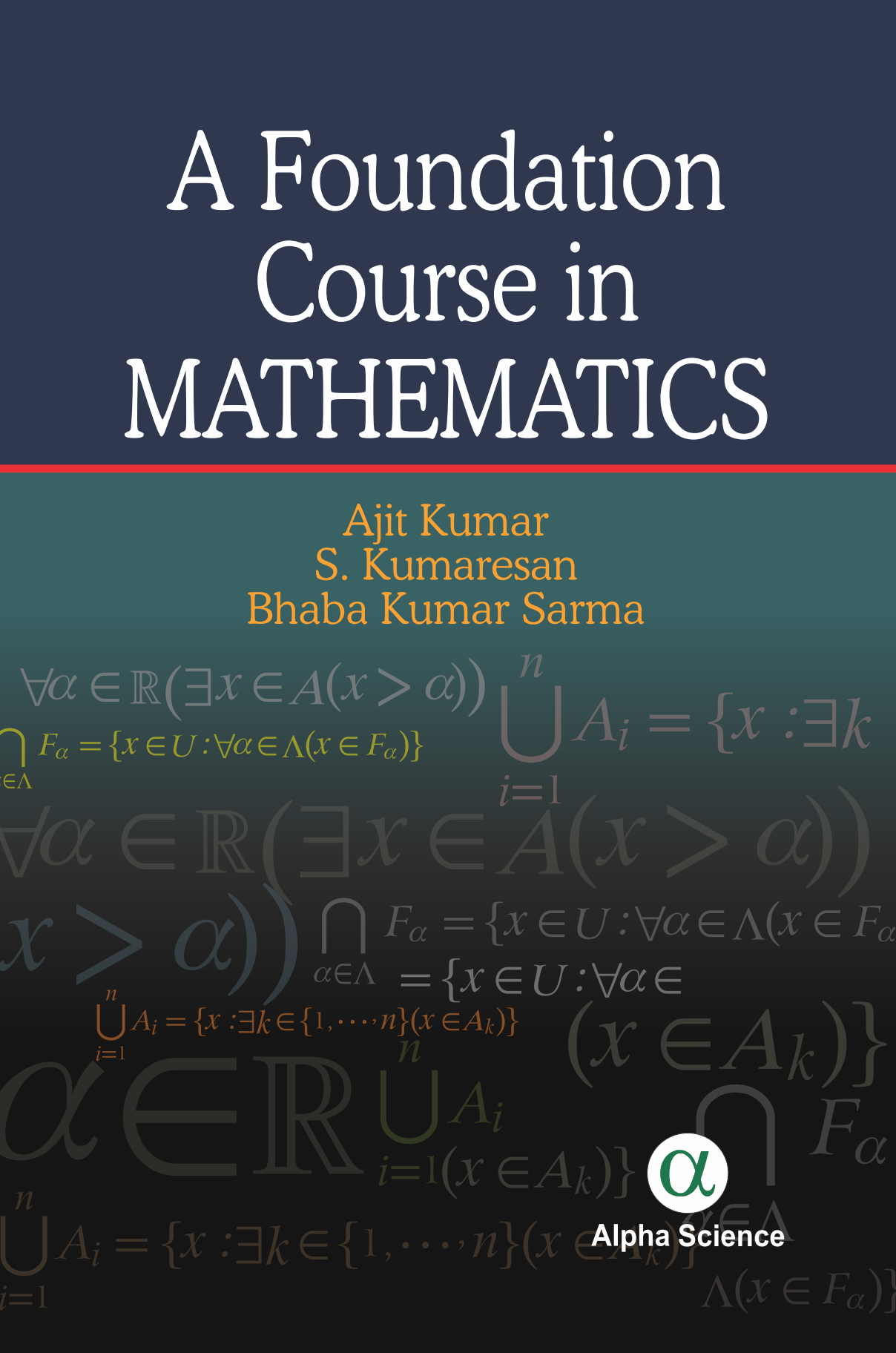
A Foundation Course in MATHEMATICS

Ajit Kumar
S. Kumaresan
Bhaba Kumar Sarma



A Foundation Course in Mathematics

A Foundation Course in Mathematics

Ajit Kumar
S. Kumaresan
Bhaba Kumar Sarma

Alpha Science International Ltd.
Oxford, U.K.

A Foundation Course in Mathematics
148 pgs.

Ajit Kumar
Institute of Chemical Technology
Mumbai

S. Kumaresan
University of Hyderabad
Hyderabad

Bhaba Kumar Sarma
Indian Institute of Technology Guwahati
Guwahati

ALPHA SCIENCE INTERNATIONAL LTD.

7200 The Quorum, Oxford Business Park North
Garsington Road, Oxford OX4 2JZ, U.K.

www.alphasci.com

ISBN 978-1-78332-358-6

E-ISBN 978-1-78332-434-7

Printed from the camera-ready copy provided by the Authors.

The book is dedicated to
all the resident faculty of MTTS Programme whose
commitment made MTTS Programme the most successful
and enduring training programme of India

Preface

This book is based on a course on "Foundations" in one of the most famous and successful mathematics training programmes titled Mathematics Training and Talent Search (MTTS) Programme which has been running in India since 1993. This course is taught in Level-O (the first level of the programme attended by second year undergraduate students) to prepare the participants for various courses in Analysis, Algebra, Topology, Geometry, etc., in the programme.

We believe that a course on foundations is not only beneficial, but also necessary. There is a conspicuous gap between the students' experience in mathematics at pre-college level and the expectation on their maturity while introducing serious mathematics like real analysis and abstract algebra to them at undergraduate level. Even students at post-graduate level are found to lack in their ability in understanding and capability of handling their courses in mathematics. The number of students who can do and learn mathematics on their own is rather small. This is why, a foundation course in some form is also given to Level-I and Level-II students in the MTTS programme who are final year undergraduate and first year postgraduate students, respectively.

The purpose of writing this book is multi-fold. The first is to introduce the fundamental concepts in logic, sets, functions and relations with due clarity. The second is to imbibe the ability in the students to understand, visualize and express mathematics with requisite rigour. The third is to train the students in problem-solving skills. A diligent reader will find that there are a variety of problems ranging from routine to challenging. Quite a few of the problems are open-ended in the sense that the reader is encouraged to form his own problems and investigate them. We believe that this type of problems promote the inquisitiveness and creativity on the part of the reader. The fourth aim is to build the writing skills of the students. In a typical MTTS camp, the writing on the board is minimal. The teacher and the students discuss the subject, problems are solved or the results are proved more often orally with minimal writing. While summarising the work, the teacher usually explains how a typical text-book solution/proof is written based on the discussion. The students are encouraged to discuss with their friends and asked to write on their own. It is remarkable that almost all students will be able to write so-called lengthier proofs on their own latest by the second week of the camp.

The target audience of this book is the first year undergraduate students who have chosen mathematics as the main subject or wish to go for higher studies in mathematics. We sincerely believe that a one semester course on the materials of the book should be given before the courses in analysis and abstract

algebra. Even otherwise, students may be encouraged to read and learn from the book, with or without assistance of teachers.

The exposition is not a formal book on logic or set theory. Rather, it has been written to equip the students to deal with fundamental concepts in mathematics. The book is written in a conversational tone and the readers may feel as if a teacher is talking to them while reading this book. The exercises are not given as bunches at the end of sections/chapters. Rather, they are interspersed inside the sections so that the student participates actively in the discussions. The students are advised to write complete solutions of the exercises, as most of the solutions are of a few lines length only. This helps the students to acquire the art of writing rigorous mathematics in a gradual manner.

At pre-college stage, students are usually exposed only to sets such as $\{1, 2, 3\}$ or $\{a, b, c\}$ and explicit maps between such sets shown pictorially by arrows connecting x and $f(x)$. For example, students invariably do not have proper understanding that subsets are defined using the so-called "axiom of separation", that is, the elements share some common properties. Sets, functions, relations which arise more naturally in mathematics are, as a rule, hardly introduced. Further, necessary understanding of the student in dealing with mathematical statements, definitions and proofs is seldom built up.

The book consists of seven chapters. In Chapter 1, we deal with logic and standard methods of proofs in mathematics in an informal way. We give special emphasis on statements involving quantifiers and their negations. We insist on writing negations of statements and definitions as it helps the student to understand the concepts and also to construct a proof of a result logically.

Chapter 2 deals with basic notions of sets and the standard operations on sets. Students at school level are usually exposed only to finite sets and sets like those of natural numbers, integers, rationals, real numbers and complex numbers, and not those defined by common properties shared by their elements. We give emphasis on identifying sets written in various forms and visualizing them geometrically. We give a rigorous account of families of sets which students usually find difficult to deal with.

In Chapter 3, we discuss various aspects of functions extensively with a large number of examples and exercises. This is required, because functions are building blocks in mathematics. We give special emphasis on images and inverse images of subsets under functions. These notions, which are usually neglected in undergraduate curriculum, are important in mathematics, particularly in analysis and topology.

In Chapter 4, we give a brief introduction of various types of relations. Equivalence relations and classes are extremely important concepts in mathematics. In fact, many distinguished sets in mathematics are identified as sets of equivalence classes. We present a detailed account of equivalence relations with a number of examples in this chapter.

Chapter 5 deals with three induction principles, namely, the induction principle in the usual or the weak form, the strong induction principle and the well-ordering principle. We demonstrate each of them as a tool for proving mathematical statements with a number of examples. Further, we provide a proof of the equivalence of these three principles.

In Chapter 6, we introduce cardinality of sets. Our emphasis is to deal with equivalence of sets in terms of countability, rather than cardinal numbers and the algebra of cardinal numbers. We give a special emphasis on countability of various familiar sets and put cardinality of finite sets on a formal standing.

In Chapter 7, we try to provide a lucid introduction of order relations on sets. Partial and total orders are concepts students often find difficult to understand. We provide a number of examples to illustrate these concepts. This chapter also deals with the notions of chain, upper bound, least upper bound, and maximal and minimal elements of subsets in partially ordered sets. We conclude with a brief introduction of the well-ordering principle, Zorn's lemma and the axiom of choice.

We believe that this book will serve as a catalyst to prepare students for understanding and doing mathematics with due rigour. It will also infuse the rudiments of mathematical thinking. Some of the MTTS faculty run a 3-day or a week long workshop based on our MTTS course on Foundations in their institutions before the start of the regular courses. Through this exercise, the understanding and the performance of the students improve significantly. In Choice Based Credit System (CBCS) as stipulated by UGC, the disciplines have to declare some courses as core courses which will also be taken by students from other disciplines. A course based on this book will be eminently suitable as it trains the students in mathematical thinking. The MTTS team will be happy to hold a 3-day workshop based on the book, if a college/institute desires.

The book is typeset in LaTeX and all pictures are drawn using Geogebra, IPE and TikZ. We thank the creators of these marvelous tools.

We will be glad to receive comments, suggestions and corrections from the readers. They may be sent to any of the authors at email addresses: ajit72@gmail.com, kumaresa@gmail.com and bhabasarma@gmail.com.

Ajit Kumar
S. Kumaresan
Bhaba Kumar Sarma

Acknowledgements

The book is based on a course on "Foundations" in MTTS programme for last several years. Authors like to wholeheartedly thank all the students and resource persons of this programme.

Authors would also like to thank the Institute of Chemical Technology, Mumbai, especially the TEQIP Phase-II programme for providing financial support to foster discussions among the authors to finish this project.

A special thank to all family members of the authors for their unequivocal support to complete this project and for allowing them to be away for the MTTS programme year after year.

Contents

List of Figures

Chapter 1

Statements and Logic

In this chapter, we discuss different types of statements one encounters in Mathematics. The aim of this chapter is to enable the students how to deal with statements which are composite or with quantifiers. Understanding such statements, being able to decide their validity (true or false), and being able to negate them are fundamental in learning Mathematics. Further, we give some ideas of standard ways of proving statements in Mathematics.

1.1 Statements

Consider the following sentences.

(a) 5 is a prime number.

(b) The sine function is not periodic.

(c) The literacy rate in India has increased since 1947.

(d) $5 + 3 = 9$.

(e) Moscow is the capital of China.

Under the usual meaning of the words and symbols, sentences in (a) and (c) are true, whereas the others are false.

In Mathematics, one deals with a large number of sentences which are strings of words and symbols having precise meanings. All such sentences are classified as being true or false and they are called *statements*. Thus, a statement is either true or false; never both or in between.

Consider the sentence: "*The sentence I am reading is false*". Is it a statement? No, because it can be neither true nor false; each assumption (true or false) leads to a contradictory conclusion.

A sentence can be a statement even if you do not know decisively whether it is true or false, or even if its being true or false depends on the context. For example, consider the following examples:

S_1: Every even integer ≥ 4 is a sum of two primes.

S_2: There are infinitely many pairs of twin primes.

S_3: $x + 3 = 7$.

S_4: The number of students in the class is fifty.

The sentences S_1 and S_2 are famous problems in Number Theory. However, it is not yet known whether they are true or false. The sentences S_3 and S_4 are context specific; they are true or false depending on the situations. For example, S_3 is true if $x = 4$, but false if $x = 3$. Nevertheless, the above sentences are statements.

Negation of a statement

Consider the statement S_4: "The number of students in the class is fifty". If you specify the context, that is, if you take a particular class, the statement is either true or false. It is true if there are fifty students in the class. When is the statement false? It is so if "the number of student in the class is not fifty", which is another statement. We call this later statement the negation of S_4.

For any given statement S, there is another statement which is defined to be the statement that is true when S is false, and false when S is true. This statement is called the *negation* of S and is denoted by **not**-S.

It is evident that for any statement S, exactly one of S and not-S is true. Further, S is the negation of not-S, that is, the statements not-(not-S) and S convey the same meaning.

Note that the negation of the statement "$5 + 3 = 9$" is "$5 + 3 \neq 9$", and not the statement "$5 + 3 = 8$". Similarly, the negation of the statement "Moscow is the capital of China" is the statement "Moscow is not the capital of China", and not the statement "Moscow is the capital of Russia".

Exercise 1.1.1. Write the negations of the following statements.

(a) I visited her place on Sunday.

(b) The market will not be open in the evening.

(c) The square of the integer n is divisible by the prime p.

(d) $x^2 + 3 = 27$.

(e) The child is cute.

(f) The apple is ripe.

1.2 Statements with quantifiers

In our everyday life we often use sentences having phrases "for some", "there is", "for all", "for every", "each of", "given any", etc. For example, look at the following sentences:

S_1: There is a rotten apple in the basket of apples.

S_2: All apples in the basket are ripe.

Note that each of these sentences asserts a statement for some or all objects in a collection. For example, the statement S_1 asserts that the statement "the apple is rotten" holds for at least one of the apples in the basket. The statement S_2, on the other hand, asserts that the statement "the apple is ripe" holds for every apple in the basket. We say that the sentences involve *quantifiers*: the first

involves the *existential quantifier* "there exists", and the second the *universal quantifier* "for every".

The two statements "There is a rotten apple in the basket", and "There are some rotten apples in the basket", convey the same meaning in Mathematics; each of them asserts at least one rotten apple in the basket.

Exercise 1.2.1. Identity the quantifiers "there exists" and "for every" that the following statements involve.

(a) Every page in this book contains at least 500 words.

(b) In this book some pages do not contain any picture.

(c) $4x = 2$ for some rational number x.

(d) There exists a student in the classroom who is at least 6 feet tall.

(e) Every student in the classroom is at least 5 feet tall.

(f) I can find a millionaire in this room.

(g) All tables in the room are dirty.

(h) There is a dirty table in the room.

In Mathematics, statements involve quantifiers very often. The symbol $\exists$ is used for the existential quantifier "there exists" and the symbol $\forall$ is used for the universal quantifier "for every". In any statement with quantifiers, the quantifiers refer to the elements of a set X, which depends on the context. This set is referred as the *set of context* for the corresponding quantifier. For example, for the statements S_1 and S_2 above, the set of context for the respective quantifiers is the set X of apples in the basket. Similarly, in the first two statements in Ex. 1.2.1 the set of context is the set of pages in this book, and in the last two it is the set of tables in the room.

There is another fundamental aspect of a statement with a quantifier. It refers to some property P that makes sense to the elements of X, the set of context, and it asserts that the property P is satisfied by either some or all elements of X. For example, the statement S_1 asserts that some apple in the basket satisfies the property of "being rotten". Similarly, S_2 asserts that all apples in the basket has the property of "being ripe".

We sum up these observations in the following.

Observation 1.2.2. In order to express or understand a statement involving quantifiers, two things are important to identify (i) the set of context and (ii) the property that is claimed to be satisfied by either some or each element of the set.

Exercise 1.2.3. Identify the sets of context and the properties involved in the statements in Exercise 1.2.1.

A statement with an existential quantifier is typically of the form

There exists an element in X for which the property P holds,

and a statement with a universal quantifier is typically of the form

For each element in X, the property P holds.

Here it is understood that P is a property which makes sense for each element

of X. Using the symbols $\exists$ and $\forall$ the statements are written respectively as

$\exists x \in X(x$ has property $P)$, and

$\forall x \in X(x$ has property $P)$.

For example, let X be the set of all pages in a given book. Then, using quantifiers the statements (a) and (b) in Exercise 1.2.1 are written respectively as

"$\forall x \in X(x$ has at least 500 words)", and

"$\exists x \in X(x$ does not have a picture)".

On the other hand, we read the two statements

"$\exists x \in \mathbb{Z}(x^2 \leq 2)$" and "$\forall x \in \mathbb{Z}(x^2 \geq 0)$"

respectively as

"There exists an integer x *such that* $x^2 \leq 2$", and

"For each integer x, $x^2 \geq 0$" (or "Given any integer x, $x^2 \geq 0$").

For each of the above two statements, the set of context is the set $\mathbb{Z}$ of integers, and the two properties are $x^2 \leq 2$ and $x^2 \geq 0$, respectively.

Exercise 1.2.4. Write the statements in Ex. 1.2.1 using quantifiers $\exists$ or $\forall$.

When is a statement involving a quantifier a true statement? For example, consider the statement "All apples in the basket are ripe". To conclude that the statement is true, we need to know (probably by checking them one by one) that each of the apples in the basket is ripe. Similarly, for concluding the statement "Each page in this book contains at least 500 words" to be true we may need to verify that the number of words in each of the pages is more than or equal to 500.

Now, consider the statement "There is a rotten apple in the basket". For concluding that the statement is true, you may check the apples one by one, and come across one instance when the apple is rotten. Note that once you come across a rotten apple, you need not continue your checking any further. Similarly, to conclude that the statement "There are pages in this book that contain no pictures" is true, you need to come across one page that does not contain a picture.

Negation of statements with quantifiers

Recall that each statement S has a negation not-S, and that exactly one of the statements S and not-S is true. How do we negate a statement with quantifiers? What does it mean to say that the statement

S_1: "All apples in the basket are ripe,"

is false? For this to be case, should it happen that none of apples in the basket is ripe? If you check the apples whether they are ripe and come across one instance that the apple is not ripe, then the statement S_1 will be false. Thus, for S_1 to be false, it is enough if we can find one apple which is not ripe. In other words, the negation of S_1 is

not-S_1: "There is an apple in the basket which is not ripe".

Similarly, the negation of the statement "Every page in the book has a picture" is the statement "There is a page in the book which does not have a picture". Notice that the negation of statement with the universal quantifier is a statement with the existential quantifier.

We conclude that the negation of the statement (with the universal quantifier)

"*For every element in X, the property P holds*,"

is the statement (with existential quantifier)

"*There exists an element x in X such that the property P* **does not** *hold for x*".

Using quantifiers, the negation of the statement

$$\forall x \in X(x \text{ has the property } P)$$

is the statement

$$\exists x \in X(x \text{ does } \textbf{not} \text{ have property } P).$$

Example 1.2.5. The statement "Each table in the room is dirty" can be written using quantifier as "$\forall x \in X(x$ is dirty)", where X denotes the set of chairs in the room. The negation of the statement using quantifier will be "$\exists x \in X(x$ is not dirty)". In plain English, the negation can be written as "There is a table in the room which is not dirty".

Exercise 1.2.6. Write down the following sentences using appropriate sets of context and quantifiers. Write the negations in plain English. Formulate their negations using quantifier symbols, and compare the two negations you have written for each sentence.

(a) Given any integer n, $n^2 > n$.

(b) Each student in the class is at least 5 feet tall.

(c) All the students in this room are bright.

(d) Every delegate in the conference was a millionaire.

(e) All tables in the room are clean.

What does it mean to say that the statement

S_2: "There is a rotten apple in the basket"

is false? You may say that it means "No apple in the basket is rotten". Well, you are right. But how do you confirm that S_2 is false? To do so, you will need to check each apple in the basket and show that each of them is "not rotten". Thus, the negation of the sentence in fact means

not-S_2: "Given any apple in the basket, it is *not rotten*".

Note that the negation has the universal quantifier "$\forall$" which has been read as "given any".

Remark 1.2.7. You may write "Every/each apple in the basket is not rotten" as the negation of the statement "There is a rotten apple in the basket". However, there may be an ambiguity due to the standard quirk in the language. When we say "Each apple in the basket is not rotten", do we mean that there

may be some apple which are not rotten and some are? For a similar instance, consider, "Every man is not a billionaire." To avoid this ambiguity, it is better to write the negation as

"Given any apple in the basket, it is not rotten".

While negating a statement which starts with the existential quantifier, we suggest you read the universal quantifier $\forall$ as "Given any".

When is the statement

"There is a book on the table which does not contain a preface"

a false statement? It is so, if you find that there is no book on the table without a preface, which actually means

"Given any book on the table, it contains a preface".

Notice that the negation of statement with the existential quantifier is a statement with the universal quantifier.

We conclude that the negation of the statement (with the existential quantifier)

"*There exists an element in X for which has property P*"

is the statement (with universal quantifier)

"*Given any element in X, the element does not have the property P*".

Using quantifiers, the negation of the statement

$$\exists x \in X(x \text{ has property } P)$$

is the statement

$$\forall x \in X(x \text{ does } \textbf{not} \text{ have property } P).$$

Example 1.2.8. The statement "A chair in this room has a broken leg" can be written using quantifier as "$\exists x \in X(x$ has a broken leg)", where X denotes the set of chairs in the room. The negation of the statement using quantifier will be "$\forall x \in X(x$ does not have a broken leg)". In words, the negation can be written as "Given any chair in the room, it does not have a broken leg".

Exercise 1.2.9. Write down the following sentences using appropriate sets of contexts and quantifiers. Write the negations in plain English. Formulate their negations using quantifier symbols, and compare the two negations you have written for each sentence.

(a) There is a dirty table in the room.

(b) In this book some pages do not contain any picture.

(c) $4x = 2$ for some rational number x.

(d) There exists a student in the classroom who is at least 6 feet tall.

(e) I can find a millionaire in this room.

(f) There exists a real number x such that $x^2 = 1$.

(g) There is a genius in this room.

Example 1.2.10. Suppose A, B are sets. We write $A \subseteq B$ (and read as A is a *subset* of B), if every element of A is also an element of B. Using quantifiers, this can be written as

$$A \subseteq B \text{ if } \forall x \in A(x \in B).$$

Note that, we have written here a definition. So, when do you write $A \not\subseteq B$, that is, A is *not a subset* of B? We say $A \not\subseteq B$, if "$\forall x \in A(x \in B)$" does not hold, that is, "$\exists x \in A(x \notin B)$" holds. In other words,

$$A \not\subseteq B \text{ if } \exists x \in A(x \notin B).$$

That is, A is not a subset of B if there is an element in A which is not in B.

Exercise 1.2.11. "The square of every real number is nonnegative." Write this in symbols, negate it and then write the negation in words.

Example 1.2.12. Suppose A is a nonempty set of real numbers and α a real number. We say that α is an *upper bound* of A, if for all $x \in A$, $x \leq \alpha$. In terms of quantifiers, we have,

$$\alpha \in \mathbb{R} \text{ is in an } \textit{upper bound} \text{ of } A \text{ if } \forall x \in A(x \leq \alpha).$$

Note that whether or not α is an upper bound of A depends on whether each element $x \in A$ has the common property, namely, of being less than or equal to α. Hence if we want to claim that α is not an upper bound of A, then we have to negate the statement $\forall x \in A(x \leq \alpha)$. That is, we must show that there exists $x \in A$ such that $x \leq \alpha$ is false. In view of the law of trichotomy, this means that $x > \alpha$. Therefore,

$$\alpha \in \mathbb{R} \text{ is } \textit{not an upper bound} \text{ of } A, \text{ if } \exists x \in A(x > \alpha).$$

Is 0.9 an upper bound of the open interval $(0, 1)$?

Exercise 1.2.13. Let $f\colon X \to Y$ be a map and $y, z \in Y$. Write the following using quantifiers:

(a) y is in the image of f.
(b) z is not in the image of f.

Statements involving multiple quantifiers

On many occasions, a statement involves quantifiers more than once. For example, consider the statements

S_1: In every shelf in the library there is a mathematics book.

S_2: There is a shelf in the library in which all books are story books.

You may notice that each of the statements involves two quantifiers. If you denote the set of shelves in the library by X, then the statement S_1 reads as "$\forall s \in X$(there is a mathematics book in s)". Note that "there is a mathematics book in s" itself is a statement with the existential quantifier. For a given shelf s, let us denote by B_s the set of books in the shelf s. Then "there is a mathematics book in s" can be written as "$\exists b \in B_s(b$ is a mathematics book)".

In a nested manner, therefore, S_1 can be written as

S_1: $\forall s \in X\big(\exists b \in B_s(b \text{ is a mathematics book})\big)$.

Similarly, the other statement can be written as

S_2: $\exists s \in X\big(\forall b \in B_s(b \text{ is a story book})\big)$.

How do you negate these statements? By our earlier discussion, S_1 is false means there is $s \in X$ for which "$\exists b \in B_s(b$ is a mathematics book)" is false. However, "$\exists b \in B_s(b$ is a mathematics book)" is false means "$\forall b \in B_s(b$ is *not* a mathematics book)". Thus, with quantifiers the negation of S_1 is the statement

not-S_1: $\exists s \in X\big(\forall b \in B_s(b \text{ is } \textit{not} \text{ a mathematics book})\big)$.

This latter statement can be read as "There is a shelf in the library in which each of the book is a non-mathematics book". At times, you will also state this as "There is a shelf in the library which does not have any mathematics book".

You will now easily convince yourself that the negation of S_2 can be written using quantifiers as

not-S_2: $\forall s \in X\big(\exists b \in B_s(b \text{ is a non-story book})\big)$,

which means "Given any shelf in the library, it has a non-story book".

Exercise 1.2.14. Express the following sentences using quantifiers. Negate each of them in plain English, write the negation using quantifiers and compare the two.

(a) There is a tree in this campus all of whose leaves are green.

(b) Every tree in this campus has at least one brown leaf.

(c) In every shelf in the library all books are mathematics books.

(d) Every tree in this campus has all of its leaves green.

(e) There is a shelf in the library in which there is a mathematics book.

Example 1.2.15. Let A be a nonempty subset of real numbers. We say that A is *bounded above* in $\mathbb{R}$, if we can find a real number which will be an upper bound of A. That is, A is *bounded above* in $\mathbb{R}$, if there exists $\alpha \in \mathbb{R}$ such that α is an upper bound of A, that is, if

$$\text{"}\exists \alpha \in \mathbb{R}(\alpha \text{ is an upper bound of } A)\text{"}.$$

Thus, when do you say A is not bounded above? It is so, if no real number is an upper bound of A. Thus, A is *not bounded above*, if

$$\text{"}\forall \alpha \in \mathbb{R}\ (\alpha \text{ is not an upper bound of } A).$$

The above two statements are interesting! If you closely look at them, you will notice that each of them uses two quantifiers to assert something. In fact, you can write them in full glory. The first statement then will read as

$$\text{"}A \text{ is } \textit{bounded above} \text{ in } \mathbb{R}, \text{ if } \exists \alpha \in \mathbb{R}\big(\forall x \in A(x \leq \alpha)\big)\text{"}.$$

and the second as

$$\text{"}A \text{ is not bounded above in } \mathbb{R} \text{ if } \forall \alpha \in \mathbb{R}\big(\exists x \in A(x > \alpha)\big)\text{"}.$$

Note that to get the negation of

$$\text{“}\exists\alpha \in \mathbb{R}\big(\forall x \in A(x \leq \alpha)\big)\text{”},$$

we need to negate a nested sentence, and we do it layer by layer to get the negation as

$$\text{“}\forall\alpha \in \mathbb{R}\big(\exists x \in A(x > \alpha)\big)\text{”}.$$

Exercise 1.2.16. There exists an integer $x \in \mathbb{Z}$ such that for any $y \in \mathbb{Z}$ we have $x + y = y$. Write this in symbols using the quantifiers $\exists$ and $\forall$, negate it and write the negation in words.

Exercise 1.2.17. $\mathbb{R}$ has Archimedean property: for each $x > 0$ and $y \in \mathbb{R}$, we can find $n \in \mathbb{N}$ such that $nx > y$. Write this in symbols, negate it and write the negation in words.

Exercise 1.2.18. Let $f\colon X \to Y$ be a map. Then f is said to be onto (or surjective) if “$\forall y \in Y(\ldots)$”. Complete the statement in symbols and negate it. Explain when you say f is not onto in plain English. Carry out an analogous exercise for one-one (or injective) maps.

In the statement “$\forall x \in X\big(\exists y \in Y(y$ has some property $P)\big)$”, the element y may depend on x.

Consider the following examples.

Example 1.2.19. In the statement

> “For every human being x, there is human being y such that y is the father of x.”

You can clearly see that y depends on x. Consider the statement

$$\forall x \in \mathbb{R} \setminus \{0\}\big(\exists y \in \mathbb{R}(xy = 1)\big).$$

The meaning of the statement in fact is that every nonzero real number has a multiplicative inverse in $\mathbb{R}$. Here, if we take $x = 4$, then $y = 1/4$, whereas for $x = -1$ we have $y = -1$.

The order in which the quantifiers appear in the statement is important. If we change the order, the meaning of the statement may change.

Example 1.2.20. Consider the first statement in the Example 1.2.19. If you write it interchanging the quantifiers, it would give another statement

> “There is human being y such that for every human being x, y is the father of x.”

Do you see how the meaning of the statement changes? The latter statement asserts that there is a human being who is the father of every human being! Again, if you change the order of the quantifiers in the second statement in Example 1.2.19, you get the statement

$$\exists y \in \mathbb{R}\big(\forall x \in \mathbb{R} \setminus \{0\}(xy = 1)\big),$$

which asserts that there is a real number which is a multiplicative inverse of every nonzero real number! Is it same as saying that every nonzero real number has a multiplicative inverse in $\mathbb{R}$?

Example 1.2.21. We say that a subset $A \subseteq \mathbb{R}$ is bounded above (in $\mathbb{R}$) if there exists $\alpha \in \mathbb{R}$ such that α is an upper bound of A. In view of the Example 1.2.15, this can be written as

$$A \subseteq \mathbb{R} \text{ is bounded above in } \mathbb{R} \text{ if } \exists \alpha \in \mathbb{R}\big(\forall x \in A(x \leq \alpha)\big).$$

Let us interchange the quantifiers as follows and define A is bounded above in $\mathbb{R}$ if for every $x \in A$, there exists $\alpha \in \mathbb{R}$ such that $x \leq \alpha$. That is,

$$A \subseteq \mathbb{R} \text{ is bounded above in } \mathbb{R} \text{ if } \forall x \in A\big(\exists \alpha \in \mathbb{R}(x \leq \alpha)\big).$$

(Believe us, this "definition" is given by a few every time this course is taught!) What is wrong with this? If this were the definition, then any subset of $\mathbb{R}$ would be bounded above! For, if $x \in A$ is given, we let $\alpha = x$ (or $\alpha = x + 1$, if we wish to have a strict inequality!). We urge you to think over this again.

Example 1.2.22. As further examples, consider the following statement which assures the existence of an additive identity in $\mathbb{R}$:

$$\exists \theta \in \mathbb{R}\big(\forall x \in \mathbb{R}\ (x + \theta = \theta + x = x)\big).$$

Consider the sentence which assures the additive inverse in $\mathbb{R}$:

$$\forall x \in \mathbb{R}\ \big(\exists y \in \mathbb{R}\ (x + y = \theta = y + x)\big).$$

Note the order in which the quantifiers appear in these two sentences. The first one says that we have the same θ for any $x \in \mathbb{R}$ satisfying the conditions $x + \theta = x = \theta + x$. The second one says if we are given $x \in \mathbb{R}$, there exists y which may depend on x (in this case, it does!) such that $x + y = \theta$.

Of course, this does not mean that y must be different for different x. For example, consider the statement: given $x \in \mathbb{R}$, there exists $y \in \mathbb{R}$ such that $y = x^2$. For $x_1 = 1$ and $x_2 = -1$, we get the same $y = 1$.

Again, consider the example of "For every human being x there exists a human being y such that y is the father of x." Clearly, y depends on x. Let x_1 and x_2 be two distinct girls with y_1 and y_2, respectively, as fathers. If x_1 and x_2 are sisters, then $y_1 = y_2$.

Example 1.2.23. Let us now deal with a slightly more complicated statement. Suppose there is an orchard, full of trees. We make the following statement.

> "In each tree in the orchard, we can find a branch on which all of the leaves are green."

How do we turn this into a mathematical sentence? Let us fix some notations. We let T to denote the set of all trees in the orchard. For a tree $t \in T$ let B_t denote the set of all branches on t. For a branch $b \in B_t$ of the tree t let L_b denote the set of all leaves on the branch b. Now we are ready to cast the statement using quantifiers.

$$\forall t \in T\Big(\exists b \in B_t\big(\ \forall \ell \in L_b\ (\ell \text{ is green})\big)\Big).$$

How do we negate it? As we said earlier, we look at the outermost layer and negate it and move to the next inner layer. Thus, we get: There exists a tree $t \in T$ which does not have the property "$\exists b \in B_t\big(\forall \ell \in L_b(\ell \text{ is green})\big)$. We negate this and so on. Finally we arrive at

$$\exists t \in T\Big(\forall b \in B_t\big(\exists \ell \in L_b\ (\ell \text{ is not green})\big)\Big).$$

Compare these two displayed statements and pay attention to the quantifiers.

Do such complicated sentences occur naturally in mathematics? Yes. For example, when we define the convergence of a sequence of real numbers. Let (x_n) be a sequence of real numbers. We say that (x_n) converges to a real number $x \in \mathbb{R}$ if

$$\forall \varepsilon > 0\ \Big(\exists N \in \mathbb{N}\ \big(\forall n \geq N\ (|x_n - x| < \varepsilon)\big)\Big).$$

A more complicated one is when we say that a sequence (x_n) is convergent. We say that (x_n) is convergent if

$$\exists x \in \mathbb{R}\ \Big(\forall \varepsilon > 0\ \big(\exists N \in \mathbb{N}\ (\forall n \geq N\ (|x_n - x| < \varepsilon))\big)\Big).$$

You may try your hand in negating each of these!

1.3 Compound statements

Consider the following statements.

(a) Jack and Jill went up the hill.

(b) The boy is intelligent and handsome.

(c) Neither my father nor I have studied History.

