Since $g(a) = g(b)$, by Theorem 10.2.7, there exists an $x \in (a, b)$ such that

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} = 0.$$

This implies that $f'(x) = \frac{f(b)-f(a)}{b-a}$. \square

Exercise 10.2.6. Let $M > 0$, and let $f : [a, b] \rightarrow \mathbf{R}$ be a function which is continuous on $[a, b]$ and differentiable on (a, b) , and such that $|f'(x)| < M$ for all $x \in (a, b)$ (i.e., the derivative of f is bounded). Show that for any

$x, y \in [a, b]$ we have the inequality $|f(x) - f(y)| \leq M|x - y|$. (Hint: apply the mean value theorem (Corollary 10.2.9) to a suitable restriction of f .) Functions which obey the bound $|f(x) - f(y)| \leq M|x - y|$ are known as **Lipschitz continuous functions** with **Lipschitz constant** M ; thus this exercise shows that functions with bounded derivative are Lipschitz continuous.

Proof. For any $x, y \in [a, b]$, we either have $x < y$ or $y < x$ ($x = y$ is trivial). We only consider $x < y$, and $y < x$ is similar to prove. Suppose that $x < y$, we can see that f is also continuous on $[x, y]$ and differentiable on (x, y) . By mean value theorem, there exists an $x_0 \in (x, y)$ such that

$$f'(x_0) = \frac{f(x) - f(y)}{x - y}.$$

Since $|f'(x)| \leq M$ for all $x \in (a, b)$ for some $M > 0$, we have $|f'(x_0)| \leq M$. This implies that

$$|f(x) - f(y)| \leq M|x - y|.$$

Because $|f(x) - f(y)| = |f(y) - f(x)|$ and $|x - y| = |y - x|$, we can similarly to prove the case of $y < x$. \square

Exercise 10.2.7. Let $f : \mathbf{R} \rightarrow \mathbf{R}$ be a differentiable function such that f' is bounded. Show that f is uniformly continuous. (Hint: use the preceding exercise.)

Proof. Let $\varepsilon > 0$ and $\delta > 0$. Suppose that $|f'(x)| \leq \varepsilon/\delta$ for all $x \in \mathbf{R}$. Then for any $x, y \in \mathbf{R}$, we have $|f(x) - f(y)| \leq \frac{\varepsilon}{\delta}|x - y|$, by Exercise 10.2.6. Then by definition, f is uniformly continuous. \square

10.3

Monotone functions and derivatives

Exercise 10.3.1. Prove Proposition 10.3.1.

Proposition 10.3.1. Let X be a subset of \mathbf{R} , let $x_0 \in X$ be a limit point of X , and let $f : X \rightarrow \mathbf{R}$ be a function. If f is monotone increasing and f is differentiable at x_0 , then $f'(x_0) \geq 0$. If f is monotone decreasing and f is differentiable at x_0 , then $f'(x_0) \leq 0$.

Proof. Since f is monotone increasing, for any $x \in X$ where $x \geq x_0$, we have $f(x) \geq f(x_0)$; and any $x \in X$ such that $x \leq x_0$, we have $f(x) \leq f(x_0)$. Since f is differentiable at x_0 , we have

$$\lim_{x \rightarrow x_0; x \in X \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

Because $f(x) - f(x_0)$ and $x - x_0$ take same symbol, $f'(x_0) \geq 0$. A similar argument show that $f'(x_0) \leq 0$ when f is monotone decreasing. \square

Exercise 10.3.2. Give an example of a function $f : (1, 1) \rightarrow \mathbf{R}$ which is continuous and monotone increasing, but which is not differentiable at 0. Explain why this does not contradict Proposition 10.3.1.

Proof. Consider the function $f : (-1, 1) \rightarrow \mathbf{R}$ which is defined as

$$f(x) := \begin{cases} x & \text{for } x \leq 0, \\ 2x & \text{for } x > 0. \end{cases}$$

We can see that $f'(0-) = 1$, $f'(0+) = 2$, and $f'(0-) \neq f'(0+)$. Proposition is not be contradicted for that differentiability of f is a sufficient condition but not a necessary condition. \square

Exercise 10.3.3. Give an example of a function $f : \mathbf{R} \rightarrow \mathbf{R}$ 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. (Hint: look at Exercise 10.2.3.)

Proof. Consider the function $f : \mathbf{R} \rightarrow \mathbf{R}$ which is defined as $f(x) := x^3$. this does not contradict Proposition 10.3.1 obviously, and $f'(x) > 0$ just a sufficient condition not necessary condition. \square

Exercise 10.3.4. Prove Proposition 10.3.3. (Hint: you do not have in-

tegrals or the fundamental theorem of calculus yet,, so these tools cannot be used. However, one can proceed via the mean-value theorem, Corollary 10.2.9.)

Proposition 10.3.3. Let $a < b$, and let $f : [a, b] \rightarrow \mathbf{R}$ be a differentiable function. If $f'(x) > 0$ for all $x \in [a, b]$, then f is strictly monotone increasing. If $f'(x) < 0$ for all $x \in [a, b]$, then f is strictly monotone decreasing. If $f'(x) = 0$ for all $x \in [a, b]$, then f is a constant function.

Proof. Suppose that $f'(x) > 0$ for all $x \in [a, b]$. For any two distinct points $x_1, x_2 \in [a, b]$ where $x_1 < x_2$. Then we have $[x_1, x_2] \subseteq [a, b]$. By Exercise 10.1.1, $f|_{[x_1, x_2]} : [x_1, x_2] \rightarrow \mathbf{R}$ is differentiable on $[x_1, x_2]$. Then by Corollary 10.1.12, $f|_{[x_1, x_2]}$ is continuous on $[x_1, x_2]$. Thus by Corollary, there exists an $x \in [x_1, x_2]$ such that $f'(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$. Since $f'(x) > 0$ for all $x \in [a, b]$, $f(x_2) - f(x_1)$ and $x_2 - x_1$ must taking same symbol. Hence f is strictly increasing. This is similar to show the remaining two statements. \square

Exercise 10.3.5. Give an example of a subset $X \subset \mathbf{R}$ and a function $f : X \rightarrow \mathbf{R}$ 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?)

Proof. Consider the function $f : \mathbf{R} \rightarrow \mathbf{R}$ which defined as $f(x) := 2x^2 + x^3$. Let $X := (-\infty, -4/3) \cup (0, +\infty)$ be a subset of \mathbf{R} . Then we can see that $f|_X$ is differentiable on X and is such that $f'(x) > 0$, but is not strictly monotone increasing. \square

10.4

Inverse functions and derivatives

Exercise 10.4.1. Let $n \geq 1$ be a natural number, and let $g : (0, \infty) \rightarrow (0, \infty)$ be the function $g(x) := x^{1/n}$.

- (a) Show that g is continuous on $(0, \infty)$. (Hint: use Proposition 9.8.3.)
- (b) Show that g is differentiable on $(0, \infty)$, and that $g'(x) = \frac{1}{n}x^{\frac{1}{n}-1}$ for all $x \in (0, \infty)$. (Hint: use the inverse function theorem and (a).)

Proof. (a) Since the function $f : (0, \infty) \rightarrow (0, \infty)$ which is defined as $f(y) := y^n$ is invertible, continuous and strictly monotone increasing. We have $f^{-1}(x) = g(x) = x^{1/n}$, we denote $f = g^{-1}$. By Proposition 9.8.3, g is also continuous.

(b) Since g^{-1} is differentiable on $(0, \infty)$, and g is continuous, by Theorem 10.4.2, we have

$$g'(x) = \frac{1}{(g^{-1})'(y)} = \frac{1}{ny^{n-1}} = \frac{1}{n} \frac{1}{x^{\frac{n-1}{n}}} = \frac{1}{n} x^{\frac{1}{n}-1}$$

as desired. □

Exercise 10.4.2. Let q be a rational number, and let $f : (0, \infty) \rightarrow \mathbf{R}$ be the function $f(x) = x^q$.

- (a) Show that f is differentiable on $(0, \infty)$ and that $f'(x) = qx^{q-1}$. (Hint use Exercise 10.4.1 and the laws of differential calculus in Theorem 10.1.13 and Theorem 10.1.15.)
- (b) Show that $\lim_{x \rightarrow 1; x \in (0, \infty) \setminus \{1\}} \frac{x^q - 1}{x - 1} = q$ for every rational number q . (Hint use part (a) and Definition 10.1.1. An alternate route is to apply L'Hôpital's rule from the next section.)

Proof. (a) Let $q = n/m$ for some natural number $n, m \in \mathbf{N}$. Then $f(x) = x^{\frac{n}{m}} = (x^{\frac{1}{m}})^n$. By Exercise 10.4.1 and Theorem 10.1.13, we can see that f is differentiable. By Theorem 10.1.15,

$$\frac{df(x)}{dx} = \frac{d(x^{\frac{1}{m}})^n}{dx} \cdot \frac{dx^{\frac{1}{m}}}{dx} = n(x^{\frac{1}{m}})^{n-1} \cdot \frac{1}{m} x^{\frac{1}{m}-1} = \frac{n}{m} x^{\frac{n-1}{m} + \frac{1-m}{m}} = qx^{q-1}.$$

(b) Given the function $f(x) = x^q$ following part (a) and consider that $f'(1)$. By Definition 10.1.1, we can see that

$$\lim_{x \rightarrow 1; x \in (0, \infty) \setminus \{1\}} \frac{f(x) - f(1)}{x - 1} = \frac{x^q - 1}{x - 1}.$$

By the part (a), we have $f'(1) = q$, as desired. \square

Exercise 10.4.3. Let α be a real number, and let $f : (0, \infty) \rightarrow \mathbf{R}$ be the function $f(x) = x^\alpha$.

- (a) Show that $\lim_{x \rightarrow 1; x \in (0, \infty) \setminus \{1\}} \frac{f(x) - f(1)}{x - 1} = \alpha$. (Hint: use Exercise 10.4.2 and the comparison principle; you may need to consider right and left limits separately. Proposition 5.4.14 may also be helpful.)
- (b) Show that f is differentiable on $(0, \infty)$ and that $f'(x) = \alpha x^{\alpha-1}$. (Hint: use (a), exponent laws (Proposition 6.7.3), and Definition 10.1.1.)