Notice that each of them consists of two statements joined by "*and*". For example, the statement (a) may be read as "Jack went up the hill *and* Jill went up the hill". The statement (b) means "The boy is intelligent *and* the boy is handsome", and the statement (c) means "My father has not studied History *and* I have not studied History".

A statement of this kind is called a *conjunction* of statements, and is of the form "S and T", where S and T are statements.

Again consider the following statements.

(d) The book belongs to either John or Tom.

(e) Either India or South Africa will make to the final.

Each of them consists of two statements joined by "*or*". For example, the statement (d) may be read as "The book belongs to John *or* the book belongs to Tom". Similarly, the statement (e) can be read as "India will make to the final *or* South Africa will make to the final".

A statement of this kind is called a *disjunction* of statements, and is of the form "S or T", where S and T are statements. A statement which is a conjunction or disjunction is referred to as a *compound* statement. In the conjunction "S and T" and the disjunction "S or T", "and" and "or" are called *connectives*.

When is a composite statement true? Let S and T be two statements. The conjunction "S and T" is true means that each of S and T is true. For example, the statement (a) is true means that each of the statements "Jack went up the hill" and "Jill went up the hill" is true.

The disjunction "S or T" is true means that at least one of the two statements S and T is true. For example, the statement (e) is true means at least one of the statements "India will make to the final" and "South Africa will make to the final" is true.

Note that the statement (e) is a true statement if both India and South Africa make to the final. In other words, if both S and T are true statements, then "at least one of S and T is true" holds, and therefore "S or T" is a true statement. Thus, if "S and T" is true then "S or T" is true. This use of "or" in a disjunction is termed as *inclusive or*.

Example 1.3.1. Let A, B be subsets of a set X, and $x \in X$. The statement "$x \in A \cap B$" is the conjunction "$x \in A$ and $x \in B$", and the statement "$x \in A \cup B$" is the disjunction "$x \in A$ or $x \in B$".

Exercise 1.3.2. Let S and T be two statements.

(a) Suppose "S and T" is false and S is true. What can you say about T?

(b) Suppose "S or T" is true and S is false. What can you say about T?

(c) Suppose "S or T" is false. What can you say about S and T?

Exercise 1.3.3. Let $a, b \in \mathbb{R}$. Assume that $a \leq b$ is true. That is, the compound statement $(a = b)$ or $a < b$ is true. Assume that $a \neq b$. What can you conclude about a and b?

Negation of compound statements

How does one negate a compound statement? Suppose you want to buy a mobile phone which has a camera and Wi-Fi hotspot. If you reject a phone which the shopkeeper shows you, what could be the reason? The handset shown to you either lacked a camera or Wi-Fi hotspot, right?

When is the statement "The boy is intelligent and handsome" not true? It is so, if either the boy is not intelligent or he is not handsome. In general, we have the following.

The negation of "S and T" is "not-S or not-T".

Example 1.3.4. Let A, B be subsets of a set X, and $x \in X$. When is $x \notin A \cap B$ true? Note that $x \notin A \cap B$ is the negation of the statement "$x \in A \cap B$" which is the conjunction: "$x \in A$ and $x \in B$". The negation of this conjunction is "$x \notin A$ or $x \notin B$". Therefore, $x \notin A \cap B$ means "$x \notin A$ or $x \notin B$".

Consider the statement "Either India or South Africa will make to the final". The statement is true when at least one of India and South Africa makes to the final. So, the statement is false means none of India and South Africa makes to the final, right? In other words, the negation of the statement is "India will not make to the final and South Africa will not make to the final".

In general, we have the following.

The negation of "S or T" is "not-S and not-T".

Exercise 1.3.5. Let A, B be subsets of a set X, and $x \in X$. Write down when $x \notin A \cup B$ is true.

Exercise 1.3.6. Write down the negations of the statements (a)–(e) given in the beginning of the section.

Example 1.3.7. Let x denote a real number and consider the statement: "$x \geq 2$". This is a disjunction of two statements: "$x > 2$ or $x = 2$". Suppose the statement is false. Then, "$x > 2$" is false and "$x = 2$" is false. The law of trichotomy of reals says that exactly one of the following is true: (i) $x = 2$, (ii) $x > 2$, (iii) $x < 2$. Thus, the negation of "$x \geq 2$" means the statement "$x < 2$".

Example 1.3.8. Let us look the following statement which you must have been using all along.

Let $a, b \in \mathbb{R}$ be such that $a \leq b$ and $b \leq a$. Then $a = b$.

Let us have a close look at the hypothesis. It is a compound statement. There are two statements at the outermost: (i) $a = b$ or $a < b$ and (ii) $b = a$ or $b < a$. The statements (i) and (ii) are combined with the connective 'and'. Thus the hypothesis is

$$[\,(a = b) \text{ or } (a < b)\,] \text{ and } [\,(b = a) \text{ or } (b < a)\,].$$

Let S be the statement $a = b$ or $a < b$, so that it is a compound statement of the form S_1 or S_2 where S_1 is $a = b$ and S_2 is $a < b$. Let T be the statement $(b = a)$ or $b < a$. Thus, T is a compound statement of the form T_1 or T_2 where T_1 is $b = a$ and T_2 is $b < a$. Thus the hypothesis means that S is true and T is true.

We wish to conclude that $a = b$. Assume $a \neq b$. Since S is true and since $a \neq b$, we conclude that $a < b$ must be true. Similarly, since T is true and $b \neq a$, we deduce that $b < a$ is true. Thus if $a \neq b$, the hypothesis leads us to conclude that both $a < b$ and $b < a$ must hold. This violates the law of trichotomy of $\mathbb{R}$. Thus we are forced to conclude that $a = b$.

Compound statements with quantifiers

We now consider some compound statements involving quantifiers. Let X be a set, and P and Q properties which make sense for elements in X. Consider the statements.

$A : \exists x \in X\big(x \text{ has property } (P \text{ and } Q)\big)$.

$B : \exists x \in X\big(x \text{ has property } (P \text{ or } Q)\big)$.

$C : \forall x \in X\big(x \text{ has property } (P \text{ and } Q)\big)$.

$D : \forall x \in X\big(x \text{ has property } (P \text{ or } Q)\big)$.

How do you negate such statements? Let us take some examples.

Example 1.3.9. Consider the statement

"There is a lady in this room who is intelligent and beautiful".

The statement is of the form (A) with X as the set of ladies in the room, P the property "being intelligent", and Q the property "being beautiful". When is the statement false? It is so exactly when there is no lady in the room who is both intelligent and beautiful, that is, when you find that given any lady in the room, she is either not intelligent or not beautiful.

Note that the negation of the statement "x has property (P and Q)" is "x does not have property P or x does not have property Q". Thus, the negation of the statement in (A) is "$\forall x \in X(x$ does not have property (P and Q))", that is

Not-A: $\forall x \in X(x$ does not have property P or x does not have property Q).

Again, consider the statement

"Each lady in this room is either intelligent or beautiful,"

With X, P and Q as above, the statement can be written in the form of

$D : \forall x \in X\big(x \text{ has property } (P \text{ or } Q)\big)$.

By our earlier discussion, the negation of this statement is "$\exists x \in X\big(x$ does not have property (P or Q))", that is,

Not-D: $\exists x \in X\big(x$ does not have property P and x does not have property $Q\big)$.

Thus, the negation of the above statement is

"There is a lady in this room who is not intelligent and not beautiful".

Exercise 1.3.10. Write down the negations of the statements B and C.

Exercise 1.3.11. With S as the students in the class, P and Q as the properties being "intelligent" and "hardworking", respectively, first write the following sentences as mathematical statements using quantifiers, then negate them, and finally write the negations in plain English.

(a) There is a student in the class who is hardworking or intelligent.

(b) Every student in the class is hardworking or intelligent.

(c) There is a student in the class who is hardworking and intelligent.

(d) Every student in the class is hardworking and intelligent.

Example 1.3.12. Given two subsets A and B of a set X, we say that $A = B$ if $A \subseteq B$ and $B \subseteq A$. This says that for A to be equal to B, two statements must be simultaneously true, namely, $A \subseteq B$ and $B \subseteq A$. Thus

$$A = B \text{ if } \big(\forall x \in A \ (x \in B)\big) \text{ and } \big(\forall y \in B \ (y \in A)\big).$$

What is the negation of $A = B$? If $A = B$ is false, then it means that the statement $\forall x \in A \ (x \in B)$ is false or the statement $\forall y \in B \ (y \in A)$ is false. That is, $A = B$ is false means $\big(\exists x \in A \ (x \notin B)\big)$ or $\big(\exists y \in B \ (y \notin A)\big)$.

Thus we have

$$A \neq B \text{ if } \big(\exists x \in A \ (x \notin B)\big) \text{ or } \big(\exists y \in B \ (y \notin A)\big).$$

Remark 1.3.13. Suppose X is a set and P and Q are properties which make sense for the elements of X. You can easily convince yourself that the statements

- $\exists x \in X\big(x \text{ has property } (P \text{ or } Q)\big)$,
- $\big(\exists x \in X(x \text{ has property } P)\big)$ or $\big(\exists x \in S(x \text{ has property } Q)\big)$,

convey the same meaning. Similarly, the statements

- $\forall x \in X\big(x \text{ has property } (P \text{ and } Q)\big)$,
- $\big(\forall x \in X(x \text{ has property } P)\big)$ and $\big(\forall x \in X(x \text{ has property } Q)\big)$,

convey the same meaning.

Remark 1.3.14. In contrast, the two statements

- $\exists x \in X\big(x \text{ has property } (P \text{ and } Q)\big)$,
- $\big(\exists x \in X(x \text{ has property } P)\big)$ and $\big(\exists x \in X(x \text{ has property } Q)\big)$,

do not convey the same meaning. For example, with X, P and Q as in Example 1.3.9, the two statements respectively mean

- "There is a lady in this room who is intelligent and beautiful", and
- "There is a lady in this room who is intelligent and there is a lady in this room who is beautiful".

Note that there may be an intelligent lady and a beautiful lady in the room so that the second statement is true, even though the first statement is false. The first statement asserts the existence of an x in X which has both the properties P and Q, whereas the latter asserts an element x in X having the property P, and an element y (possibly different from x) in X having the property Q. It is evident that if the first statement is true, then the second is true. In case the latter statement is true and the two elements x and y are always distinct, then the first statement is not true.

Exercise 1.3.15. Let S be the set of all students is a class. Consider the following statements.

(a) There is a student in S who is 6 feet tall and who owns a BMW car.

(b) There is student in S who is 6 feet tall and there is student in S who owns a BMW car.

Do the two statements convey the same meaning? If the first statement is true, will the second be true?

Exercise 1.3.16. Consider the following two statements.

1. Every human being is either a male or a female.

2. Every human being is a male or every human being is a female.

Do they convey the same meaning?

Exercise 1.3.17. Argue that the following statements do not convey the same meaning.

(i) $\forall x \in S(x$ has property $(P$ or $Q))$,

(ii) $\big(\forall x \in S(x$ has property $P)\big)$ or $\big(\forall x \in S(x$ has property $Q)\big)$.

Suppose the second statement is true. Is the first statement true?

The upshot of all the discussions above is that you should be able to do the following exercise.

Exercise 1.3.18. Negate the statement

"Every street in the city has at least one house in which we can find a person who is either rich and beautiful or highly educated and kind".

We suggest that you first write this using quantifiers choosing appropriate notations and then negate it.

1.4 Implications

Consider the following statements.

(i) If the apple is red, then it is ripe.

(ii) In case the bakery is open, I will buy a cake for you.

(iii) The mobile handset has a camera implies that it has Wi-Fi hot-spot.

(iv) The two boys stay in the same house, if they are brothers.

(v) The boy is rich, if he owns a BMW car.

Notice that each of them consists of two statements, and asserts that one of the two statements implies the other. For example, the statement (i) consists of the statements "The apple is red" and "The apple is ripe", and asserts that the first implies the second, that is, (i) asserts that if "The apple is red" is a true statement, then "The apple is ripe" is a true statement.

Given two statements S and T, the statement "if S then T" is called an *implication* (or a *conditional statement*). The statement "if S then T" is one which asserts that "if S is a true statement, then T is a true statement".

When is an implication a true statement, and when is it false? Consider the statement "If the bakery is open, then I will buy a cake for you". Clearly, the statement is false only when the bakery was open and I did not buy any cake for you. What if the bakery is closed (not open)? As per the statement, I was not bound to buy a cake for you, since the condition that the bakery is open was not fulfilled. Thus, a statement "if S then T" is false only in the case "S is true and T is false". In all other cases, the statement is true. This also means that the negation of "if S then T" is the statement "S and not-T". Thus, the statement "If the apple is red, then it is ripe" is false means "The apple is red and it is not ripe".

An implication can be expressed in several different ways. For example, the following statements convey the same meaning.

(a) If the student is good in mathematics, then he is humble.

(b) The student is humble, if he is good in mathematics.

(c) The student is good in mathematics implies that he is humble.

(d) The student is good in mathematics only if he is humble.

(e) To be humble is necessary for the student to be good in mathematics.

(f) The student's being good in mathematics is sufficient to conclude that he is humble.

Observation 1.4.1. For two statements S and T the following are implications which convey the same meaning:

(a) If S then T.

(b) T if S.

(c) S implies T.

(d) S only if T.

(e) T is necessary for S.

(f) S is sufficient for T.

Exercise 1.4.2. Write each of the statements (i)–(v) (at the beginning of the section) in different forms of implications as in Observation 1.4.1.

Converse of implications

Consider the statements

(i) If the apple is red, then it is ripe.

(ii) If the apple is ripe, then it is red.

Whatever the first implication asserts, the reverse is asserted by the second. If the first is written as "if S then T", then the second is "if T then S". Each of the two statements are said to be the converse of the other.

If a statement S is "if P then Q", then the statement "if Q then P" is called the *converse* of S.

Consider the statement:

If the mobile handset has a camera, then it has Wi-Fi hot-spot.

The statement asserts that the mobile handset cannot have a camera without having Wi-Fi hot-spot. On the other hand, the converse of the statement, namely,

If the mobile handset has Wi-Fi hot-spot, then it has a camera,

asserts that the mobile handset cannot have Wi-Fi hot-spot without having a camera.

Exercise 1.4.3. Write the converse of the implications (i)–(v) in different forms as in Observation 1.4.1.

Example 1.4.4. For real numbers x and $a > 0$, consider the statements "$|x| < a$" and "$x \in (-a, a)$". Then the two statements "if $|x| < a$, then $x \in (-a, a)$" and "if $x \in (-a, a)$, then $|x| < a$" are converses of each other. Note that the two statements can also be written as

$$\text{“}|x| < a \Rightarrow x \in (-a, a)\text{” and “}x \in (-a, a) \Rightarrow |x| < a\text{”},$$

respectively.

'If and only if' statements

Consider the two implications

(a) If the student is sincere, then he is humble.

(b) If the student is humble, then he is sincere.

Note that the two statements are converses of each other. What is the conjunction of the two implications? In view of Observation 1.4.1 the two statements can be written as

(a) The student is humble, *if* he is sincere.

(b) The student is humble, *only if* he is sincere.

The conjunction of (a) and (b) is written as

The student is humble *if and only if* he is sincere.

What does this conjunction mean? It means that the student is humble exactly when he is sincere. The conjunction is an example of an "if and only if" statement.

Consider the statement "The apple is ripe *if and only if* it is red". It is the conjunction of the two statements, that is,

(The apple is ripe, if it is red) and (the apple is ripe, only if it is red).

The first part of this conjunction is the implication "If the apple is red, then it is ripe" and the second is the implication "If the apple is ripe, then it is red". The conjunction is an "if and only if" statement which means that the apple is ripe exactly when it is red. In other words, the apple being red and being ripe are one and the same thing.

For two statements S and T, the conjunction of the implication "if S then T" and its converse "if T then S" is written as "*T if and only if S*".

Example 1.4.5. Consider the statement "A man is a mammal with two eyes". What does the statement assert? If you look closely, it means that "A mammal is a man if and only if it has two eyes".

First, it says, "A mammal is a man only if it has two eyes". In other words, for a mammal to be a man, it is necessary that it has two eyes, that is, a mammal is a man implies that it has two eyes.

Second, the statement says that "A mammal is a man if it has two eyes". In other words, for a mammal to be a man, it is sufficient that it has two eyes, that is, a mammal is a man is implied by it's having two eyes.

Thus, the statement "*S if and only if T*" ("*S iff T*", in short) is conjunction of two statements which may be termed as the two parts of the "if and only if" statement, namely,

(i) ('Only if part', 'necessary part' or '$\Rightarrow$ part') "S only if T", since it means "T is necessary for S", that is, "S implies T".

(ii) ('If part', 'sufficient part' or '$\Leftarrow$ part') "S if T", since it means "T is sufficient for S", that is, "S is implied by T";

Example 1.4.6. For real numbers x and $a > 0$, consider the statements

$$\text{"}|x| < a \text{ if and only if } x \in (-a, a)\text{"}.$$

The 'only if' (or 'necessary' or '$\Rightarrow$') part of the statement is "If $|x| < a$, then $x \in (-a, a)$" and the 'if' (or 'sufficient' or '$\Leftarrow$') part of the statement is "if $x \in (-a, a)$, then $|x| < a$".

At times, one writes this statement as "$|x| < a \Leftrightarrow x \in (-a, a)$".

The statement "S if and only if T" is also expressed as "T is necessary and sufficient for S".

In Mathematics, you often come across definitions. For example, consider the following statement:

"An integer $p > 1$ is a *prime*, if the only positive divisors of p are 1 and p"

The statement defines when an integer is a prime. If $p > 1$, and 1 and p are the only positive divisors of p, then p is a prime. It also means that if p is a prime, then the only positive divisors of p are 1 and p. Thus, the definition actually means, for a positive integer $p > 1$.

"p is a *prime*, if and only if the only positive divisors of p are 1 and p".

In fact, every definition in mathematics is an 'if and only if' statement.

Suppose X is a set, and P and Q are properties which make sense to the elements of X. Consider the statement

$$\forall x \in X(\text{ if } x \text{ has the property } P, \text{ then } x \text{ has the property } Q).$$

What is the negation of this statement? Note that we can also write the statement as

$$\forall y \in \{x \in X : x \text{ has the property } P\}(y \text{ has the property } Q).$$

Therefore, the negation of the statement is

$$\exists y \in \{x \in X : x \text{ has property } P\} \ \big(y \text{ does not have the property } Q\big),$$

that is,

$$\exists x \in X\big(x \text{ has property } P \text{ and } x \text{ does not have the property } Q\big).$$

Exercise 1.4.7. Write the negations of the following statements

1. For each book in the library, if it has a preface, then it has a bibliography.
2. For any integer n, if n^2 is divisible by 10, then n is divisible by 10.

Contrapositive of implications

Suppose that "If the boy owns a BMW car, then he is rich" is a true statement. Now, suppose that the boy is not rich. Can he own a BMW car? No. Thus, "If the boy is not rich, then he does not own a BMW car" is a true statement.

Conversely, suppose that "If the boy is not rich, then he does not own a BMW car" is a true statement. Suppose the boy owns a BMW car. Can he be not rich? No. Thus "If the boy owns a BMW car, then he is rich" is a true statement. We therefore conclude that the two statements

1. "If the boy owns a BMW car, then he is rich", and
2. "If the boy is not rich, then he does not own a BMW car",

are equivalent, that is, each one implies the other. In other words, each of the two statements is true if and only if the other is true. We say that they are *contrapositive* of each other.

In general, consider an implication "if S then T". Suppose that the implication is true. If T is a false statement, then S cannot be a true statement. Thus, in this case "if not-T then not-S" is a true implication. On the other hand, suppose "if not-T then not-S" is a true implication. Then, "if S then T"

is a true statement. So, the two implications "if P then Q" and "if not-Q then not-P" are equivalent statements, that is, either both are true or both are false.

The implication "if not-Q then not-P" is called the *contrapositive* of "if P then Q". You may note that "if P then Q" is the contrapositive of "if not-Q then not-P".

Exercise 1.4.8. Write the contrapositives of the implications considered in Exercise 1.4.3.

Example 1.4.9. Let z be a complex number. If z is real, then $z^2 \geq 0$. Suppose I have a complex number z whose square does not satisfy $z^2 \geq 0$. Can z be real? Thus, if z^2 is not nonnegative, then z is not real. The two statements "For a complex number z, if z is real, then z^2 is nonnegative" and "For a complex number z, if z^2 is negative (i.e., not nonnegative), then z is not real" are contrapositive of each other. Note that they are equivalent statements.

Exercise 1.4.10. Write contrapositive of the following statements.

1. For an integer n, if $n^2 < 20$, then $n < 5$.
2. For an integer x, if $x^2 - 6x + 5$ is even, then x is odd.
3. For an integer m, if m^2 is not divisible by 4, then m is odd.
4. For real numbers x and y, if xy is an irrational number then x is irrational or y is irrational.

Since contrapositive of a statement is equivalent to the statement, proving the contrapositive of a given statement amounts to proving the statement. This is a very useful technique in Mathematics.

1.5 Proofs in Mathematics

In this section, we will acquaint you with some standard ways of reasoning to produce proofs of mathematical statements.

Most often, a statement that mathematicians prove is one of the forms

(i) $\forall x \in X(x$ has property $P)$,

(ii) $\exists x \in X(x$ has property $P)$.

Frequently, it happens that the property P in the statement (i) is of the form "if Q, then R", that is, "if x has the property Q, then x has the property R". Note that such a statement also can be written as

$$\forall x \in \{y : y \text{ has the property } Q\}(x \text{ has the property } R).$$

Thus, if we let $A := \{y \in X : y$ has property $Q\}$, then the statement can be written as

$$\forall x \in A(x \text{ has property } R).$$

Let us consider the statement. All students of our class who are hardworking are intelligent. Here X is the set of students in our class. Being 'hardworking' is P while Q is being intelligent. Hence the statement is

$$\forall x \in X(\text{if } x \text{ has } P, \text{ then } x \text{ has } Q).$$

Thus if we let $A := \{x \in X : x \text{ is hardworking }\}$, then the statement can be written as

$$\forall x \in A(x \text{ is intelligent}).$$

Many a time, a statement is also expressed as

$$\text{"There is no } x \in X \text{ which satisfy the property } P\text{"},$$

because it is easy to comprehend in that form. We understand that the statement means

$$\text{"}\forall x \in X(x \text{ does not satisfy the property } P).$$

How do you prove a statement of the form "$\forall x \in X(x$ has property $P)$". Consider the statement

$$\text{"Every integer in the set } X = \{2, 3, \ldots, 10\} \text{ has a prime divisor"}.$$

To prove this you can simply demonstrate one prime divisor for each of the integers in X. However, the same strategy will not work for the above statement with $X = \{m \in \mathbb{Z} : m \geq 2\}$, simply because you cannot exhaust showing prime divisors for each of $m \in X$, since X is infinite. What would you do? When the set X is of the form $\{m \in \mathbb{Z} : m \geq k\}$ for some integer k, one very useful method is to use the *principle of induction* which will be discussed in Chapter 5.

In general, when the set X is large (e.g., infinite) the usual way of proving the statement (i) is to assume x to be an arbitrary element in X and then to argue that x has the property P. Since x is taken to be arbitrary, the claim of the statement is proved. Consider the following example.

Example 1.5.1. To prove the statement "$\forall x \in (0,1)\big(x^2 < x\big)$", assume x to be an arbitrary real number x with $0 < x < 1$, that is, $0 < x$ and $x < 1$. Since x is a positive real number, multiplying both sides of $x < 1$ by x, we get $x^2 < x$. This completes the proof.

When you prove a statement of the form $\exists x \in X(x$ has property $P)$, there are two strategies. At times, it is possible to identify or construct an $x \in X$ satisfying the property P. In many cases, you put valid argument in support of the existence of such x, though you do not construct it explicitly.

Look at the following examples.

Example 1.5.2. To prove "$\exists a \in \mathbb{Z}(\forall b \in \mathbb{Z}(a + b = b))$" you would consider $a = 0 \in \mathbb{Z}$ and then say that a has the required property. Here you have identified an integer a.

Example 1.5.3. Suppose $\mathbb{R}^+$ denotes the set of positive real numbers and we need to prove the statement

$$\forall x \in \mathbb{R}^+\big(\exists y \in \mathbb{R}(0 < y < x)\big),$$

that is, "For each positive real number x, there is a real number y such that $0 < y < x$". Here we need to prove the existence of y for each $x \in \mathbb{R}^+$, and expect that y can be different for different x. As there are infinitely many $x \in \mathbb{R}^+$, we consider x to be an arbitrary positive real number. For the given x, consider then $y = x/2 \in \mathbb{R}$ and show that $0 < y < x$. Here, you have constructed y for x.

Example 1.5.4. Suppose we need to prove the statement "The set of natural numbers is not bounded above in $\mathbb{R}$" (see Example 1.2.15 for definitions). In terms of quantifiers, this statement can be written as

$$\forall x \in \mathbb{R}\big(\exists m \in \mathbb{N}(m > x)\big).$$

Let x be an arbitrary real number. How do we find a natural number m such that $m > x$? Imagine x on the real line. If $x \leq 0$, then we can choose $m = 1$. Let $x > 0$. You may recall, the greatest integer function $[x]$ on $\mathbb{R}$, where $[x]$ is the greatest integer less than or equal to x. Now it is easy to see that $m = [x] + 1$ will do the trick.

In Real Analysis, you will find that the given statement, known as the Archimedean property of $\mathbb{R}$, is shown to be a fallout of the least upper bound (LUB) property (or the order completeness axiom) of $\mathbb{R}$, and the existence of the function $[x]$, as a consequence of the Archimedean property.

Proving implications

For proving an implication $S \Rightarrow T$, we assume the validity of the statement S and using logical reasonings and mathematical facts, we show the statement T is true. Based on discussions we had in the previous sections, it is evident that there are different ways of proving the statement $S \Rightarrow T$. The standard ways of proving such a statement include *direct proof*, *proof by contradiction* and proving the *contrapositive* of the given statement.

Direct proofs

The basis of a direct proof of an implication is the following fundamental assumption in reasoning.

> If "S implies T" and "T implies R" are true statements, then "S implies R" is a true statement.

Thus, a proof of an implication "if S then T" amounts to producing finitely many statements $S_0, S_1, \ldots, S_k$ such that $S = S_0, T = S_k$ and "$S_0, S_1, \ldots, S_{i-1}$ together imply S_i" for $i = 1, 2, \ldots, k$. Providing such a finite sequence of statements is a direct proof of the statement "if S then T". Let us take a few examples.

Example 1.5.5. Suppose you are to prove the statement

"Square of an even number is divisible by 4".

We realize that we are going to prove the statement

$$\text{“}\forall n \in \mathbb{Z}(\text{if } n \text{ is even, then } n^2 \text{ is divisible by } 4)\text{”},$$

As in earlier examples, we take a generic representative of these integers and show that its square is divisible by 4. Thus your proof will be something like the following.

Proof. Let n be an even integer. Then, there is an integer k such that $n = 2k$. This gives $n^2 = 4k^2$. Since k^2 is an integer, we get 4 divides n^2. □

Exercise 1.5.6. Write the statements in Examples 1.5.1 and 1.5.3 as implications and write direct proofs.

Example 1.5.7. Suppose you want to prove the following familiar statement in your real analysis course:

“If a function is differentiable at a point, then it is continuous at that point”.

Note that the statement concerns a class (set) of functions defined on different possible domains; and also concerns different possible points in the domain. Suppose the statement refers to real functions of a single real variable.

How do you handle this? You will consider an arbitrary function $f : A \to B$ on arbitrary domain A and codomain B in $\mathbb{R}$ and an arbitrary point in A at which differentiability makes sense. Now you have made your stage ready with the statement “if f is differentiable at $x = a$, then f is continuous at $x = a$”. You will assume that the statement “f is differentiable at $x = a$” to be true and then produce a series of steps that lead to the conclusion that “f is continuous at $x = a$”. Compare the details in any analysis book.

Proving statements to be equivalent.

At times you need to prove a multiple of statements to be equivalent. For example, you can easily convince yourself the following:

Example 1.5.8. For given sets A and B the following are equivalent.

(i) $A \subseteq B$.
(ii) $A \cap B = A$.
(iii) $A \cup B = B$.

How do you prove equivalence of multiple statements? A standard practice is the following: arrange the statements in some suitable order: $S_1, S_2, \ldots, S_k$ and then prove the implications $S_1 \Rightarrow S_2, S_2 \Rightarrow S_3, \ldots, S_{k-1} \Rightarrow S_k$ and $S_k \Rightarrow S_1$. Then, in view of the principle listed at the beginning of this section, any two statements S_i and S_j are equivalent.

Exercise 1.5.9. Prove the assertion in Example 1.5.8 by arguing for (i) ⇒ (ii), (ii) ⇒ (iii) and (iii) ⇒ (i). (See Exercise 2.2.29 on Page 40)

Sometimes you need to prove statements of the form “if S, then (T or R)”. It is easy to see that the statement is equivalent to the statement “if (S and not-T), then R”. Consider the following example.

Example 1.5.10. To prove the statement

"For real numbers x and y, if $xy = 0$, then $x = 0$ or $y = 0$",

assume that $xy = 0$ and $x \neq 0$. Since $\frac{1}{x}$ is a real number, conclude that $y = \frac{1}{x}(xy) = \frac{1}{x} \cdot 0 = 0$.

Indirect Proofs

A. Proof using contrapositive.

We noted earlier that a statement "if S, then T" and its contrapositive "if not-T, then not-S" are equivalent statements. Thus, to prove a statement "$S \Rightarrow T$", we may prove its contrapositive. That is, we assume that T is a false statement and prove that S is false. The following examples illustrates this method.

Example 1.5.11. Suppose you are to prove the statement

"For an integer n, if $n^3 - 1$ is even, then n is odd".

Let n be a given integer. We need to prove that "if $n^3 - 1$ is even, then n is odd".

Suppose $n^3 - 1$ is even. Then there exists an integer k such that $n^3 - 1 = 2k$, that is, $n^3 = 2k + 1$. This gives that n^3 is odd. How do you see that n is odd?

So, the problem at your hand now is to prove "If n^3 is odd, then n is odd". What about proving its contrapositive "If n is even, then n^3 is even"? It is easy, right? Because, if $n = 2r$ for some integer, then $n^3 = 8r^3$.

What is the contrapositive of the original statement? It is "If n is even, then $n^3 - 1$ is odd". The above discussion suggests that we try to prove this contrapositive straightway. Then your proof runs as follows:

Proof. Suppose n is even. Then there exists integer r such that $n = 2r$. Consequently, $n^3 - 1 = 8r^3 - 1 = 2(4r^3) - 1$. This means $n^3 - 1$ is an odd integer. Thus, if n is even, then $n^3 - 1$ is odd. This proves that if $n^3 - 1$ is even, then n is odd. □

Example 1.5.12. Consider the statement

"For a real number $a \geq 0$, if for each $\epsilon > 0$ we have $a \leq \epsilon$, then $a = 0$".

Here, we are given a real number $a \geq 0$. The statement that we need to prove reads

$$\text{If } \forall \epsilon > 0 (a \leq \epsilon), \text{ then } a = 0.$$

What is its contrapositive? What does $a \neq 0$ mean here? Since $a \geq 0$, it means $a > 0$. Thus, the contrapositive of our statement is

$$\text{If } a > 0, \text{ then } \exists \epsilon > 0 (a > \epsilon).$$

To prove this statement, assume $a > 0$. You need to produce $\epsilon > 0$ such that $a > \epsilon$. How do you produce such an ϵ? If $a > 0$ then $a/2 > 0$. We can choose $\epsilon = a/2$.

Now a formal proof can be written as follows:

Proof. We prove the contrapositive of the statement. Suppose $a \neq 0$. Since $a \geq 0$, we have $a > 0$. Let $\epsilon = a/2$. We see that $a > \epsilon$. Thus, we have produced an $\epsilon > 0$ such that $a > \epsilon$. This completes the proof. □

Exercise 1.5.13. Give proofs using contrapositive of the following statements.

1. For integers x, y, if $x + y$ is even, then x and y are both odd or both even.
2. For an integer n, if $n^2 < 20$, then $n < 5$.
3. For an integer x, if $x^2 - 6x + 5$ is even, then x is odd.
4. For an integer m, if m^2 is not divisible by 4, then m is odd.
5. For real numbers x and y, if xy is an irrational number then x is irrational or y is irrational.

B. Proof by contradiction.

When is a statement is true? It is so, if its negation is false. This is the essence of many proofs in Mathematics. Such a proof assumes the statement which is to be proved as a false statement, that is, its negation as a true statement, and then produces a statement which is always false under the context.

We provide here a couple of examples.

Example 1.5.14. We prove that there exists no $\alpha \in \mathbb{R}$ such that the following holds: $\forall x \in \mathbb{R}(x \leq \alpha)$.

We shall prove this by contradiction. Assume such an $\alpha \in \mathbb{R}$ exists. Consider $y := \alpha + 1$. Then $y \in \mathbb{R}$ and we have $y > \alpha$. This contradicts our assumption that for each $x \in \mathbb{R}$, we have $x \leq \alpha$. Hence our assumption that such an α exists is wrong. Hence we conclude that no such real number exists.

A classic example is the following result which was proved in Euclid's *Elements*, Book IX (Proposition 20), more than 2400 years ago.

Theorem 1.5.15. *There are infinitely many prime numbers.*

The proof assumes the fact that every integer which is greater than 1 has a prime divisor. (We will see a proof of this fact in Chapter 5, see Example 5.2.5). We assume that there are only finitely many prime numbers and produce an integer having no prime divisor, a contradiction!

Proof. Suppose there are only finitely many prime numbers, say $p_1, \ldots, p_k$. Consider the integer $m + 1$, where $m = p_1 p_2 \cdots p_k$. Clearly, each p_i divides m. None of the primes $p_1, \ldots, p_k$ divides $m + 1$, because if p_i divides $m + 1$ for some i, then p_i divides $1 = (m + 1) - m$, which is not true. Since $p_1, \ldots, p_k$ are the only primes, we have $m + 1 > 1$ and $m + 1$ has no prime divisor, a contradiction. □

Let us take another example.

Example 1.5.16. Prove the following statement.

There is no rational number x such that $x^2 = 15$.

Note that the conclusion of the statement means that $\forall x \in \mathbb{Q}(x^2 \neq 15)$. Clearly, it is impossible to verify the conclusion of the statement considering all different rational numbers x. A natural way is to assume the negation of the statement and then arrive at a contradiction. So, you assume that there is a rational number x such that $x^2 = 15$. Then there are integers a, b such that $x = a/b$ and there is no common divisor of a and b other than ± 1. But then $a^2/b^2 = 15$, that is, $a^2 = 15b^2$. Since the integer 3 divides $15b^2$, 3 must divide a^2. This implies that, 3 being a prime number, 3 divides a, that is $a = 3k$ for some integer k. However, this gives $9k^2 = 15b^2$, that is, $3k^2 = 5b^2$. This implies that 3 divides $5b^2$, and therefore, 3 divides b. Thus, 3 is a common divisor of a and b, contrary to our assumption. Hence, there cannot be a rational number x such that $x^2 = 15$.

In fact, you can replace 15 by any positive integer m which is not a perfect square (that is, not a square of any integer) and give a similar proof. Use the fact that any integer $m \neq \pm 1$ has a prime divisor p.

Exercise 1.5.17. Prove the following statements by contradiction.

1. There is no greatest integer.
2. For $a, b \in \mathbb{Z}$, $a^2 - 4b \neq 2$.
3. For any real number $x \in [0, \pi/2]$, $\sin x + \cos x \geq 1$.

Proof by induction

There is another oft-used standard method of proving a statement, namely, the induction principle. Suppose that you have a situation where the set of context is $S = \{k \in \mathbb{Z} : k \geq m\}$, where m is a fixed integer, and the statement is "$\forall k \in S(k$ has property $S)$". A standard way of proving such statements is the *method of induction*. We will discuss this method in detail in Chapter 5.

Counterexamples

On many occasions you would like to disprove a statement of the form "$\forall x \in S(x$ has property $S)$". Note that the negation of the given statement is "$\exists x \in S(x$ does not have property $S)$". So, to disprove the given statement we produce an $x \in S$ for which the property S does not hold. Such an x is called a *counterexample* which disproves the given statement.

Let us take an examples.

Example 1.5.18. Consider the statement

> For integers a, b, c, if a divides bc, then a divides b or a divides c.

To disprove the statement we need to produce one instance of a, b, c such that a divides bc, but divides none of b and c. We have ample such triplets, e.g., $a = 4, b = c = 2$. Give two more counterexamples.

Exercise 1.5.19. Produce counterexamples to disprove the following statements.

1. For any $x \in \mathbb{R}$, $x^2 \geq x$.
2. For $x, y \in \mathbb{R}$, $x^2 = y^2$ implies $x = y$.
3. For $x, y \in \mathbb{R}$, $|a| > |b|$ if $a > b$.
4. For $x, y \in \mathbb{R}$, $x^2 + y^2 > 2xy$.
5. If $f : \mathbb{R} \to \mathbb{R}$ is continuous at a point $c \in \mathbb{R}$, then f is differentiable at c.

Chapter 2

Sets

In this chapter, we present an informal discussion on sets, and various notions pertaining to sets. The sets are most basic of all the mathematical objects. The aim of this chapter is to identify sets written in various forms, prove rigorously some results which you may already be aware of.

2.1 Basic terminologies

A *set* is a collection of objects known as *elements* or *members*. Elements of a set can be anything, such as numbers, lines, fishes, or even sets.

A set can be thought of as a box that can contain several elements (that is, the objects inside the box). As a box may be empty, we have the *empty set*, denoted by $\emptyset$, containing no element. A set with a single element is called a *singleton* set.

Usually, uppercase letters are used to label sets, and lowercase letters to label elements in a set. Let A be a set and x an object. Then, we write $x \in A$ (read as x belongs to A or x is in A) if x is an element of A. If x is not an element of A, then we write $x \notin A$ (read as x does not belong to A).

Since a set is determined by its elements, one way to write a set is to list all its elements and enclose them within curly braces. This is called the *roster method* of writing a set, and the list is known as a *roster*. The braces signify that a set has been defined. For example, the set of all integers strictly between 1 and 10 is

$$A = \{2, 3, 4, 5, 6, 7, 8, 9\}$$

Another example is

$$B = \{\text{the set of Mathematics books in a library, the set of English newspapers in India}\}.$$

Note that B is a set having two elements and each element is a set.

If the roster is too long, one uses ellipses (...). For example, if we wish to display the set of all even integers between 1 and 2000, we may write it as

$\{2, 4, 6, \ldots, 1998, 2000\}$. As another example, the set of positive integers which are perfect squares can be written as

$$\{1, 4, 9, 16, \ldots\}.$$

Using this strategy, we can write the sets like those of natural numbers and integers in roster form as

$$\mathbb{N} = \{1, 2, 3, \ldots\} \text{ and } \mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

We cannot expect to write all sets as rosters. In fact, most of the sets that we encounter in mathematics cannot be written as rosters.

Consider the set $\mathbb{Q}$ of rational numbers. Since the rational numbers are defined as ratio of integers, we may try to list them like

$$\ldots, \frac{-2}{3}, -2, \ldots, \frac{-1}{3}, \frac{-1}{2}, -1, 0, 1, \frac{1}{2}, \frac{1}{3}, \ldots, 2, \frac{2}{3}, \ldots$$

However, there are complications with the pattern and we are not quite sure that the list will exhaust all rational numbers. We will see that writing the set $\mathbb{R}$ of real numbers and many of its subsets as rosters is out of question.

The mostly used way of writing sets is by describing the properties of its elements which are satisfied only by them. This form of a set is usually called *set-builder form*. Consider the set of all real numbers that are strictly between 1 and 2. The set is denoted by $(1, 2)$, an *open interval*. The set is consisting of real numbers x which satisfy two properties, namely, $x > 1$ and $x < 2$. In set-builder form

$$(1, 2) = \{x \in \mathbb{R} : 1 < x \text{ and } x < 2\} = \{x \in \mathbb{R} : 1 < x < 2\}.$$

We read this as the set of all real numbers x such that $x > 1$ and $x < 2$.

The set of rational numbers can be written as

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z} \text{ and } q \neq 0 \right\}.$$

Note that while writing a rational number as $\frac{p}{q}$, we agree to the fact that $\frac{p'}{q'}$ is the same element as $\frac{p}{q}$ of $\mathbb{Q}$, if $p'q = q'p$.

Exercise 2.1.1. Show that the set of rational numbers can also be written as

$$\begin{aligned} \mathbb{Q} &= \left\{ \frac{p}{q} : p \in \mathbb{Z} \text{ and } q \in \mathbb{N} \right\} \\ &= \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}, \text{ and } p \text{ and } q \text{ have no common factors} \right\}. \end{aligned}$$

Example 2.1.2. Suppose S is the set of all three letter words in English uppercase alphabet. Note that there are 26^3 elements in the set, and it would be very cumbersome to write the set as a roster. However, in set-builder from you can write

$$S = \{xyz : x, y, z \in \{A, B, \ldots, Z\}\}.$$

The empty set

The set that has no element is called the *empty set* (or *null set*). It is denoted by $\{\}$ or $\emptyset$. It is preferable to refer to this set as "empty set", as the word "null set" has a different meaning in the context of measure theory, a branch of mathematics.

Example 2.1.3. Each of the following is the empty set.

(i) The set of people having 10 heads.

(ii) The set of real numbers whose square is negative.

(iii) Natural numbers which are both even and odd.

A set A is said to be *nonempty*, if A has at least one element.

Remark 2.1.4. The set $\{\emptyset\}$ is not same as $\{\}$ or $\emptyset$. For, $\{\emptyset\}$ is a set which contains an element, namely, the empty set.

Subsets

Suppose A, B are two sets. We say that A is a *subset* of B, if every element of A is also an element of B. In that case, we write $A \subseteq B$.

In other words,

$$A \subseteq B \text{ iff } \forall x \in A(x \in B).$$

If A is a subset of B, then B is called a *superset* of A and we write $B \supseteq A$.

Exercise 2.1.5. Let A, B be sets. When is A not a subset of B?

Equality of sets

Two sets A and B are said to be *equal* if they have the same elements. In other words, $A = B$ iff

$$\forall x \in A(x \in B) \text{ and } \forall x \in B(x \in A),$$

that is, iff

$$A \subseteq B \text{ and } B \subseteq A.$$

If $A \subseteq B$ and $A \neq B$, then we say that A is a *proper subset* of B, and we write $A \subset B$ or $A \subsetneq B$. Note that in this case every element of A is in B and there is an element of B which is not in A.

Example 2.1.6. Express $\mathbb{Z}$ as a proper subset of $\mathbb{Q}$.

How do you see this? If x is a rational number, then $x = p/q$ for some $p \in \mathbb{Z}, q \in \mathbb{N}$. When is x an integer k? It is, if $p = kq$. So, consider the subset

$$S = \{x \in \mathbb{Q} : \exists p \in \mathbb{Z}, q \in \mathbb{N} \text{ and } k \in \mathbb{Z} \text{ such that } x = p/q, p = kq\}$$

of $\mathbb{Q}$. If $k \in \mathbb{Z}$, then with $p = k, q = 1$ we have $k = k/1 \in S$, that is, $\mathbb{Z} \subseteq S$. If $x \in S$, then for some $q, k \in \mathbb{Z}$, $x = kq/q = k \in \mathbb{Z}$, that is, $S \subseteq \mathbb{Z}$. Thus, $\mathbb{Z} = S$, a subset of $\mathbb{Q}$.

Moreover, $\mathbb{Z} \neq \mathbb{Q}$, since (for example) $1/2 \in \mathbb{Q}$ and $1/2 \notin S = \mathbb{Z}$. Hence, $\mathbb{Z}$ is a proper subset of $\mathbb{Q}$.

Exercise 2.1.7. Show that $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$.

Example 2.1.8. Can you see that the empty set is a subset of any set A?

You need to show that "$\forall x \in \emptyset\,(x \in A)$". The last statement is false if and only if its negation "$\exists x \in \emptyset\,(x \notin A)$" is true. Can the negation be true? Surely not, since $\emptyset$ has no element. Therefore, "$\forall x \in \emptyset\,(x \in A)$", that is, $\emptyset \subseteq A$, is a true statement. Statements like $\emptyset \subseteq A$ are said to be *vacuously true* .

Example 2.1.9. Express the empty set as a subset of $\mathbb{Q}$ in two different ways.

What we are looking for is a set of rational numbers with some property, which is not satisfied by any rational number. We can think of many such properties which is not satisfied by any rational number. One such property is that there is no rational solution of the equation $x^2 = 2$. Thus we have

$$\emptyset = \{x \in \mathbb{Q} : x^2 = 2\} \subseteq \mathbb{Q}.$$

Think of some more ways of expressing the empty set as a subset of $\mathbb{Q}$.

Exercise 2.1.10. Express the empty set as a subset of $\mathbb{R}$ in (at least) two different ways.

Exercise 2.1.11. Is $\{h : h \text{ is human being with 5 legs}\} = \{x \in \mathbb{R} : x^2 + 1 = 0\}$?

In mathematics, most often the sets are described in set-builder form. Therefore it is important that you can "identify" the set given by a set-builder construction in a more concrete way. For example, we have seen the empty set as a subset of different sets in Example 2.1.3 and Ex.2.1.11. Consider the following example.

Example 2.1.12. Identify the set $A := \{x \in \mathbb{R} : x + \frac{1}{x} \geq 2\}$.

In such cases, the first thing to do would be to check whether some elements are there or not. Does $0 \in A$? No, since $1/x$ should make sense. Does $-1 \in A$? No, in fact, we observe that if $x < 0$, then so is $1/x$ and hence their sum is also negative. So, an obvious conclusion is that A is a subset of the set of positive real numbers. Does $1 \in A$? Yes. How about $1/2$? Yes. More generally how about $1/n$ with $n \in \mathbb{N}$? Yes, they all lie in A. How about $x \geq 2$? Yes, it belongs to A.

Now, we take $x > 0$. Then

$$x + \frac{1}{x} \geq 2 \;\Leftrightarrow x^2 + 1 \geq 2x \;\Leftrightarrow x^2 - 2x + 1 \geq 0.$$

The last inequality is $(x-1)^2 \geq 0$, which is anyway satisfied by x. Therefore, if $x > 0$, $x \in A$.

We have shown that A is a subset of the set of positive real numbers and that the set of positive real numbers is a subset of A. We therefore conclude that $A = \{x \in \mathbb{R} : x > 0\}$.

As we have seen above, it is very much possible to "describe" the same set in two different ways in set builder form. A frequent problem is to show that the two describe the same set.

Example 2.1.13. Consider the set S of solutions of $ax+by = 0$ with $a \neq 0 \neq b$. The set can be described in two ways:

$$\begin{aligned} S_1 &:= \{(x, y) \in \mathbb{R}^2 : y = -(a/b)x\} \\ S_2 &:= \{(x, y) \in \mathbb{R}^2 : x = -(b/a)y\}. \end{aligned}$$

Note how the sets S_1 and S_2 are described. When does $(u, v) \in \mathbb{R}^2$ lie in S_1? It is so, if $u \in \mathbb{R}$ and $v = -(a/b)u$. Thus you get an element (x, y) in S_1 corresponding to each real numbers x and choosing $y = -(a/b)x$. Similarly, you get elements (x, y) in S_2 corresponding to different real numbers y and choosing $x = -(b/a)y$.

To see that $S = S_1$ first suppose $(x, y) \in S$, that is, $(x, y) \in \mathbb{R}^2$ such that $ax + by = 0$. Then $by = -ax$. Since $b \neq 0$, we have $y = -(a/b)x$, and therefore $(x, y) \in S_1$. This shows that $S \subseteq S_1$. Conversely, if $(x, y) \in S_1$, then $y = -(a/b)x$, i.e., $by = -ax$, i.e., $ax + by = 0$. Thus, $(x, y) \in S$. This shows that $S_1 \subseteq S$. We therefore conclude that $S = S_1$. Similarly you can show that $S = S_2$.

However, our concern is whether you can see that $S_1 = S_2$ without referring to S. Indeed, because $a \neq 0 \neq b$, for $(x, y) \in \mathbb{R}^2$, we have

$$(x, y) \in S_1 \;\Leftrightarrow\; y = -(a/b)x \;\Leftrightarrow\; x = -(b/a)y \;\Leftrightarrow\; (x, y) \in S_2.$$

This implies that $S_1 \subseteq S_2$ and $S_2 \subseteq S_1$. Consequently, $S_1 = S_2$.

Exercise 2.1.14. Identify the following sets and justify:

(a) $\{x \in \mathbb{R} : \exists y \in \mathbb{R}$ such that $y^2 = x\}$. Describe the set in words.

(b) $\{x \in \mathbb{R} : \exists y \in \mathbb{R}$ such that $y^{2n} = x\}$, where n is a fixed natural number.

(c) $\{x \in \mathbb{R} : \exists y \in \mathbb{R}$ such that $y^{2n+1} = x\}$, where $n > 1$ is a fixed natural number.

(d) $\{\frac{m}{n} \in \mathbb{Q} : m, n$ have the same sign and n is a divisor of $m\}$.

(e) $S \subseteq \mathbb{N}$ such that $1 \in S$ and if $k \in S$, then $k + 1 \in S$.

(f) $\{x \in \mathbb{Q} : \exists a, b \in \mathbb{Z}$ such that $a > 0$ and $ax - b = 0\}$.

(g) $\{x \in \mathbb{Q} : \exists a, b \in \mathbb{Z}$ such that $a > 0$ and $x^2 + ax + b = 0\}$.

(h) $\{x \in \mathbb{R} : e^x = 0\}$.

Hints: (a), (b) and (c) require some background in real analysis (you may refer to [1]). The set in (a) is the set of nonnegative real numbers. The set in (c) is $\mathbb{R}$. (e) $S = \mathbb{N}$. It is a restatement of the induction principle. (See Section 5.1.)

Exercise 2.1.15. Express $\mathbb{Z}$ as a subset of $\mathbb{R}$ using some trigonometric functions.
Hint: When is $\sin x = 0$?

Example 2.1.16. Let $a, b \in \mathbb{R}$. The following sets are called *intervals* in $\mathbb{R}$.

(a) $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ (a closed interval).

(b) $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ (an open interval).

(c) $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$ (a closed-open interval).

(d) $[a, \infty) = \{x \in \mathbb{R} : x \geq a\}$ (a closed ray).

Exercise 2.1.17. For what choices of a and b the above intervals are nonempty? Define similarly an open-closed interval $(a, b]$, a closed ray $(-\infty, a]$ and open rays $(-\infty, a)$ and $(a, -\infty)$. How many conditions should a real number x satisfy to belong to (a, b)? How many conditions should x satisfy to lie in (a, ∞)?

Exercise 2.1.18. Suppose $a, b \in \mathbb{R}$ and $a \geq b$. Identify the following subsets of $\mathbb{R}$ on the real line: $(a, b), (a, b]$ and $[a, b]$.

Example 2.1.19. Identify the set $A = \{x \in \mathbb{R} : \big||x-2| - |x-4|\big| = 3\}$.

Note that the expressions inside the outer modulus changes sign in the intervals at 2 and 4. So, we consider three cases: (i) $x \leq 2$, (ii) $2 < x < 4$ and (iii) $x \geq 4$.

Case 1: $x \leq 2$. We have $|x-2| = -x+2$ and $|x-4| = -x+4$. Therefore,

$$|x-2| - |x-4| = -x+2+x-4 = -2,$$

which gives $\big||x-2| - |x-4|\big| = 2 \neq 3$. Thus, no such x belongs to A.

Case 2: $2 < x < 4$. We have $|x-2| = x-2$ and $|x-4| = -x+4$. Therefore,

$$|x-2| - |x-4| = x-2+x-4 = 2x-6.$$

Since $2 < x < 4$, we have

$$-2 = 2 \cdot 2 - 6 < 2x - 6 < 2 \cdot 4 - 6 = 2,$$

and therefore $|2x-6| < 2$. Thus, no such x belongs to A.

Case 3: $x \geq 4$. We have $|x-2| = x-2$ and $|x-4| = x-4$. Therefore,

$$|x-2| - |x-4| = x-2-x+4 = 2,$$

which gives $\big||x-2| - |x-4|\big| = 2 \neq 3$. Thus, no such x belongs to A.

We therefore conclude that A is the empty set.

Exercise 2.1.20. Identify the set $B = \{x \in \mathbb{R} : \big||x-2| - |x-4|\big| = 2\}$.

Hint. See the previous example. $B = \{x \in \mathbb{R} : x < 2\} \cup \{x \in \mathbb{R} : x > 4\}$.

Exercise 2.1.21. Identify the set $C = \{x \in \mathbb{R} : \big||x-2| - |x-4|\big| = 1\}$.

Hint. See the previous example. $C = \{3/2, 7/2\}$

Exercise 2.1.22. Identify the sets $A = \{x \in \mathbb{R} : \big||x-2| + |x-4|\big| = 1\}$, $B = \{x \in \mathbb{R} : |x-2| + |x-4| = 2\}$ and $C = \{x \in \mathbb{R} : |x-2| + |x-4| = 3\}$.

Exercise 2.1.23. Let $a, b \in \mathbb{R}$ with $a < b$. Let $c \in \mathbb{R}$. Identify the sets $A = \{x \in \mathbb{R} : |x-a| + |x-b| = c\}$ and $B = \{x \in \mathbb{R} : ||x-a| - |x-b|| = c\}$. *Hint:* The last few exercises should alert you about the various possibilities and the answers which depend on the relation between $b-a$ and c. Do you recall how the ellipses and hyperbolas are defined? Thinking geometrically may help.

Exercise 2.1.24. Describe the following sets explicitly and mark them on the real line:
(a) $A_0 := \{x \in \mathbb{R} : x < 0\}$.
(b) $A_1 := \{x \in \mathbb{R} : x(x-1) < 0\}$.
(c) $A_2 := \{x \in \mathbb{R} : x(x-1)(x-2) < 0\}$.
(d) $A_3 := \{x \in \mathbb{R} : x(x-1)(x-2)(x-3) < 0\}$.

Hint: (c) Whenever you need to identify a set in set-builder form, always try to find some elements in the set and find some elements which are not in the set. It is easy to see that 0, 1, and 2 are not in A_2. Now, look at what happens to real numbers other than 0, 1, and 2. If x is any such real number, then there are four possibilities (i) $x < 0$, that is, $x \in (-\infty, 0)$, (ii) $x \in (0, 1)$, (iii) $x \in (1, 2)$ and (iv) $x > 2$, that is, $x \in (0, \infty)$. Now, look at each of the possibilities and conclude that $A_2 = (-\infty, 0) \cup (1, 2)$ (see Figure 2.1).

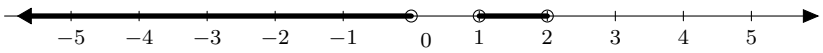


Figure 2.1: $A_2 := \{x \in \mathbb{R} : x(x-1)(x-2) < 0\}$

The next exercise is the generalization of the previous one.

Exercise 2.1.25. Let $a_1 < a_2 < \cdots < a_n$ be real numbers. Describe the set

$$\{x \in \mathbb{R} : (x - a_1) \cdots (x - a_n) < 0\}.$$

Exercise 2.1.26. Identify the set $\{(x, y) \in \mathbb{R}^2 : xy \neq 0 \text{ and } \frac{x}{y} + \frac{y}{x} \geq 2\}$ and plot them in the xy-plane.
Hint: The inequality above should remind you a similar inequality in $\mathbb{R}$. What is the inequality? Have you seen it in an earlier example? Use it to solve the problem in $\mathbb{R}^2$.

Exercise 2.1.27. Let $S := \{(x, y) \in \mathbb{R}^2 : (1 - x)(1 - y) \geq 1 - x - y\}$. Give a simple description of S which involves signs of x and y.

Exercise 2.1.28. Let $A := \{x \in \mathbb{R} : x^2 > x + 6\}$ and $B := \{x \in \mathbb{R} : x > 3\}$. Which of the following is true? (i) $A \subseteq B$, (ii) $B \subseteq A$?

Exercise 2.1.29. Identify the set $S := \{(x, y) \in \mathbb{R}^2 : |x| \leq |y|\}$. Draw its picture as a subset of $\mathbb{R}^2$.

2.2 Operations on sets

We now use connectives to define different set operations which allow us to generate more sets.

Union and intersection of sets

Let A, B be two sets. The *union* of A and B is the set defined by

$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

That is, $x \in A \cup B$ if and only if $x \in A$ or $x \in B$.

The union of two sets is the set obtained by collecting all the elements of both the sets.

The *intersection* of A and B is the set defined by

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

That is, $x \in A \cap B$ if and only if $x \in A$ and $x \in B$.

The intersection of two sets is the collection of all elements which are common to the sets.

Exercise 2.2.1. If $A \subseteq B$, then what are $A \cup B$ and $A \cap B$?

Exercise 2.2.2. What do you mean by saying (i) $x \notin A \cup B$ and (ii) $x \notin A \cap B$?

Exercise 2.2.3. Let $A := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ and $B := \{(x, y) \in \mathbb{R}^2 : x = 1\}$. Find $A \cap B$. Do you understand this geometrically?

Exercise 2.2.4. Let A be the set of integers which are divisible by 2, and B the set of integers which are divisible by 3. Find $A \cap B$.

Exercise 2.2.5. Let A be the set of integers which are divisible by 4, and B the set of integers which are divisible by 6. Find $A \cap B$.

Exercise 2.2.6. Let $A = \{x \in \mathbb{R} : x^2 \geq 0\}$ and $B = \{x \in \mathbb{R} : x^3 \geq 0\}$. Find $A \cup B$ and $A \cap B$.

Exercise 2.2.7. Let A be the set of $n \times n$ real symmetric matrices, and B the set of $n \times n$ real skew-symmetric matrices. Find $A \cap B$.

Exercise 2.2.8. Suppose A, B and C are three sets. Prove the following associativity properties:
(a) $A \cup (B \cup C) = (A \cup B) \cup C$.
(b) $A \cap (B \cap C) = (A \cap B) \cap C$.

Exercise 2.2.9. Suppose $B \subseteq C$. Prove that for any set A
(i) $A \cup B \subseteq A \cup C$.
(ii) $A \cap B \subseteq A \cap C$.

Exercise 2.2.10. (i) If $A \subseteq C$ and $B \subseteq C$, then show that $A \cup B \subseteq C$.
(ii) If $A \subseteq B$ and $A \subseteq C$, then show that $A \subseteq B \cap C$.

Theorem 2.2.11. *The intersection of sets distributes over the union of sets. In other words, for any sets A, B and C*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Proof. To prove that the sets on two sides are equal, we need to show
(i) $A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$, and
(ii) $(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)$.
For (i) note that

$$\begin{aligned} x \in A \cap (B \cup C) &\Rightarrow x \in A \text{ and } x \in B \cup C \\ &\Rightarrow x \in A \text{ and } (x \in B \text{ or } x \in C) \\ &\Rightarrow (x \in A \text{ and } x \in B) \text{ or } (x \in A \text{ and } x \in C) \\ &\Rightarrow x \in A \cap B \text{ or } x \in A \cap C \\ &\Rightarrow x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

Do you think that the implications above can be reversed? If yes, what do you get? What is its relevance to (ii)? Nevertheless, we shall give a direct proof of (ii).

Note that $B \subseteq B \cup C$ and $C \subseteq B \cup C$. Taking intersection by A on both sides we get $A \cap B \subseteq A \cap (B \cup C)$ and $A \cap C \subseteq A \cap (B \cup C)$. Now taking unions, we get

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C).$$

Note that we have used the Exercises 2.2.9 and 2.2.10. □

Using similar arguments, one can prove the following theorem.

Theorem 2.2.12. *The union of sets distributes over the intersection of sets. In other words, for any sets A, B and C*

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

We can generalize the definitions of union and intersection to a finite number of sets $A_1, A_2, \ldots, A_n$ as follows.

The *union* of A_i, $1 \leq i \leq n$, is the set

$$\bigcup_{i=1}^{n} A_i = \{x : \exists k \in \{1, \ldots, n\}(x \in A_k)\}.$$

In other words, $x \in \bigcup_{i=1}^{n} A_i$ iff $x \in A_k$ for some $k \in \{1, \ldots, n\}$.

The intersection of $A_1, A_2, \ldots, A_n$ is the set

$$\bigcap_{i=1}^{n} A_i = \{x : \forall k \in \{1, \ldots, n\}(x \in A_k)\}.$$

In other words, $x \in \bigcap_{i=1}^{n} A_i$ iff $x \in A_k$ for each $k \in \{1, \ldots, n\}$.

Exercise 2.2.13. For sets $A_1, \ldots, A_n$, show that (i) $A_i \subseteq \bigcup_{i=1}^{n} A_i$ for each i, (ii) $\bigcap_{i=1}^{n} A_i \subseteq A_i$ for each i, and (iii) $\bigcap_{i=1}^{n} A_i \subset \bigcup_{i=1}^{n} A_i$.

Exercise 2.2.14. For $k \in \{1, 2, \ldots, 10\}$ define $A_k = \{m \in \mathbb{Z} : m^2 \leq 2k\}$. Find $A_1 \cup A_2 \cup \cdots \cup A_{10}$ and $A_1 \cap A_2 \cap \cdots \cap A_{10}$.

Exercise 2.2.15. For $k \in \{1, 2, \ldots, 10\}$ define $B_k = \{m \in \mathbb{Z} : m \text{ divides } k\}$. Find $B_1 \cup B_2 \cup \cdots \cup B_{10}$ and $B_1 \cap B_2 \cap \cdots \cap B_{10}$.

Disjoint sets

Two sets A and B are said to be *disjoint* if $A \cap B = \emptyset$, that is, if A and B do not have any element in common.

For example, $\{m \in \mathbb{Z} : \exists k \in \mathbb{Z} \text{ such that } m = 2k\}$, the set of even integers, and $\{m \in \mathbb{Z} : \exists k \in \mathbb{Z} \text{ such that } m = 2k+1\}$, the set of odd integers, are disjoint sets.

Similarly, the set of positive real numbers and the set of negative real numbers are disjoint.

Example 2.2.16. $A = [1, 3)$ and $B = [3, 4]$ are disjoint whereas $[-1, 1]$ and $[0, 2]$ are not disjoint.

Exercise 2.2.17. Suppose m and n are distinct positive integers. Let A and B be sets of all integral multiples of m and n, respectively. Are A and B disjoint?

Exercise 2.2.18. Consider $A := \{(x, y) \in \mathbb{R}^2 : x + y = 0\}$ and $B := \{(x, y) \in \mathbb{R}^2 : x + y = 1\}$. Are they disjoint?

Exercise 2.2.19. Let $A := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ and $B := \{(x, y) \in \mathbb{R}^2 : xy = 1\}$. Prove that they are disjoint.

Set difference

For sets A and B the *set difference* of B from A is defined to be the set

$$A \setminus B = \{x \in A : x \notin B\}.$$

Thus, $x \in A \setminus B$ means that $x \in A$ and $x \notin B$. The shaded portion in Figure 2.2 represents $A \setminus B$.

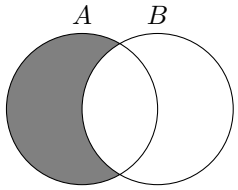


Figure 2.2: Set difference $A \setminus B$

Exercise 2.2.20. For a set A, what is $A \setminus A$? What is $A \setminus B$, if A and B are disjoint? What is $A \setminus B$, if $A \subseteq B$?

Exercise 2.2.21. Let $A = (-5, 2)$ and $B = [-1, 4]$. Find (i) $A \cup B$, (ii) $A \cap B$ (iii) $A \setminus B$ (iv) $B \setminus A$ (v) $(A \cup B) \setminus (A \cap B)$ (vi) $(A \setminus B) \cup (B \setminus A)$. Did you notice that the last two sets are equal? See Ex. 2.2.24 below.

Exercise 2.2.22. Let $A = \mathbb{R}^2$ and $B = \{(x, y) \in \mathbb{R}^2 : |x| < 1\}$. Draw the picture of $A \setminus B$.

Exercise 2.2.23. Let A be the set of $n \times n$ real symmetric matrices, and B the set of $n \times n$ real skew-symmetric matrices. Describe the elements of $A \setminus B$.

Exercise 2.2.24. For any two sets A and B, show that

$$(A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A).$$

Exhibit the set $(A \cup B) \setminus (A \cap B)$ as a shaded region in $A \cup B$.

Symmetric difference of sets

For sets A and B the *symmetric difference* of A and B is defined to be the set

$$A \Delta B = (A \setminus B) \cup (B \setminus A).$$

Thus, $x \in A\Delta B$ if it belongs to one of A and B, but not both. In view of Exercise 2.2.24, we have $A\Delta B = (A \cup B) \setminus (A \cap B)$.

Exercise 2.2.25. Draw the figure for $A\Delta B$ similar to Figure 2.2 and prove the following:

(i) $A\Delta B = B\Delta A$.

(ii) $A\Delta(B\Delta C) = (A\Delta B)\Delta C$.

(iii) $A \cup B = (A\Delta B)\Delta(A \cup B)$.

Exercise 2.2.26. Let $A := \{(x, y) \in \mathbb{R}^2 : x \leq 1\}$, $B := \{(x, y) \in \mathbb{R}^2 : y \leq 1\}$, $C := \{(x, y) \in \mathbb{R}^2 : x \geq -1\}$ and $D := \{(x, y) \in \mathbb{R}^2 : y \geq -1\}$. Draw the pictures of these sets. Find $A \cap C$ and draw its picture. Show that the union of these four sets is $\mathbb{R}^2$. Can you choose two of these sets whose union is $\mathbb{R}^2$? What is $A \setminus B$? What is their symmetric difference?

Complement of a set

Whenever we consider a set, its elements are chosen from some set which we call the *universe* or the *universal set.* In general, the universal set is considered as the totality of elements under consideration. For example, in calculus, universal set is the set of real numbers. In number theory, universal set most often is the set of integers, etc.

If A is a part of the universe U, that is, if $A \subseteq U$, then the complement of A (in U), written as A^c, is defined to be

$$A^c := U \setminus A = \{x \in U : x \notin A\}.$$

If $A \subset S$, then we call the set $S \setminus A = \{x \in S : x \notin A\}$ the *complement* of A in S.

Exercise 2.2.27. Let $A = (-1, 1)$, $B = [1, 1]$ and $C = (0, \infty)$. Find A^c, B^c and C^c. (It follows from the context that the universe is $\mathbb{R}$.)

The next result deals with the relationship between complements of the union and intersection of sets. Suppose A and B are two sets. What is meaning of $x \in (A \cap B)^c$? This means $x \notin A \cap B$, that is, the negation of the compound statement "$x \in A$ and $x \in B$", that is, "$x \notin A$ or $x \notin B$". Similarly, $x \in (A \cup B)^c$ means "$x \notin A$ and $x \notin B$".

Theorem 2.2.28 (De Morgan's Law)**.** *Let A and B be subsets of a universal set U. Then*
(a) $(A \cap B)^c = A^c \cup B^c$
(b) $(A \cup B)^c = A^c \cap B^c$.

Proof. We prove (a) and leave (b) as an exercise.

To prove (a), we need to show that $(A \cap B)^c \subseteq A^c \cup B^c$ and $A^c \cup B^c \subseteq (A \cap B)^c$.

Let $x \in (A \cap B)^c$. This implies $x \notin A \cap B$ which means "$x \notin A$ or $(x \notin B)$". This implies "$x \in A^c$ or $x \in B^c$", that is, $x \in A^c \cup B^c$. Hence $(A \cap B)^c \subseteq A^c \cup B^c$.

Next, $x \in A^c \cup B^c$ means "$x \in A^c$ or $x \in B^c$", that is, "$x \notin A$ or $x \notin B$" which is the negation of "$x \in A$ and $x \in B$" (that is, $x \in A \cap B$)). Thus, $x \in A^c \cup B^c$ implies $x \in (A \cap B)^c$, and therefore, $A^c \cup B^c \subseteq (A \cap B)^c$. □

We take up a problem posed in the previous chapter. Refer to Example 1.5.8 and the subsequent discussion.

Example 2.2.29. Let A, B be sets. Show that the following are equivalent.

(a) $A \subseteq B$.

(b) $A \cup B = B$.

(b) $A \cap B = A$.

(d) $B^c \subseteq A^c$.

Solution: What has been claimed here is that any two of the above statements are equivalent, that is, $(r) \Leftrightarrow (s)$ for $r, s \in \{a, b, c, d\}$. We prove the result by showing $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$.

$(a) \Rightarrow (b)$. Assume (a), that is, $A \subseteq B$. To prove (b), we need to prove $B \subseteq A \cup B$ and $A \cup B \subseteq B$. Note that $B \subseteq A \cup B$ is always true.

Let $x \in A \cup B$. This implies $x \in A$ or $x \in B$. If $x \in B$, then we are done. If $x \in A$, then $x \in B$, since $A \subseteq B$. Thus, $x \in A \cup B$ implies $x \in B$. Hence $A \cup B \subseteq B$. This proves $A \cup B = B$.

$(b) \Rightarrow (c)$. Assume $A \cup B = B$. We need to prove that $A \cap B = A$. Note that $A \cap B \subseteq A$.

To prove $A \subseteq A \cap B$, we let $x \in A$. Then $x \in A \cup B = B$. Hence $x \in A \cap B$. Thus, $A \subseteq A \cap B$, and we have $A \cap B = A$.

$(c) \Rightarrow (d)$. Assume $A \cap B = A$. Let $x \in B^c$. We claim that $x \in A^c$. Suppose this is not true, then $x \in A$. This implies $x \in A = A \cap B$, which means $x \in B$. This is a contradiction, to our assumption $x \in B^c$.

$(d) \Rightarrow (a)$. Assume $B^c \subseteq A^c$. Let $x \in A$. We want to show $x \in B$. Suppose this is not true. Then $x \in B^c$. Since $B^c \subseteq A^c$, we have $x \in A^c$, which is contradiction to the assumption that $x \in A$. Hence, $x \in B$, and therefore $A \subseteq B$. □

We urge the reader to supply proofs for the other implications directly, and also to write shorter proofs using Exercises 2.2.9 and 2.2.10.