Proof. (a) Let $\varepsilon > 0$. By Proposition 5.4.14, for any $\alpha, \alpha - \varepsilon \in \mathbf{R}$ and $\alpha, \alpha + \varepsilon \in \mathbf{R}$, there exists a rational such that

$$\alpha - \varepsilon < p < \alpha, \quad \alpha < q < \alpha + \varepsilon,$$

respectively. Thus there are rationals such that $p < \alpha < q$.

For $0 < x < 1$, we have

$$\frac{x^p - 1}{x - 1} < \frac{x^\alpha - 1}{x - 1} < \frac{x^q - 1}{x - 1}.$$

By Exercise 10.4.2,

$$\lim_{x \rightarrow 1; x \in (0, 1)} \frac{x^p - 1}{x - 1} \leq \lim_{x \rightarrow 1; x \in (0, 1)} \frac{x^\alpha - 1}{x - 1} \leq \lim_{x \rightarrow 1; x \in (0, 1)} \frac{x^q - 1}{x - 1}$$

implies that $\alpha - \varepsilon < \lim_{x \rightarrow 1; x \in (0, 1)} \frac{x^\alpha - 1}{x - 1} < \alpha + \varepsilon$.

For $x > 1$, we still have

$$\frac{x^p - 1}{x - 1} < \frac{x^\alpha - 1}{x - 1} < \frac{x^q - 1}{x - 1}.$$

With similar process, $\alpha - \varepsilon < \lim_{x \rightarrow 1; x \in (1, \infty)} \frac{x^\alpha - 1}{x - 1} < \alpha + \varepsilon$.

Thus

$$\alpha - \varepsilon < \lim_{x \rightarrow 1; x \in (0, \infty) \setminus \{1\}} \frac{x^\alpha - 1}{x - 1} < \alpha + \varepsilon.$$

Since this holds for every $\varepsilon > 0$, we can know that $\lim_{x \rightarrow 1; x \in (0, \infty) \setminus \{1\}} \frac{x^\alpha - 1}{x - 1} = \alpha$.

(b) Since

$$\frac{x^\alpha - x_0^\alpha}{x - x_0} = x_0^{\alpha-1} \frac{\left(\frac{x}{x_0}\right)^\alpha - 1}{\frac{x}{x_0} - 1}.$$

We compute the limit, by the part (a)

$$\lim_{x \rightarrow x_0; x \in (0, \infty) \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = x_0^{\alpha-1} \lim_{x \rightarrow x_0; x \in (0, \infty) \setminus \{x_0\}} \frac{\left(\frac{x}{x_0}\right)^\alpha - 1}{\frac{x}{x_0} - 1} = \alpha x_0^{\alpha-1}.$$

Thus f is differentiable and its derivative is $\alpha x^{\alpha-1}$. \square

10.5

L'Hôpital's rule

Exercise 10.5.1. *Prove Proposition 10.5.1. (Hint: to show that $g(x) \neq 0$ near x_0 , you may wish to use Newton's approximation (Proposition 10.1.7). For the rest of the proposition, use limit laws, Proposition 9.3.14.)*

Proposition 10.5.1 (L'Hôpital's rule I). Let X be a subset of \mathbf{R} , let $f : X \rightarrow \mathbf{R}$ and $g : X \rightarrow \mathbf{R}$ be functions, and let $x_0 \in X$ be a limit point of X . Suppose that $f(x_0) = g(x_0) = 0$, that f and g are both differentiable at x_0 , but $g'(x_0) \neq 0$. Then there exists a $\delta > 0$ such that $g(x) \neq 0$ for all $x \in (X \cap (x_0 - \delta, x_0 + \delta)) \setminus \{x_0\}$, and

$$\lim_{x \rightarrow x_0; x \in (X \cap (x_0 - \delta, x_0 + \delta)) \setminus \{x_0\}} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}.$$

Proof. Since g is differentiable at $x_0 \in X$ and $g'(x_0) \neq 0$, by the Newton's approximation, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|g(x) - (g(x_0) + g'(x_0)(x - x_0))| = |g(x) - g'(x_0)(x - x_0)| \leq \varepsilon |x - x_0|.$$

whenever $x \in X$ such that $|x - x_0| \leq \delta$. In particular, for all $x \in (X \cap (x_0 - \delta, x_0 + \delta)) \setminus \{x_0\}$, we have $|x - x_0| > 0$. Then by triangle inequality

$$|g'(x_0)(x - x_0)| - |g(x)| \leq |g(x) - g'(x_0)(x - x_0)|$$

implies that

$$|g'(x_0)(x - x_0)| - \varepsilon |x - x_0| \leq |g(x)|.$$

Take $\varepsilon = |g'(x_0)(x - x_0)|/2$, then $g(x) > 0$ for that both of $g'(x_0)$ and $|x - x_0|$ are strictly greater than zero.

Now we consider the limit

$$\begin{aligned}
 & \lim_{x \rightarrow x_0; x \in (X \cap (x_0 - \delta, x_0 + \delta)) \setminus \{x_0\}} \frac{f(x)}{g(x)} \\
 &= \lim_{x \rightarrow x_0; x \in (X \cap (x_0 - \delta, x_0 + \delta)) \setminus \{x_0\}} \frac{(f(x) - f(x_0))/(x - x_0)}{(g(x) - g(x_0))/(x - x_0)} \\
 &= \frac{f'(x_0)}{g'(x_0)}.
 \end{aligned}$$

□

Exercise 10.5.2. *Explain why Example 1.2.12 does not contradict either of the propositions in this section.*

Example 1.2.12 (L'Hôpital's rule). ... But even when $f(x)$ and $g(x)$ do go to zero as $x \rightarrow x_0$ there is still a possibility for an incorrect conclusion. For instance, consider the limit

$$\lim_{x \rightarrow 0} \frac{x^2 \sin(x^{-4})}{x}.$$

Proof. Since $\sin(x^{-4})$ also is not differentiable at x_0 , the example doesn't satisfy the condition of L'Hôpital's rule. □

CHAPTER XI

The Riemann integral

11.1

Partitions

Exercise 11.1.1. *Prove Lemma 11.1.4. (Hint: in order to show that (a) implies (b) in the case when X is non-empty, consider the supremum and infimum of X .)*

Lemma 11.1.4. Let X be a subset of the real line. Then the following two statements are logically equivalent:

- (a) X is bounded and connected.
- (b) X is a bounded interval.

Proof. (a) \Rightarrow (b). If $X = \emptyset$, (b) is vacuously true. While if $X \neq \emptyset$. Suppose that X is bounded, then there exists an $M > 0$ such that $X \subseteq [-M, M]$. Consider the supremum and the infimum of X , denoted by $\sup(X)$ and $\inf(X)$, by definition, we have $X \subseteq [\inf(X), \sup(X)]$. This is easy to prove that $\sup(X)$ and $\inf(X)$ is an adherent point of X . Thus by Definition 9.1.15, we have $\sup(X) \in X$ and $\inf(X) \in X$. Since X is connected, we

can find $\sup(X), \inf(X) \in X$ such that $[\inf(X), \sup(X)] \subseteq X$. Thus $X = [\inf(X), \sup(X)]$. Let $N = \max(|\inf(X)|, |\sup(X)|)$, then X is bounded interval by N .

(b) \Rightarrow (a). X can be represented by $X := [\inf(X), \sup(X)]$ or $X := (\inf(X), \sup(X))$. Consider former, for every $x, y \in X$ where $x < y$, we must have $\inf(X) \leq x < y \leq \sup(X)$. Then $[x, y] \subseteq [\inf(X), \sup(X)]$. This is similar to show the case of $(\inf(X), \sup(X))$. Thus X is bounded and connected. \square

Exercise 11.1.2. *Prove Corollary 11.1.6. (Hint: use Lemma 11.1.4, and explain why the intersection of two bounded sets is automatically bounded, and why the intersection of two connected sets is automatically connected.)*

Corollary 11.1.6. If I and J are bounded intervals, then the intersection $I \cap J$ is also a bounded interval.

Proof. By Lemma 11.1.4, I and J are bounded and connected. Assume that I and J is bounded by M and N respectively. Then for every $x \in I$ and $y \in J$, we have $|x| \leq M$ and $|y| \leq N$. Thus for $z \in I \cap J$, we must have $|z| \leq \min(N, M)$. Thus $I \cap J$ is bounded. Since I and J are connected. Let $x, y \in I \cap J$ where $x < y$. Because $x, y \in I$ and $x, y \in J$ at the time, we have $[x, y] \subseteq I$ and $[x, y] \subseteq J$. Thus $[x, y] \subseteq I \cap J$. This means that $I \cap J$ is connected. By Lemma 11.1.4 again, $I \cap J$ is bounded interval. \square

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, \dots, 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 $a \leq c \leq b$. (Hint: prove by contradiction. First show that if I_j is **not** of the form (c, b) or $[c, b)$ for any $a \leq c \leq b$, then $\sup I_j$ is **strictly** less than b .)*

Proof. Suppose for sake of contradiction that for all partition I_j is not of the form (c, b) or $[c, b)$ for any $a \leq c \leq b$. Then we either have $I_j \subsetneq [c, b)$ or $[c, b) \subsetneq I_j \cup [c, b)$ for any $a \leq c \leq b$. Latter case would not hold for that there exists an $x \in I_j$ but $x \notin I$, a contradiction. Consider former one, suppose

that $I_j = [c, \sup(I_j)]$ or $I_j = [c, \sup(I_j))$ where $c \leq \sup(I_j) < b$. Then we have $(\sup(I_j), b) = [c, b) \setminus I_j$ or $(\sup(I_j), b) = [c, b) \setminus I_j$, a contradiction. Thus for any partition I_i is not of the form (c, b) or $[c, b)$, we always can find a partition $I_j := [c, b) \setminus I_i$ with the form (c, b) or $[c, b)$. \square

Exercise 11.1.4. Prove Lemma 11.1.18.