2.3 Family of sets

Many a time in mathematics we deal with sets whose elements are sets, for example set of intervals, set of lines, set of triangles, etc. Suppose we have 30 sets. Then it is natural to denote these sets as $A_1, A_2, \ldots, A_{30}$. We say that the sets are *labeled* or *indexed* by the numbers from the set $\{1, 2, \ldots, 30\}$.

Suppose we consider intervals of the form $I_n = [n, n+1]$ for natural numbers n. Then, the sets are labeled or indexed by the set of natural numbers.

We may consider intervals of the form $F_r = [r, r+1]$ for real numbers r. The sets are labeled or indexed by the real numbers.

Let U_{India} denote the set of universities in India, U_{USA} denote the set of universities in USA. In general, suppose U_x is the set of universities in the county x. Here, the sets of universities are labeled or indexed by the set with elements as the countries.

In each of the above cases, we have a family of sets indexed by a set, called the *index set*.

Suppose Λ is a nonempty set, and for each $\alpha \in \Lambda$ there is a set A_α. Then, we have a *family of sets indexed by* Λ which is written as

$$\{A_\alpha : \alpha \in \Lambda\}.$$

Here, the set Λ is called the *index set* for the family.

Union of a family of sets

Let $\{F_\alpha : \alpha \in \Lambda\}$ be a family of sets indexed by a nonempty set Λ. Let us assume that each of the sets in the family is a subset of some universal set U. Recall how we defined the union and the intersection of a finite collection of sets. In a similar manner, we define the union of the family of sets by

$$\bigcup_{\alpha\in\Lambda} F_\alpha = \{x \in U : \exists \alpha \in \Lambda \text{ such that } x \in F_\alpha\}.$$

That is, $\bigcup_{\alpha\in\Lambda} F_\alpha$ is the set of all elements of U which belong to at least one of F_α's.

Example 2.3.1. Let us consider sets $F_k = \{-k, k\}$ for $k \in \mathbb{N}$. This family is indexed by the natural numbers $\mathbb{N}$ and each member in the family is a subset of $\mathbb{Z}$. What is $\bigcup_{k\in\mathbb{N}} F_k$?

Let $F = \bigcup_{k\in\mathbb{N}} F_k$. Note that for each $k \in \mathbb{N}$, $F_k \subseteq \mathbb{Z}$. Therefore, we have $F \subseteq \mathbb{Z}$. Further, for each $k \in \mathbb{N}$, $0 \notin \{-k, k\} = F_k$, and therefore, $F \subseteq \mathbb{Z} \setminus \{0\}$. Conversely, if $n \in \mathbb{Z} \setminus \{0\}$, then $n \in F_{|n|}$. We, therefore, conclude that $F = \mathbb{Z} \setminus \{0\}$.

Example 2.3.2. Consider the family

$$F_r = \{(x, y) \in \mathbb{R}^2 : x = r, y \in \mathbb{R}\}$$

of subsets of $\mathbb{R}^2$ indexed by $\Lambda = \mathbb{R}$. What is $\bigcup_{r\in\mathbb{R}} F_r$?

Note that F_r is the vertical line $x = r$ in the plane $\mathbb{R}^2$. If $(x, y) \in \mathbb{R}^2$, then $(x, y) \in F_x$. This implies that $\mathbb{R}^2 \subseteq \bigcup_{r\in\mathbb{R}} F_r$. Can you conclude now that $\bigcup_{r\in\mathbb{R}} F_r = \mathbb{R}^2$?

Exercise 2.3.3. Express $\mathbb{N}$ as the union of an infinite number of finite sets I_n indexed by $n \in \mathbb{N}$.

Exercise 2.3.4. Express $\mathbb{R}$ as the union of an infinite number of intervals J_n of finite length, indexed by $n \in \mathbb{N}$.

Exercise 2.3.5. Express $\mathbb{R}$ as the union of an infinite number of intervals J_n of infinite length, indexed by $n \in \mathbb{N}$.

Exercise 2.3.6. Express $\mathbb{R}$ as the union of an infinite number of intervals J_x of finite length, indexed by $x \in \mathbb{R}_+$, the set of positive reals.

Exercise 2.3.7. Express $\mathbb{R}$ as the union of an infinite number of intervals J_x of infinite length, indexed by $x \in \mathbb{R}_+$, the set of positive reals.

Intersection of a family of sets

Let $\{F_\alpha : \alpha \in \Lambda\}$ be a family of sets indexed by a nonempty set Λ. Let us assume that each of these sets is a subset of some universal set U. Then the intersection of this family is defined by

$$\bigcap_{\alpha\in\Lambda} F_\alpha = \{x \in U : \forall\alpha \in \Lambda(x \in F_\alpha)\}.$$

That is, $\bigcap_{\alpha\in\Lambda} F_\alpha$ is the set of all elements of U which belong to all of F_α's.

Example 2.3.8. Keeping the same notation, we prove that

$$\bigcap_{\alpha\in\Lambda} F_\alpha \subseteq \bigcup_{\alpha\in\Lambda} F_\alpha.$$

Let $x \in \bigcap_{\alpha\in\Lambda} F_\alpha$. Since $\Lambda \neq \emptyset$, there exists $\alpha \in \Lambda$. Since x belongs to the intersection, $x \in F_\alpha$. Hence, we have $x \in \bigcup_{\alpha\in\Lambda} F_\alpha$. Notice that we made use of the fact that the index set is nonempty.

Example 2.3.9. Let us consider the family of sets $F_k = \{-k, 0, k\}$ for $k \in \mathbb{N}$. This family is indexed by the natural numbers $\mathbb{N}$ and each member in the family is a subset of $\mathbb{Z}$. What is $\bigcap_{k\in\mathbb{N}} F_k$?

Clearly $0 \in F_k$ for each k. If $k \neq 0$ is an integer, then $k \notin F_{|k|+1}$. This shows that $\bigcap_{k\in\mathbb{N}} F_k = \{0\}$.

The next few examples and exercises need Archimedean property. We shall state this result without proof. For a proof refer to [1].

Archimedean Property

Given any real number x, there exists a natural number n such that $n > x$. In other words, the set $\mathbb{N}$ is not bounded above in $\mathbb{R}$.

Example 2.3.10. Let the universal set be $\mathbb{R}$, and the index set $\mathbb{N}$. For a natural number n, define $J_n = (0, 1/n)$. Identify the set $\bigcap_{n\in\mathbb{N}} J_n$.

Draw the intervals $J_1 = (0,1), J_2 = (0,1/2), J_n = (0,1/3)$, etc., on the real line and observe what is happening. You might have observed that the interval is shrinking towards 0. If we let $x \in \bigcap_{n\in\mathbb{N}} J_n$, then we have $0 < x < 1/n$ for each natural number n. In particular, $x > 0$. However, we claim that there exists no positive real number x such that $\forall n \in \mathbb{N}(x < 1/n)$. This will imply that $\bigcap_{n\in\mathbb{N}} J_n = \emptyset$.

Suppose x is a positive real number. Then, $1/x \in \mathbb{R}$. By Archimedean property, there exists a natural number m such that $m > 1/x$, and hence $1/m < x$. Thus, there is an $m \in \mathbb{N}$ such that $x \notin (0,1/m) = J_m$. Therefore $x \notin \bigcap_{n\in\mathbb{N}} J_n$. This proves our claim.

Example 2.3.11. Let the universal set be $\mathbb{R}$, and the index set $I = (1,2)$. For $t \in I$, let $A_t := \{x \in \mathbb{R} : x < t\}$. What is the union $\bigcup_{t\in I} A_t$?

We are sure that you would conclude the union to be $(-\infty, 2) = \{x \in \mathbb{R} : x < 2\}$. Do you see that this requires a proof? Note that $2 \notin I$!

Let us supply the details. It is clear that the union is a subset of $(-\infty, 2)$. (Why?) Let $x \in (0,2)$. Since $x < 2$, the mid point $t := (x+2)/2$ satisfies $x < \frac{x+2}{2} < 2$. Hence $x \in A_t$.

What is $\bigcap_{t\in I} A_t$? Almost all students say it is $(-\infty, 1)$. Do you see that $1 \in A_t$ for each $t \in I$?

Exercise 2.3.12. Express the singleton set $\{0\}$ as $\bigcap_{n\in\mathbb{N}} J_n$, where each J_n is an open interval. Also $\{0\}$ as $\bigcap_{n\in\mathbb{N}} J_n$, where each J_n is a closed interval.

Exercise 2.3.13. Express $[0,1]$ as $\bigcap_{n\in\mathbb{N}} J_n$, where J_n is an open interval. Can you find open intervals J_n such that $[0,1] = \bigcup_{n\in\mathbb{N}} J_n$?

Pairwise disjoint family of sets

Let $\{A_\alpha : \alpha \in \Lambda\}$ be a family of sets indexed by a set Λ. Then, the family is said to be *disjoint* if

$$\bigcap_{\alpha\in\Lambda} A_\alpha = \emptyset.$$

The family is said to be a *pairwise disjoint* if for every pair of elements $\alpha, \beta \in \Lambda$, with $\alpha \neq \beta$, we have $A_\alpha \cap A_\beta = \emptyset$.

Among the above two notions, which one is stronger? See Figure 2.3.

Exercise 2.3.14. Show that if a family of sets is pairwise disjoint, then it is disjoint.

Example 2.3.15. The families of sets in Examples 2.3.1 and 2.3.2 are disjoint (in fact, pairwise disjoint). The family of sets in Example 2.3.9 is not disjoint.

Exercise 2.3.16. Let $\mathcal{C}$ be the set of all circles in $\mathbb{R}^2$. Can you write it as a family indexed by $\Lambda = \mathbb{R}^2 \times (0,\infty)$? Is the family disjoint? Pairwise disjoint?

Exercise 2.3.17. Give an example of a family of sets which is disjoint but not pairwise disjoint.

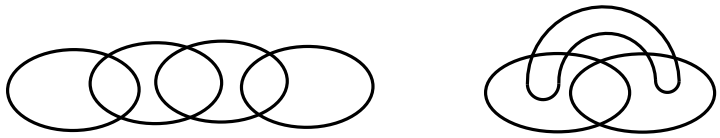


Figure 2.3: Disjoint family of sets which is not pairwise disjoint

Exercise 2.3.18. Let X be a finite set of $n \geq 3$ elements. Let $\mathcal{A}$ be the collection of subsets of X having $n-1$ elements. How many elements are there in $\mathcal{A}$? Prove that any $n-1$ subsets in $\mathcal{A}$ have a nonempty intersection while the family $\mathcal{A}$ is disjoint.

Example 2.3.19. Let $M(n,\mathbb{R})$ denote the set of $n \times n$ real matrices. For each real number $r \in \mathbb{R}$, define $\mathcal{M}_r = \{A \in M(n,\mathbb{R}) : \det(A) = r\}$.

We have $\mathcal{M}_r \subseteq M(n,\mathbb{R})$, and for $A \in M(n,\mathbb{R})$, $A \in \mathcal{M}_{\det(A)}$. Therefore, we get $M(n,\mathbb{R}) = \bigcup_{r\in\mathbb{R}} \mathcal{M}_r$. Let us check whether the family is pairwise disjoint. Suppose $\mathcal{M}_r \cap \mathcal{M}_s \neq \emptyset$ and $A \in \mathcal{M}_r \cap \mathcal{M}_s$. Then $r = \det(A) = s$, that is, $r = s$. Therefore, if $r \neq s$, then $\mathcal{M}_r \cap \mathcal{M}_s = \emptyset$. (Note that we proved the contrapositive.)

Thus, we have expressed $M(n,\mathbb{R})$ as the union of a pairwise disjoint family of sets $\mathcal{M}_r$ indexed by the set $\mathbb{R}$.

Exercise 2.3.20. Express $\mathbb{R}^2$, the xy-plane, as a pairwise disjoint union of a family of lines indexed by $\mathbb{R}$.

Exercise 2.3.21. Express $\mathbb{R}^2 \setminus \{(0,0)\}$ as a pairwise disjoint union of a family of circles.

2.4 Power sets

From a given set, there is a natural way to construct a family of sets. One can simply consider the family consisting of all subsets of the given set.

For a set S, the *power set* of S is defined to be the family of all subsets of S, and is denoted by $P(S)$. Thus

$$P(S) = \{A : A \subseteq S\}.$$

Exercise 2.4.1. Can you see that $P(A)$ is always nonempty? What is the power set $P(\emptyset)$ of the empty set?

Exercise 2.4.2. Let X be the set of all prime numbers between 1 and 10. Write down explicitly $P(X)$.

Exercise 2.4.3. If the set A contains n elements, then show that $P(A)$ contains 2^n elements. (*Hint.* Use induction.)

2.5 Cartesian product of sets

The *Cartesian product* of two sets A and B, denoted by $A \times B$, is defined as the set $\{(a, b) : a \in A,\ b \in B\}$. Here (a, b) is called an *ordered pair*. It is ordered in the sense that (a, b) is not same as (b, a), unless $a = b$, since a and b have been written in different order. (An *unordered pair* is a set $\{a, b\}$.)

Example 2.5.1. Let $A = \{1, 2, 3\}$ and $B = \{a, b\}$. The

$$A \times B = \{(1, a), (1, b), (2, a), (2, b), (3, a), (3, b)\}$$

and

$$B \times A = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\}.$$

Note that $A \times B \neq B \times A$.

We can generalize Cartesian product to finitely many sets $A_i, 1 \leq i \leq n$, as follows:

$$A_1 \times A_2 \cdots \times A_n := \{(a_1, a_2, \ldots, a_n) : a_i \in A_i, 1 \leq i \leq n\}.$$

The Cartesian product of n copies of $\mathbb{R}$ is denoted by $\mathbb{R}^n$. That is

$$\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : \text{ for each } i \in \{1, 2, \ldots, n\}, x_i \in \mathbb{R}\}.$$

Let $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ be elements in $A_1 \times A_2 \cdots \times A_n$. Then

$$x = y \text{ iff } \forall\ i \in \{1, 2, \ldots, n\}(x_i = y_i)$$

and

$$x \neq y \ \text{ iff } \ \exists\ i \in \{1, 2, \ldots, n\}(x_i \neq y_i).$$

Exercise 2.5.2. When is $A \times B = \emptyset$?

Given nonempty sets X and Y, we form the Cartesian product $X \times Y$. How do we create subsets of $X \times Y$? One easy way is to start with subsets $A \subseteq X$ and $B \subseteq Y$ and form the Cartesian product $A \times B$. Note that

$$A \times B = \{((x, y) \in X \times Y : x \in A \text{ and } y \in B\}.$$

Thus, $A \times B$ is a subset of $X \times Y$. Let us look at some examples.

Example 2.5.3. Let $X = \mathbb{R} = Y$.

1. Take $A = \{1\}$ and $B = \mathbb{R}$. The picture of $A \times B$ as a subset of $\mathbb{R}^2$ is given in Figure 2.4. Draw the pictures of $B \times A$.

2. Take $A = [0, \infty)$ and $B = (-1, 1)$. The picture of $A \times B$ as a subset of $\mathbb{R}^2$ is given in Figure 2.5. Draw the pictures of $B \times A$.

3. Take $A = \mathbb{Z}$ and $B = \mathbb{R}$. The picture of $A \times B$ as a subset of $\mathbb{R}^2$ is given in Figure 2.6. Draw the picture of $B \times A$.

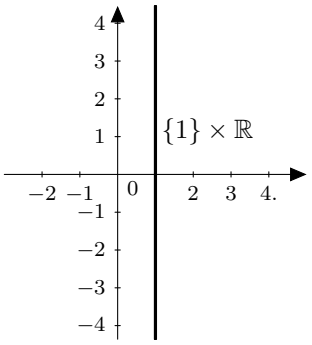


Figure 2.4: $\{1\} \times \mathbb{R}$

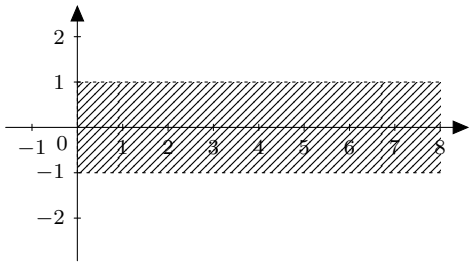


Figure 2.5: $[0, \infty) \times (-1, 1)$

Exercise 2.5.4. Let $X = \mathbb{R} = Y$. Draw the pictures of $A \times B$ for the following subsets A and B.
(a) $A = [-1, 1], B = [2, 3]$
(b) $A = (-1, 1), B = (2, 3)$
(c) $A = [-1, 1), B = (2, 3]$.

Exercise 2.5.5. Given a map $f\colon X \to Y$, think of a "natural" subset of $X \times Y$ associated with f.

Exercise 2.5.6. Given a subset $A \times B \subset X \times Y$ and elements $(x_1, y_1), (x_2, y_2) \in A \times B$, can you think of a few other (possibly new) elements in $A \times B$?

While construction of subsets of the form $A \times B$ is important, the students should be aware of the fact that there are subsets of $X \times Y$ that are not of the form $A \times B$. Let us look at an example.

Example 2.5.7. Let $C = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$. Then $C \subseteq \mathbb{R}^2$ and it is not of the form $A \times B$ for any subsets $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}$.

To see this draw the circle C. Suppose there exist subsets A and B of $\mathbb{R}$ such that $C = A \times B$. Think of some obvious points on C. For example, take two points $(1, 0), (0, 1) \in C = A \times B$. Since $(1, 0) \in A \times B$, $1 \in A$ and $0 \in B$. Similarly, $(0, 1) \in A \times B$ implies $0 \in A$ and $1 \in B$. Thus $0 \in A$ and $0 \in B$. Therefore $(0, 0) \in A \times B = C$, which is not true.

Exercise 2.5.8. Let $L = \{(x, y) \in \mathbb{R}^2 : x = y\}$. Show that L is not of the form $A \times B$ for any subsets $A, B \subseteq \mathbb{R}$.

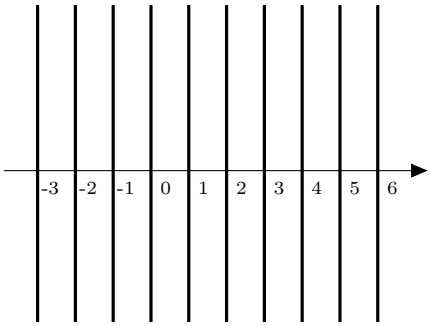


Figure 2.6: $\mathbb{Z} \times \mathbb{R}$

To acquire confidence in working with Cartesian product, the reader may solve some of the exercises below.

Exercise 2.5.9. Let $A, C \subseteq X$ and $B, D \subseteq Y$.
(a) True or false? $A \times B \subseteq C \times D$ iff $A \subseteq C$ and $B \subseteq D$.
(b) True or false? $(A \cap C) \times (B \cap D) = (A \times B) \cap (C \times D)$.
(c) True or false? $(A \times C) \cup (B \times D) = (A \cup B) \times (C \cup D)$.

Exercise 2.5.10. Let A, B and C be three sets. Show that the following hold.
(i) $A \times (B \cup C) = (A \times B) \cup (A \times C)$.
(ii) $A \times (B \cap C) = (A \times B) \cap (A \times C)$.
(iii) $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.

Exercise 2.5.11. Prove or disprove: If $A \times B = A \times C$, then $B = C$.

Exercise 2.5.12. Under what conditions $A \times B = C \times D$ implies $A = C$ and $B = D$?

Exercise 2.5.13. Prove or disprove: For any sets A and B, $P(A \times B) = P(A) \times P(B)$.

Chapter 3

Functions

This chapter deals with basic notions of functions which are building blocks of mathematics. We are sure that students have learned some functions and various concepts related to functions in school. Our aim is to revise these concepts and put them on a rigorous footing.

3.1 Basic definitions

Definition 3.1.1 (Function)**.** Let X and Y be nonempty sets. A *function* (or *map* or *mapping*) f from X to Y is a rule (or correspondence, association) that assigns to each element in X, a unique element in Y.

If f is a function from X to Y, we write it as $f\colon X \to Y$ or as $X \xrightarrow{f} Y$.

Note that for a correspondence or association from X to Y to be a function two things are important: (i) each element in X must be associated to an element in Y and (ii) no element in X should be associated to more than one element in Y.

Let us look at an example. Let $X = Y$ be the set of all human beings. If the 'rule' f associates each $x \in X$ to the father of x, then it is a function. On the other hand, if g associates x to a parent of x, it is not a function, since $g(x)$ could be the father of x or the mother of x!

An example closer to Mathematics is the following. Let X be the set of positive reals and $Y = \mathbb{R}$. For each $x \in X$, let f be the 'rule' which sends x to its square root. Is f a function? No, if $x = 4$, $f(4)$ could be either 2 or -2. But on the other hand, if g is the rule which associates to each x in X, the positive square root, then g is a function.

If f is function from X to Y, then X is called the *domain* of f and Y is called the *codomain* of f. If $x \in X$ is associated to the element y in Y, we say that y is the *image* of x under f and write $y = f(x)$. If $y = f(x)$, then x is called a *preimage* of y.

The *range* of $f\colon X \to Y$ is defined to be the subset

$$R(f) := \{y \in Y : \exists x \in X \text{ such that } y = f(x)\}.$$

of Y, that is, the set of images of all elements of X.

In the sequel, when we write $f\colon X \to Y$, we mean that X and Y are nonempty sets and f is a function from X to Y. When $X = Y$, we call a function $f\colon X \to X$ a *function* on X.

Some examples

Example 3.1.2. Let X, Y be nonempty sets. Fix $y_0 \in Y$. The function $f\colon X \to Y$ defined by $f(x) = y_0$ for each $x \in X$ is called a *constant* function.

For example, $f\colon \mathbb{R} \to \mathbb{R}$ with $f(x) = 0$ for $x \in \mathbb{R}$ is a constant function, called the *zero function* on $\mathbb{R}$.

Example 3.1.3. For any nonempty set X, the function $f\colon X \to X$ defined by $f(x) = x$ for $x \in X$ is called the *identity* function on X. The identity function on X is denoted by Id_X.

Example 3.1.4. Let X be a nonempty set and $X \subseteq Y$. Then we have a function $f\colon X \to Y$ defined by $f(x) = x$ for $x \in X$. The function f is called the *inclusion* function (also called *embedding*) of X in Y.

Example 3.1.5. For a nonnegative integer n, and real numbers $a_0, a_1, \ldots, a_n$, $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is called a polynomial in $\mathbb{R}$. If $p(x)$ is a polynomial, we have a function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = p(x)$, called a *polynomial* function. Note that the zero function, the identity function and the functions given by $x^2, x^3, \ldots$ are polynomial functions on $\mathbb{R}$.

Example 3.1.6. The trigonometric (also called circular) functions $\sin, \cos$ are functions on $\mathbb{R}$. What are the domains for the functions given by $\tan, \cot, \sec$ and cosec? What are the domains for the inverse circular functions given by $\arcsin, \arccos$ and $\arctan$?

Example 3.1.7. The function on $\mathbb{R}$ given by e^x is called the *exponential* function and the function from $(0, \infty)$ to $\mathbb{R}$ given by $\log x$ (or sometimes $\ln x$) is called the *natural logarithm*.

Example 3.1.8. Let X be the set of students of your class. Can you think of a function from X to $\mathbb{N}$? How about the age of a student as the number of days since he/she is born? Or, how about the total marks in his/her last board examination?

If A is the set of all members of the parliament, can you think of a function from A to $\mathbb{N}$ which involves "votes"?

For functions $f\colon X \to Y$ and $g\colon U \to V$, it is natural to ask when f and g are equal.

Definition 3.1.9 (Equality of functions)**.** Let X, Y, Z, W be nonempty sets, and $f\colon X \to Y$ and $g\colon Z \to W$ be two functions. We say that f is *equal* to g, and write $f = g$, if

(i) $X = Z$ and $Y = W$, that is, domains and codomains of f and g are equal, and

(ii) for each $x \in X = Z$, we have $f(x) = g(x)$.

Thus, if f and g are functions from X to Y, then $f = g$ if and only if for each $x \in X$, we have $f(x) = g(x)$.

Let us look at some examples.

Example 3.1.10. Consider the functions $f\colon \mathbb{Z} \to \mathbb{Z}$ and $g\colon \mathbb{N} \to \mathbb{Z}$ defined by $f(x) = g(x) = x^2$. Is $f = g$? The answer is no, as the domains of f and g are not equal.

Example 3.1.11. Let $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. Let $g\colon \mathbb{R} \to [0, \infty)$ be defined by $g(x) = x^2$. Are f and g equal? If we go by our definition, then they are not equal, since their codomains are different!

Exercise 3.1.12. Let $f\colon \mathbb{Z} \to \mathbb{Z}$ be given by $f(m) = m + 1$. What is the range of f?

Let $g\colon \mathbb{N} \to \mathbb{N}$ be given by $g(n) = n + 1$. What is its range?

Example 3.1.13. Let $X = \{-1, 0, 1\}$ and $f, g\colon X \to \mathbb{R}$ be defined by $f(x) = x$ and $g(x) = x^3$ for $x \in X$. Then $f = g$, since $f(x) = g(x)$ for each $x \in X$.

Exercise 3.1.14. Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be defined as $f(x) := (x - 2)^2 + 1$ and $g(x) := x^2 - 4x + 5$. Check if $f = g$.

Exercise 3.1.15. Let $f, g\colon X \to Y$. If $R(f) = R(g)$. Can we conclude $f = g$?

Definition 3.1.16 (Graph of a function)**.** Let $f\colon X \to Y$ be a function. Then the *graph* $G(f)$ of f is the subset of $X \times Y$ defined by

$$G(f) := \{(x, y) \in X \times Y : y = f(x)\}.$$

You must have dealt with real valued functions of real numbers earlier. Graph of such functions are subsets of the Cartesian plane $\mathbb{R}^2$. We are certain that you must have plotted graphs of some of the standard functions such as $f(x) = x, 3x + 5, x^2, |x|, \sin x, \cos x, e^x$ and $\log x$.

Suppose we are given a subset S of $\mathbb{R}^2$. Can we find $X \subseteq \mathbb{R}$ and a function $f\colon X \to \mathbb{R}$, such that $S = G(f)$? This is not true in general. However, if every vertical line in $\mathbb{R}^2$ intersects S at most once, then S is indeed the graph of a function. This is called the *vertical line test* of function.

Exercise 3.1.17. If $S \subseteq \mathbb{R}^2$ satisfies the above property, how do you find the function? What is its domain?

Exercise 3.1.18. Verify the following.

(i) The circle $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ is not graph of any function. (Look at Figure 3.1.)

(ii) The semi-circle $\{(x, y) \in \mathbb{R}^2 : y \geq 0 \text{ and } x^2 + y^2 = 1\}$ is the graph of the function $f\colon [-1, 1] \to \mathbb{R}$ defined by $f(x) = \sqrt{1 - x^2}$.

(iii) The set $\{(x, y) \in \mathbb{R}^2 : x = |y|\}$ is not graph of any function. (Look at Figure 3.2.)

(iv) The set $\{(x, y) \in \mathbb{R}^2 : x^2 + 4y = 0\}$ is the graph of a function.

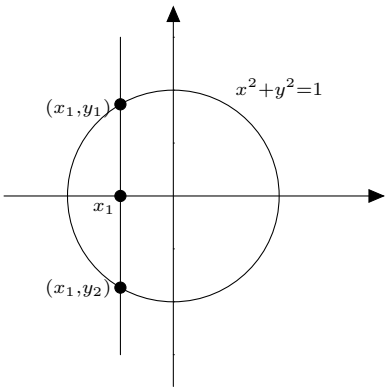


Figure 3.1: $x^2 + y^2 = 1$ is not a graph.

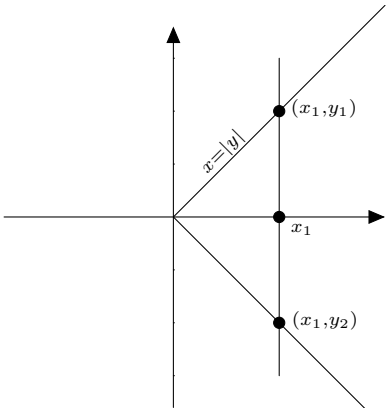


Figure 3.2: $x = |y|$ is not a graph.

Exercise 3.1.19. For any real number x, define $[x]$ to be the integer m such that $m \leq x < m+1$. (You need analysis to show that there exists such a unique integer m. See [1], for example.) Note that $[x]$ is the greatest integer less than or equal to x. Draw the graph of $[x]$ on $[-5, 6)$.

Definition 3.1.20. Let X be a nonempty set and $f, g\colon X \to \mathbb{R}$. The *maximum* $\max\{f, g\}$ and the *minimum* $\min\{f, g\}$ of f and g are functions from X to $\mathbb{R}$ defined as follows:

$$\max\{f, g\}(x) := \max\{f(x), g(x)\} \quad \min\{f, g\}(x) := \min\{f(x), g(x)\}.$$

Exercise 3.1.21. Plot the graph of $\max\{f, g\}$ and $\min\{f, g\}$, where
(i) $f(x) = x, g(x) = -x$ defined on $[-2, 2]$.
(ii) $f(x) = x^2, g(x) = x^3$ defined on $[-3, 3]$.
(iii) $f(x) = \sin x, g(x) = \cos x$ defined on $[-2\pi, 2\pi]$.

Example 3.1.22. Recall the definition of the *modulus* function on $\mathbb{R}$:

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

Can you draw the graph of the function?

If you have plotted the graph of the function in Exercise 3.1.21 (i), you must have noticed that $\max\{f, g\}(x) = |x|$ for $x \in [-2, 2]$. This is not by accident.

Define $h\colon \mathbb{R} \to \mathbb{R}$ by $h(x) = \max\{x, -x\}$. Then, $h(x) = |x|$ for $x \in \mathbb{R}$ (verify). In other words,

$$\text{For } x \in \mathbb{R},\ |x| = \max\{x, -x\}. \tag{3.1}$$

(3.1) gives way to an alternative definition of the modulus function on $\mathbb{R}$. This is very useful in proving several inequalities involving the modulus function. (See [1], for details.)

3.2 One-one, onto functions and bijections

You must have come across the notions of one-one and onto functions. When is a function said to be one-one? Roughly speaking, it is so if different elements of the domain have different images. What does this mean? If we take two elements x_1 and x_2 in the domain, and $x_1 \neq x_2$, then we must have the images of x_1 and x_2 are different. Let us write this formally.

Definition 3.2.1. A function $f\colon X \to Y$ is said to be *one-one* or *one to one* if for each pair of points $x_1, x_2 \in X$ with $x_1 \neq x_2$, we have $f(x_1) \neq f(x_2)$. One-one functions are also known as *injective* functions.

In terms of quantifiers, it can be written as follows.

$$f\colon X \to Y \text{ is one-one if } \forall x_1, x_2 \in X, x_1 \neq x_2 \big(f(x_1) \neq f(x_2)\big).$$

This actually means

$$f\colon X \to Y \text{ is one-one if } \forall (x_1, x_2) \in S\,(f(x_1) \neq f(x_2))\,. \tag{3.2}$$

where $S := \{(x_1, x_2) \in X \times X : x_1 \neq x_2\} \subseteq X \times X$.

As an immediate example, you can see that the identity function on any nonempty set is one-one.

Example 3.2.2. Consider the function $f\colon (0,1) \to \mathbb{R}$, defined by $f(x) = \frac{1}{x}$. For any $x_1, x_2 \in (0,1)$, if $x_1 \neq x_2$, then we have $\frac{1}{x_1} \neq \frac{1}{x_2}$, that is, $f(x_1) \neq f(x_2)$. Therefore, f is one-one.

Example 3.2.3. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$. We show that f is one-one. Suppose $x, y \in \mathbb{R}$ be such that $x \neq y$. We claim that $f(x) \neq f(y)$, i.e., $x^3 \neq y^3$. This is equivalent to showing that $x^3 - y^3 \neq 0$. Note that $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$.

We consider different cases. If x and y are both positive or both negative, then $x^2 + xy + y^2 > 0$. Since $x - y \neq 0$, this implies $x^3 - y^3 \neq 0$.

If one of x and y is zero, then the other is nonzero. Thus, one of x^3 and y^3 is zero and the other nonzero, and therefore they are unequal.

If x and y are of opposite signs, so are x^3 and y^3 and hence $x^3 \neq y^3$.

Thus our claim holds, and therefore f is one-one.

Suppose $f\colon X \to Y$ is one-one. Then, for any elements $x_1, x_2 \in X$ following holds:

$$\text{If } x_1 \neq x_2, \text{ then } f(x_1) \neq f(x_2). \tag{3.3}$$

Note that (3.3) is equivalent to its contrapositive, namely,

$$\text{If } f(x_1) = f(x_2), \text{ then } x_1 = x_2. \tag{3.4}$$

This means that if f is one-one, then by definition (3.3) holds for every $x, y \in X$, and therefore (3.4) holds. On the other hand, if for every $x, y \in X$ (3.4) holds, then (3.3) holds, that is, f is one-one.

We list this observation in the following proposition.

Proposition 3.2.4. *A function $f\colon X \to Y$ is one-one if and only if*

$$\forall x_1, x_2 \in X,\ f(x_1) = f(x_2) \text{ implies } x_1 = x_2. \qquad \square$$

Proposition 3.2.4 is very useful in proving a function to be one-one.

Example 3.2.5. Consider $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = 3x + 2$. For $x_1, x_2 \in \mathbb{R}$, we have

$$f(x_1) = f(x_2) \ \Rightarrow 3x_1 + 2 = 3x_2 + 2 \ \Rightarrow x_1 = x_2.$$

Therefore, f is one-one.

Example 3.2.6. Consider $f\colon [0, \pi) \to \mathbb{R}$ given by $f(x) = \cos x$. We show that f is one-one. Suppose $x, y \in [0, \pi)$ and $f(x) = f(y)$, i.e., $\cos x = \cos y$. We assume $x \geq y$. Now, we have

$$0 = \cos x - \cos y = -2 \sin \frac{x - y}{2} \sin \frac{x + y}{2}.$$

This implies $\sin \frac{x-y}{2} = 0$ or $\sin \frac{x+y}{2} = 0$. Since $x, y \in [0, \pi)$ and $x \geq y$, we have $0 \leq \frac{x-y}{2},\ \frac{x+y}{2} < \pi$. Because $\sin \theta > 0$ for $0 < \theta < \pi$, we get $\frac{x-y}{2} = 0$ or $\frac{x+y}{2} = 0$, that is $x - y = 0$ or $x + y = 0$. Because $x, y \geq 0$, $x + y = 0$ implies $x = y = 0$. Hence, $f(x) = f(y)$ implies $x = y$ and we are done.

Exercise 3.2.7. Show that the following functions are one-one, both ways, that is, using the definition and using Proposition 3.2.4.

(i) $f\colon [0, \infty) \to \mathbb{R}$, defined by $f(x) = x^2$.

(ii) $f\colon \mathbb{Z} \to \mathbb{Z}$, defined by $f(k) = 3k + 7$.

(iii) $f\colon \mathbb{N} \to \mathbb{N}$, defined by $f(n) = n + 2$, if n is odd and $f(n) = 2n$, if n is even.

Exercise 3.2.8. Using Proposition 3.2.4 prove that the functions in Examples 3.2.2 and 3.2.3 are one-one.

When is a function not one-one? From (3.2) we can see readily that

$$f\colon X \to Y \text{ is } \textbf{not} \text{ one-one if } \exists (x_1, x_2) \in S\,(f(x_1) = f(x_2))\,.$$

where $S := \{(x_1, x_2) \in X \times X : x_1 \neq x_2\} \subseteq X \times Y$. Thus, $f\colon X \to Y$ is not one-one if we can find elements $x_1, x_2 \in X$, such that $x_1 \neq x_2$ and $f(x_1) = f(x_2)$.

Example 3.2.9. Consider the function $f(x) = x^2$ defined on $\mathbb{R}$. It is easy to see that f is not a one-one function. How do you show this? Well, take $1, -1 \in \mathbb{R}$. We have $f(-1) = f(1) = 1$.

(Many times the answer we get from the student is as follows: take any x and $-x$. This is not proper for two reasons. First, if we take $x = 0$, the desired pair is not obtained. Second, it is your responsibility to give a pair, not that of the questioner! There are situations where you will need to argue for existence of a pair rather than providing one explicitly. But in situations like the present one give an explicit pair, may be something like $\{5, -5\}$ or $\{-\pi, \pi\}$ or $\{10^{-27}, -10^{-27}\}$, if you want.)

Exercise 3.2.10. Let $M(2,\mathbb{R})$ denote the set of all 2×2 matrices over $\mathbb{R}$. Consider the function $f\colon M(2,\mathbb{R}) \to \mathbb{R}$ defined by $f(A) = \det(A)$. Show that f is not one-one.

When do you say a function $f\colon X \to Y$ to be onto? It is so, if the codomain of f equals its range, that is, if every element in Y is the image of an element, that is, if every element $y \in Y$ has a preimage. What does it mean to say that y has a preimage? Well, it means that there is an element $x \in X$ such that $y = f(x)$. Let us write the definition now.

Definition 3.2.11 (Onto function)**.** A function $f\colon X \to Y$ is said to be *onto* if for each $y \in Y$, there exists $x \in X$ such that $y = f(x)$. In other words,

$$f\colon X \to Y \text{ is onto, if } \forall y \in Y\big(\exists x \in X\big(y = f(x)\big)\big). \tag{3.5}$$

An onto function is also called a *surjective* functions. Sometimes, we say f is a function from X **onto** Y to mean that f is surjective.

Exercise 3.2.12. Suppose $f\colon X \to Y$ is onto. Justify (from the above definition) that $R(f) = Y$.

Example 3.2.13. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 3x + 2$. It is easy to see that f is onto. Let $y \in \mathbb{R}$. We need to find $x \in \mathbb{R}$ such that $f(x) = 3x + 2 = y$. The last equation is satisfied when $x = (y-2)/3$. So, choose $x = (y-2)/3$. Then $x \in \mathbb{R}$ and $y = f(x)$.

When is a function $f\colon X \to Y$ not onto? It follows from (3.5) that

$$f\colon X \to Y \text{ is } \textbf{not} \text{ onto, if } \exists y \in Y\big(\forall x \in X\big(y \neq f(x)\big)\big). \tag{3.6}$$

In other words, f is not onto, if there is an element $y \in Y$ which does not have a preimage, that is, each element of X is not a preimage of y. Alternately, f is not onto, if $R(f) \subsetneq Y$. That is, the range $R(f)$ of f is a proper subset of Y.

Example 3.2.14. The function $f\colon \mathbb{Z} \to \mathbb{N}$ defined by $f(m) = m^2$ is not onto. For example, $2 \in \mathbb{N}$ does not have a preimage. How? You need to show that for any integer m, $f(m) = m^2 \neq 2$. Let us argue this way. If $|m| \geq 2$, then $f(m) = m^2 = |m|^2 \geq 4 > 2$. Also, for $m \in \{-1, 0, 1\}$, $f(m) \in \{0, 1\}$. Thus $f(m) \neq 2$ for each integer m.

Example 3.2.15. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 + x + 1$. We show that $0 \notin R(f)$ and therefore f is not onto. Note that $f(x) = x^2 + x + 1 = x^2 + 2x + 1 - x = (x+1)^2 - x$. Thus, $f(x) = 0$ implies $(x+1)^2 = x$. That is $x \geq 0$. However, if $x \geq 0$, then $f(x) \geq 1$, a contradiction.

Exercise 3.2.16. Which of the following functions are surjective?
(a) $f\colon \mathbb{R} \to \mathbb{R}$ where $f(x) = e^x$.
(b) $f\colon (0,\infty) \to \mathbb{R}$ where $f(x) = \log x$.
(c) $f\colon \mathbb{R} \to \mathbb{R}$ where $f(x) = \sin x$.

Definition 3.2.17 (Bijection)**.** A function $f\colon X \to Y$ which is both one-one and onto is called a *bijection* or a *one-one onto* function.

Thus, a bijection $f\colon X \to Y$ is a one-one function from X onto Y. A bijection is also called a *one-one correspondence*. Sometimes we say that sets X and Y are *bijective* to mean that there is a bijection from X to Y.

Example 3.2.18. The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 2$ is a bijection. (Refer to Examples 3.2.5 and 3.2.13.)

Example 3.2.19. Consider $\mathbb{E} = \{n \in \mathbb{N} : n = 2k \text{ for some } k \in \mathbb{N}\}$, the set of even natural numbers, and $\mathbb{O} = \{n \in \mathbb{N} : n = 2k - 1 \text{ for some } k \in \mathbb{N}\}$, the set of odd natural numbers. Is there a bijection between $\mathbb{N}$ and $\mathbb{O}$? Between $\mathbb{N}$ and $\mathbb{E}$? Between $\mathbb{E}$ and $\mathbb{O}$? Consider $\mathbb{N}$ and $\mathbb{O}$. You may try to map elements of $\mathbb{N}$ to elements in $\mathbb{O}$ using arrows. For example, you may choose

$$1 \to 1, 2 \to 3, 3 \to 5, 4 \to 7, 5 \to 9, \cdots$$

(see Figure 3.3). Can we produce now a function? What does the natural number n mapped to? If you think for a while, you will get the answer; $f(n) = 2n - 1$. Now it is routine to show that f is a bijection.

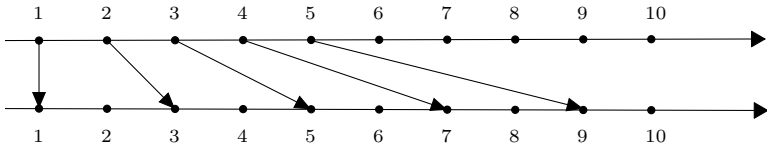


Figure 3.3: Bijection between $\mathbb{N}$ and $\mathbb{O}$

Similarly, $f(n) = 2n$ is a bijection between $\mathbb{N}$ and $\mathbb{E}$. Give a bijection between $\mathbb{E}$ and $\mathbb{O}$.

Exercise 3.2.20. Exhibit a bijection between $\mathbb{N}$ and $\{0\} \cup \mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

Exercise 3.2.21. Exhibit a bijection between $\mathbb{N}$ and

$$\{-3, -2, -1, 0, 1, 2, 3, \ldots\} = \{-3, -2, -1, 0\} \cup \mathbb{N}.$$

Can you now exhibit a bijection between $\mathbb{N}$ and $\{m \in \mathbb{Z} : m \geq m_0\}$, where m_0 is a fixed integer?

Exercise 3.2.22. Exhibit a bijection between the set of even natural numbers $\mathbb{E}$ and $\{0\} \cup \mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

Exercise 3.2.23. Let $\mathbb{O}$ be the set of all odd natural numbers and $\mathbb{E}$ be the set of all even natural numbers. Give an example of a function from $\mathbb{O}$ to $\mathbb{E}$ which is

(a) one-one, not onto,
(b) onto, not one-one,
(c) neither one-one nor onto,
(d) one-one and onto.

Exercise 3.2.24. Find a one-one map from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ and a one-one map from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$.
Hint: Consider $(m, n) \mapsto 2^m 3^n$ from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$.

Exercise 3.2.25. Give a one-one map from the set of rational numbers $\mathbb{Q}$ to $\mathbb{Z} \times \mathbb{N}$.
Hint: Recall how we described the set $\mathbb{Q}$ in Section 2.1.

Example 3.2.26. Consider a function $h\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = ax^2 + bx + c$, $a \neq 0$. Let us find out whether this map is one-one or onto.

> *Strategy:* How do we attack this problem? Since $a \neq 0$, f is a quadratic. Let us assume that $a > 0$. Hence its graph will be a parabola opening upward and it will have vertex at some point, say (α, β). Hence, the graph will lie above the horizontal line $y = \beta$. Choose $\gamma < \beta$. Then γ is not in the range of f. This means f is not onto. (See Figure 3.4.)
>
> On the other hand, for $\delta > \beta$, the horizontal line $y = \delta$ will intersect the graph at two points, say (x_1, δ) and (x_2, δ). Do you see some relation between the two points? Their positions are symmetric about the vertical line $x = \alpha$. In fact, if we take any real numbers x_1, x_2 which are equidistant from α, then we should have $f(x_1) = f(x_2)$. Hence f is not one-one.

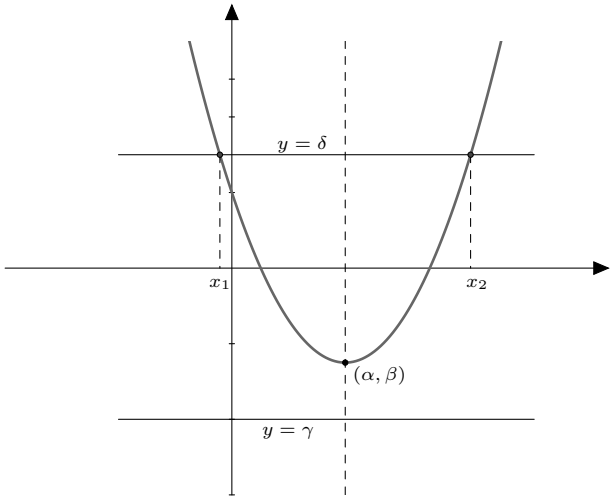


Figure 3.4: Figure for Example 3.2.26

Let us find the vertex of the parabola $y = ax^2 + bx + c$, $a > 0$. Note that f is differentiable, and $f'(x) = 0$ iff $x = -b/(2a)$. Further, $f''(x) = 2a > 0$. Thus, $x = -b/(2a)$ is a point of global minimum. Hence, the vertex of this parabola is at

$$(\alpha, \beta) = \left(-\frac{b}{2a}, f\left(-\frac{b}{2a}\right)\right) = \left(-\frac{b}{2a}, -\frac{b^2}{4a} + c\right).$$

Suppose $\gamma < \beta$. We claim that $f(x) \neq \gamma$ for any real number x. Let $\gamma = \beta - h$, where $h > 0$. Then, we have

$$\begin{aligned} f(x) - \gamma = ax^2 + bx + c - \gamma &= ax^2 + bx + \frac{b^2}{4a} + h \\ &= a\left(x^2 + 2\frac{b}{2a}\,x + \left(\frac{b}{2a}\right)^2\right) + h \\ &= a\left(x + \frac{b}{2a}\right)^2 + h > 0. \end{aligned}$$

This proves our claim, and therefore f is not onto.

Next, we claim that f is not one-one. Choose $r > 0$ and consider $x_1 = -\frac{b}{2a} - r$ and $x_2 = -\frac{b}{2a} + r$. Then, $x_1 \neq x_2$. However, $f(x_1) = f(x_2) = ar^2 - \frac{b^2}{4a} + c$ (verify). This proves our claim.

To prove f is not one-one you may also take any real number $\gamma > \beta$ and show that $f(x) = \gamma$ has two distinct real roots.

The case when a is negative is similar and is left as an exercise.

Example 3.2.27. Let $a, b, c, d \in \mathbb{R}$ be such that $ad - bc \neq 0$. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f(x, y) = (ax + by, cx + dy).$$

Here, $f(x, y)$ means the image of (x, y) under f. Let us find out whether this map is one-one or onto.

Let us first explore whether f is onto. Let (u, v) be a point in $\mathbb{R}^2$. We want to check if there exists a point $(x, y) \in \mathbb{R}^2$ such that $f(x, y) = (u, v)$. This amounts to solving the system of linear equations

$$ax + by = u,\ cx + dy = v.$$

Since $ac - bd \neq 0$, this system has a solution, namely,

$$x = \frac{du - bv}{ad - bc},\ y = \frac{av - cu}{ad - bc}.$$

With these values of x and y, we have $f(x, y) = (u, v)$. Hence, f is onto.

To check whether f is one-one, we take $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ and assume $f(x_1, y_1) = f(x_2, y_2)$, that is,

$$(ax_1 + by_1, cx_1 + dy_1) = (ax_2 + by_2, cx_2 + dy_2).$$

This gives us a system of homogeneous equations,

$$\begin{aligned} a(x_1 - x_2) + b(y_1 - y_2) &= 0, \\ c(x_1 - x_2) + d(y_1 - y_2) &= 0, \end{aligned}$$

with $x_1 - x_2$ and $y_1 - y_2$ as the variables. Since $ac - bd \neq 0$, the system has only the zero solution, that is, $x_1 - x_2 = 0, y_1 - y_2 = 0$. Thus, $(x_1, y_1) = (x_2, y_2)$. This implies that f is one-one.

Exercise 3.2.28. Find out whether the following functions are one-one and/or onto.

(a) $f : \mathbb{N} \to \mathbb{N}$, $f(n) = n$ if n is odd and $f(n) = 2n$ if n is even.
(b) $f : [0,\ 1] \to [0,\ 1]$, $f(x) = (1 - x)/(1 + x)$.
(c) $f : [0,\ 1] \to [a,\ b]$, $f(x) = bx + a(1 - x)$.
(d) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{2}\,(x + |x|)$.
(e) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x + [x]$, where $[x]$ denotes the greatest integer less than or equal to x.
(f) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2 + x + 1$ if $x \geq 0$ and $f(x) = x + 1$ for $x < 0$.
(g) $f : [0,\ 2\pi) \to D = \{(x,\ y)\ :\ x^2 + y^2 = 1\}$, $f(x) = (\cos x,\ \sin x)$.
(h) $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.
(i) $g\colon \mathbb{R} \to \mathbb{R}$ given by $g(x) = x^3$.

Example 3.2.29. Given any function $f\colon X \to Y$, there exists a natural bijection between X and $G(f) = \{(x, y) \in X \times Y : y = f(x)\}$, the graph of f.

We want to define a one-one function from X to $G(f)$. If $x \in X$, then under the function f, it corresponds to a unique $f(x) \in Y$. That is, for each $x \in X$, we have a unique $(x, f(x)) \in G(f)$. This suggests that we define $H\colon X \to G(f)$ by $H(x) = (x, f(x))$. We claim that H is a bijection.

Suppose $x_1, x_2 \in X$ such that $H(x_1) = H(x_2)$, that is, $(x_1, f(x_1)) = (x_2, f(x_2))$. This implies $x_1 = x_2$. Hence, in view of Proposition 3.2.4, H is one-one.

Next, let $(x, y) \in G(f)$. Then, by definition of $G(f)$, $x \in X$ and $y = f(x)$. Thus, $(x, y) = (x, f(x)) = H(x)$. This shows that H is onto, and therefore a bijection.

Example 3.2.30. Suppose $f\colon \mathbb{N} \to A$ and $g\colon \mathbb{N} \to B$ be maps. If they are onto, should there be an onto map $h\colon \mathbb{N} \to A \cup B$?

We can use the following idea. Take a positive integer. If it is the k-th odd number, then map it to $f(k)$, and if it is the k-th even number, then map it to $g(k)$.

So, let n be a positive integer. If n is the k-th odd number, then $n = 2k - 1$, and we have $k = (n + 1)/2$. Similarly, if n is the k-th even number, then $k = n/2$. Therefore, the map $h\colon \mathbb{N} \to A \cup B$ can be defined as

$$h(n) = \begin{cases} f\big((n+1)/2\big), & \text{if } n \text{ is odd,} \\ g\big(n/2\big), & \text{if } n \text{ is even.} \end{cases} \tag{3.7}$$

Can you show that h is onto? Since f is onto, for $a \in A$ there is $k \in \mathbb{N}$ such that $a = f(k) = h(2k - 1)$. Similarly, for $b \in B$ we have $k \in \mathbb{N}$ such that $b = g(k) = h(2k)$. Therefore, h is onto.

When is h one-one? Clearly, for that to be the case f and g must be one-one. Will that be sufficient? Suppose m and n are distinct positive integers. If they are both odd or both even (that is, if they are of same *parity*), then $h(m) \neq h(n)$. So, if $h(m) = h(n) = c$, then one of m and n is odd and the other is even (that is, they are of opposite *parity*). Suppose m is odd and n is even. Then $h(m) = f((m + 1)/2) \in A$ and $h(n) = g(n/2) \in B$, that is, $c \in A \cap B$. Thus, if h is not one-one, then $A \cap B \neq \emptyset$. Conversely, if $c \in A \cap B$, then

$c = f(k) = h(2k-1)$ for some $k \in \mathbb{N}$ and $c = g(\ell) = h(2\ell)$ for some $\ell \in \mathbb{N}$, and therefore h is not one-one.

Hence, $h\colon \mathbb{N} \to A \cup B$ is one-one if and only if f and g are one-one, and $A \cap B = \emptyset$. In particular, we have the following.

> If $f\colon \mathbb{N} \to A$ and $g\colon \mathbb{N} \to B$ are bijections and $A \cap B = \emptyset$, then the map $h\colon \mathbb{N} \to A \cup B$ given by (3.7) is a bijection.

Exercise 3.2.31. Exhibit a bijection between $\mathbb{N}$ and $\mathbb{Z}$.

Understanding a function graphically

Once we are given a reasonably simple function $f\colon \mathbb{R} \to \mathbb{R}$, we can plot its graph either by hand or using some computer software. Recall that a subset $S \subseteq \mathbb{R}^2$ is graph of some function f if and only if every vertical line in $\mathbb{R}^2$ intersects S at most once. This is called the *vertical line test* of function.

Suppose, we are given the graph of a function f. Can we find what type of function f is?

First, let us assume that f is onto. In this case, for every $\beta \in \mathbb{R}$, there exists $\alpha \in \mathbb{R}$ such that $\beta = f(\alpha)$. This is same as saying that for $\beta \in \mathbb{R}$, there exist $\alpha \in R$ such the (α, β) lies on the graph. That is, the horizontal line $y = \beta$ intersects the graph at (α, β). Thus we have the following graphical characterization of onto functions.

> $f\colon \mathbb{R} \to \mathbb{R}$ is onto if and only if every horizontal line intersects the graph of f at least once.

Next, suppose that a horizontal line $y = \beta$ intersects the graph at two distinct points, say (x_1, β) and (x_2, β). Then we have $x_1 \neq x_2$ and $f(x_1) = f(x_2) = \beta$. Thus, in this case, f is not one-one. We conclude that if f is one-one, then no horizontal line should intersect the graph more than once. However, a horizontal line may not intersect the graph (why?). This leads to the following graphical characterization of one-one functions.

> $f\colon \mathbb{R} \to \mathbb{R}$ is one-one if and only if every horizontal line intersects the graph of f at most once.

(Note that the above result is true if the domain of f is any subset of $\mathbb{R}$.)

Combining the above observations, we have the following:

> $f\colon \mathbb{R} \to \mathbb{R}$ is a bijection if and only if every horizontal line intersects the graph of f exactly once.

Definition 3.2.32 (Increasing function)**.** Let $I \subset \mathbb{R}$ be any nonempty subset. A function $f\colon I \to \mathbb{R}$ is said to be *increasing* if for every $x, y \in I$, $x < y$ implies $f(x) \leq f(y)$.

If for every $x, y \in I$, $x < y$ implies $f(x) < f(y)$, then we say that f is *strictly increasing.*

Note that in some textbooks an increasing function is called a *nondecreasing* function, and a strictly increasing function is called an *increasing* function.

Definition 3.2.33 (Decreasing function)**.** A function $f\colon I \to \mathbb{R}$ is said to be *decreasing* if for every $x, y \in I$, $x < y$ implies $f(x) \geq f(y)$.

If for every $x, y \in I$, $x < y$ implies $f(x) > f(y)$, then we say that f is *strictly decreasing.*

Definition 3.2.34. A function which is either increasing or decreasing is called a *monotone function.*

Exercise 3.2.35. Show that $f(x) = x^2$ is decreasing in $(-\infty, 0]$ and increasing in $[0, \infty)$.

Exercise 3.2.36. Let $f\colon I \to \mathbb{R}$ be strictly increasing. Then show that f is injective. What about strictly decreasing functions?

Exercise 3.2.37. Let $a, b \in \mathbb{R}$ and $f(x) = ax + b$ for $x \in \mathbb{R}$. Show that f is strictly increasing if $a > 0$ and is strictly decreasing if $a < 0$. When is f injective?

Example 3.2.38. In calculus, you must have seen that e^x, $\log x$, and x^n (for odd n) are increasing functions on their usual domains. Therefore, they are one-one.

Exercise 3.2.39. Let $f\colon I \to \mathbb{R}$ be both decreasing and increasing. Prove that f is a constant function.

3.3 Composition of functions

Definition 3.3.1. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be two functions. Then the *composition* of f and g is the function $g \circ f\colon X \to Z$ defined by $(g \circ f)(x) = g(f(x))$. (See Figure 3.5.)

Note that $(g \circ f)(x)$ is defined for every $x \in X$ because the image set of f is contained in the domain of g. Suppose functions $f\colon X \to Y$ and $g\colon W \to Z$ are given. If $Y \neq W$ can you define $g \circ f$? Yes, if $R(f) \subseteq W$. In that case, $g(f(x))$ is defined for every $x \in X$, and therefore $g \circ f\colon X \to Z$ is defined.

Exercise 3.3.2. Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = x^2$ and $g(x) = 3x + 2$. Find $g \circ f$ and $f \circ g$. Are they same?

Example 3.3.3. Consider the functions $f\colon \mathbb{R} \to [0, \infty)$, $f(x) = x^2$, and $g\colon [0, \infty) \to \mathbb{R}$, $g(x) = \sqrt{x}$, the unique nonnegative square root of x. We get two compositions $g \circ f\colon \mathbb{R} \to \mathbb{R}$ and $f \circ g\colon [0, \infty) \to [0, \infty)$. Since their domains of $g \circ f$ and $f \circ g$ are different, $g \circ f \neq f \circ g$. However,

$$(g \circ f)(x) = g(x^2) = \sqrt{x^2} = |x|, \text{ for } x \in \mathbb{R}, \text{ and}$$
$$(f \circ g)(x) = (\sqrt{x})^2 = |x| = x \text{ for } x \in [0, \infty),$$

and therefore $(g \circ f)(x) = (f \circ g)(x)$ for $x \in [0, \infty)$.

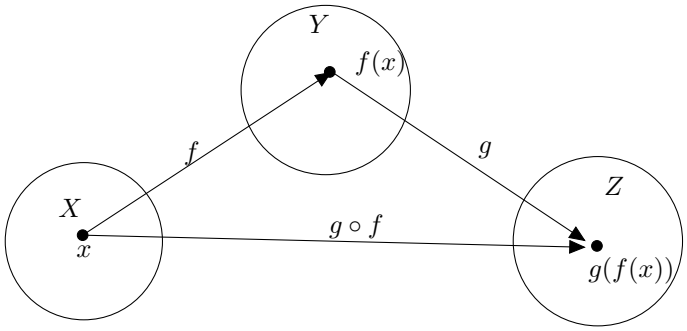


Figure 3.5: Composition of functions

Exercise 3.3.4. Let $f, g\colon \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^3 + 2$ and $g(x) = \sqrt[3]{x}$. Find $f \circ g$ and $g \circ f$. Is $f \circ g = g \circ f$?

Exercise 3.3.5. Let $f\colon M(2,\mathbb{R}) \to M(2,\mathbb{R})$ be defined by $f(A) = A^T A$, and $g\colon M(2,\mathbb{R}) \to \mathbb{R}$ be defined by $g(A) = \text{trace(A)}$. Find $(g \circ f)(A)$ for $A \in M(2,\mathbb{R})$ in terms of entries a_{ij} of A.

Example 3.3.6. Let $f\colon \mathbb{R} \to \mathbb{Z}$ and $g\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be defined by $f(x) = [x]$, the greatest integer less than or equal to x, and $g(x) = 1/x$. Note that in this case, $g \circ f$ is not defined (why?). However, $f \circ g\colon \mathbb{R} \setminus \{0\} \to \mathbb{Z}$ is defined. You can verify that the composite is given by

$$(f \circ g)(x) = \left[\frac{1}{x}\right] = \begin{cases} 0, & \text{if } x \in (1, \infty), \\ n, & \text{if } x \in \left(\frac{1}{n+1}, \frac{1}{n}\right], n \in \mathbb{N}, \\ -(n+1), & \text{if } x \in \left(\frac{-1}{n}, \frac{-1}{n+1}\right], n \in \mathbb{N}, \\ -1, & \text{if } x \in (-\infty, -1). \end{cases}$$

How do we know that we have defined $(f \circ g)(x)$ for all $x \in \mathbb{R} \setminus \{0\}$? Recall $\cup_n [1/n, 1) = (0, 1)$. This needs Archimedean property of $\mathbb{R}$. See Ex. 1.3.8 in [1]

Consider two functions $f\colon X \to Y$ and $g\colon Y \to Z$. Suppose f and g are one-one. What about $g \circ f$? Let $x_1, x_2 \in X$ and $x_1 \neq x_2$. Since f is one-one, $f(x_1) \neq f(x_2)$. Again, since g is one-one, $g(f(x_1)) \neq g(f(x_2))$, that is, $g \circ f(x_1) \neq g \circ f(x_2)$. Thus, $g \circ f$ is one-one. Similarly, if f and g are onto, then $g \circ f$ is onto. The proof is similar and is left as an exercise. We list the results in the following theorem.

Theorem 3.3.7. *Let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions.*

(a) If f and g are injective (one-one), then $g \circ f$ is injective.

(b) If f and g are surjective (onto), then $g \circ f$ is surjective.

(c) If f and g are bijections, then $g \circ f$ is a bijection. □

Suppose you are given that the composite $g \circ f$ of two functions $f\colon X \to Y$ and $g\colon Y \to Z$ is one-one. What can you say about f and g? What is given to you is that

$$x_1, x_2 \in X, x_1 \neq x_2 \implies g(f(x_1)) \neq g(f(x_2)). \tag{3.8}$$

Suppose $x_1, x_2 \in X$ and $x_1 \neq x_2$. Can you have $f(x_1) = f(x_2)$? No, otherwise we will have $g(f(x_1)) = g(f(x_2))$ contradicting (3.8). This means that f must be one-one. What about g? From (3.8) we see that if $y_1, y_2 \in R(f)$ and $y_1 \neq y_2$, then $g(y_1) \neq g(y_2)$. That is, g is one-one on the range $R(f)$. But g may not be one-one on its domain Y. For example, let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = e^x$ and $g\colon \mathbb{R} \to \mathbb{R}$ be defined by $g(x) = x^2$. Then $(g \circ f)(x) = e^{2x}$. Note that $g \circ f$ is one-one, but although g is one-one on the range of f, it is not so on its entire domain.

Suppose $g \circ f$ is given to be surjective. You can see that g is necessarily surjective, though f need not be. We list these observations as a theorem and urge you to write a formal proof.

Theorem 3.3.8. *Let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions.*

(a) If $g \circ f$ is injective, then f is injective.

(b) If $g \circ f$ is surjective, then g is surjective. □

Exercise 3.3.9. Give examples of functions $f, g\colon \mathbb{Z} \to \mathbb{Z}$ such that

(a) g is not injective but $g \circ f$ is.

(b) f is not surjective but $g \circ f$ is.

Exercise 3.3.10. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be such that $g \circ f$ is injective and f is surjective. Is g injective?

Exercise 3.3.11. Let $f\colon X \to X$ be function such that $f \circ f$ is bijective. Should f be bijective?

Exercise 3.3.12. Suppose $f\colon X \to Y$ and $g\colon Y \to X$ be such that $g \circ f = Id_X$. Show that f is one-one and g is onto.

Example 3.3.13. Consider the function $h\colon \mathbb{N} \to A \cup B$ in Example 3.2.30 obtained from maps $f\colon \mathbb{N} \to A$ and $g\colon \mathbb{N} \to B$. The map h can be described by compositions of certain maps. Let $\mathbb{O}$ and $\mathbb{E}$ denote the sets of odd and even positive integers, respectively. If $\phi\colon \mathbb{O} \to \mathbb{N}$ is defined by $\phi(n) = (n+1)/2$, and $\psi\colon \mathbb{E} \to \mathbb{N}$ is defined by $\psi(n) = n/2$, then

$$h(n) = \begin{cases} f\big((n+1)/2\big) = (f \circ \phi)(n), & \text{if } n \in \mathbb{O}, \\ g\big(n/2\big) = (g \circ \psi)(n), & \text{if } n \in \mathbb{E}. \end{cases}$$

Exercise 3.3.14. Suppose $f\colon X \to Y$ and $g\colon Y \to X$ be such that $g \circ f = Id_X$ and $f \circ g = Id_Y$. Show that f and g are bijections?

Proposition 3.3.15. *For functions $f\colon X \to Y$, $g\colon Y \to Z$ and $h\colon Z \to W$*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

What is the meaning of the assertion here? We have two functions $h \circ (g \circ f)$ and $(h \circ g) \circ f$. When are they equal? They are so, if they have same domains and codomains, and the images of any element in the domain are equal under the two functions. Let us give a systematic proof of the assertion.

Proof. Since $g \circ f \colon X \to Z$ and $h \colon Z \to W$ we have $h \circ (g \circ f) \colon X \to W$. Similarly, $(h \circ g) \circ f \colon X \to W$. Therefore, the two functions have same domains and codomains.

Let $x \in X$. We need to show that $\big(h \circ (g \circ f)\big)(x) = \big((h \circ g) \circ f\big)(x)$. Assume $y = f(x) \in Y$, $z = g(y) \in Z$ and $w = h(z) \in W$. Then $(g \circ f)(x) = g(f(x)) = g(y) = z$, and therefore $\big(h \circ (g \circ f)\big)(x) = h\big((g \circ f)(x)\big) = h(z) = w$. Again, $(h \circ g)(y) = h\big(g(y)\big) = h(z) = w$, and therefore $\big((h \circ g) \circ f\big)(x) = (h \circ g)(f(x)) = (h \circ g)(y) = w$. We have $\big(h \circ (g \circ f)\big)(x) = w = \big((h \circ g) \circ f\big)(x)$. This completes the proof. □

3.4 Inverse of a function

Consider the function $f \colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 2$. Let $x \in \mathbb{R}$ and $f(x) = y \in \mathbb{R}$. Then $y = 3x + 2$ and therefore $x = (y - 2)/3$. Note that for $y \in \mathbb{R}$, $(y - 2)/3$ is a unique real number in $\mathbb{R}$. Therefore, we get a function $g \colon \mathbb{R} \to \mathbb{R}$ given by $g(y) = (y - 2)/3$. You can easily check that if $f(x) = y$, then $g(y) = x$, and if $g(y) = x$, then $f(x) = y$. Thus, the function g reverses the process of f. We call g an inverse of f.

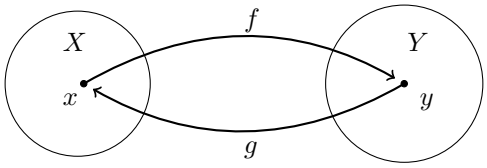


Figure 3.6: Inverse Function

For a function $f \colon X \to Y$, an inverse of f is a function $g \colon Y \to X$ that reverses the process of f. That is, if f maps x to y, then g should map y back to x, and if g maps y to x, then f should map x back to y (see Figure 3.6).

Suppose $f \colon X \to Y$ has an inverse $g \colon Y \to X$. What kind of map f should be? If f maps two distinct elements x_1 and x_2 to the same element $y \in Y$, where will g map y to? Thus to make g a function we need to assume that f is one-one.

Also the domain of g needs to be all of Y, and if $g(y) = x$ we should have $f(x) = y$. Thus, for each $y \in Y$ there must exist an $x \in X$ such that $f(x) = y$. This forces us to require f to be an onto function. We conclude that we can "reverse" the process of f only if f is a bijection.

For example, the function we considered above, that is, $f \colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3x + 2$, is a bijection (see Example 3.2.5) and we could find an inverse of f.

Suppose $f\colon X \to Y$ is a bijection. Does it have an inverse? Suppose $y \in Y$. Since f is onto, there exists $x \in X$ such that $f(x) = y$. Note that this x is unique, since f is one-one. Thus, we can define a map $g\colon Y \to X$ by $g(y) = x$ if $f(x) = y$. Then, by definition we also have $f(x) = y$ if $g(y) = x$. Therefore, g reverses the process of f, that is, g is an inverse of f.

Let us now define an inverse of a function formally.

Definition 3.4.1. Let $f\colon X \to Y$ be a bijection. An *inverse* of f is a map $g\colon Y \to X$ such that if $f(x) = y$, then $g(y) = x$.

Example 3.4.2. For $x \in [0, \infty)$ you can easily see that $x/(x+1) \in [0, 1)$. Therefore we have a map $f\colon [0, \infty) \to [0, 1)$ defined by $f(x) = x/(x+1)$. Verify that f is one-one. Suppose $y \in [0, 1)$. If $y = x/(x+1)$ then we should have $x = y/(1-y)$. Note that for $0 \leq y < 1$, $y/(1-y)$ is a nonnegative real number. This shows two things: (i) the map f is onto, and (ii) the map $g\colon [0, 1) \to [0, \infty)$ defined by $g(y) = y/(1-y)$ is an inverse of f.

Suppose $f\colon X \to Y$ is a bijection. Can f have more than one inverse? No. If g_1 and g_2 are inverses of f, then they both have domain Y and codomain X. Further, for $y \in Y$ we have $x \in X$ such that $f(x) = y$. Then by definition of an inverse we have $g_1(y) = x = g_2(y)$. Therefore we have $g_1 = g_2$. We conclude that if f is a bijection, then f has a unique inverse.

The inverse of a bijection $f\colon X \to Y$ is denoted by f^{-1} which is a function from Y to X.

Using the notation we write the inverse of the function $f\colon [0, \infty) \to [0, 1)$, which is defined by $f(x) = x/(x+1)$, as the function $f^{-1}\colon [0, 1) \to [0, \infty)$ defined by $f^{-1}(y) = y/(1-y)$.

Exercise 3.4.3. Let X be a nonempty set and Id_X be the identity function on X. What is its inverse?

Example 3.4.4. The function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f(x, y) = (ax + by, cx + dy).$$

is a bijection if $ad - bc \neq 0$ (see Example 3.2.27). The inverse $f^{-1}\colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$f^{-1}(u, v) = \left(\frac{du - bv}{ad - bc}, \frac{av - cu}{ad - bc}\right), \ (u, v) \in \mathbb{R}^2.$$

Exercise 3.4.5. Show that the following functions are bijections and find their inverses.