Lemma 11.1.18. Let I be a bounded interval and let \mathbf{P} and \mathbf{P}' be to partitions of I . Then $\mathbf{P} \# \mathbf{P}'$ is also a partition of I , and is both finer than \mathbf{P} and finer than \mathbf{P}' .

Proof. This is clear that $K \cap J$ contains in I for any $K \in \mathbf{P}$ and $J \in \mathbf{P}'$. Since \mathbf{P} and \mathbf{P}' both are a partition of I , for any $x \in I$, there exists a K and an J such that $x \in K$ and $x \in J$, thus $x \in K \cap J$. By Definition 11.1.10, $K \cap J$ is also a partition of I . For any $K \cap J \in \mathbf{P} \# \mathbf{P}'$ there exists $K \in \mathbf{P}$ and $J \in \mathbf{P}'$ such that $K \cap J \subseteq K$ and $K \cap J \subseteq J$. By Definition 11.1.14, $\mathbf{P} \# \mathbf{P}'$ is both finer than \mathbf{P} and finer than \mathbf{P}' . \square

11.2

Piecewise constant functions

Exercise 11.2.1. Prove Lemma 11.2.7.

Lemma 11.2.7. Let I be a bounded interval, let \mathbf{P} be a partition of I , and let $f : I \rightarrow \mathbf{R}$ be a function which is piecewise constant with respect to \mathbf{P} . Let \mathbf{P}' be a partition of I which is finer than \mathbf{P} . Then f is also piecewise constant with respect to \mathbf{P}' .

Proof. Since f is piecewise constant with respect to \mathbf{P} , by Definition 11.2.3, for every $K \in \mathbf{P}$, f is constant on K . Since \mathbf{P}' is finer than \mathbf{P} , for every $J \in \mathbf{P}'$ there exists a $K \in \mathbf{P}$ such that $J \subseteq K$. Thus by Definition 11.2.1, f is constant on J , so that f is also piecewise constant with respect to \mathbf{P}' . \square

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 .)

Lemma 11.2.8. Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ be piecewise constant functions on I . Then the functions $f + g$, $f - g$, $\max(f, g)$ and fg are also piecewise constant functions on I . Here of course $\max(f, g) : I \rightarrow \mathbf{R}$ is the function $\max(f, g)(x) := \max(f(x), g(x))$. If g does not vanish anywhere on I (i.e., $g(x) \neq 0$ for all $x \in I$) then f/g is also a piecewise constant function on I .

Proof. Suppose that $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ to be piecewise constant with respect to \mathbf{P} and \mathbf{P}' respectively. By Lemma 11.1.18, $\mathbf{P} \# \mathbf{P}'$ is finer than \mathbf{P} and \mathbf{P}' ; and by Lemma 11.2.7, f and g is piecewise constant with respect to $\mathbf{P} \# \mathbf{P}'$. Thus $f + g$, $f - g$, $\max(f, g)$, fg , $\max(f, g)$ and f/g ($g(x) \neq 0$ for all $x \in I$) are also piecewise constant functions on I are piecewise constant functions on I by Definition 11.2.5.

Exercise 11.2.3. Prove Proposition 11.2.13. (Hint: first use Theorem 11.1.13 to show that both integrals are equal to p.c. $\int_{[\mathbf{P} \# \mathbf{P}']} f$.)

Proposition 11.2.13 (Piecewise constant integral is independent of partition). Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ be a function. Suppose that \mathbf{P} and \mathbf{P}' are partitions of I such that f is piece-wise constant both with respect to \mathbf{P} and with respect to \mathbf{P}' . Then $p.c. \int_{[\mathbf{P}]} f = p.c. \int_{[\mathbf{P}']} f$.

Proof. By Lemma 11.2.7, f is piecewise constant with respect to $\mathbf{P} \# \mathbf{P}'$. Clearly, $\mathbf{P} \# \mathbf{P}'$ refines \mathbf{P} . Let $J \in \mathbf{P} \# \mathbf{P}'$ and $K \in \mathbf{P}$, let c_K be the constant value of f on \mathbf{P} . By Definition 11.2.9, we have

$$p.c. \int_{[\mathbf{P}]} f := \sum_{K \in \mathbf{P}} c_K |K|.$$

Note that for every $J \in \mathbf{P} \# \mathbf{P}'$ there exists $K \in \mathbf{P}$ such that $J \subseteq K$. Thus for every $K \in \mathbf{P}$, $\{J \in \mathbf{P} \# \mathbf{P}' : J \subseteq K\}$ is a partition of K . By Theorem

11.1.13

$$|K| = \sum_{J \in \mathbf{P} \# \mathbf{P}': J \subseteq K} |J|.$$

Hence

$$p.c. \int_{[\mathbf{P}]} f := \sum_{K \in \mathbf{P}} c_K \left(\sum_{J \in \mathbf{P} \# \mathbf{P}': J \subseteq K} |J| \right).$$

Since $f|_{J \subseteq K} = c_J = c_K$, we have

$$\begin{aligned} \sum_{K \in \mathbf{P}} c_K \left(\sum_{J \in \mathbf{P} \# \mathbf{P}': J \subseteq K} |J| \right) &= \sum_{K \in \mathbf{P}} \sum_{J \in \mathbf{P} \# \mathbf{P}': J \subseteq K} c_J |J| \\ &= \sum_{J \in \mathbf{P} \# \mathbf{P}'} c_J |J| = p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f. \end{aligned}$$

This shows that $p.c. \int_{[\mathbf{P}]} f = p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f$. This is similar to show that $p.c. \int_{[\mathbf{P}']} f = p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f$. Thus $p.c. \int_{[\mathbf{P}]} f = p.c. \int_{[\mathbf{P}']} f$. \square

Exercise 11.2.4. Prove Theorem 11.2.16. (Hint: you can use earlier parts of the theorem to prove some of the later parts of the theorem. See also the hint to Exercise 11.2.2.)

Theorem 11.2.16 (Laws of integration). Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ be piecewise constant functions on I .

- (a) We have $p.c. \int_I (f + g) = p.c. \int_I f + p.c. \int_I g$.
- (b) For any real number c , we have $p.c. \int_I (cf) = c(p.c. \int_I f)$.
- (c) We have $p.c. \int_I (f - g) = p.c. \int_I f - p.c. \int_I g$.
- (d) If $f(x) \geq 0$ for all $x \in I$, then $p.c. \int_I f \geq 0$.
- (e) If $f(x) \geq g(x)$ for all $x \in I$, then $p.c. \int_I f \geq p.c. \int_I g$.
- (f) If f is the constant function $f(x) = c$ for all x in I , then $p.c. \int_I f = c|I|$.

- (g) Let J be a bounded interval containing I (i.e., $I \subseteq J$) and let $F : J \rightarrow \mathbf{R}$ be the function

$$F(x) := \begin{cases} f(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases}$$

Then F is piecewise constant on J , and $p.c. \int_J F = p.c. \int_I f$.

- (h) Suppose that $\{J, K\}$ is a partition of I into two intervals J and K . Then the functions $f|_J : J \rightarrow \mathbf{R}$ and $f|_K : K \rightarrow \mathbf{R}$ are piecewise constant on J and K respectively, and we have

$$p.c. \int_I f = p.c. \int_J f|_J + p.c. \int_K f|_K.$$

Proof. Suppose that $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ to be piecewise constant with respect to \mathbf{P} and \mathbf{P}' respectively. By Lemma 11.1.18, $\mathbf{P} \# \mathbf{P}'$, denoted by \mathcal{P} , is finer than \mathbf{P} and \mathbf{P}' ; and by Lemma 11.2.7, f and g is piecewise constant with respect to \mathcal{P} .

- (a) Let c_J and c'_J be the constant value of f and g respectively on $J \in \mathcal{P}$. Then by Lemma 11.2.8, $c_J + c'_J$ is the constant value of $f + g$ on $J \in \mathcal{P}$. By Definition 11.2.9 and Definition 11.2.14, we have

$$\begin{aligned} p.c. \int_I (f + g) &= \sum_{J \in \mathcal{P}} (c_J + c'_J) |J| \\ &= \sum_{J \in \mathcal{P}} c_J |J| + \sum_{J \in \mathcal{P}} c'_J |J| \\ &= p.c. \int_I f + p.c. \int_I g. \end{aligned}$$

- (b) Let c_J be the constant value of f on $J \in \mathcal{P}$. Then by Lemma 11.2.8, $c \times c_J$ is the constant value of cf on $J \in \mathcal{P}$. By Definition 11.2.9 and Definition 11.2.14, we have

$$p.c. \int_I (cf) = \sum_{J \in \mathcal{P}} (c \times c_J) |J| = c \left(\sum_{J \in \mathcal{P}} c_J |J| \right) = c \left(p.c. \int_I f \right).$$

- (c) Consider $p.c. \int_I (f + cg)$ where $c = -1$, then by Theorem 11.2.16(a) and (b), we have $p.c. \int_I (f - g) = p.c. \int_I f - p.c. \int_I g$.

- (d) $f(x) \geq 0$ for all $x \in I$ implies that $c_J \geq 0$, which is the constant value of f on $J \in \mathcal{P}$. Thus by Definition 11.2.9 and Definition 11.2.14

$$p.c. \int_I f = p.c. \int_{[\mathcal{P}]} f = \sum_{J \in \mathcal{P}} c_J |J| \geq 0.$$

- (e) Let c_J be the constant value of f on $J \in \mathcal{P}$. $f(x) \geq g(x)$ for all $x \in I$ implies that $f(x) - g(x) \geq 0$, then by Theorem 11.2.4(c) and (d), $p.c. \int_I (f - g) \geq 0$ implies $p.c. \int_I f \geq p.c. \int_I g$.