(a) $f\colon \mathbb{R} \to \mathbb{R}, \ f(x) = 5x - 7.$
(b) $f\colon [0, 1] \to [0, 1], \ f(x) = (1-x)/(1+x).$
(c) $f\colon [0, 1] \to [a, b], \ (a < b) \ f(x) = a(1-x) + bx.$
(d) $f\colon [0, \infty) \to [0, 1), \ f(x) = x^2/(1+x^2).$

Exercise 3.4.6. Define $f\colon \mathbb{N} \to \mathbb{N}$ by $f(m) = m-1$, if m is even, and $f(m) = m+1$, if m is odd. Show that f is a bijection and find its inverse.

Remark 3.4.7. Suppose $f\colon X \to Y$ is a bijection. Then by definition f^{-1} has the property "if $f(x) = y$, then $f^{-1}(y) = x$". Can you see that "if $f^{-1}(y) = x$, then $f(x) = y$" holds as well? Suppose $y \in Y$ and $f^{-1}(y) = x \in X$. Since f is onto, there exists $x' \in X$ such that $f(x') = y$. However, this implies $f^{-1}(y) = x'$. Thus, $x' = x$, that is, $f(x) = y$.

Suppose $f\colon X \to Y$ is a bijection and $f^{-1}\colon Y \to X$ is its inverse. Is f^{-1} a bijection? For $x \in X$, let $y = f(x)$. Then by definition $x = f^{-1}(y)$. This shows that f^{-1} is onto. Further, for $y_1, y_2 \in Y$ if $f^{-1}(y_1) = f^{-1}(y_2) = x \in X$, then in view of Remark 3.4.7 we have $y_1 = y_2 = f(x)$. This implies that f^{-1} is one-one. Therefore, f^{-1} is a bijection.

What is the inverse of f^{-1}? Suppose $y \in Y$. If $f^{-1}(y) = x \in x$, then by Remark 3.4.7 we have $f(x) = y$. Therefore, we have the following.

The inverse of $f^{-1}\colon Y \to X$ is $f\colon X \to Y$, that is, $\left(f^{-1}\right)^{-1} = f$.

Example 3.4.8. For the function $f\colon [0,\infty) \to [0,1)$ defined by $f(x) = x/(x+1)$, the inverse is $f^{-1}\colon [0,1) \to [0,\infty)$ which is defined by $f^{-1}(y) = y/(1-y)$. Therefore the inverse of $g\colon [0,1) \to [0,\infty)$ defined by $g(y) = y/(1-y)$ is the function $f\colon [0,\infty) \to [0,1)$ defined by $f(x) = x/(x+1)$.

Suppose $f\colon X \to Y$ is a bijection and $g\colon Y \to X$ is its inverse. Note that the maps $g \circ f\colon X \to X$ and $f \circ g\colon Y \to Y$ are defined. What are these maps? For $x \in X$ we have $(g \circ f)(x) = g(f(x)) = x$, and therefore $g \circ f = Id_X$, the identity function on X. Similarly, for $y \in Y$ we have a unique $x \in X$ such that $y = f(x)$. Hence $g(y) = x$. Using this, we see that $(f \circ g)(y) = f(g(y)) = y$, and therefore $f \circ g = Id_Y$. Thus we have the following.

Proposition 3.4.9. *If $f\colon X \to Y$ is a bijection, then $f^{-1} \circ f = Id_X$ and $f \circ f^{-1} = Id_Y$.* □

Let us look at the converse. Suppose $f\colon X \to Y$ and $g\colon Y \to X$ are such that $g \circ f = Id_X$ and $f \circ g = Id_Y$. What can you say about f and g? Note that identity maps are bijections. Should f be a bijection? Yes. Since $g \circ f$ is one-one, f must be one-one, and since $f \circ g$ is onto, f must be onto (see Theorem 3.3.8, or quickly go through the argument in your mind). Thus, f is a bijection. Suppose $x \in X$ and $f(x) = y \in Y$. Then $g(y) = g(f(x)) = (g \circ f)(x) = Id_X(x) = x$. This shows that $g = f^{-1}$. Thus we have proved the following result.

Proposition 3.4.10. *Suppose $f\colon X \to Y$ and $g\colon Y \to X$ are such that $g \circ f = Id_X$ and $f \circ g = Id_Y$. Then f is a bijection and $g = f^{-1}$.* □

Example 3.4.11. Consider the function $f\colon \mathbb{N} \to \mathbb{N}$ defined in Exercise 3.4.6. It is easy to see that $f \circ f = Id_{\mathbb{N}}$. Indeed, if $m \in \mathbb{N}$ is even, then $(f \circ f)(m) = f(f(m)) = f(m-1) = (m-1)+1 = m$, and if $m \in \mathbb{N}$ is odd, then $(f \circ f)(m) = f(f(m)) = f(m+1) = (m+1)-1 = m$. In view of Proposition 3.4.10, f is a bijection and $f^{-1} = f$.

Remark 3.4.12. It is important that both the conditions $g \circ f = Id_X$ and $f \circ g = Id_Y$ (in Proposition 3.4.10) are satisfied to conclude that f is a bijection. For instance, let $X = Y$ be the set of all real sequences. Define $f((x_1, x_2, \ldots, x_n, \ldots)) = (0, x_1, x_2, x_3, \ldots, x_n, \ldots)$ and $g((x_1, x_2, \ldots, x_n, \ldots)) = (x_2, x_3, \ldots, x_n, \ldots)$. Then $g \circ f = Id_X$ while $f \circ g \neq Id_X$. You can verify that f is one-one but not onto.

Theorem 3.4.13. *If $f\colon X \to Y$ and $g\colon Y \to Z$ are bijections, then $g \circ f$ is a bijection and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Proof. That $g \circ f$ is a bijection was seen in Theorem 3.3.7. Put $\phi = g \circ f\colon X \to Z$ and $\psi = f^{-1} \circ g^{-1}\colon Z \to X$. We need to show that $\psi = \phi^{-1}$. In view of Proposition 3.4.10, this is the case if $\psi \circ \phi = Id_X$ and $\phi \circ \psi = Id_Z$. We have

$$\psi \circ \phi = (f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ f = Id_X.$$

Did you notice that we used the result (Theorem 3.3.15) on the associativity of composition of functions? Similarly, $\phi \circ \psi = Id_Z$. (Verify.) Thus, $\phi^{-1} = \psi$. □

Exercise 3.4.14. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be such that $g \circ f$ is surjective and g is injective. Is f surjective?
Hint. Observe that g is a bijection and $f = g^{-1} \circ (g \circ f)$.

Graph of inverse of a function

Suppose $f\colon \mathbb{R} \to \mathbb{R}$ is a bijection. Suppose we know the graph of f. How does the graph of f^{-1} look like? Suppose that $(a, b) \in G(f)$. Then $b = f(a)$ and $f^{-1}(b) = a$. This means that the point $(b, a) \in G(f^{-1})$ (see Figure 3.7).

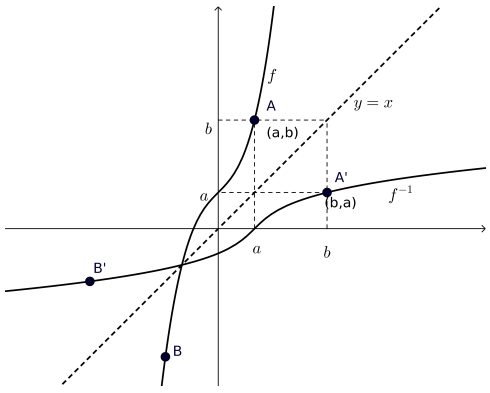


Figure 3.7: Graphs of f and f^{-1}

How are the points (a, b) and (b, a) related geometrically in $\mathbb{R}^2$? If you flip the coordinates of all the points in $\mathbb{R}^2$, it amounts to the flipping of $\mathbb{R}^2$ about the line $y = x$. This suggests that the graph of f^{-1} can be obtained by taking the mirror image of the graph of f about the line $y = x$.

In calculus you must have come across the graphs of the functions e^x and $\log x$. Note that the two functions are inverse of each other. Compare their graphs in view of the above discussion.

Exercise 3.4.15. Draw the graph of following functions. Check whether the functions are bijections and, if so, draw the graphs of the inverses.

(i) $f\colon [0,\infty) \to (0,\infty)$ defined by $f(x) = x^2$.
(ii) $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$.
(iii) $f\colon \mathbb{R} \to (0,\infty)$, $f(x) = e^x$,
(iv) $f\colon [-\pi/2, \pi/2] \to [-1,1]$, $f(x) = \sin x$.

Bijections between intervals

Our aim is to explore if there exists a bijection between two intervals $[a,b]$ and $[c,d]$, where $a < b$ and $c < d$. Before we embark on the general case, let us find a bijection between $[0,1]$ and $[0,4]$ and a bijection between $[0,1]$ and $[6,7]$. One way to get this is as follows: Imagine a rubber string of length 1, tied at one end say at 0. This string can be thought of as the interval $[0,1]$. We want to map this to $[0,4]$. How do we do this? We can simply stretch this rubber string increasing its length to four times to reach up to 4. This gives a one-one onto function from $[0,1]$ to $[0,4]$ which can be written as $f(x) = 4x$. It is easy to check that $f\colon [0,1] \to [0,4]$ is a bijection. In general $f(x) = ax$ is a bijection from $[0,1]$ to $[0,a]$ (if $a > 0$) or $[a,0]$ (if $a < 0$).

Next, how do we map the rubber string tied between 0 and 1 to $[6,7]$. We can simply translate the string by 6, that is, $f(x) = x + 6$ is the map. Now it is easy to check that $f(x) = x + 6$ is a bijection between $[0,1]$ and $[6,7]$.

Now, if we combine the above processes, that is, first stretch the rubber string by 4 and then translate it by 6. What do we get? The string will be mapped to the interval $[6,10]$. Thus

$$f(x) = 4x + 6 \text{ is a bijection from } [0,1] \to [6,10].$$

In general, for $a < b$ we can map the string $[0,1]$ to $[a,b]$ by first stretching by a factor $b - a$ and then translate it by a. That is,

$$f(x) = (b-a)x + a \text{ is a bijection from } [0,1] \to [a,b].$$

Can you write a bijection from $[a,b] \to [0,1]$? We can reverse the process, that is, first bring one end at a to 0 by translating by $-a$ and then compress it by a factor $1/(b-a)$. Thus,

$$g(y) = (y-a)/(b-a) \text{ is a bijection from } [a,b] \to [0,1].$$

Next, how do we set up a bijection between $[a,b]$ to $[c,d]$? First we can bring the rubber string tied between a and b to $[0,1]$ and the bring $[0,1]$ to $[c,d]$ using the above process. Thus, the required map is the composite $f \circ g$ of two bijections: $g\colon [a,b] \to [0,1]$ given by $g(y) = (y-a)/(b-a)$ and $f\colon [0,1] \to [c,d]$ given by $f(x) = (d-c)x + c$. Verify that

$$(f \circ g)(x) = \frac{x-a}{b-a} \times (d-c) + c. \tag{3.9}$$

This is a bijection from $[a,b] \to [c,d]$.

Geometric way of setting bijections between $[a, b]$ and $[c, d]$

Suppose we want $f\colon [a, b] \to [c, d]$ to be a bijection with $f(a) = c$ and $f(b) = d$. In that case, the graph of f meets the points (a, c) to (b, d). Any strictly increasing continuous function whose graph contains these two points will work. The simplest curve that will do this job is the straight line joining (a, c) to (b, d) (see Figure 3.8).

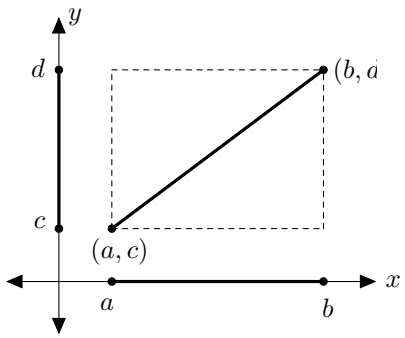


Figure 3.8: Bijection between $[a, b]$ and $[c, d]$

Now you can write down the equation of the line joining points (a, c) and (b, d) and check that the function is given by

$$f(x) = y = \frac{x - a}{b - a} \times (d - c) + c.$$

Exercise 3.4.16. Does there exist a bijection between (a, b) and (c, d)?

Exercise 3.4.17. Construct a function $f\colon (0, 1) \to \mathbb{R}$ which is one-one and onto.
Hint: Recall that $\tan x$ is a bijection from $(-\pi/2, \pi/2)$ to $\mathbb{R}$.

Exercise 3.4.18. Let $f\colon \mathbb{R} \to (-1, 1)$ be given by $f(x) = \frac{x}{\sqrt{1+x^2}}$. Show that f is a bijection. Also find the inverse of f.

Exercise 3.4.19. Is $f\colon (0, 1) \to (0, 1)$ given by $f(x) = x^2$ a bijection? How about $g(x) = x^n$ for a fixed $n \in \mathbb{N}$? *Hint:* Analysis is needed!

Example 3.4.20. Does there exist a bijection f between $[0, 1]$ and $(0, 1)$? You cannot have a continuous map between these two intervals which is a bijection (why?). However, there are many bijections between these two intervals, and we produce one here. The idea is to place the extra two points 0 and 1 inside $(0, 1)$ and keep adjusting the occupied places in sequential manner.

Consider the sequence of points $0, 1, \frac{1}{2}, \frac{1}{3}, \ldots$ in the interval $[0, 1]$ and the sequence of points $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots$ in the interval $(0, 1)$. Note that if you delete these points from $[0, 1]$ and $(0, 1)$, respectively, you will be left with the disjoint intervals $\left(\frac{1}{n+1}, \frac{1}{n}\right), n \in \mathbb{N}$. The idea is to map the points in the first sequence to the corresponding points in the second and the remaining points to themselves.

Thus, define $f\colon [0, 1] \to (0, 1)$ as follows: $f(0) = \frac{1}{2}, f(1) = \frac{1}{3}, f(\frac{1}{n}) = \frac{1}{n+2}$ for $n = 2, 3, 4, \ldots$, and $f(x) = x$ if $x \neq 0, 1/n$. Check that f is a bijection.

Exercise 3.4.21. Use the function in Example 3.4.20 to find a bijection between $[a, b]$ and (c, d).

Exercise 3.4.22. Show that $f(x) := x^{2k} + x^{2k-1} + \cdots + x + 1$ defined on $\mathbb{R}$ is not surjective by showing that 0 is not in the image of f. More generally, show that the only $(x, y) \in \mathbb{R}^2$ such that

$$x^{2k} + x^{2k-1}y + \cdots + xy^{2k-1} + y^{2k} = 0$$

is (0,0). *Hint:* What is $(x^n - 1)/(x - 1)$ and hence $(a^n - b^n)/(a - b)$?

Exercise 3.4.23. Give an example of a cubic polynomial function on $\mathbb{R}$ which is not injective.

Exercise 3.4.24. Show that any cubic polynomial function on $\mathbb{R}$ is surjective. *Hint:* Requires analysis!

Exercise 3.4.25. Discuss one-one, onto, bijective nature of the following functions. Also, indicate what modifications either on the domain, or on the range or on both is needed to make the function one-one, or onto or bijective. The problem is open ended so that the students can investigate as thoroughly as possible.

(a) $f\colon \mathbb{Q} \to \mathbb{Q}$ given by $f(x) = x^2$.

(b) $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$.

(c) $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$.

(d) $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$.

(e) $f\colon \mathbb{C} \to \mathbb{C}$ given by $f(z) = P(z)$ where $P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ is a non-constant polynomial, that is, $a_n \neq 0$ and $n \geq 1$.

(f) $f(x) = ax^2 + bx + c$ where $a, b, c \in \mathbb{R}$, $a \neq 0$ and $x \in \mathbb{R}$.

(g) Let X be the set of all people on the earth. Let Y be the subset of all men. Let $f\colon X \to Y$ be defined by $f(x)$ to be the father of x for $x \in X$.

Exercise 3.4.26. Find a "natural" bijection between the two sets X and Y given below.

(a) X is the set of all lines in $\mathbb{R}^2$ parallel to the x-axis and $Y = \mathbb{R}$.

(b) X is the set of all maps from $\{1, 2\}$ to $\mathbb{R}$ and Y is $\mathbb{R}^2$.

(c) X is the set of all natural numbers that leave 2 as a remainder when divided by 3 and $Y = \mathbb{N}$.

(d) X is the set of all circles $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = r^2, r \in \mathbb{R}\}$ and Y is the set of non negative real numbers.

Exercise 3.4.27. Generalize the Exercise 3.4.26 (b).

Exercise 3.4.28. Let $f, g, h\colon \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) := \frac{x}{1 + x^2}, \quad g(x) := \frac{x^2}{1 + x^2}, \quad h(x) := \frac{x^3}{1 + x^2}.$$

(a) Determine which of them are injective.

(b) Show that f and g are not surjective.
Hint: Recall the notion of $\lim_{x\to\pm\infty}$ of a rational function. What is the sign of $g(x)$ for $x \in \mathbb{R}$?

Exercise 3.4.29. Let $a, b \in \mathbb{R}$. Consider $f(x) := ax+b$ for $x \in \mathbb{R}$. Under what conditions on a, b
(i) $f\colon \mathbb{R} \to \mathbb{R}$ is an injective map?
(ii) $f\colon \mathbb{Z} \to \mathbb{Z}$ is an injective map?
(iii) $f\colon \mathbb{N} \to \mathbb{N}$ is an injective map?

Exercise 3.4.30. Let X be any nonempty set and let $P(X)$ be its power set, that is, the set of all subsets of X. Give an one-one map $f\colon X \to P(X)$. Can you give another?

Can we find an onto function from X to $P(X)$?

If X is a finite set, say, with n elements, we know that $P(X)$ will have 2^n elements. Since $2^n > n$ for any $n \in \mathbb{Z}_+$, we 'believe' that we cannot find a function $f\colon X \to P(X)$ which is onto. For example, when $X = \{1\}$, then $P(X) = \{\emptyset, X\}$. Assume that there exists $f\colon X \to P(X)$ which is onto. Now $f(1)$ must be one of $\emptyset$ or X. Assume without loss of generality that $f(1) = \emptyset$. Since f is onto, there exists an $x \in X$ such that $f(x) = X$. Since the only element in X is 1, it means that $f(1) = X$. But then f is not a function at all!

The next theorem, known as Cantor's theorem, say that our intuition for finite sets X holds true for any set.

Theorem 3.4.31 (Cantor's Theorem)**.** *Let X be any set and $P(X)$ the power set of X. Then, there does not exist an onto function from X to $P(X)$.*

> *Motivation:* We shall prove this theorem using contradiction. The proof is motivated by the famous Russell's paradox about a barber. "There is a barber in a village, who shaves all those and only those men who do not shave themselves. Now, the question is who shaves the barber." Either the barber shaves himself or he does not. You will soon realize that either of the assumptions gives a contradiction.

Proof. Consider any function $f\colon X \to P(X)$. We show that f is not onto. Given any $x \in X$, $f(x) \in P(X)$, that is, $f(x)$ is a subset of X. Hence, one of $x \in f(x)$ and $x \notin f(x)$ holds. Consider the subset

$$B = \{x \in X : x \notin f(x)\}$$

of X. Then, B is an element of $P(X)$. We claim that B does not have a preimage under f. Suppose, if possible, there exists $b \in X$ such that $f(b) = B$. We have two possibilities: $b \in B$ or $b \notin B$. Note that exactly one of them must be true.

Case (i): $b \in B$. Then, by definition of B, $b \notin f(b) = B$. Thus we get $b \notin B$, a contradiction.

Case (ii): $b \notin B$. Since $B = f(b)$ we get $b \notin f(b)$. However, by definition of B, this gives $b \in B$, again a contradiction.

Thus we see that both cases lead to contradictions. We are therefore forced to conclude that our assumption on the existence of an onto function from X to $P(X)$ is false. □

Cantor's theorem implies that for any set X there cannot be a bijection between X and its power set $P(X)$.

3.5 Image of subsets under functions

Let $f\colon X \to Y$ and $A \subseteq X$. Then we define the *image* $f(A)$ of A under f to be the subset

$$f(A) := \big\{y \in Y : \exists a \in A\big(f(a) = y\big)\big\}$$

of Y (see Figure 3.9). In other words, $f(A)$ is the set of all images of elements in A.

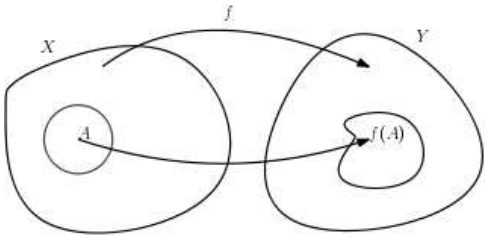


Figure 3.9: Image of a set

One can also write $f(A)$ as

$$f(A) = \{f(a) : a \in A\}.$$

It is evident that the range $R(f)$ of $f\colon X \to Y$ is the set $f(X)$. Therefore, the function f is onto if and only if $f(X) = Y$.

Note that the image of the empty set under any map f is the empty set, that is, $f(\emptyset) = \emptyset$. Can you prove this? What does it mean to say $y \in f(\emptyset)$?

Example 3.5.1. Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Then, we have

(i) $f(\{-1, 1\}) = \{1\}$,
(ii) $f([0, 2]) = [0, 4]$
(iii) $f([-4, 3]) = [0, 16]$.
(iv) $f([-2, 3)) = [0, 9]$.
(v) $f([-3, -1] \cup [0, 2]) = [0, 9]$.

Solution: We shall verify (iii) and leave the remaining as exercise.

If $y \in f([-4, 3])$, then there exists $x \in [-4, 3]$ such that $y = x^2$. First note that $y \geq 0$. Also image of end points are $f(-4) = 16$ and $f(3) = 9$. For $x \in [-4, 3]$, $0 \leq x^2 \leq 16$. That is, $f([-4, 3] \subseteq [0, 16]$.

Let $y \in [0, 16]$. If $y \leq 9$ then there exists $x \in [0, 3]$ such that $y = x^2$. Hence $y \in f([-4, 3])$. If $9 < y \leq 16$, then there exists $x \in [-4, -3)$ such that $y = x^2$. This means $y \in f([-4, 3])$. Thus $[0, 16] \subseteq f([-4, 3])$.

Exercise 3.5.2. Let $f\colon X \to Y$. Suppose $A \subseteq X$ and $x \in X$ such that $f(x) \in f(A)$. Is it true that $x \in A$? (*Hint:* Look at the last example.)

Exercise 3.5.3. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = |x| + 2$. Find the images of the following sets under f.

$$\{-2, -1\}, [-2, 1], (-1, 4],\ \mathbb{R},\ \mathbb{R} \setminus \{0\}.$$

Next, our aim is to look at what happens to the images if we take various set operations.

Lemma 3.5.4. *Let $f\colon X \to Y$ and $A, B \subseteq X$. If $A \subseteq B$, then $f(A) \subseteq f(B)$.*

Proof. Let $y \in f(A)$, we need to show that $y \in f(B)$. Since $y \in f(A)$, by definition there exists $a \in A$ such that $y = f(a)$. Since $A \subseteq B$, $a \in B$. Hence $y = f(a) \in f(B)$. □

Let us first explore the relationship between $f(A \cup B)$ and $f(A) \cup f(B)$ for subsets A and B of X. It follows from Lemma 3.5.4 that $f(A) \cup f(B) \subseteq f(A \cup B)$ (why?). Can we show that $f(A \cup B) \subseteq f(A) \cup f(B)$?

Let $y \in f(A \cup B)$. Then there exists $x \in A \cup B$ such that $y = f(x)$. We have $x \in A$ or $x \in B$. If $x \in A$, then $y = f(x) \in f(A) \subseteq f(A) \cup f(B)$. Hence $y \in f(A) \cup f(B)$. Similarly, if $x \in B$, then $y = f(x) \in f(B) \subseteq f(A) \cup f(B)$. That is, $y \in f(A) \cup f(B)$. Hence, $y \in f(A \cup B)$ implies $y \in f(A) \cup f(B)$. Thus, we have proved the following.

Theorem 3.5.5. *Let $f\colon X \to Y$ be a function. Then, for all subsets A and B of X, $f(A \cup B) = f(A) \cup f(B)$.* □

Now we look at what is the relationship between $f(A \cap B)$ and $f(A) \cap f(B)$. The following result follows from Lemma 3.5.4. We leave its proof as an exercise.

Theorem 3.5.6. *Let $f\colon X \to Y$ be a function. Then, for all subsets A and B of X, $f(A \cap B) \subseteq f(A) \cap f(B)$.* □

It is natural to ask if the inclusion in Theorem 3.5.6 is an equality. Let us try to prove $f(A) \cap f(B) \subseteq f(A \cap B)$ and see what we get. Let $y \in f(A) \cap f(B)$. Then $y \in f(A)$ and $y \in f(B)$. This implies there exist $x_1 \in A$ and $x_2 \in B$ such that $f(x_1) = y$ and $f(x_2) = y$. However, we want $x \in A \cap B$ such that $f(x) = y$. This may not hold always, right? When can one be sure of such an x? Rather, when is one sure that $x_1 = x_2$? Can you see that it is the case if f is one-one?

What if f is not one-one? The above observation gives you a hint to produce an example where the inclusion is proper. If f is not one-one, there exist $x_1, x_2 \in X$, $x_1 \neq x_2$, such that $f(x_1) = f(x_2) = y \in Y$. Choose $A = \{x_1\}$, $B = \{x_2\}$. Then, $A \cap B = \emptyset$, and so $f(A \cap B) = \emptyset \subsetneq \{y\} = f(A) \cap f(B)$.

The investigations above lead us to the following characterization of a one-one function.

Theorem 3.5.7. *A function $f\colon X \to Y$ is one-one if and only if $f(A \cap B) = f(A) \cap f(B)$ holds for all subsets A and B of X.*

Proof. Let $f\colon X \to Y$ be a one-one, and $A, B \subseteq X$. We show that $f(A \cap B) = f(A) \cap f(B)$. In view of Theorem 3.5.6, we need to prove that $f(A) \cap f(B) \subseteq f(A \cap B)$.

Let $y \in f(A) \cap f(B)$. This implies $y \in f(A)$ and $y \in f(B)$. This means, there exist $x_1 \in A$ and $x_2 \in B$ such that $f(x_1) = y = f(x_2)$. Since f is one-one, we have $x_1 = x_2$. Therefore, $x_1 \in A \cap B$, which means $y = f(x_1) \in f(A \cap B)$. Consequently, $f(A) \cap f(B) \subseteq f(A \cap B)$.

Conversely, suppose for any two subsets $A, B \subseteq X$, $f(A \cap B) = f(A) \cap f(B)$. We show that f is one-one. Consider $x_1, x_2 \in X$, $x_1 \neq x_2$. Let $A = \{x_1\}, B = \{x_2\}$. Then $A \cap B = \emptyset$. Therefore, $\{f(x_1)\} \cap \{f(x_2)\} = f(A) \cap f(B) = f(A \cap B) = \emptyset$. This implies that $f(x_1) \neq f(x_2)$. Consequently, f is one-one. □

Remark 3.5.8. Go through the second part of the proof above. Assume that the hypothesis (of a theorem) has the universal quantifier (in this case, $\forall A, B \subseteq X$). When we try to deduce the conclusion (in this case f is one-one), we have to make intelligent/appropriate choice of elements from the set of the context for the quantifier $\forall$. In the proof above, we chose $A = \{x_1\}$ and $B = \{x_2\}$. Be on the lookout for such a trick in future. For example, see the first part of the proof of Theorem 3.5.12, and the second part of Theorem 3.6.26.

Exercise 3.5.9. Show that $f\colon X \to Y$ is injective if and only if for all $A, B \subseteq X$, $A \cap B = \emptyset$ implies $f(A) \cap f(B) = \emptyset$.

Example 3.5.10. Let $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 1$ for all $x \in \mathbb{R}$. Let $A = (-\infty, 0)$ and $B = (0, \infty)$. Then $f(A \cap B) = \emptyset$. But $f(A) \cap f(B) = \{1\}$. Thus $f(A \cap B)$ is a proper subset of $f(A) \cap f(B)$.

Example 3.5.11. Let $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ for all $x \in \mathbb{R}$. Let $A = [-2, 1]$ and $B = [0, 2]$. Then $f(A \cap B) = [0, 1]$. But $f(A) \cap f(B) = [0, 4]$.

Next, we explore what happens to the image of the complement of a set. Let $f\colon X \to Y$ and $A \subseteq X$. We wish to find relationship between $f(X \setminus A)$ and $Y \setminus f(A)$.

When does $Y \setminus f(A) \subseteq f(X \setminus A)$ hold for an arbitrary set $A \subseteq X$? Note that if $y \in Y \setminus f(A)$, then $y \notin f(A)$, that is, for each $x \in A$ we have $f(x) \neq y$. Does it follow from this that $y = f(x)$ for some $x \in X \setminus A$ so that $y \in f(X \setminus A)$? When will this be surely the case?

The following result gives an answer to this questions.

Theorem 3.5.12. *Let $f\colon X \to Y$. Then, f is onto if and only if $f(X \setminus A) \supseteq Y \setminus f(A)$ for any set $A \subseteq X$.*

Proof. First, assume that $f(X \setminus A) \supseteq Y \setminus f(A)$ for any set $A \subseteq X$. We wish to show that f is onto. That is, we need to show that $f(X) = Y$. We take $A = \emptyset$. Then $f(A) = \emptyset$, and we have $Y = Y \setminus f(A) \subseteq f(X \setminus A) = f(X)$. Since $f(X) \subseteq Y$, we have $f(X) = Y$.

Next, assume that f is onto. We want to show that $f(X \setminus A) \supseteq Y \setminus f(A)$ for any set $A \subseteq X$. Let $A \subseteq X$ and $y \in Y \setminus f(A)$. Since f is onto, there exists $x \in X$ such that $y = f(x)$. We claim that $x \notin A$. If $x \in A$, then $y = f(x) \in f(A)$ which is a contradiction. Thus, if $y \in Y \setminus f(A)$, then $y \in f(X \setminus A)$. □

We may now ask the 'dual' question. When does $f(X \setminus A) \subseteq Y \setminus f(A)$ hold for an arbitrary set $A \subseteq X$? Note that if $y \in f(X \setminus A)$, then there exists $x \in X \setminus A$ such that $y = f(x)$. That is, there exists $x \notin A$ such that $y = f(x)$.

Does it follow from this that $f(x) \notin f(A)$ so that $y = f(x) \in Y \setminus f(A)$? When will this be surely the case?

The next result answers this. The proof is left as an easy exercise.

Theorem 3.5.13. *Let $f\colon X \to Y$. Then, f is one-one if and only if $f(X \setminus A) \subseteq Y \setminus f(A)$ for all sets $A \subseteq X$.* □

Combining Theorem 3.5.12 and Theorem 3.5.13, we get a characterization of a bijection.

Theorem 3.5.14. *A function $f\colon X \to Y$ is bijection if and only if $f(X \setminus A) = Y \setminus f(A)$ for all sets $A \subseteq X$.* □

Theorems 3.5.5 and 3.5.6 can be generalized to union and intersection of family of sets. We leave the proof as an exercise.

Theorem 3.5.15. *Let $f\colon X \to Y$ and $\{F_\alpha : F_\alpha \subseteq X, \alpha \in \Lambda\}$ is a family of subsets of X indexed by the set Λ. Then*

(a) $f\left(\bigcup_{\alpha\in\Lambda} F_\alpha\right) = \bigcup_{\alpha\in\Lambda} f(F_\alpha)$.

(b) $f\left(\bigcap_{\alpha\in\Lambda} F_\alpha\right) \subseteq \bigcap_{\alpha\in\Lambda} f(F_\alpha)$. *Equality holds if and only if f is one-one.* □

3.6 Inverse image of subsets under functions

Let $f\colon X \to Y$ and $B \subseteq Y$. We define the *inverse image* of B under f to be the subset

$$f^{-1}(B) := \{x \in X : f(x) \in B\}$$

of X, that is, the set of all preimages of the elements in B.

Note that $f^{-1}(B)$ is not the image of B under f^{-1}. This is merely a notation. It should be read as the inverse image of B under f, and not as the image of B under f^{-1}. In fact, $f^{-1}(B)$ makes sense for any function $f\colon X \to Y$ and any subset $B \subseteq Y$, even if f^{-1} does not exist.

If $B = \{y\}$ is a singleton set, it is customary to denote $f^{-1}(B)$ by $f^{-1}(y)$ rather than by $f^{-1}(\{y\})$.

What is meaning of $x \in f^{-1}(B)$? It simply means $f(x) \in B$. Also, note that $f^{-1}(\emptyset) = \emptyset$.

Example 3.6.1. Let $f\colon X \to Y$ be a constant map. That is, there exists $y_0 \in Y$ such that $f(x) = y_0$ for all $x \in X$. Let $B \subseteq Y$. Then $f^{-1}(B) = X$ if $y_0 \in B$ and is $\emptyset$ otherwise.

Example 3.6.2. Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Then

(i) $f^{-1}(1) = \{\pm 1\}$ and $f^{-1}(-1) = \emptyset$,
(ii) $f^{-1}([0,1]) = [-1,1]$,
(iii) $f^{-1}([4,16]) = [-4,-2] \cup [2,4]$,

(iv) $f^{-1}((0,1)) = (-1,1)$,
(v) $f^{-1}([-1,1]) = [-1,1]$,
(vi) $f^{-1}([-4,4]) = [-2,2]$,
(vii) $f^{-1}((0,4)) = (-2,2)$.

Solution: (ii) Let $x \in f^{-1}([0,1])$. Then, by definition $f(x) = x^2 \in [0,1]$. That is, $0 \leq x^2 \leq 1$. What does it say about x? Clearly in this case $x \in [-1,1]$. Therefore, $f^{-1}([0,1]) \subseteq [-1,1]$. On the other hand, if $x \in [-1,1]$ then $f(x) = x^2 \in [0,1]$. That is, $[-1,1] \subseteq f^{-1}[0,1])$.

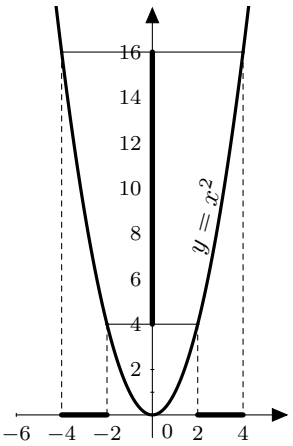


Figure 3.10: $f^{-1}[4,16]$ for $f(x) = x^2$

(iii) Let $x \in f^{-1}([4,16])$. Then $4 \leq x^2 \leq 16$. This means $x^2 \geq 4$ and $x^2 \leq 16$. The first one implies $|x| \geq 2$, that is, $x \in (\infty,-2] \cup [2,\infty)$. The second one implies $x \in [-4,4]$. Combining the two together, we have $f^{-1}([4,16]) = [-4,-2] \cup [2,4]$.

You may also try to see this geometrically (see Figure 3.10).

(v) Let $x \in f^{-1}([-1,1])$. Then by definition $f(x) = x^2 \in [-1,1]$. Since $x^2 \geq 0$, it is same as saying $f(x) \in [0,1]$. Now the result follows from (ii).

Exercise 3.6.3. Let $f\colon (0,\infty) \to (0,\infty)$ be given by $f(x) = 1/x$. What are $f^{-1}((0,1))$ and $f^{-1}((1,\infty))$?

Exercise 3.6.4. Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) := \sum_{k=0}^{n} a_k x^k$. Show that there exists a natural number N such that for any $c \in \mathbb{R}$, the number of elements in $f^{-1}(c)$ is at most N.

Exercise 3.6.5. Let $f\colon [-2\pi, 2\pi] \to \mathbb{R}$ be given by $f(x) = \sin x$. Find $f^{-1}([0,1])$.

Exercise 3.6.6. Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \cos x$. Find $f^{-1}(1)$ and $f^{-1}(-1/2)$.

Example 3.6.7. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = x^2 + y^2$. Then

(a) $f^{-1}(0) = \{(0,0)\}$.

(b) $f^{-1}(r) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = r^2\}$ if $r > 0$ and is empty set if $r < 0$.

(c) $f^{-1}([0, 4]) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 4\}$, the closed disk centered at the origin $(0, 0)$ with radius 2.

(d) $f^{-1}([1, 4)) = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 < 4\}$, an annulus without outer boundary and with inner boundary (see Figure 3.11).

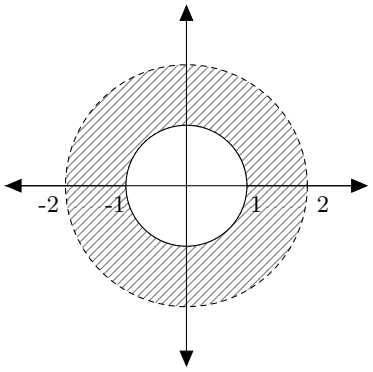


Figure 3.11: $1 \leq x^2 + y^2 < 4$

Example 3.6.8. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x$. Then

(a) $f^{-1}(r) = \{(r, y) : y \in \mathbb{R}\}$, the vertical line at $x = r$.

(b) $f^{-1}([a, b]) = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, y \in \mathbb{R}\}$, the vertical strip between the lines $x = a$ and $x = b$ (see Figure 3.12).

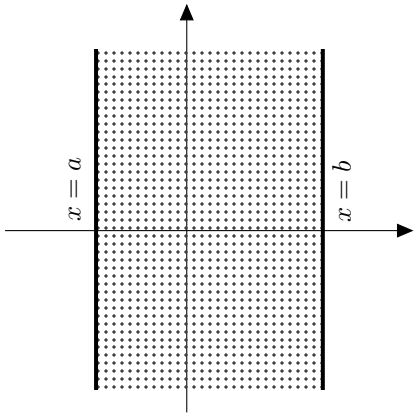


Figure 3.12: $f^{-1}[a, b]$ for $f(x, y) = x$

Exercise 3.6.9. Let $f\colon \mathbb{R} \to \mathbb{R}$ be given by

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

This function is known as *Dirichlet's function*. Find $f^{-1}([0,1/2])$, $f^{-1}([1/2,2])$, $f^{-1}([-1,2])$ and $f^{-1}([2,3])$.

Exercise 3.6.10. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = xy$. What are $f^{-1}(r)$ for $r \in \mathbb{R}$? Draw pictures of these inverse images.

Exercise 3.6.11. Let $f\colon M(n,\mathbb{R}) \to \mathbb{R}$ be given by $f(X) = \det(X)$. Identify the sets $f^{-1}(0)$ and $f^{-1}(\mathbb{R}^*)$, where $\mathbb{R}^*$ denotes the set of nonzero real numbers.

Exercise 3.6.12. Let $f\colon M(n,\mathbb{R}) \to M(n,\mathbb{R})$ be given by $f(X) = XX^T$. Identify the sets $f^{-1}(I)$.

Exercise 3.6.13. Let $f,g\colon M(n,\mathbb{R}) \to M(n,\mathbb{R})$ be given by $f(X) = X + X^T$ and $g(X) = X - X^T$. Identify the sets $f^{-1}(0)$ and $g^{-1}(0)$.

Exercise* 3.6.14. Let $f\colon M(n,\mathbb{R}) \to M(n,\mathbb{R})$ be given by $f(X) = X^n$. Identify the set $f^{-1}(0)$.

Exercise* 3.6.15. Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous and strictly increasing. Assume that $\alpha < \beta$ are in the image of f. What is $f^{-1}([\alpha,\beta])$?

Answer the same question if f is strictly decreasing.

Lemma 3.6.16. *Let $f\colon X \to Y$ and $g\colon Y \to Z$ be maps. Let $C \subseteq Z$. Then $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.*

Proof. Let $x \in (g \circ f)^{-1}(C)$. Then $(g \circ f)(x) \in C$. By definition, $(g \circ f)(x) = g(f(x))$. Hence $f(x) \in g^{-1}(C) \subset Y$. It follows that $x \in f^{-1}(g^{-1}(C))$. Thus we have shown that $(g \circ f)^{-1}(C) \subset f^{-1}(g^{-1}(C))$. The reverse inclusion is similar. Let $x \in f^{-1}(g^{-1}(C))$. This means that $f(x) \in g^{-1}(C)$, which in turn says that $g(f(x)) \in C$. That is, $(g \circ f)(x) \in C$. Hence $x \in (g \circ f)^{-1}(C)$. □

Theorem 3.6.17. *Suppose $f\colon X \to Y$ is a bijection and g is the inverse of f. Then, $f^{-1}(B) = g(B)$ for any subset $B \subseteq Y$.*

That is, the inverse image of B under a bijective map f is the image of B under the inverse of f.

Proof. First, let $x \in f^{-1}(B)$. By definition $f(x) \in B$. We have to show $x \in g(B)$. That is, there exists y in B such that $g(y) = x$. Since g is the inverse of f, $g(f(x)) = x$. This suggests us to consider $y = f(x)$. Then $y \in B$ and $g(y) = x$. Hence, $f^{-1}(B) \subseteq g(B)$.

Conversely, suppose $x \in g(B)$. Then, there exists $y \in B$ such that $g(y) = x$. By the definition of g, this is true if and only if $f(x) = y$. Since $y \in B$, this means that $x \in f^{-1}(B)$. Hence, for any $x \in g(B)$, we have $x \in f^{-1}(B)$, that is, $g(B) \subseteq f^{-1}(B)$. □

Exercise 3.6.18. Let $f\colon X \to Y$ be a map. Let y_1, y_2 be two distinct elements in Y. Show that $f^{-1}(y_1) \cap f^{-1}(y_2) = \emptyset$. Can you generalize this?

Exercise 3.6.19. For any function $f\colon X \to Y$, let $A_y := f^{-1}(y)$. Show that $\{A_y : y \in Y\}$ is a pairwise disjoint family of subsets of X and $X = \bigcup_{y\in Y} A_y$.

Exercise 3.6.20. Let $f\colon X \to Y$ be a function and $B_1 \subseteq B_2 \subseteq Y$. Show that $f^{-1}(B_1) \subseteq f^{-1}(B_2)$.

The inverse images of subsets behave well with respect to set theoretic operations unlike the images of subsets.

Theorem 3.6.21. *Let $f\colon X \to Y$ be a map. Let B_1, B_2 be subsets of Y. Then*

(a) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$.

(b) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.

Proof. (a) Let $x \in f^{-1}(B_1 \cup B_2)$, we wish to show that $x \in f^{-1}(B_1) \cup f^{-1}(B_2)$. We have

$$\begin{aligned} x \in f^{-1}(B_1 \cup B_2) &\implies f(x) \in B_1 \cup B_2 \\ &\implies f(x) \in B_1 \text{ or } f(x) \in B_2 \\ &\implies x \in f^{-1}(B_1) \text{ or } x \in f^{-1}(B_2) \\ &\implies x \in f^{-1}(B_1) \cup f^{-1}(B_2). \end{aligned}$$

Thus, we have $f^{-1}(B_1 \cup B_2) \subseteq f^{-1}(B_1) \cup f^{-1}(B_2)$.

Do you think you can reverse the implications above and conclude that $f^{-1}(B_1) \cup f^{-1}(B_2) \subseteq f^{-1}(B_1 \cup B_2)$? Justify.

You can also prove $f^{-1}(B_1) \cup f^{-1}(B_2) \subseteq f^{-1}(B_1 \cup B_2)$ using Exercise 3.6.20.

The proof of (b) follows on similar lines and is left to the reader. □

Can you formulate a generalization of Theorem 3.6.21 for arbitrary union and intersection of subsets of Y?

Theorem 3.6.22. *Let $f\colon X \to Y$ and $\mathcal{F} = \{A_\alpha \subseteq Y : \alpha \in I\}$ a family of subsets of Y indexed by I. Then*

(a) $f^{-1}\left(\bigcup_{\alpha\in I} A_\alpha\right) = \bigcup_{\alpha\in I} f^{-1}(A_\alpha)$,

(b) $f^{-1}\left(\bigcap_{\alpha\in I} A_\alpha\right) = \bigcap_{\alpha\in I} f^{-1}(A_\alpha)$.

Proof. Let us prove (b) and leave (a) as an exercise.

Let $x \in f^{-1}(\bigcap_{\alpha\in I} A_\alpha)$, that is, $f(x) \in \bigcap_{\alpha\in I} A_\alpha$. Then for each $\alpha \in I$ we have $f(x) \in A_\alpha$. This implies that $x \in f^{-1}(A_\alpha)$ for each $\alpha \in I$, and therefore $x \in \bigcap_{\alpha\in I} f^{-1}(A_\alpha)$. Hence we have $f^{-1}\left(\bigcap_{\alpha\in I} A_\alpha\right) \subseteq \bigcap_{\alpha\in I} f^{-1}(A_\alpha)$.

Conversely, let $x \in \bigcap_{\alpha\in I} f^{-1}(A_\alpha)$. Then for each $\alpha \in I$, we have $x \in f^{-1}(A_\alpha)$. This implies that for each $\alpha \in I$, $f(x) \in A_\alpha$. Hence, $f(x) \in \bigcap_{\alpha\in I} A_\alpha$, which implies $x \in f^{-1}\left(\bigcap_{\alpha\in I} A_\alpha\right)$. This proves $\bigcap_{\alpha\in I} f^{-1}(A_\alpha) \subseteq f^{-1}\left(\bigcap_{\alpha\in I} A_\alpha\right)$. □

Next, we prove that the inverse image of subsets also obeys the set complements.

Theorem 3.6.23. *Let $f\colon X \to Y$ be a map and $B \subseteq Y$. Then $f^{-1}(B^c) = (f^{-1}(B))^c$. In other words, $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$.*

Proof. Let $x \in f^{-1}(B^c)$. This implies $f(x) \in B^c$. We claim that $x \notin f^{-1}(B)$. For, if $x \in f^{-1}(B)$ then $f(x) \in B$, a contradiction. Hence we have proved that if $x \in f^{-1}(B^c)$, then $x \in (f^{-1}(B))^c$. That is, $f^{-1}(B^c) \subseteq (f^{-1}(B))^c$.

Next let $x \in (f^{-1}(B))^c$. Then $x \notin f^{-1}(B)$. This implies $f(x) \notin B$, that is $f(x) \in B^c$ and hence $x \in f^{-1}(B^c)$. This proves $(f^{-1}(B))^c \subseteq f^{-1}(B^c)$. □

Let $f\colon X \to Y$ be a map. Let $A \subseteq X$ and $B := f(A) \subseteq Y$. Then we can think of $f^{-1}(B) = f^{-1}(f(A))$. We want to look at how A and $f^{-1}(f(A))$ related? If you think for a while, you will realize that the inverse image of $f(A)$ will contain elements of A. Indeed, if $x \in A$, then $f(x) \in f(A)$, which implies $x \in f^{-1}(f(A))$. However, $f^{-1}(f(A))$ may contain A as a proper subset, as the following examples show.

Example 3.6.24. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$.

(a) Let $A = \{1\}$. Then $f^{-1}(f(A)) = \{\pm 1\}$.

(b) Let $A = [0, 1]$. Then $f(A) = [0, 1]$, and therefore $f^{-1}(f(A)) = [-1, 1]$.

The example suggests a way to generate more such examples.

Example 3.6.25. Let X be a set with at least two elements, and $f\colon X \to Y$ be a constant function $f(x) = y_0$. Let $a \in X$ and $A = \{a\}$. Then $f(A) = \{y_0\}$ and $f^{-1}(f(A)) = X \neq A$.

The preceding examples suggest that $f^{-1}(f(A)) = A$ may hold for each $A \subseteq X$ if f is one-one. Indeed, this is the case. We have another characterization of a one-one function.

Theorem 3.6.26. *A function $f\colon X \to Y$ is one-one if and only if $A = f^{-1}(f(A))$ for each $A \subseteq X$.*

Proof. Let us assume that f is one-one and show that $A = f^{-1}(f(A))$ for each $A \subseteq X$.

Let $A \subseteq X$. We have noted that $A \subseteq f^{-1}(f(A))$ is true for any function. We need to show that $f^{-1}(f(A)) \subseteq A$.

Let $x \in f^{-1}(f(A))$. Then $y = f(x) \in f(A)$. By definition of $f(A)$ there exists $a \in A$ such that $f(a) = y$. Thus we have $x, a \in X$ such that $f(x) = y = f(a)$. Since f is one-one, we have $x = a$. This proves that $x \in A$.

Conversely, assume that $A = f^{-1}(f(A))$ for each $A \subseteq X$. We wish to prove that f is one-one. Suppose $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$. Consider $A = \{x_1\}$. Then $f(A) = \{f(x_1)\}$ and $x_2 \in f^{-1}(f(A))$. By the hypothesis, $f^{-1}(f(A)) = A = \{x_1\}$. This implies $x_2 \in \{x_1\}$, that is, $x_1 = x_2$. Hence, f is one-one. □

The following theorem characterizes an onto function. The proof is similar to that of Theorem 3.6.26 and is left as an exercise.

Theorem 3.6.27. *A function $f\colon X \to Y$ is onto if and only if $f(f^{-1}(B)) = B$ for each $B \subseteq Y$.* □

Chapter 4

Relation

In this chapter, we discuss relations defined between and on sets, and deal with many concrete examples. We introduce an important class of relations on sets, namely, equivalence relations, which appear frequently in many branches of mathematics.

4.1 Relations on sets

Let X be the set of all human beings in the world. Let us say that a human x is related to another human y if x is the father of y. What do we observe about this relation? The first thing to observe is that the order of x and y is important. The second thing is that given two human beings x and y, x may or may not be related to y. The first one suggests that we consider ordered pairs $(x, y) \in X \times X$ and the second one says that the ordered pairs (x, y) such that x is related to y form a subset of $X \times X$, which may be proper.

Definition 4.1.1. A *relation* R on a nonempty set X is a subset of $X \times X$. Given such a subset $R \subseteq X \times X$, we say that x is related to y under R if $(x, y) \in R$. In that case, we write xRy.

In practice, most often we shall define the relation on X very directly and explicitly, and not through the *defining subset* R. In fact, sometimes we define the relation directly on X and ask you to identify the defining set R. See, for instance, Example 4.1.10 and Example 4.1.12.

Before we look at some concrete examples in mathematics, we shall generalize this notion. Let M be the set of men and W be the set of women. Then we say $x \in M$ is related to $y \in W$ if x is the husband of y. A closer look as above yields the following definition.

Definition 4.1.2. A *relation between* nonempty sets X and Y is a subset $R \subseteq X \times Y$. We say that $x \in X$ is related to $y \in Y$ if $(x, y) \in R$. If that be the case, we shall denote it by xRy.

At times, a relation between X and Y is also called a *binary relation* between X and Y.

Example 4.1.3. Let $X = \{a, b, c\},\ Y = \{1, 2, 3\}$. Consider the subset R of $X \times Y$ as

$$R = \{(a, 1), (a, 3), (b, 1), (b, 2)\}.$$

In this relation, a is related to 1 and 3, since $(a, 1), (a, 3) \in R$. Similarly b is related to 1 and 2, since $(b, 1), (b, 2) \in R$. Note that c is not related to any element of Y.

Exercise 4.1.4. Let $X = \{a, b, c\},\ Y = \{1, 2, 3\}$. Consider the subset S of $X \times Y$ as

$$S = \{(a, 1), (c, 1), (c, 2)\}.$$

In this relation, is a related to 2? Is b related to any element of Y? What are elements in X to which 3 is related under S?

Example 4.1.5. Let X and Y be nonempty sets. (i) Consider $R = \emptyset \subseteq X \times Y$. Under this relation, no element of X is related to any element of Y.
(ii) Suppose $R = X \times Y$. Given any $x \in X$ and $y \in Y$, x is related to y under R.

Exercise 4.1.6. On $\mathbb{R}$ consider the relation defined by $R = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$, the unit circle centered at the origin in $\mathbb{R}^2$.
(i) Is 1 related to 1 via R?
(ii) Find all real numbers y such that 0 is related to y.
(iii) Suppose $x \in [-1, 1]$, and $y \in \mathbb{R}$ with $|y| > 1$. Can xRy hold?
(iv) Find all real numbers x such that x is related to 1.

Example 4.1.7. Consider the relation on $\mathbb{R}$ with defining set $R := \{(x, y) \in \mathbb{R} \times \mathbb{R} : x < y\}$. Here $(x, y) \in R$ iff $x < y$. Thus, the subset R yields the standard 'less than' relation on $\mathbb{R}$. (This relation can be defined on $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ in an analogous manner.) The relation R is denoted by $<$.

Example 4.1.8. In the last example, let us replace $<$ by $\leq$ in the defining set R. Then xRy iff $x \leq y$. The relation is called the *standard order* on $\mathbb{R}$, denoted by $\leq$.

Exercise 4.1.9. Let $R \subseteq \mathbb{R} \times \mathbb{R}$ be the defining set that corresponds to the standard order $\leq$ on $\mathbb{R}$. Draw the picture of the subset R.

Example 4.1.10. Let S be a nonempty set and $X = P(S)$, the power set of S. We define a relation on X by setting ARB iff $A \subseteq B$. What is the defining subset $R \subseteq X \times X$ for this relation?

Example 4.1.11. Let $R := \{(m, km) : m \in \mathbb{N}, k \in \mathbb{N}\}$. Do you recognize the relation on $\mathbb{N}$? We have mRn holds iff n is a multiple of m, that is, iff m divides n.

Example 4.1.12. Let $X = \mathbb{Z}^*$, the set of nonzero integers. We define a relation R on X by setting mRn iff m divides n. Identify the defining subset $R \subseteq \mathbb{Z}^* \times \mathbb{Z}^*$.

Example 4.1.13. Let X be a nonempty set and let $\Delta(X) = \{(x, x) : x \in X\}$, called the *diagonal subset* in $X \times X$. The relation $R = \Delta(X)$ is called the *equality* or *identity* relation on X, since xRy iff $x = y$.

Example 4.1.14. Let $X = \mathbb{R}$ and let $R = \{(x, y) \in \mathbb{R}^2 : xy = 0\}$. When is $x \in \mathbb{R}$ related to $y \in \mathbb{R}$? We note that xRy iff one of x, y is 0.

Example 4.1.15. Let $f\colon X \to Y$ be a function. Let R be the relation defined by the graph of f, that is, $R = \{(x, f(x)) : x \in X\}$. Given $x \in X$ and $y \in Y$, x is related to y iff $y = f(x)$.

Suppose $f\colon X \to X$, and R is the graph of f. Then xRx holds iff x is a fixed point of f, that is, $x = f(x)$.

Function as a relation

At times, a function is defined as a relation. Let X and Y be two nonempty sets. Recall that if f is a function from X to Y, then each element of X corresponds to a unique element in Y. That is, for each $x \in X$, we have a unique pair (x, y). Thus, we can define a function as a relation as follows:

A relation, say R, from a set X to a set Y is a function if and only if for each element $x \in X$ there exists a unique $y \in Y$ such that $(x, y) \in R$. In such a case, we define $f\colon X \to Y$ by setting $f(x) = y$ if $(x, y) \in R$. For example, the relation in Example 4.1.13 corresponds to the identity function on X.

Note that not every relation from X to Y can be viewed as a function. For example, the relation defined in Example 4.1.3 is not a function. (Why?) Similarly, the relation defined on $\mathbb{R}$ by the unit circle $R = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1\}$ is not a function (see Ex. 4.1.6).

4.2 Types of relations

Definition 4.2.1. Let R be a relation on a nonempty set X. Then R is said to be

- **Reflexive**, if $\forall x \in X(xRx)$, that is, each $x \in X$ is related to itself.
- **Symmetric**, if $\forall x, y \in X(xRy$ implies $yRx)$, that is, if x is related to y, then y is related to x.
- **Anti-symmetric**, if $\forall x, y \in X((xRy$ and $yRx)$ implies $x = y)$. In other words, if x is related to y and y is related to x, then $x = y$.
- **Transitive**, if $\forall x, y, z \in X((xRy$ and $yRz)$ implies $xRz)$. That is, if x is related to y and y is related to z, then x is related to z.

Example 4.2.2. The equality relation on a nonempty set X (i.e., xRy iff $x = y$) is reflexive, symmetric, anti-symmetric and transitive.

Example 4.2.3. On a nonempty set X consider the relation defined by $R = X \times X$, that is xRy for every $x, y \in X$. The relation is called the *universal relation* on X. This relation is reflexive, symmetric, and transitive. However, it is not anti-symmetric unless X is a singleton.

Example 4.2.4. Let us look at the relations defined in previous examples and examine if they are reflexive, symmetric, anti-symmetric or transitive.

(i) The relation defined in Example 4.1.8 is reflexive, anti-symmetric and transitive, but not symmetric.

(ii) The relation defined in Example 4.1.7 is not reflexive, not symmetric, but transitive.

It is anti-symmetric. You need logic to prove this. If it is not anti-symmetric, then there exist x and y such that $x < y$ and $y < x$ hold but $x \neq y$. The law of Trichotomy says that there do not exist $x, y \in \mathbb{R}$ such that $x < y$ and $y < x$ hold simultaneously. (Recall how we proved that the empty set is a subset of any set!)

(iii) The relation defined in Example 4.1.11 is reflexive, anti-symmetric and transitive. However it is not symmetric.

(iv) The relation defined in Example 4.1.12 is reflexive and transitive. However, it is not symmetric and not anti-symmetric. (Why?) Compare this relation with the one in (iii).

(v) The relation defined in Ex. 4.1.6 is symmetric, but not reflexive, not anti-symmetric and not transitive.

(vi) The relation defined in Example 4.1.14 is symmetric, but not reflexive, not anti-symmetric and not transitive.

Exercise 4.2.5. Consider the relation given by the graph R of a function $f\colon X \to X$. (See Example 4.1.15.) Show with an example that the relation R may neither be reflexive, nor symmetric nor transitive.

Example 4.2.6. Given two real numbers $x, y \in \mathbb{R}$ we say that x is related to y if they are nonzero and are of the same sign. What is the defining set $R \subseteq \mathbb{R}\times\mathbb{R}$ which yields this relation? It is given by $R := \{(x, y) \in \mathbb{R} \times \mathbb{R} : xy > 0\}$.

The relation R is reflexive, symmetric and transitive.

Exercise 4.2.7. Let R and S be relations on sets X and Y, respectively. Then there exists a natural relation T on $X \times Y$ defined by $(x_1, y_1)T(x_2, y_2)$ iff $x_1 R x_2$ and $y_1 S y_2$. What properties of R and S are shared by T?

Example 4.2.8 (Lexicographic or Dictionary Relation)**.** Let us recall how we locate words in an English dictionary. If the words start with different letters, the word whose first letter precedes that of the other word appears before the other word. For example, if we are given 'Son' and 'Tin', we know that 'Son' will precede 'Tin' in the dictionary. On the other hand, if we are given 'Sun' and 'Son', since their first letters are the same, we move to the second letter and decide which one precedes the other in the dictionary. In the case on hand, 'Son' appears before 'Sun' in the dictionary. This reasoning shows us the way to define a relation on the Cartesian product $X \times Y$ where X and Y have relations defined on them.

Given relations R on X and S on Y, we have a relation T on $X \times Y$ by setting

$$(x_1, y_1)T(x_2, y_2) \quad \text{iff} \quad x_1 R x_2 \text{ and if } x_1 = x_2, \text{ then } y_1 S y_2.$$

This is known as the *lexicographic or dictionary relation* on $X \times Y$.

For example, take $X = Y = \mathbb{R}$ with the standard order relation $\leq$. Then the dictionary order on $\mathbb{R} \times \mathbb{R}$ is the relation $\leq$ given by:

$$(x_1, y_1) \leq (x_2, y_2) \text{ iff } x_1 \leq x_2 \text{ and if } x_1 = x_2 \text{ then } y_1 \leq y_2.$$

The relation can also be expressed as

$$(x_1, y_1) \leq (x_2, y_2) \text{ iff either } x_1 < x_2 \text{ or } (x_1 = x_2 \text{ and } y_1 \leq y_2).$$

We shall revisit this example in Chapter 7.

Exercise 4.2.9. Consider $\mathbb{Z} \times \mathbb{Z}$ with dictionary order, denoted by $\leq$. What is the relation between the elements (i) $(-1, 0)$ and $(2, -20)$ and (ii) $(1, 10)$ and $(1, -10)$?

Example 4.2.10. Let $X = \mathbb{C}$, the set of complex numbers. We write $z \in \mathbb{C}$ as $z = x + iy$ with $x, y \in \mathbb{R}$. If $z = x + iy$ and $w = u + iv$ are complex numbers, we say that zRw if either $x < u$ or ($x = u$ and $y < v$). Note that this is 'essentially' the same as the dictionary order on $\mathbb{R}^2$, as discussed in Example 4.2.8, if we agree to 'identify' $z = x + iy \in \mathbb{C}$ with $(x, y) \in \mathbb{R}^2$.

Investigate whether R is reflexive, symmetric, anti-symmetric, or transitive.

Exercise 4.2.11. Investigate what common properties of R and S are inherited by the lexicographic relation T in Example 4.2.8.

Exercise 4.2.12. We define a relation $\preceq$ on $\mathbb{N}$ as follows. We say that $a \preceq b$ iff $a \leq 2b$. Is this relation reflexive, symmetric, anti-symmetric, transitive?

Definition 4.2.13. Given a relation R between X and Y, the *inverse* R^{-1} of R is the relation between Y and X defined by $R^{-1} := \{(y, x) \in Y \times X : (x, y) \in R\}$. Similarly, the *inverse* of a relation R on X is defined by $R^{-1} := \{(y, x) \in X \times X : (x, y) \in R\}$.

Exercise 4.2.14. Write down the inverses of the relations discussed in Section 4.1. What properties of R are shared by its inverse?

Exercise 4.2.15. Let R be a relation on X induced by a function X to X. (See Example 4.1.15.) When is R^{-1} induced by a function?

4.3 Equivalence relations

Definition 4.3.1. Let X be a nonempty set. A relation $\sim$ on X is said to be an *equivalence relation* if it is reflexive, symmetric and transitive.

Thus, a relation $\sim$ on X is an equivalence relation if for every $x, y, z \in X$ (i) $x \sim x$, (ii) $x \sim y$ implies $y \sim x$, and (iii) ($x \sim y$ and $y \sim z$) implies $x \sim z$.

Example 4.3.2. On any nonempty set X, the equality relation (see Example 4.2.2) and the universal relation (see Example 4.2.3) are equivalence relations on X.

Example 4.3.3. The relation $\leq$ defined on $\mathbb{R}$ (that is, xRy iff $x \leq y$) is not an equivalence relation as it is not symmetric.

Definition 4.3.4. Suppose $\sim$ is an equivalence relation on a nonempty set X. For $x \in X$ define

$$[x] := \{y \in X : x \sim y\}$$

The subset $[x]$ of X is called the *equivalence class of* x for $\sim$. It is the collection of all those elements in X which are related to x for the relation $\sim$.

For any $x \in X$ the class $[x]$ is nonempty, since $x \in [x]$.

Example 4.3.5. Let X be a nonempty set and $x \in X$. For the equality relation on X, $[x] = \{x\}$, and for the universal relation on X, $[x] = X$.

Definition 4.3.6. Let $\sim$ be an equivalence relation on X. Let $[x]$ be an equivalence class. If $a \in [x]$, then a is called a *representative* of $[x]$. A subset S of X is called a *transversal* for $\sim$, if S contains precisely one representative from each equivalence class, that is, if for each equivalence class $[x]$, we have $S \cap [x]$ is a singleton.

Example 4.3.7. Consider the relation on $\mathbb{R}$ with the defining set

$$R := \{(x, y) \in \mathbb{R} \times \mathbb{R} : xy > 0\} \cup \{(0, 0)\}.$$

It is easy to see that R is an equivalence relation. (Compare with Example 4.2.6.) It has three equivalence classes: $[0] = \{0\}$, $[1] = \{x \in \mathbb{R} : x > 0\}$ and $[-1] = \{x \in \mathbb{R} : x < 0\}$. Thus, $\{-1, 0, 1\}$ is a transversal of R.

Example 4.3.8. Fix $k \in \mathbb{N}$. For integers m and n let us say that $m \sim n$ iff $m - n$ is a integral multiple of k. (What is the defining subset $R \subseteq \mathbb{Z} \times \mathbb{Z}$ for this relation?)

Let us show that this relation is an equivalence relation. For any integer $m \in \mathbb{Z}$, $m - m = 0 = 0 \times k$. Therefore, $m \sim m$, and so $\sim$ is reflexive.

Let $m \sim n$ and $m - n = kd$, where $d \in \mathbb{Z}$. Then, $n - m = k(-d)$, that is, $n \sim m$. Hence $\sim$ is symmetric.

Next, let $m \sim n$ and $n \sim p$. Then there exist integers d and d' such that $m - n = kd$ and $n - p = kd'$.

$$m - p = (m - n) + (n - p) = kd - kd' = k(d - d'),$$

yielding $m \sim p$. Hence, $\sim$ is transitive.

Thus $\sim$ is an equivalence relation on $\mathbb{Z}$. This relation is called the *congruence modulo* k, and $m \sim n$ is written as $m \equiv n (\mathrm{mod}\ k)$.

Now, let us find the equivalence class $[d]$ of an integer d. If $m \in [d]$, we have $m \equiv d (\mathrm{mod}\ k)$, that is, $m - d = kr$ for some integer r. Therefore, $m = d + kr$, where $r \in \mathbb{Z}$. On the other hand, for $s \in \mathbb{Z}$, $d + ks \in [d]$. Hence, the equivalence class of d is given by

$$[d] = \{d + ks : s \in \mathbb{Z}\}.$$

How many equivalence classes are there? Can you find a natural transversal of this relation?

Example 4.3.9. Let us take a special case of the relation defined in Example 4.3.8 with $k = 6$. Thus, $m, n \in \mathbb{Z}$ are related iff $m - n = 6k$ for some integer k. Let us list some equivalence classes.

$$\begin{array}{lcll}
[0] & = & \{\ldots, -18, -12, -6, 0, 6, 12, 18, 24, \ldots\} & = \{6n : n \in \mathbb{Z}\}. \\
[1] & = & \{\ldots, -17, -11, -5, 1, 7, 13, 19, 25, \ldots\} & = \{1 + 6n : n \in \mathbb{Z}\}. \\
[2] & = & \{\ldots, -16, -10, -4, 2, 8, 14, 20, 26, \ldots\} & = \{2 + 6n : n \in \mathbb{Z}\}. \\
[3] & = & \{\ldots, -15, -9, -3, 3, 9, 15, 21, 27, \ldots\} & = \{3 + 6n : n \in \mathbb{Z}\}. \\
[4] & = & \{\ldots, -14, -8, -2, 4, 10, 16, 22, 28, \ldots\} & = \{4 + 6n : n \in \mathbb{Z}\}. \\
[5] & = & \{\ldots, -13, -7, -1, 5, 11, 17, 23, 29, \ldots\} & = \{5 + 6n : n \in \mathbb{Z}\}. \\
[6] & = & \{\ldots, -12, -6, 0, 6, 12, 18, 20, 24, \ldots\} & = \{6 + 6n : n \in \mathbb{Z}\}. \\
[7] & = & \{\ldots, -11, -5, 1, 7, 13, 19, 25, 31, \ldots\} & = \{7 + 6n : n \in \mathbb{Z}\}.
\end{array}$$

You may write some more equivalence classes, for example, $[-2]$, $[-1]$, $[8]$, $[13]$ etc. Note that $[1] = [7] = [-5]$.

Now, choose any equivalence class C, and an arbitrary element say $d \in C$. Write the equivalence class of d. Observe that $[d] = C$.

Next, select any two equivalence classes randomly, and observe that either they are same or they are disjoint.

There are only 6 distinct equivalence classes, namely [0], [1], [2], [3], [4] and [5]. Why? In view of the division algorithm, if $n \in \mathbb{Z}$, then we have $n = 6q + r$ for some $q \in \mathbb{Z}$ and $0 \leq r < 6$. We then have $[n] = [r]$.

Thus, a transversal for this equivalence relation is $S = \{0, 1, 2, 3, 4, 5\}$.

Further, we claim that $\mathbb{Z} = \bigcup_{r \in S}[r]$. Since each equivalence class is a subset of $\mathbb{Z}$, the union in the right hand side is a subset of $\mathbb{Z}$. On the other hand, if $n \in \mathbb{Z}$, then $[n] = [r]$ for some $r \in S$, and we have $n \in [r]$. This proves the claim. □

Exercise 4.3.10. In Example 4.3.8, show that there are k distinct equivalence classes, namely, $[0], [1], \ldots, [k-1]$. In particular, a transversal for this relation is $S = \{0, 1, \ldots, k-1\}$. Further, show that $\mathbb{Z} = \bigcup_{r \in S}[r]$.

Example 4.3.11. On $\mathbb{R}^2$, define $(x_1, y_1) \sim (x_2, y_2)$, iff $x_1^2 + y_1^2 = x_2^2 + y_2^2$.

Verify that $\sim$ is an equivalence relation. Let us find some equivalence classes. What is the equivalence class of the origin $(0, 0)$? Note that if $(x, y) \in [(0, 0)]$, then $x^2 + y^2 = 0^2 + 0^2 = 0$. This happens only when $x = y = 0$. Thus, $[(0, 0)] = \{(0, 0)\}$, a singleton set.

What is the equivalence class of the point $(3, 4)$? Suppose, $(x, y) \in [(3, 4)]$, then $x^2 + y^2 = 3^2 + 4^2 = 25 = 5^2$. This means, (x, y) lies on the circle centered at the origin and with radius 5. Clearly any point on the circle is related to $(3, 4)$. Thus

$$[(3, 4)] = \{(x, y) : x^2 + y^2 = 25\}.$$

In general, the equivalence class of a point $(a, b) \in \mathbb{R}^2$ is given by

$$[(a, b)] = \{(x, y) : x^2 + y^2 = a^2 + b^2\},$$

the circle centered at the origin with radius $\sqrt{a^2 + b^2}$.

Thus, the equivalence classes are concentric circles centered at the origin in $\mathbb{R}^2$. Again, notice that any two equivalence classes (any two concentric circles) are either distinct or identical.