- (f) By Definition 11.2.9, Definition 11.2.14 and Theorem 11.1.13

$$p.c. \int_I f = \sum_{J \in \mathcal{P}} c |J| = c \left(\sum_{J \in \mathcal{P}} |J| \right) = c |I|.$$

- (g) Let \mathbf{U} be a partition of I , and \mathbf{V} be a partition of $J \setminus I$, then $\mathbf{U} \cup \mathbf{V}$ is a partition of J . Let c_K be the constant value of F on $K \in \mathbf{U} \cup \mathbf{V}$. Then

$$\begin{aligned} p.c. \int_J F &= \sum_{K \in \mathbf{U} \cup \mathbf{V}} c_K |K| \\ &= \sum_{K \in \mathbf{U}} c_K |K| + \sum_{K \in \mathbf{V}} 0 \times |J \setminus I| \\ &= \sum_{K \in \mathbf{U}} c_K |K| \\ &= p.c. \int_I f. \end{aligned}$$

- (h) We can see that I is a bounded interval containing J and K , i.e., $J \subseteq I$ and $K \subseteq I$. Since for every $x \in I$ either $x \in J$ or $x \in K$, we have

$$f(x) = \begin{cases} f|_J(x) & \text{if } x \in J \\ f|_K(x) & \text{if } x \in K \end{cases}$$

By (g), we have $p.c. \int_I f = p.c. \int_J f|_J + p.c. \int_K f|_K$. □

11.3

Upper and lower Riemann integrals

Exercise 11.3.1. Let $f : I \rightarrow \mathbf{R}$, $g : I \rightarrow \mathbf{R}$, and $h : I \rightarrow \mathbf{R}$ 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.

Proof. Since f majorizes g and g majorizes h , by Definition 11.3.1, we have $f(x) \geq g(x)$ and $g(x) \geq h(x)$ for all $x \in I$. Thus $f(x) \geq h(x)$ for all $x \in I$. By Definition 11.3.1 again, f majorizes h . If f and g majorize each other, we have $f(x) \geq g(x)$ and $g(x) \geq f(x)$ for all $x \in I$. Thus $f(x) = g(x)$ for all $x \in I$. \square

Exercise 11.3.2. Let $f : I \rightarrow \mathbf{R}$, $g : I \rightarrow \mathbf{R}$, and $h : I \rightarrow \mathbf{R}$ be functions. If f majorizes g , is it true that $f + h$ majorizes $g + h$? Is it true that $f \cdot h$ majorizes $g \cdot h$? If c is a real number, is it true that cf majorizes cg ?

Proof. $f + h$ majorizes $g + h$ is hold by addition. Let h to be a function by $h(x) = -1$ for all $x \in I$, then $f \cdot h$ minorizes $g \cdot h$. Let $c = -1$, then cf minorizes cg . \square

Exercise 11.3.3. Prove Lemma 11.3.7.

Lemma 11.3.7. Let $f : I \rightarrow \mathbf{R}$ be a piecewise constant function on a bounded interval I . Then f is Riemann integrable, and $\int_I f = p.c. \int_I f$.

Proof. Since f is a piecewise constant function on I which majorizes and minorizes it self. By Definition 11.3.2, we have

$$\overline{\int}_I f \leq p.c. \int_I f \leq \underline{\int}_I f.$$

Other hand, by Lemma 11.3.3

$$\underline{\int}_I f \leq \overline{\int}_I f.$$

Thus we have

$$p.c. \int_I f = \overline{\int_I f} = \underline{\int_I f}.$$

By Definition 11.3.4, f is Riemann integrable, and $\int_I f = p.c. \int_I f$. \square

Exercise 11.3.4. Prove Lemma 11.3.11.

Lemma 11.3.11. Let $f : I \rightarrow \mathbf{R}$ be a bounded function on a bounded interval I , and let g be a function which majorizes f and which is piecewise constant with respect to some partition \mathbf{P} of I . Then

$$p.c. \int_I g \geq U(f, \mathbf{P}).$$

Similarly, if h is a function which minorizes f and is piecewise constant with respect to \mathbf{P} , then

$$p.c. \int_I h \leq L(f, \mathbf{P}).$$

Proof. Since g majorizes f , we have $g(x) \geq f(x)$ for all $x \in I$. Let c_J be a constant value of g on J , so that $c_J \geq \sup_{x \in J} f(x)$ for all $x \in J \subseteq I$. By definition

$$p.c. \int_I g = \sum_{J \in \mathbf{P}: J \neq \emptyset} c_J |J| \geq \sum_{J \in \mathbf{P}: J \neq \emptyset} \left(\sup_{x \in J} f(x) \right) |J| = U(f, \mathbf{P}).$$

This is similar to show the second assertion. \square

Exercise 11.3.5. Prove Proposition 11.3.12. (Hint: you will need Lemma 11.3.11, even though this Lemma will only do half of the job.)

Proposition 11.3.12. Let $f : I \rightarrow \mathbf{R}$ be a bounded function on a bounded interval I . Then

$$\overline{\int_I f} = \inf \{ U(f, \mathbf{P}) : \mathbf{P} \text{ is a partition of } I \}$$

and

$$\underline{\int_I f} = \sup \{ L(f, \mathbf{P}) : \mathbf{P} \text{ is a partition of } I \}.$$

Proof. Let g be any piecewise constant function majorizing f , and let

h be any piecewise constant function minorizing f . Taking infima in g and $U(f, \mathbf{P})$, by Lemma 11.3.11, we obtain that $\overline{\int}_I f \geq \inf(U(f, \mathbf{P}))$. Taking suprema in h and $L(f, \mathbf{P})$, by Lemma 11.3.11, we obtain that $\underline{\int}_I f \leq \sup(L(f, \mathbf{P}))$. This is the half of our work.

For the remaining half part. Suppose for sake of contradiction that $\overline{\int}_I f > \inf(U(f, \mathbf{P}))$. Then there exists a \mathbf{P}' such that

$$p.c. \int_I g \geq \overline{\int}_I f > \sum_{J \in \mathbf{P}': J \neq \emptyset} (\sup_{x \in J} f(x)) |J|.$$

Since $\sup_{x \in J} f(x) \geq f(x)$ for all $x \in J$, by Definitions 11.2.9, 11.2.14 and 11.3.2, we have

$$\sum_{J \in \mathbf{P}': J \neq \emptyset} (\sup_{x \in J} f(x)) |J| \geq \overline{\int}_I f,$$

a contradiction. This is similar to show that $\underline{\int}_I f \geq \sup(L(f, \mathbf{P}))$. Together our conclusions, we have $\overline{\int}_I f = \inf(U(f, \mathbf{P}))$ and $\underline{\int}_I f = \sup(L(f, \mathbf{P}))$. \square

11.4

Basic properties of the Riemann integral

Exercise 11.4.1. *Prove Theorem 11.4.1. (Hint: you may find Theorem 11.2.16 to be useful. For part (b): First do the case $c > 0$. Then do the case $c = -1$ and $c = 0$ separately. Using these cases, deduce the case of $c < 0$. You can use earlier parts of the theorem to prove later ones.)*

Theorem 11.4.1 (Laws of Riemann integration). Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ be Riemann integrable functions on I .

- (a) The function $f + g$ is Riemann integrable, and we have $\int_I (f + g) = \int_I f + \int_I g$.
- (b) For any real number c , the function cf is Riemann integrable, and we have $\int_I (cf) = c(\int_I f)$.

- (c) The function $f + g$ is Riemann integrable, we have $\int_I (f - g) = \int_I f - \int_I g$.
- (d) If $f(x) \geq 0$ for all $x \in I$, then $\int_I f \geq 0$.
- (e) If $f(x) \geq g(x)$ for all $x \in I$, then $\int_I f \geq \int_I g$.
- (f) If f is the constant function $f(x) = c$ for all x in I , then $\int_I f = c|I|$.
- (g) Let J be a bounded interval containing I (i.e., $I \subseteq J$) and let $F : J \rightarrow \mathbf{R}$ be the function

$$F(x) := \begin{cases} f(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases}$$

Then F is Riemann integrable on J , and $\int_J F = \int_I f$.

- (h) Suppose that $\{J, K\}$ is a partition of I into two intervals J and K . Then the functions $f|_J : J \rightarrow \mathbf{R}$ and $f|_K : K \rightarrow \mathbf{R}$ are Riemann integrable on J and K respectively, and we have

$$\int_I f = \int_J f|_J + \int_K f|_K.$$

Proof. Let $\varepsilon > 0$. Since f and g are Riemann integrable, there exists piecewise constant functions $\overline{f}, \overline{g}$ which majorize f, g respectively on I such that

$$\int_I \overline{f} \leq \int_I f + \varepsilon, \quad \int_I \overline{g} \leq \int_I g + \varepsilon.$$

Similarly we can find piecewise constant functions $\underline{f}, \underline{g}$ which minorize f, g respectively on I such that

$$\int_I \underline{f} \geq \int_I f - \varepsilon, \quad \int_I \underline{g} \geq \int_I g - \varepsilon.$$

- (a) By Exercise 11.3.2, we can see that $\overline{f} + \overline{g}$ and $\underline{f} + \underline{g}$ are piecewise constant functions on I , and majorizes and minorizes $f + g$ respectively.

Thus

$$\int_I (\underline{f} + \underline{g}) \leq \int_I (f + g) \leq \int_I (\overline{f} + \overline{g}),$$

and so

$$0 \leq \overline{\int_I (f + g)} - \underline{\int_I (f + g)} \leq \int_I (\overline{f} + \overline{g}) - \int_I (\underline{f} + \underline{g}).$$

By Theorem 11.2.16(a), we have

$$\int_I (\overline{f} + \overline{g}) = \int_I \overline{f} + \int_I \overline{g} \leq \int_I f + \int_I g + 2\varepsilon$$

and

$$\int_I (\underline{f} + \underline{g}) = \int_I \underline{f} + \int_I \underline{g} \geq \int_I f + \int_I g - 2\varepsilon.$$

This implies that

$$\int_I (\overline{f} + \overline{g}) - \int_I (\underline{f} + \underline{g}) \leq 4\varepsilon.$$

To summarize, we have shown that for every $\varepsilon > 0$

$$0 \leq \overline{\int_I (f + g)} - \underline{\int_I (f + g)} \leq 4\varepsilon.$$

Since $\overline{\int_I (f + g)} - \underline{\int_I (f + g)}$ does not depend on ε , we thus see that

$$\overline{\int_I (f + g)} - \underline{\int_I (f + g)} = 0$$

and hence that $f + g$ is Riemann integrable. Since we have

$$\int_I f + \int_I g - 2\varepsilon \leq \int_I (f + g) \leq \int_I f + \int_I g + 2\varepsilon,$$

for every $\varepsilon > 0$, we thus see that

$$\int_I (f + g) = \int_I f + \int_I g.$$

- (b) Let $c > 0$. By Exercise 11.3.2, we can see that $c\overline{f}$ and $c\underline{f}$ are piecewise constant functions on I , and majorizes and minorizes cf respectively. Thus

$$\int_I c\underline{f} \leq \int_I cf \leq \int_I c\overline{f},$$

and so

$$0 \leq \overline{\int_I (cf)} - \underline{\int_I (cf)} \leq \int_I (c\overline{f}) - \int_I (c\underline{f}).$$

By Theorem 11.2.16(b), we have

$$\int_I (c\bar{f}) = c \int_I \bar{f} \leq c \int_I f + c\varepsilon$$

and

$$\int_I (c\underline{f}) = c \int_I \underline{f} \geq c \int_I f - c\varepsilon.$$

This implies that

$$\int_I (c\bar{f}) - \int_I (c\underline{f}) \leq 2c\varepsilon$$

To summarize, we have shown that for every $\varepsilon > 0$

$$0 \leq \overline{\int_I (cf)} - \underline{\int_I (cf)} \leq 2c\varepsilon.$$

Since $\overline{\int_I (cf)} - \underline{\int_I (cf)}$ does not depend on ε , we thus see that

$$\overline{\int_I (cf)} - \underline{\int_I (cf)} = 0$$

and hence that cf is Riemann integrable when $c > 0$. Since we have

$$c \int_I f - c\varepsilon \leq \int_I (cf) \leq c \int_I f + c\varepsilon,$$

for every $\varepsilon > 0$, we thus see that

$$\int_I (cf) = c \int_I f.$$

In particular, when $c = 0$, $cf = 0$ which is a piecewise constant function. Thus cf is Riemann integrable when $c = 0$. And

$$\int_I (0 \times f) = 0 \times \int_I f.$$

When $c = -1$, $-\bar{f}$ and $-\underline{f}$ are piecewise constant functions on I , and $-\bar{f}$ minorizes and $-\underline{f}$ majorizes $-f$ respectively. Thus

$$\int_I (-\bar{f}) \leq \underline{\int_I (-f)} \leq \overline{\int_I (-f)} \leq \int_I (-\underline{f}),$$

and so

$$0 \leq \overline{\int_I (-f)} - \underline{\int_I (-f)} \leq \int_I (-\bar{f}) - \int_I (-\underline{f}).$$

By Theorem 11.2.16(b), we have

$$\int_I (-\bar{f}) = - \int_I \bar{f} \geq - \int_I f - \varepsilon$$

and

$$\int_I (-\underline{f}) = - \int_I \underline{f} \leq - \int_I f + \varepsilon.$$

This implies that

$$\int_I (-\bar{f}) - \int_I (-\underline{f}) \leq 2\varepsilon$$

To summarize, we have shown that for every $\varepsilon > 0$

$$0 \leq \overline{\int_I} (-f) - \underline{\int_I} (-f) \leq 2\varepsilon.$$

Since $\overline{\int_I} (-f) - \underline{\int_I} (-f)$ does not depend on ε , we thus see that

$$\overline{\int_I} (-f) - \underline{\int_I} (-f) = 0$$

and hence that cf is Riemann integrable when $c = -1$. Since we have

$$- \int_I f - \varepsilon \leq \int_I (-f) \leq - \int_I f + \varepsilon,$$

for every $\varepsilon > 0$, we thus see that

$$\int_I (-f) = - \int_I f.$$

Use conclusions above, we have $-cf$ for all $c > 0$ is Riemann integrable.

Therefore, cf is Riemann integrable for any real number c , and for same reason, we have $\int_I(cf) = c(\int_I f)$.

(c) Consider $f + cg$, where $c = -1$. Then by Theorem 11.4.1(a) and (b), $f - g$ is Riemann integrable, and $\int_I(f - g) = \int_I f - \int_I g$.

(d) Since $f(x) \geq 0$ for all $x \in I$, we have $\sup f(x) \geq 0$ and $\inf f(x) \geq 0$ for all $x \in I$. Thus $U(f, \mathbf{P})$ and $L(f, \mathbf{P})$ are also greater than 0. f is Riemann integrable, thus

$$\int_I f = \overline{\int_I} f = \underline{\int_I} f \geq 0.$$

(e) By Theorem 11.4.1(c) and (d), we have $f(x) - g(x) \geq 0$ implies that $\int_I f - \int_I g \geq 0$.

- (f) In this case, f is a piecewise constant function, thus by Theorem 11.2.16(f) and Lemma 11.3.7, $\int_I f = c|I|$.
- (g) Since \overline{f} and \underline{f} majorizes and minorizes f respectively on I , we can see that \overline{F} and \underline{F} defined as

$$\overline{F}(x) := \begin{cases} \overline{f}(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases} \quad \underline{F}(x) := \begin{cases} \underline{f}(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases}$$

majorizes and minorizes F respectively on J . Because \overline{f} and \underline{f} are piecewise constant functions, so do \overline{F} and \underline{F} . By Theorem 11.2.16(g), we have $\int_J \overline{F} = \int_I \overline{f}$ and $\int_J \underline{F} = \int_I \underline{f}$. Thus

$$\int_I \underline{f} \leq \int_{\underline{J}} F \leq \int_J \overline{F} \leq \int_I \overline{f}.$$

This implies that

$$0 \leq \int_J \overline{F} - \int_{\underline{J}} F \leq \int_I \overline{f} - \int_I \underline{f} \leq 2\varepsilon.$$

Since $\int_J \overline{F} - \int_{\underline{J}} F$ does not depend on ε , we thus see that

$$\int_J \overline{F} - \int_{\underline{J}} F = 0$$

and hence that F is Riemann integrable. From the inequalities

$$\int_I \underline{f} - \varepsilon \leq \int_{\underline{J}} F \leq \int_J \overline{F} \leq \int_I \overline{f} + \varepsilon,$$

we have $\int_J F = \int_I f$,

- (h) Let $F : I \rightarrow \mathbf{R}$ be the function

$$f(x) := \begin{cases} f|_J(x) & \text{if } x \in J \\ f|_K(x) & \text{if } x \in K \end{cases}$$

This is similar to show that f is Riemann integrable, and by Theorem 11.4.1(g), $\int_I f = \int_J f|_J + \int_K f|_K$. \square

Exercise 11.4.2. Let $a < b$ be real numbers, and let $f : [a, b] \rightarrow \mathbf{R}$ be a continuous, non-negative function (so $f(x) \geq 0$ for all $x \in [a, b]$). Suppose

that $\int_{[a,b]} f = 0$. Show that $f(x) = 0$ for all $x \in [a, b]$. (Hint: argue by contradiction.)

Proof. Suppose for sake of contradiction that there exists an $c \in [a, b]$ such that $f(x) > 0$. Since f is continuous on $[a, b]$, for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(c)| \leq \varepsilon$ for all $x \in [a, b]$ such that $|x - c| < \delta$. This means that $f(c) - \varepsilon \leq f(x) \leq f(c) + \varepsilon$ for all $x \in [a, b] \cap (c - \delta, c + \delta)$. This is easy to see that $[a, b]$ can be partitioned into $X := [a, b] \cap (c - \delta, c + \delta)$ and $Y := [a, b] \setminus (c - \delta, c + \delta)$. Since $f(x)$ either greater than 0 or equal to 0. We can define the function F by following rule

$$f(x) := \begin{cases} f|_X(x) & \text{if } x \in X \\ f|_Y(x) & \text{if } x \in Y \end{cases}$$

Since $f|_Y(x) = 0$ for all $x \in Y$ and $f|_X(x) > 0$ for all $x \in X$, by Theorem 11.4.1(d) and (h), we have

$$\int_{[a,b]} f = \int_X f|_X + \int_Y f|_Y = \int_X f|_X \geq 0,$$

a contradiction. □

Exercise 11.4.3. Let I be a bounded interval, let $f : I \rightarrow \mathbf{R}$ be a Riemann integrable function, and let \mathbf{P} be a partition of I . Show that

$$\int_I f = \sum_{J \in \mathbf{P}} \int_J f.$$

Proof. Let $P(n)$ be the property that whenever I is a bounded interval, and whenever P is a partition of I with cardinality n , that $\int_I f = \sum_{J \in \mathbf{P}} \int_J f$. We prove the assertion by induction on n .

The base case $P(0)$ is trivial. The case $P(1)$ means that \mathbf{P} is a singleton set $\{J\}$ and $I = J$. Thus $\sum_{J \in \mathbf{P}} \int_J f = \int_I f$.

Now suppose inductively that $P(n)$ is true for all $n \geq 1$, and now we prove $P(n+1)$. Let I be a bounded interval, and let \mathbf{P} be a partition of I of cardinality $n+1$. There exists a $K \in \mathbf{P}$ such that $\mathbf{P}' = \mathbf{P} \setminus \{K\}$ to be a partition of $I \setminus K$ with cardinality n . By induction hypothesis

$$\int_{I \setminus K} f = \sum_{J \in \mathbf{P}'} \int_J f.$$

And of course

$$\int_K f = \sum_{J \in \{K\}} \int_J f.$$

By Theorem 11.4.1(h), $\{I \setminus K, K\}$ is a partition of I , thus

$$\int_I f = \int_{I \setminus K} f + \int_K f = \sum_{J \in \mathbf{P}'} \int_J f + \sum_{J \in \{K\}} \int_J f = \sum_{J \in \mathbf{P}} \int_J f.$$

□

Exercise 11.4.4. *Without repeating all the computations in the above proofs give a short explanation as to why the remaining cases of Theorem 11.4.3 and Theorem 11.4.5 follow automatically from the cases presented in the text. (Hint: from Theorem 11.4.1 we know that if f is Riemann integrable, then so is $-f$.)*

Proof. For Theorem 11.4.3, let $\min(f, g)(x) := -\max(-f(x), -g(x))$. Then remaining case follows. For Theorem 11.4.5, let $f_+(-g)_+$, $(-f)_+g_+$, $(-f)_+(-g)_+$, then remaining case follows. □

11.5

Riemann integrability of continuous functions

Exercise 11.5.1. *Prove Proposition 11.5.6. (Hint: use Theorem 11.4.1(a) and (g).)*

Proposition 11.5.6. Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ be both piecewise continuous and bounded. Then f is Riemann integrable.

Proof. Since f is piecewise continuous, by Definition 11.5.4, there exists a partition \mathbf{P} of I such that $f|_J$ is continuous on J for all $J \in \mathbf{P}$. Thus by Proposition 11.5.3, f is Riemann integrable on J for all $J \in \mathbf{P}$. By Theorem 11.4.1(a) and (g), we have

$$\int_I f = \sum_{J \in \mathbf{P}} \int_J f.$$

Thus f is Riemann integrable. □

11.6

Riemann integrability of monotone functions

Exercise 11.6.1. Use Proposition 11.6.1 to prove Corollary 11.6.3. (Hint: adapt the proof of Proposition 11.5.3.)

Corollary 11.6.3. Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ be both monotone and bounded. Then f is Riemann integrable on I .

Proof. If I is a point or an empty set then the claim is trivial; if I is a closed interval the claim follows from Proposition 11.6.1. So let us assume that I is of the form $(a, b]$, (a, b) , or $[a, b)$ for some $a < b$.

Since f is bounded, for some M we have $-M \leq f(x) \leq M$ for all $x \in I$. Let $\varepsilon > 0$. The function f when restricted to the interval $[a + \varepsilon, b - \varepsilon]$ is closed and bounded, and hence Riemann integrable by Proposition 11.6.1. In particular, we can find a piecewise constant function $h : [a + \varepsilon, b - \varepsilon] \rightarrow \mathbf{R}$ which majorizes f on $[a + \varepsilon, b - \varepsilon]$ such that

$$\int_{[a+\varepsilon, b-\varepsilon]} h \leq \int_{[a+\varepsilon, b-\varepsilon]} f + \varepsilon.$$

Define $\tilde{h} : I \rightarrow \mathbf{R}$ by

$$\tilde{h} := \begin{cases} h(x) & \text{if } x \in [a + \varepsilon, b - \varepsilon], \\ M & \text{if } x \in I \setminus [a + \varepsilon, b - \varepsilon]. \end{cases}$$

Clearly \tilde{h} is piecewise constant on I and majorizes f ; by Theorem 11.2.16 we have

$$\int_I \tilde{h} = \varepsilon M + \int_{[a+\varepsilon, b-\varepsilon]} h + \varepsilon M \leq \int_{[a+\varepsilon, b-\varepsilon]} f + (2M + 1)\varepsilon.$$

In particular we have

$$\overline{\int_I f} \leq \int_{[a+\varepsilon, b-\varepsilon]} f + (2M + 1)\varepsilon.$$

A similar argument gives

$$\underline{\int_I f} \geq \int_{[a+\varepsilon, b-\varepsilon]} f - (2M + 1)\varepsilon$$

and hence

$$\overline{\int_I f} - \underline{\int_I f} \leq (2M + 1)\varepsilon.$$

Since ε is arbitrary, f is Riemann integrable. \square

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.*

Proof. First of all we give the definition of piecewise monotone function:

Definition 1 (Piecewise monotone function). Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$. We say that f is piecewise monotone iff there exists a partition \mathbf{P} of I such that $f|_J$ is monotone on J for all $J \in \mathbf{P}$.

Following proposition would be hold:

Proposition 1. Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ be both piecewise monotone and bounded. Then f is Riemann integrable.

Proof of Proposition 1. By Corollary 11.6.3, $f|_J$ is Riemann integrable on J for all $J \in \mathbf{P}$. By Theorem 11.4.1(a) and (h),

$$\int_I f = \sum_{J \in \mathbf{P}} \int_J f|_J$$

is Riemann integrable. \square

Exercise 11.6.3. *Prove Proposition 11.6.4 (Hint: what is the relationship between the sum $\sum_{n=1}^N f(n)$, the sum $\sum_{n=0}^{N-1} f(n)$, and the integral $\int_{[0,N]} f$?)*

Proposition 11.6.4 (Integral test). Let $f : [0, \infty) \rightarrow \mathbf{R}$ be a monotone decreasing function which is non-negative (i.e., $f(x) \geq 0$ for all $x \geq 0$). Then the sum $\sum_{n=0}^{\infty} f(n)$ is convergent if and only if $\sup_{N>0} \int_{[0,N]} f$ is finite.

Proof. Suppose that $\sum_{n=0}^{\infty} f(n)$ is convergent. Then there is a real number L such that $L = \sum_{n=0}^{\infty} f(n)$. Then we have $\sum_{n=1}^N f(n) = L - f(0)$.

Let $N > 0$ be an integer, and partition into N half-open intervals $\{[n, n+1) : 0 \leq n \leq N-1\}$ of length 1, together with the point $\{N\}$. Then

by Proposition 11.3.12 we have

$$\overline{\int}_{[0,N]} f \leq \sum_{n=0}^N \left(\sup_{x \in [n, n+1)} f(x) \right).$$

Since f is monotone decreasing, we have

$$\overline{\int}_{[0,N]} f \leq \sum_{n=0}^{N-1} f(n).$$

Notice that the part $f(N)|\{N\}|$ in sum equals to zero. Similarly we have

$$\underline{\int}_{[0,N]} f \geq \sum_{n=0}^N \left(\inf_{x \in [n, n+1)} f(x) \right),$$

and

$$\underline{\int}_{[0,N]} f \geq \sum_{n=0}^N f(n+1) = \sum_{n=1}^N f(n).$$

By Proposition 11.6.1 f is Riemann integrable and $\int_{[0,N]} f = \underline{\int}_{[0,N]} f = \overline{\int}_{[0,N]} f$. Thus have

$$\sum_{n=1}^N f(n) \leq \int_{[0,N]} f \leq \sum_{n=0}^{N-1} f(n).$$

This is hold for all $N > 0$, thus $L - f(0) \leq \sup_{N>0} \int_{[0,N]} f \leq L$. Because L is real and $f(0)$ cannot be infinity, hence $\sup_{N>0} \int_{[0,N]} f$ is finite.

For the converse, suppose that $\sup_{N>0} \int_{[0,N]} f$ is finite. Since we have $L - f(0) \leq \sup_{N>0} \int_{[0,N]} f$, this implies that $L - f(0) < \infty$. Notice that $f(0)$ cannot be ∞ for that $\int_{[0,1]} f$ is finite, by Definition 11.3.12, it requires f to be bounded on $[0, N]$. Thus $\sum_{n=0}^{\infty} f(n)$ is convergent. \square

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.

Proof. Define the function $f : \mathbf{R}_+ \rightarrow \{0, 1\}$ as

$$f(x) = \begin{cases} 1, & \text{if } x \notin \mathbf{N}, \\ 0, & \text{if } x \in \mathbf{N}. \end{cases}$$

Then $\sum_{n=1}^{\infty} = 0$ and $\int_{[0,N]} f = N$, so that $\sup_{N>0} \int_{[0,N]} f = \infty$.

Define the function $f : \mathbf{N} \rightarrow \{-1, 1\}$ as

$$f(x) = \begin{cases} 1, & \text{if } x = n, \\ -1, & \text{if } x = 2n. \end{cases}$$

This is easy to see that $\sum_{n=1}^{\infty} f$ is divergent and $\sup_{N>0} \int_{[0,N]} f = 0$.

Exercise 11.6.5. Use Proposition 11.6.4 to prove Corollary 11.6.5.

Corollary 11.6.5. Let p be a real number. Then $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges absolutely when $p > 1$ and diverges when $p \leq 1$.

Proof. Since $n \geq 1$ and $\frac{1}{n^p} \geq 0$, we only need to show that $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges, then it automatically be absolutely convergent. Let $N > 1$ be an integer, and partition $[0, N]$ into N half-open intervals $\{[n, (n+1)) : 1 \leq n \leq N-1\}$ of length 1, together with the point $\{N\}$.

When $p > 1$, $\frac{1}{n^p}$ is monotone decreasing. Then we have

$$\int_{[1,N]} \frac{1}{n^p} \leq U(1/n^p, \mathbf{P}) = \sum_{n=1}^N \sup_{x \in [n, n+1)} \left(\frac{1}{x^p} \right) = \sum_{n=1}^N \frac{1}{n^p} < N.$$

Thus for any $N > 1$, we have $\sup_{N>1} \int_{[1,N]} \frac{1}{n^p} < \infty$ which is finite. Thus by Proposition 11.6.4 $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is absolutely convergent when $p > 1$.

Now we consider $0 < p \leq 1$. Since we have

$$\int_{[1,N]} \frac{1}{n^p} \geq L(1/n^p, \mathbf{P}) = \sum_{n=1}^N \inf_{x \in [n, n+1)} \left(\frac{1}{x^p} \right) = \sum_{n=1}^N \frac{1}{(n+1)^p} \geq N^{1-p}.$$

Thus for any $N > 1$, we have $\sup_{N>1} \int_{[1,N]} \frac{1}{n^p} \geq \infty$ which is infinite. By Proposition 11.6.4 $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is divergent.

When $p < 0$, $\lim_{n \rightarrow \infty} \frac{1}{n^p} \neq 0$, thus $\sum_{n=1}^{\infty} \frac{1}{n^p}$ also is divergent. Thus $\sum_{n=1}^{\infty} \frac{1}{n^p}$ diverges when $p \leq 1$. \square

11.7

A non-riemann integrable function

THERE IS NO EXERCISES IN SECTION 11.7.

11.8

The Riemann-Stieltjes integral

Exercise 11.8.1. Prove Lemma 11.8.4. (Hint: modify the proof of Theorem 11.1.13.)

Lemma 11.8.4. Let I be a bounded interval let $\alpha : X \rightarrow \mathbf{R}$ be a function defined on some domain X which contains I , and let \mathbf{P} be a partition of I . Then we have

$$\alpha[I] = \sum_{J \in \mathbf{P}} \alpha[J].$$

Proof. We prove this by induction on n . More precisely, we let $P(n)$ be the property that whenever I is a bounded interval, and whenever \mathbf{P} is a partition of I with cardinality n , that $\alpha[I] = \sum_{J \in \mathbf{P}} \alpha[J]$.

The base case $P(0)$ is trivial; the only way that I can be partitioned into an empty partition is if I is itself empty, at which point claim is easy. The case $P(1)$ is also very easy; the only way that I can be partitioned into a singleton set $\{J\}$ is if $J = I$, at which point the claim is again very easy.

Now suppose inductively that $P(n)$ is true for some $n \geq 1$ and now we prove $P(n+1)$. Let I be a bounded interval, and let \mathbf{P} be a partition of I of cardinality $n+1$.

If I is the empty set or a point, then all the intervals in \mathbf{P} must also be either the empty set or a point, and so every interval has length zero and the claim is trivial. Thus we will assume that I is an interval of the form (a, b) , $(a, b]$, $[a, b)$, or $[a, b]$.

Let us first suppose that $b \in I$, i.e., I is either $(a, b]$ or $[a, b]$. Since $b \in I$, we know that one of the intervals K in \mathbf{P} contains b . Since K is contained in I , it must therefore be of the form $(c, b]$, $[c, b]$, or $\{b\}$ for some real number c , with $a \leq c \leq b$ (in the latter case of $K = \{b\}$, we set $c := b$). In particular, this means that the set $I \setminus K$ is also an interval of the form $[a, c]$, (a, c) , $(a, c]$, $[a, c)$ when $c > a$, or a point or empty set when $a = c$.

Either way, we easily see that

$$\alpha[I] = \alpha(b) - \alpha(a)$$

and

$$\alpha[K] + \alpha[I \setminus K] = \alpha(b) - \alpha(c) + \alpha(c) - \alpha(a) = \alpha(b) - \alpha(a).$$

Thus $\alpha[I] = \alpha[K] + \alpha[I \setminus K]$.

On the other hand, since \mathbf{P} forms a partition of I , we see that $\mathbf{P} \setminus \{K\}$ forms a partition of $I \setminus K$. By the induction hypothesis, we thus have

$$\alpha[I - K] = \sum_{J \in \mathbf{P} \setminus \{K\}} \alpha[J].$$

Combining these two identities (and using the laws of addition for finite sets, see Proposition 7.1.11) we obtain

$$\alpha[I] = \sum_{J \in \mathbf{P}} \alpha[J]$$

as desired.

Now suppose that $b \notin I$, i.e., I is either (a, b) or $[a, b)$. Then one of the intervals K also is of the form (c, b) or $[c, b)$ (see Exercise 11.1.3). In particular, this means that the set $I \setminus K$ is also an interval of the form $[a, c]$, (a, c) , $(a, c]$, $[a, c)$ when $c > a$, or a point or empty set when $a = c$. The rest of the argument then proceeds as above. \square

Exercise 11.8.2. *State and prove a version of Proposition 11.2.13 for the Riemann-stieltjes integral*

Proposition (P.c. Riemann-Stieltjes integral is independent of partition). Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ 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 with respect to \mathbf{P}' . Then

$$p.c. \int_{[\mathbf{P}]} f d\alpha = p.c. \int_{[\mathbf{P}']} f d\alpha.$$

Proof. By Lemma 11.2.7, f is piecewise constant with respect to $\mathbf{P} \# \mathbf{P}'$. Clearly, $\mathbf{P} \# \mathbf{P}'$ refines \mathbf{P} . Let $J \in \mathbf{P} \# \mathbf{P}'$ and $K \in \mathbf{P}$, let c_K be the constant value of f on \mathbf{P} . By Definition 11.8.5, we have

$$p.c. \int_{[\mathbf{P}]} f d\alpha := \sum_{K \in \mathbf{P}} c_K \alpha[K].$$

Note that for every $J \in \mathbf{P} \# \mathbf{P}'$ there exists $K \in \mathbf{P}$ such that $J \subseteq K$. Thus for every $K \in \mathbf{P}$, $\{J \in \mathbf{P} \# \mathbf{P}' : J \subseteq K\}$ is a partition of K . By Lemma 11.8.4,

$$\alpha[K] = \sum_{J \in \mathbf{P} \# \mathbf{P}' : J \subseteq K} \alpha[J].$$

Hence

$$p.c. \int_{[\mathbf{P}]} f d\alpha := \sum_{K \in \mathbf{P}} c_K \left(\sum_{J \in \mathbf{P} \# \mathbf{P}' : J \subseteq K} \alpha[J] \right).$$

Since $f|_{J \subseteq K} = c_J = c_K$, we have

$$\begin{aligned} \sum_{K \in \mathbf{P}} c_K \left(\sum_{J \in \mathbf{P} \# \mathbf{P}' : J \subseteq K} \alpha[J] \right) &= \sum_{K \in \mathbf{P}} \sum_{J \in \mathbf{P} \# \mathbf{P}' : J \subseteq K} c_J \alpha[J] \\ &= \sum_{J \in \mathbf{P} \# \mathbf{P}'} c_J \alpha[J] = p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f d\alpha. \end{aligned}$$

This shows that $p.c. \int_{[\mathbf{P}]} f d\alpha = p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f d\alpha$. This is similar to show that $p.c. \int_{[\mathbf{P}']} f d\alpha = p.c. \int_{[\mathbf{P} \# \mathbf{P}']} f$. Thus $p.c. \int_{[\mathbf{P}]} f d\alpha = p.c. \int_{[\mathbf{P}']} f d\alpha$. \square

Exercise 11.8.3. *State and prove a version of Theorem 11.2.16 for the Riemann-Stieltjes integral.*

Theorem (Laws of integration). Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ be piecewise constant functions on I .

- (a) We have $p.c. \int_I (f + g) d\alpha = p.c. \int_I f d\alpha + p.c. \int_I g d\alpha$.
- (b) For any real number c , we have $p.c. \int_I (cf) d\alpha = c(p.c. \int_I f d\alpha)$.
- (c) We have $p.c. \int_I (f - g) d\alpha = p.c. \int_I f d\alpha - p.c. \int_I g d\alpha$.
- (d) If $f(x) \geq 0$ for all $x \in I$, then $p.c. \int_I f d\alpha \geq 0$.
- (e) If $f(x) \geq g(x)$ for all $x \in I$, then $p.c. \int_I f d\alpha \geq p.c. \int_I g$.
- (f) If f is the constant function $f(x) = c$ for all x in I , then $p.c. \int_I f d\alpha = c\alpha[I]$.

- (g) Let J be a bounded interval containing I (i.e., $I \subseteq J$) and let $F : J \rightarrow \mathbf{R}$ be the function

$$F(x) := \begin{cases} f(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I \end{cases}$$

Then F is piecewise constant on J , and $p.c. \int_J F d\alpha = p.c. \int_I f d\alpha$.

- (h) Suppose that $\{J, K\}$ is a partition of I into two intervals J and K . Then the functions $f|_J : J \rightarrow \mathbf{R}$ and $f|_K : K \rightarrow \mathbf{R}$ are piecewise constant on J and K respectively, and we have

$$p.c. \int_I f d\alpha = p.c. \int_J f|_J d\alpha + p.c. \int_K f|_K d\alpha.$$

Proof. Suppose that $f : I \rightarrow \mathbf{R}$ and $g : I \rightarrow \mathbf{R}$ to be piecewise constant with respect to \mathbf{P} and \mathbf{P}' , respectively. By Lemma 11.1.18, $\mathbf{P} \# \mathbf{P}'$, denoted by \mathcal{P} , is finer than \mathbf{P} and \mathbf{P}' ; and by Lemma 11.2.7, f and g is piecewise constant with respect to \mathcal{P} .

- (a) Let c_J and c'_J be the constant value of f and g respectively on $J \in \mathcal{P}$. Then by Lemma 11.2.8, $c_J + c'_J$ is the constant value of $f + g$ on $J \in \mathcal{P}$. By Definition 11.8.5, we have

$$\begin{aligned} p.c. \int_I (f + g) d\alpha &= \sum_{J \in \mathcal{P}} (c_J + c'_J) \alpha[J] \\ &= \sum_{J \in \mathcal{P}} c_J \alpha[J] + \sum_{J \in \mathcal{P}} c'_J \alpha[J] \\ &= p.c. \int_I f d\alpha + p.c. \int_I g d\alpha. \end{aligned}$$

- (b) Let c_J be the constant value of f on $J \in \mathcal{P}$. Then by Lemma 11.2.8, $c \times c_J$ is the constant value of cf on $J \in \mathcal{P}$. By Definition 11.8.5, we have

$$p.c. \int_I (cf) d\alpha = \sum_{J \in \mathcal{P}} (c \times c_J) \alpha[J] = c \sum_{J \in \mathcal{P}} c_J \alpha[J] = c \left(p.c. \int_I f d\alpha \right).$$

- (c) Consider $p.c. \int_I (f + cg)$ where $c = -1$, then from (ab), we have $p.c. \int_I (f - g) d\alpha = p.c. \int_I f d\alpha - p.c. \int_I g d\alpha$.

- (d) $f(x) \geq 0$ for all $x \in I$ implies that $c_J \geq 0$, which is the constant value of f on $J \in \mathcal{P}$. Thus by Definitions 11.8.1 and 11.8.5,

$$p.c. \int_I f d\alpha = p.c. \int_{[\mathcal{P}]} f d\alpha = \sum_{J \in \mathcal{P}} c_J \alpha[J] \geq 0.$$

- (e) Let c_J be the constant value of f on $J \in \mathcal{P}$. $f(x) \geq g(x)$ for all $x \in I$ implies that $f(x) - g(x) \geq 0$, then from (cd), $p.c. \int_I (f - g) d\alpha \geq 0$ implies $p.c. \int_I f d\alpha \geq p.c. \int_I g d\alpha$.

- (f) By Definition 11.8.5 and Exercise 11.8.2, we have

$$p.c. \int_I f d\alpha = \sum_{J \in \mathcal{P}} c_J \alpha[J] = c \left(\sum_{J \in \mathcal{P}} \alpha[J] \right) = c \alpha[I].$$

- (g) Let \mathbf{U} be a partition of I , and \mathbf{V} be a partition of $J \setminus I$, then $\mathbf{U} \cup \mathbf{V}$ is a partition of J . Let c_K be the constant value of F on $K \in \mathbf{U} \cup \mathbf{V}$. Then

$$\begin{aligned} p.c. \int_J F d\alpha &= \sum_{K \in \mathbf{U} \cup \mathbf{V}} c_K \alpha[K] \\ &= \sum_{K \in \mathbf{U}} c_K \alpha[K] + \sum_{K \in \mathbf{V}} 0 \times \alpha[K] \\ &= \sum_{K \in \mathbf{U}} c_K \alpha[K] \\ &= p.c. \int_I f d\alpha. \end{aligned}$$

- (h) We can see that I is a bounded interval containing J and K , i.e., $J \subseteq I$ and $K \subseteq I$. Since for every $x \in I$ either have either $x \in J$ or $x \in K$, we have

$$f(x) = \begin{cases} f|_J(x) & \text{if } x \in J \\ f|_K(x) & \text{if } x \in K \end{cases}$$

By (g), we have $p.c. \int_I f d\alpha = p.c. \int_J f|_J d\alpha + p.c. \int_K f|_K d\alpha$. □

Exercise 11.8.4. *State and prove a version of Theorem 11.5.1 for the Riemann-Stieltjes integral. (Hint: one has to be careful with the proof; the*

problem here is that some of the references to the length of $|J_k|$ should remain unchanged and other references to the length of $|J_k|$ should be changed to the α -length $\alpha(J_k)$ - basically, all of the occurrences of $|J_k|$ which appear inside a summation should be replaced with $\alpha(J_k)$, but the rest should be unchanged.

Theorem. Let I be a bounded interval, and let f be a function which is uniformly continuous on I . Then f is Riemann-Stieltjes integrable.

Proof. From Proposition 9.9.15 we see that f is bounded. Now we have to show that $\int_I f = \overline{\int}_I f$.

If I is a point or the empty set then the theorem is trivial, so let us assume that I is one of the four intervals $[a, b]$, (a, b) , $(a, b]$, or $[a, b)$ for some real numbers $a < b$.

Let $\varepsilon > 0$ be arbitrary. By uniform continuity, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in I$ are such that $|x - y| < \delta$. By the Archimedean principle, there exists an integer $N > 0$ such that $(b - a)/N < \delta$.

Note that we can partition I into N intervals J_1, \dots, J_N , each of length $(b - a)/N$. By definition of upper and lower Riemann-Stieltjes integrals, we thus have

$$\overline{\int}_I f \leq \sum_{k=1}^N \left(\sup_{x \in J_k} f(x) \right) \alpha[J_k]$$

and

$$\underline{\int}_I f \geq \sum_{k=1}^N \left(\inf_{x \in J_k} f(x) \right) \alpha[J_k],$$

so in particular

$$\overline{\int}_I f - \underline{\int}_I f \leq \left(\sup_{x \in J_k} f(x) - \inf_{x \in J_k} f(x) \right) \alpha[J_k].$$

However, we have $|f(x) - f(y)| < \varepsilon$ for all $x, y \in J_k$, since $|J_k| = (b - a)/N < \delta$. In particular we have

$$f(x) < f(y) + \varepsilon \text{ for all } x, y \in J_k.$$

Taking suprema in x , we obtain

$$\sup_{x \in J_k} f(x) \leq f(y) + \varepsilon \text{ for all } y \in J_k,$$

and then taking infima in y we obtain

$$\sup_{x \in J_k} f(x) \leq \inf_{y \in J_k} f(y) + \varepsilon.$$

Inserting this bound into our previous inequality, we obtain

$$\overline{\int_I} f - \underline{\int_I} f \leq \sum_{k=1}^N \varepsilon \alpha[J_k],$$

but by Lemma 11.8.4 we thus have

$$\overline{\int_I} f - \underline{\int_I} f \leq \varepsilon \alpha(b-a).$$

But $\varepsilon > 0$ was arbitrary, while $(b-a) \geq 0$ and α is monotone increasing, we have $\alpha(b-a) \geq 0$. Thus $\overline{\int_I} f d\alpha - \underline{\int_I} f d\alpha$ cannot be positive. By the definition of Riemann-Stieltjes integrability we thus have that f is Riemann-Stieltjes integrable. \square

Exercise 11.8.5. Let $\text{sgn} : \mathbf{R} \rightarrow \mathbf{R}$ be the signum function

$$\text{sgn}(x) := \begin{cases} 1 & \text{when } x > 0 \\ 0 & \text{when } x = 0 \\ -1 & \text{when } x < 0. \end{cases}$$

Let $f : [-1, 1] \rightarrow \mathbf{R}$ be a continuous function. Show that f is Riemann-Stieltjes integrable with respect to sgn , and that

$$\int_{[-1,1]} f d\text{sgn} = 2f(0).$$

(Hint: for every $\varepsilon > 0$, find piecewise constant functions majorizing and minorizing f whose Riemann-Stieltjes integral is ε -close to $2f(0)$).

Proof. Since f continuous on $[-1, 1]$, by Exercise 11.8.4 and Theorem 9.9.16, f is Riemann-Stieltjes integrable, and we have a bound M such that $-M \leq f(x) \leq M$ for all $x \in [-1, 1]$.

Let $\varepsilon > 0$. Then for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$|f(x) - f(0)| < \varepsilon$ for every $x \in (-\delta, \delta) \cap [-1, 1]$. Define $h : [-1, 1] \rightarrow \mathbf{R}$ by

$$h(x) := \begin{cases} f(0) + \varepsilon & \text{if } x \in (-\delta, \delta) \cap [-1, 1] \\ M & \text{if } x \in [-1, 1] \setminus (-\delta, \delta) \end{cases}$$

which majorizing f . Then by Exercise 11.8.3, we have

$$\begin{aligned} \int_{[-1,1]} f \, d\text{sgn} &\leq \int_{[-1,1]} h \, d\text{sgn} \\ &= -(1-\delta)M \int_{(-\delta,\delta) \cap [-1,1]} h \, d\text{sgn} + (1-\delta)M \\ &\leq \sum_{x \in (-\delta,\delta) \cap [-1,1]} (f(0) + \varepsilon) \cdot \text{sgn}(x) \end{aligned}$$

Let $\varepsilon > 0$ to be enough small such that $\delta < 1$, then we have

$$\begin{aligned} \int_{[-1,1]} f \, d\text{sgn} &\leq \int_{[-1,1]} h \, d\text{sgn} \\ &\leq \sum_{x \in (-\delta,\delta) \cap [-1,1]} (f(0) + \varepsilon) \cdot \text{sgn}(x) \\ &= (f(0) + \varepsilon) \cdot [(\text{sgn}(0) - \text{sgn}(-\delta)) \\ &\quad + (\text{sgn}(\delta) - \text{sgn}(0)) + \text{sgn}(0)] \\ &= 2(f(0) + \varepsilon). \end{aligned}$$

A similar argument gives

$$\int_{[-1,1]} f \, d\text{sgn} \geq 2(f(0) - \varepsilon).$$

Since ε is arbitrary, we have $\int_{[-1,1]} f \, d\text{sgn} = 2f(0)$ as desired. \square

11.9

The two fundamental theorems of calculus

Exercise 11.9.1. Let $f : [0, 1] \rightarrow \mathbf{R}$ be the function in Exercise 9.8.5. Show that for every rational number $q \in \mathbf{Q} \cap [0, 1]$, the function $F : [0, 1] \rightarrow \mathbf{R}$ defined by the formula $F(x) := \int_0^x f(y) \, dy$ is not differentiable at q .

Proof.

Exercise 11.9.2. Prove Lemma 11.9.5. (Hint: apply the mean-value theorem, Corollary 10.2.9, to the function $F - G$. One can also prove this lemma using the second Fundamental theorem of calculus (how?), but one has to be careful since we do not assume f to be Riemann integrable.

Lemma 11.9.5. Let I be a bounded interval, and let $f : I \rightarrow \mathbf{R}$ be a function. Let $F : I \rightarrow \mathbf{R}$ and $G : I \rightarrow \mathbf{R}$ be two antiderivatives of f . Then there exists a real number C such that $F(x) = G(x) + C$ for all $x \in I$.

Proof. Since F and G are antiderivatives of f , by Definition 11.9.3, F and G are differentiable on I . Thus $F - G$ is also differentiable on I . Then by mean value theorem, for every $x \in I$, there exists a $y \in (a, b)$ and $x \neq y$ such that

$$\begin{aligned} (F - G)'(x) &= F'(x) - G'(x) \\ &= \frac{F(x) - G(x) - F(y) + G(y)}{x - y} \\ &= f(x) - f(x) \\ &= 0. \end{aligned}$$

Thus we have $F(x) - G(x) = F(y) - G(y)$. Since $F \neq G$, if $F(y) - G(y) \neq 0$, there exists a $C := F(y) - G(y)$ such that $F(x) = G(x) + C$. Otherwise, we have $F(x) = G(x)$. \square