For transversal, we may choose, a representative from distinct concentric circles. For, example, $S = \{(0, r) : r \geq 0\}$ is a transversal. You may choose any ray from the origin, fix a nonzero point $(a, b) \in \mathbb{R}^2$ and obtain a transversal $\{(ra, rb) \in \mathbb{R}^2 : r \geq 0\}$. We leave it as an exercise to see that $\mathbb{R}^2 = \bigcup_{r \geq 0}[(0, r)] = \bigcup_{r \geq 0}[(ra, rb)]$.

Definition 4.3.12. Let $f : X \to Y$ be a map. Define a relation $\sim$ on X by $x_1 \sim x_2$ iff $f(x_1) = f(x_2)$. Then $\sim$ is an equivalence relation on X. (Verify.) The relation $\sim$ on X is said to be the *relation induced* by f.

For example, the relation in Example 4.3.11 is induced by the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = x^2 + y^2$.

Exercise 4.3.13. Identify the equivalence classes and a transversal for the relation on $\mathbb{R}$ induced by the function $f(x) = x^2$.

Exercise 4.3.14. Consider $f, g\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = x + y$ and $g(x, y) = |x| + |y|$. Describe the equivalence classes geometrically and think of some "natural" transversals for the relations induced by these functions.

Exercise 4.3.15. Let R and S be equivalence relations on X and Y, respectively, and T be the relation defined on $X \times Y$ as in Ex. 4.2.7. Is T an equivalence relation? If so, can you identify the equivalence classes for T in terms of the equivalence classes for R and S?

Example 4.3.16. Let R and S be equivalence relations on X and Y, respectively, and T be the lexicographic relation on $X \times Y$ defined in Example 4.2.8. The relation T need not be transitive. For example, consider R to be the universal relation on $X = \{a, b\}$ and S to be the equality relation on $Y = \{c, d\}$. Then $(a, c)T(b, d)$ and $(b, d)T(a, d)$ holds but $(a, c)T(a, d)$ does not!

Thus, the lexicographic relation of two equivalence relations is not an equivalence relation, in general.

Exercise 4.3.17. Let $M(n, \mathbb{R})$ denote the set of all $n \times n$ matrices over $\mathbb{R}$. Define the relation $\sim$ on $M(n, \mathbb{R})$ as follows: for $A, B \in M(n, \mathbb{R})$, $A \sim B$ iff there exists an invertible matrix P such that $P^{-1}AP = B$. Show that $\sim$ is an equivalence relation.

4.4 Equivalence classes and partitions of a set

In the previous examples, we noticed that if we take two equivalence classes, then either they are identical or they are disjoint. We will see now that this is true for any equivalence relation.

Lemma 4.4.1. *Let X be a nonempty set and $\sim$ be an equivalence relation on X. If $y \in [x]$, then $[x] = [y]$.*

Proof. Let $y \in [x]$. Then, by definition $x \sim y$. Suppose $z \in [x]$. Then $x \sim z$. Since $x \sim y$, by symmetry and transitivity, we have $y \sim z$. This implies $z \in [y]$. This proves that $[x] \subseteq [y]$.

Next, let $w \in [y]$. Then $y \sim w$. Since $x \sim y$, by transitivity we have $x \sim w$. This implies, $w \in [x]$. Hence $[y] \subseteq [x]$. This proves $[x] = [y]$. □

Theorem 4.4.2. *Let X be a nonempty set and $\sim$ an equivalence relation on X. Let $x, y \in X$. Then exactly one of the following is true.*
(i) $[x] \cap [y] = \emptyset$.
(ii) $[x] = [y]$.

Proof. If (i) holds, then it is clear that (ii) is not true.

Suppose (i) is not true. We show that (ii) must hold. Let $z \in [x] \cap [y]$. Then by the last lemma, we have $[z] = [x]$ and $[z] = [y]$. Hence $[x] = [y]$. □

Theorem 4.4.2 says that any two equivalence classes are either identical or disjoint.

Definition 4.4.3. A *partition* of a nonempty set X is a pairwise disjoint collection of subsets of X whose union is X.

In other words, a family of subsets $\{A_\alpha \subseteq X : \alpha \in \Lambda\}$ of X is a partition of X if (i) $A_\alpha \cap A_\beta = \emptyset$ for every $\alpha, \beta \in \Lambda,\ \alpha \neq \beta$, and (ii) $X = \bigcup_{\alpha\in\Lambda} A_\alpha$. (See Figure 4.1.)

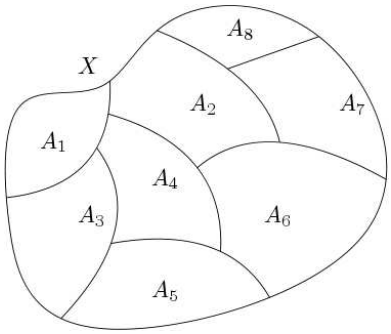


Figure 4.1: Partition of a set

Suppose X is a nonempty set and $\sim$ an equivalence relation on X. Suppose S is a transversal for the relation. For $x \in X$ we have $x \in [s]$ for some $s \in S$. Therefore, $X = \cup_{s\in S}[s]$. Further, if $s, t \in S$ and $s \neq t$, then from Theorem 4.4.2 we get $[s]\cap[t] = \emptyset$. Therefore, the collection $\{[s] \subseteq X : s \in S\}$ of the equivalence classes for $\sim$ is a partition of the set X.

Let $\{A_\alpha : \alpha \in \Lambda\}$ be a partition of X. Define $\sim$ on X as follows:

$$\text{for } x, y \in X, x \sim y \text{ iff } \exists\alpha \in \Lambda(x, y \in A_\alpha),$$

that is, $x \sim y$ iff x and y lie in the same set of the partition.

Convince yourself that $\sim$ is an equivalence relation on X. Let $x \in X$. Then there exists $\alpha \in \Lambda$ such that $x \in A_\alpha$. You can easily verify that $[x] = A_\alpha$. Thus, we see that any partition $\{X_\alpha : \alpha \in \Lambda\}$ of X induces an equivalence relation on X for which the equivalence classes are the sets in the partition.

Let us summarise the discussion in the form of a theorem.

Theorem 4.4.4. *(a) The equivalence classes of an equivalence relation on a set X induces a partition of X.*

(b) Given a partition $\{A_\alpha : \alpha \in \Lambda\}$ of a set X, the relation "to be in the same subset A_α" is an equivalence relation. Further, the equivalence classes for this relation are the subsets A_α. □

Example 4.4.5. Let X be the of $\mathbb{R}^2 \setminus$ the coordinate axes. Then the four different quadrants give rise to a partition of X. Therefore, the quadrants are the equivalence classes of an equivalence relation on X. How do you describe the relation? Note that for $(x, y) \in X$, x, y are nonzero. When do (x_1, y_1) and (x_2, y_2) lie in the same quadrant? It is so provided $x_1x_2 > 0$ and $y_1y_2 > 0$. Therefore, the relation $\sim$ is given by

$$(x_1, y_1) \sim (x_2, y_2) \text{ if and only if } x_1x_2 > 0 \text{ and } y_1y_2 > 0.$$

Let us look at a few more examples of equivalence relations.

Exercise 4.4.6. Show that the following relations are equivalence relations. Identify the equivalence classes and provide suitable transversals.

(a) On $\mathbb{R}$, $x \sim y$, if $x - y$ is an integer.

(b) On $\mathbb{R}$, $x \sim y$ if $\lfloor x \rfloor = \lfloor y \rfloor$ where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.

(c) On $\mathbb{R}^2$, $(x_1, y_1) \sim (x_2, y_2)$, if $x_1 = x_2$.

(d) On $\mathbb{R}^2 \setminus \{(0,0)\}$, $(x_1, y_1) \sim (x_2, y_2)$, if there exists $\alpha \in \mathbb{R}$, $\alpha \neq 0$, such that $x_1 = \alpha x_2$ and $y_1 = \alpha y_2$.

The next two examples and the subsection following them are very demanding. They are not used elsewhere in the text. They may be omitted on first reading. You will still enjoy reading the rest of the book without any lacuna. However, if you are a serious student of mathematics, we urge you to come back to these later and learn them well.

Example 4.4.7. Let $X = \mathbb{N} \times \mathbb{N}$. Define

$$(m, n) \sim (p, q) \text{ if and only if } m + q = p + n.$$

That the relation is reflexive and symmetric, is evident. For transitivity, assume $(m, n) \sim (p, q)$ and $(p, q) \sim (r, s)$. Then $m + q = p + n$ and $p + s = r + q$, which give $(m + q) + (p + s) = (p + n) + (r + q)$, that is, $m + s = r + n$. Hence $(m, n) \sim (r, s)$. Thus $\sim$ is an equivalence relation on X.

What are the equivalence classes for this relation? Let us look at the set X as a subset in the xy-plane. The elements of X are the points with positive integral coordinates. See Figure 4.4.7.

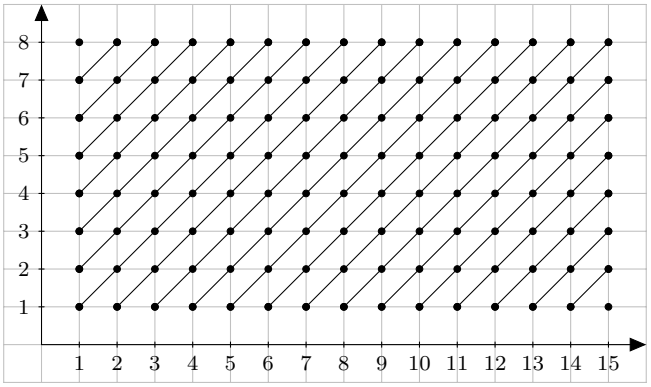


Figure 4.2: Equivalence classes of Example 4.4.7

Let us understand what the relation means. One way to see this is as follows: $(m, n) \sim (p, q)$ holds if and only if $m - n = p - q$ as integers. Thus, points

(m, n) are those lying on the line $x - y = p - q$. However, let us avoid integers at the moment.

Suppose $(m, n) \sim (p, q)$, that is, $m + q = p + n$. If $m = p$ then $n = q$, that is, $(m, n) = (p, q)$. If $m > p$, then $m = p + k$ for some $k \in \mathbb{N}$. Since $m + q = p + n$, it follows that $n = q + k$. That is, $(m, n) = (p + k, q + k)$ for some $k \in \mathbb{N}$. Finally, if $m < p$, then $p = m + k$ for some $k \in \mathbb{N}$. We then get $q = n + k$, that is, $(m + k, n + k) = (p, q)$ for some $k \in \mathbb{N}$. Thus, $(m, n) \sim (p, q)$ if and only if either $(m, n) = (p, q)$ or $(m, n) = (p + k, q + k)$ for some $k \in \mathbb{N}$ or $(m + k, n + k) = (p, q)$ for some $k \in \mathbb{N}$. Thus, the points (m, n) are those which lie on the line passing through (p, q) and making an angle 45° with the x-axis.

What is the equivalence class of the point $(1, 1)$? From the above discussion, we have

$$[(1, 1)] = \{(1, 1), (2, 2), (3, 3), (4, 4), \ldots\} = \{(m, m) : m \in \mathbb{N}\}.$$

Similarly,

$$[(4, 3)] = \{(2, 1), (3, 2), (4, 3), (5, 4), \ldots\} = \{(m + 1, m) : m \in \mathbb{N}\} = [(2, 1)].$$

In general, what is the equivalence class of a point (p, q) of $\mathbb{N} \times \mathbb{N}$? Look at the figure. What are the points (m, n) on the line passing through (p, q) and making an angle 45° with the x-axis? Assume that $p \geq q$. If $(m, n) \in [(p, q)]$, then we have $m = n + (p - q) \in \mathbb{N}$. Thus, for any $n \in \mathbb{N}$ we see that $(n + (p - q), n) \in [(p, q)]$. Therefore, for $p \geq q$, we conclude that

$$[(p, q)] = \{(n + (p - q), n) : n \in \mathbb{N}\} = [(1 + (p - q), 1)].$$

Note that, in the case when $p \geq q$, the equivalence class $[(p, q)]$ is of the form $[(n, 1)]$, where $n = 1 + (p - q) \in \mathbb{N}$.

On the other hand, if $p < q$, proceeding as above, we see that

$$[(p, q)] = \{(m, m + (q - p)) : n \in \mathbb{N}\} = [(1, 1 + (q - p))].$$

Note that in this case the equivalence class $[(p, q)]$ for (p, q) is of the form $[(1, n)]$, where $n = 1 + (q - p) > 1$.

Can you now identify the above two types of equivalence classes in Figure 4.2? You can easily see now that a transversal for this relation is given by

$$S = \{(n, 1) : n \in \mathbb{N}\} \cup \{(1, n) : n \in \mathbb{N}, n > 1\}.$$

Example 4.4.8. Let $X := \mathbb{Z} \times \mathbb{Z}^*$, where $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$. Define $\sim$ on X by $(p, q) \sim (r, s)$ iff $ps = rq$. As in Example 4.4.7, you can easily show that $\sim$ is an equivalence relation. We want to find the equivalence classes and a suitable transversal.

Note that for any $(p, q) \in X$ and $k \in \mathbb{Z}^*$, we have $(p, q) \sim (pk, qk)$. Let $(r, s) \in X$. If $d = \gcd(r, s)$, $r = pd$ and $s = qd$, then we have $\gcd(p, q) = 1$ and $(p, q) \sim (r, s)$. This shows that each element in X lies in the equivalence class of some (p, q) with $\gcd(p, q) = 1$.

Suppose $(p, q) \in X$ such that $\gcd(p, q) = 1$. What is the equivalence class of (p, q)? Let $(r, s) \sim (p, q)$, that is $rq = ps$. Then ps is divisible by q. This

implies that s is divisible by q, since $\gcd(p, q) = 1$. Since $s \neq 0$, $s = kq$ for some $k \in \mathbb{Z}^*$. Further, it follows from $rq = ps$ that $r = kp$, that is, $(r, s) = (kp, kq)$. We can conclude that the equivalence class of (p, q) is

$$[(p, q)] = \{(kp, kq) \in X : k \in \mathbb{Z}^*\}.$$

Now, consider the set

$$S = \{(p, q) : p, q \in \mathbb{Z}, q \neq 0, \gcd(p, q) = 1\}.$$

Does S form a transversal for the relation $\sim$ on X? In view of the above discussion, this will be the case, provided that the distinct elements in S lie in distinct equivalent classes.

Suppose $(p, q), (r, s) \in S$ are such that $(p, q) \sim (r, s)$. We show that $(p, q) = (r, s)$. Since $(r, s) \in [(p, q)]$, we have $(r, s) = (kp, kq)$ for some $k \in \mathbb{Z}^*$. Similarly, $(p, q) = (k'r, k's)$ for some $k' \in \mathbb{Z}^*$. Consequently, $q = k's = k'kq$, that is, $k'k = 1$, and therefore, $k = 1$ or -1. However, $q > 0, s > 0$ implies $k = 1$. This yields $(p, q) = (r, s)$. Thus, S forms a transversal for the relation $\sim$ on X.

Significance of equivalence relation

We have proved that any equivalence relation $\sim$ on a nonempty set X gives rise to a partition of X into equivalence classes. We can identify each of these equivalence classes as one entity and it is represented by one of its elements in a given transversal for $\sim$.

Definition 4.4.9. Let X be a nonempty set and $\sim$ an equivalence relation on X. The set of equivalence classes of $\sim$ is called the *quotient set* or the *cosets modulo* $\sim$, and is denoted by $X/\sim$.

Example 4.4.10. (a) The quotient set in Example 4.3.8 is given by

$$\mathbb{Z}/\sim \; = \{[0], [1], \ldots, [k-1]\}.$$

(In abstract algebra, the quotient set is denoted by $\mathbb{Z}_k$ and it belongs to important classes of groups and rings.)

(b) The quotient set in Example 4.3.11 is given by

$$\mathbb{R}^2/\sim \; = \{C_r : r \geq 0\},$$

where $C_r = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = r\}$. We can set up a natural bijection between the quotient set $\mathbb{R}^2/\sim$ and the set $\mathbb{R}^+$ of nonnegative reals. In other words, the quotient set can be identified with $\mathbb{R}^+$.

Example 4.4.11. Consider the equivalence relation on $\mathbb{N} \times \mathbb{N}$ defined by $(m, n) \sim (p, q)$ iff $m + q = p + n$. (See Example 4.4.7.) The set

$$S = \{(n, 1) : n \in \mathbb{N}\} \cup \{(1, n) : n \in \mathbb{N}, n > 1\}$$

is a transversal for this relation. Therefore, the quotient set of $\sim$ is given by

$$\mathbb{N} \times \mathbb{N}/\sim \; = \{[(a, b)] : (a, b) \in S\}.$$

Define a map $f\colon S \to \mathbb{Z}$ by $f(a,b) = a - b$. Note that

$$f\Big(\big\{(n,1) : n \in \mathbb{N}\big\}\Big) = \{m \in \mathbb{Z} : m \geq 0\}, \text{ and}$$
$$f\Big(\big\{(1,n) : n \in \mathbb{N}, n > 1\big\}\Big) = \{m \in \mathbb{Z} : m < 0\}.$$

In fact, f is a bijection. (Verify.) Note that we have a bijection g between $\mathbb{N} \times \mathbb{N}/\sim$ and S defined by $g([(a,b)]) = (a,b)$, $(a,b) \in S$. Therefore, we have a bijection $f \circ g$ between $\mathbb{N} \times \mathbb{N}/\sim$ and $\mathbb{Z}$. This shows that the set $\mathbb{Z}$ of integers can be identified with the quotient set $\mathbb{N} \times \mathbb{N}/\sim$.

The example exhibits a way to construct the integers from the natural numbers, as equivalence classes in $\mathbb{N} \times \mathbb{N}$. (Recall that we avoided using integers in our discussion in Example 4.4.7.)

We can view this natural identification of $\mathbb{Z}$ with $\mathbb{N} \times \mathbb{N}/\sim$ another way. Define a map $g\colon \mathbb{N} \times \mathbb{N} \to \mathbb{Z}$ by $g(a,b) = a - b$. Then the relation on $\mathbb{N} \times \mathbb{N}$ induced by the function g coincides with the relation $\sim$. It is evident that (i) g is onto, and (ii) if $m \in \mathbb{Z}$ and $g(a,b) = m$, then $g^{-1}(m) = [(a,b)]$. Therefore, the equivalence classes for $\sim$ correspond to the integers in a bijective manner.

Exercise 4.4.12. Consider the equivalence relation on $X = \mathbb{Z} \times \mathbb{Z}^*$ defined in Example 4.4.8, and the transversal S.

(a) Show that there is a bijection between S and $\mathbb{Q}$, the set of rational numbers.

(b) Deduce that $\mathbb{Q}$ can be identified with the quotient set $\mathbb{Z} \times \mathbb{Z}^*/\sim$.

(c) Define $g\colon \mathbb{Z} \times \mathbb{Z}^* \to \mathbb{Q}$ by $g(m,n) = \frac{m}{n}$. Show that

(i) $\sim$ is induced by g.

(ii) g is onto.

(iii) if $x \in \mathbb{Q}$ and $g(a,b) = x$, then $g^{-1}(x) = [(a,b)]$.

(The example exhibits a way to construct the rational numbers from the integers, as equivalence classes in $\mathbb{Z} \times \mathbb{Z}^*$.)

Exercise 4.4.13. For each of the relations given in Exercise 4.4.6 do the following.

(i) Write the quotient set.

(ii) Find a suitable set that can be identified with the quotient set.

Chapter 5

Induction Principles

In this chapter, we deal with three principles widely used in mathematics, namely, the induction principle, the strong induction principle and the well-ordering principle. These principles essentially state properties enjoyed by the set $\mathbb{N}$ of positive integers. We give a number of examples to illustrate applications of each of these principles. At the end we show that $\mathbb{N}$ has one of these properties if and only if it enjoys the other two.

5.1 The Induction Principle

The Induction Principle (Form 1)

Let $A \subseteq \mathbb{N}$. Assume that (i) $1 \in A$, and (ii) for $k \geq 1$, $k \in A$ implies that $k+1 \in A$. Then $A = \mathbb{N}$.

The induction principle can be restated in the following form which is more familiar to the students.

The Induction Principle (Form 2)

Suppose for each $n \in \mathbb{N}$, a statement $P(n)$ is given. Assume that (i) $P(1)$ is true, and (ii) for $k \geq 1$, $P(k)$ is true implies $P(k+1)$ is true. Then, $P(n)$ is true for each $n \in \mathbb{N}$.

How are the above two formulations of the induction principle same?

Suppose $A = \{n \in \mathbb{N} : P(n) \text{ is true}\}$. Then the conclusion of the Form 2 is equivalent to saying that $A = \mathbb{N}$. Now, (i) "$P(1)$ is true" means "$1 \in A$", and (ii) for $k \geq 1$, "$P(k)$ is true implies that $P(k+1)$ is true" is same as saying "$k \in A$ implies $k+1 \in A$".

Example 5.1.1. For any positive integer n, we show that

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}. \tag{5.1}$$

Let $P(n)$ be the statement (5.1) for a given $n \in \mathbb{N}$. Then the claim here is that "$P(n)$ is true for each $n \in \mathbb{N}$".

(i) Observe that $P(1)$ is true, as both sides of (5.1) is 1, if $n = 1$.

(ii) Let $k \geq 1$ and assume that $P(k)$ is true. Thus we have $1+\cdots+k = k(k+1)/2$. We now show that $P(k+1)$ is true. Adding $k+1$ to both sides in the last equation, we get

$$1 + \cdots + k + (k+1) = \frac{k(k+1)}{2} + (k+1).$$

The right side of this equation is

$$\frac{k(k+1) + 2(k+1)}{2} = \frac{(k+1)(k+2)}{2}.$$

This shows that $P(k+1)$ is true.

Hence, by the induction principle, (5.1) holds for all $n \in \mathbb{N}$. □

Remark 5.1.2. While proving "the statement $P(n)$ is true for each $n \in \mathbb{N}$" by induction, one first proves that $P(1)$ is true. This is called the *base case*. Then one assumes that $P(k)$ is true for some $k \geq 1$. This assumption is called the *induction hypothesis*. Proving that $P(k+1)$ is true using the induction hypothesis is sometimes called the *inductive leap*.

The next two exercises are similar to Example 5.1.1 and we encourage the reader to write a proof with as much details as in Example 5.1.1.

Exercise 5.1.3. For any positive integer n, we have

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}. \tag{5.2}$$

Exercise 5.1.4. For any positive integer n, we have

$$\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}. \tag{5.3}$$

Example 5.1.5 (Bernoulli's Inequality)**.** Let $x \in \mathbb{R}$ such that $x > -1$. Then for each $n \in \mathbb{N}$

$$(1+x)^n \geq 1 + nx. \tag{5.4}$$

For $n \in \mathbb{N}$, let $P(n)$ be the statement (5.4).

It is clear that $P(1)$ is true.

Assume that $k \geq 1$ and $P(k)$ is true. That is, we have $(1+x)^k \geq 1 + kx$. Since $1 + x > 0$, multiplying both sides of the inequality by $1 + x$, we get

$$\begin{aligned}(1+x)^{k+1} = (1+x)(1+x)^k &\geq (1+x)(1+kx) \\ &\geq 1 + (k+1)x + kx^2 \\ &\geq 1 + (k+1)x.\end{aligned}$$

This shows that $P(k+1)$ is true.

Hence, by the induction principle, (5.4) holds true for all $n \in \mathbb{N}$. □

Recall that two integers a and b are said to be *relatively prime* if the only positive integral divisor of a and b is 1, that is, if $\gcd(a, b) = 1$ (see Definition 5.3.4).

Two integers a and b are relatively prime if and only if there exist integers x and y such that $ax + by = 1$. (See Corollary 5.3.6).

Example 5.1.6. Let $a_1, a_2, \ldots, a_n$ be positive integers. Assume that each a_i is relatively prime to an integer b. Then, the product $a_1 a_2 \cdots a_n$ is relatively prime to b.

What is $P(n)$ here? We let $P(n)$ to be the statement that if we are given n integers each of which is relatively prime to b, then so is their product.

For $n = 1$ the result is true. This follows by observing that product of a single integer is the integer itself.

Assume that $k \geq 1$ and $P(k)$ is true. Suppose we are given positive integers $a_1, a_2, \ldots, a_{k+1}$, each of which is relatively prime to b. We want to show that $a_0 = a_1 a_2 \cdots a_{k+1}$ is relatively prime to b.

Suppose $a = a_1 a_2 \cdots a_k$. Since each of the positive integers $a_1, \ldots, a_k$ is relatively prime to b, by the induction hypothesis, a is relatively prime to b. Therefore, we have positive integers a and a_{k+1} each of which is relatively prime to b. We therefore need to show that $a_0 = a a_{k+1}$ is relatively prime to b. But, this amounts to showing that $P(2)$ is a true statement.

Let us therefore prove that $P(2)$ is true. Let a_1 and a_2 be relatively prime to b. Then there exist integers x_j, y_j, $j = 1, 2$ such that $a_1 x_1 + b y_1 = 1$ and $a_2 x_2 + b y_2 = 1$. Multiplying these equations, we get

$$a_1 a_2 (x_1 x_2) + b(a_1 x_1 y_2 + a_2 x_2 y_1) = 1.$$

This proves that $a_1 a_2$ is relatively prime to b, that is, $P(2)$ is true.

> Another way to prove P_2 is true: Since a_1 and b are relatively prime, there exist integers x, y such that $a_1 x + by = 1$. Multiply both sides of this equation by a_2 to obtain $a_1 a_2 x + a_2 by = a_2$. If d is a positive common divisor of $a_1 a_2$ and b, then d divides the LHS of the equation $a_1 a_2 x + a_2 by = a_2$ and hence it divides a_2. Since it already divides b, we conclude that it is a common divisor of a_2 and b. As a_2 and b are relatively prime, we conclude that $d = 1$. That is, any common divisor of $a_1 a_2$ and b is 1 and hence they are relatively prime.

From the above discussion we have for $k \geq 1$, $P(k)$ is true implies $P(k + 1)$ is true.

Hence by the induction principle, $P(n)$ holds true for all $n \in \mathbb{N}$. □

Remark 5.1.7. The crucial step in the proof of Example 5.1.6 is to show independently that $P(2)$ holds. Proving independently that "$P(2)$ holds" is not an essential part of the induction principle. However, it is proved and used to get the induction leap from the induction hypothesis on many occasions.

In an induction proof, it is crucial to verify the result for the base case. Look at the following exercise, in which base case is not true, however we can prove $P(n + 1)$, if we assume $P(n)$.

Exercise 5.1.8. Consider a statement $P(n)$ as $2+4+\cdots+2n = (n+2)(n-1)$. Show that if $P(k)$ is true then $P(k+1)$ is also true. However, the base case $P(1)$ fails.

Note carefully the condition "(ii) for $k \geq 1$, $k \in A$ implies that $k+1 \in A$" in the induction principle. For the induction hypothesis, the case $k = 1$ is included, because it is essential. In fact, this is the base case from where you start making inductive leap.

Exercise 5.1.9. Find the fallacious step in the following argument. We prove by induction that any n things are the same. If $n = 1$ this is clear. Assume that any k things are the same. Let $a_1, a_2, \ldots, a_k, a_{k+1}$ be given. By induction hypothesis (that is another way of saying $P(k)$ is true) applied to the k things $a_1, \ldots, a_k$ we find that $a_1 = a_2 = \cdots = a_k$. Similarly, we conclude that $a_2 = a_3 = \cdots = a_k = a_{k+1}$. Hence $a_1 = a_2 = \cdots = a_k = a_{k+1}$. Hence by induction principle, given any n things, they are always the same.

Did we consider the possibility $k = 1$ in the argument?

Exercise 5.1.10. Derive the formula for the sum of first n terms of an arithmetic progression:

$$a + (a+d) + (a+2d) + \cdots + (a+(n-1)d) = \frac{n(2a+(n-1)d)}{2}.$$

Exercise 5.1.11. Derive the formula for the sum of first n terms of a geometric progression: $\displaystyle\sum_{k=1}^{n-1} ar^k = \frac{a(r^{n-1})}{r-1}, r \neq 1.$

Exercise 5.1.12. Using induction principle, prove that for each $n \in \mathbb{N}$

(i) $(n+1)^2 + (n+2)^2 + \cdots + (2n)^2 = \dfrac{n(2n+1)(7n+1)}{6}$.

(ii) $1 \cdot n + 2(n-1) + 3(n-2) + \cdots + n \cdot 1 = \dfrac{n(n+1)(n+2)}{6}$.

(iii) $n(n+1)(n+2)(n+3)$ is divisible by 24.

(iv) $n^3 + 2n$ is divisible by 3.

(v) $3^n > n^2$. *Hint.* $3^k > k^2$ implies $3^{k+1} > 3 \cdot k^2 > k^2 + 2k + 1 = (k+1)^2$ for $k \geq 2$. So, $P(2)$ needs to be proved independently.

(vi) for any real number t, $(\cos t + i \sin t)^n = \cos nt + i \sin nt$. (Here $i^2 = -1$.)

Exercise 5.1.13. Use induction principle to prove that for any $n \in \mathbb{Z}, r \in \mathbb{N}$ $n(n+1)(n+2)\cdots(n+r-1)$ is divisible by $r! := 1 \cdot 2 \cdots r$. *Hint:* Fix $n \in \mathbb{Z}$ and use induction on r.

Exercise 5.1.14. Let k be a fixed positive integer. Recall the relation "congruence modulo k" on $\mathbb{Z}$ defined in Example 4.3.8, that is, $a \equiv b \pmod{k}$ iff $a - b$ is a multiple of k.

(i) Prove that if $a \equiv b \pmod{k}$ and $c \equiv d \pmod{k}$, then $ac \equiv bd \pmod{k}$.

(ii) Use induction to deduce that if $a \equiv b \pmod{k}$, then $a^n \equiv b^n \pmod{k}$, for each $n \in \mathbb{N}$.

Exercise 5.1.15. Show that $4^n \equiv 3n + 1 (\text{mod } 9)$ for each $n \in \mathbb{N}$.

Suppose N is a fixed integer (possibly negative), and a statement $P(n)$ is given for each integer $n \geq N$. Suppose that the following hold:

(i) The statement $P(N)$ is true, and

(ii) For integer $k \geq N$, $P(k)$ is true implies $P(k + 1)$ is true.

What can you conclude? The conditions imply that $P(n)$ holds for each integer $n \geq N$.

To see this, first note that if n is an integer with $n \geq N$, then n equals $(N - 1) + m$ for some $m \in \mathbb{N}$. Therefore, consider the set $A = \{m \in \mathbb{N} : P((N - 1) + m) \text{ is true}\}$. Note that our conclusion holds if and only if $A = \mathbb{N}$.

Now, the given conditions mean that (i) $1 \in A$, and (ii) for $r \geq 1$, $r \in A$ implies $r + 1 \in A$. Therefore, by the induction principle, $A = \mathbb{N}$.

Thus we have another formulation of the induction principle which is as follows.

The Induction Principle (arbitrary base case)

Let $A \subseteq \mathbb{Z}$ and $N \in \mathbb{Z}$. Assume that (i) $N \in A$, and (ii) for $k \geq N$, $k \in A$ implies $k + 1 \in A$. Then $\{n \in \mathbb{Z} : n \geq N\} \subseteq A$.

With this formulation, $n = N$ is the base case. Note that with $N = 1$ you get Form 1 of the principle.

Exercise 5.1.16. Prove that $n! > 2^n$ for all positive integers $n \geq 4$. (The base case here is 4.)

5.2 The Strong Induction Principle

In proofs by induction, while establishing the inductive leap $P(k+1)$, so far we used $P(k)$ as the induction hypothesis. However, in many cases we may need to use $P(r)$ for some/all $1 \leq r \leq k$. That is, we need more than one predecessor at times. This leads to what is known as the strong induction principle, which we deal with in this section.

The Strong Induction Principle

Let $B \subseteq \mathbb{N}$. Assume that (i) $1 \in B$ and (ii) $\{1, 2, \ldots, k\} \subseteq B$ implies $k + 1 \in B$. Then $B = \mathbb{N}$.

Remark 5.2.1. Compare the induction and the strong induction principles. The strong induction principle has strong or more stringent hypothesis than the "standard" induction principle, but the conclusions in both the principles are the same.

Example 5.2.2. Assume that $x_1 = 1, x_2 = 2, x_3 = 3$. For $n \in \mathbb{N}, n \geq 4$, define $x_n = x_{n-1} + x_{n-2} + x_{n-3}$. Show that $x_n < 2^n$ for each $n \in \mathbb{N}$.

Clearly, $x_n < 2^n$ for $n = 1, 2, 3$. Let $k \geq 3$ and assume that the result is true for $1 \leq r \leq k$. That is, $x_r < 2^r$ for $1 \leq r \leq k$. We need to show that $x_{k+1} < 2^{k+1}$.

We have $x_{k+1} = x_k + x_{k-1} + x_{x-2}$. By the assumption, we have $x_k < 2^k, x_{k-1} < 2^{k-1}$ and $x_{k-2} < 2^{k-2}$, hence

$$\begin{aligned} x_{k+1} = x_k + x_{k-1} + x_{x-2} &< 2^k + 2^{k-1} + 2^{k-2} \\ &= 2^{k-2}(2^2 + 2 + 1) \;\; = \;\; 2^{k-2} \times 7 \\ &< 2^{k-2} \cdot 2^3 = 2^{k+1}. \end{aligned}$$

Thus, by the strong induction principle, we have $x_n < 2^n$ for each $n \in \mathbb{N}$. □

Example 5.2.3. For $n \in \mathbb{N}$ define a_n as follows: $a_1 = 1, a_2 = 8$ and $a_n = a_{n-1} + 2a_{n-2}$ for $n \geq 3$. Prove that for each $n \in \mathbb{N}$, $a_n = 3 \cdot 2^{n-1} + 2(-1)^n$.

For $n \in \mathbb{N}$, let $P(n)$ be the statement $a_n = 3 \cdot 2^{n-1} + 2(-1)^n$.

Verify that $P(1)$ and $P(2)$ are true. Assume that $k \geq 2$ and that $P(r)$ is true for $r = 1, 2, \ldots, k$. Then

$$\begin{aligned} a_{k+1} &= a_k + 2a_{k-1} \\ &= (3 \cdot 2^{k-1} + 2(-1)^k) + 2(3 \cdot 2^{k-2} + 2(-1)^{k-1}) \\ &= 3(2^{k-1} + 2^{k-1}) + 2((-1)^k + 2(-1)^{k-1}) \\ &= 3 \cdot 2 \cdot 2^{k-1} + 2(-1)^{k-1}(-1 + 2) \\ &= 3 \cdot 2^k + 2(-1)^{k+1}, \end{aligned}$$

since $(-1)^{k-1} = (-1)^{k+1}$. This shows that $P(k+1)$ is true. Hence, by the strong induction principle, $P(n)$ is true for each $n \in \mathbb{N}$. □

Exercise 5.2.4. Fibonacci numbers. We define a sequence of numbers as follows: $f_1 = 1$, $f_2 = 1$ and $f_n = f_{n-2} + f_{n-1}$ for all $n \geq 3$. The number f_n is called the n-th Fibonacci number.

Prove that, for each $n \in \mathbb{N}$, the following hold:

(i) $f_2 + f_4 + \cdots + f_{2n} = f_{2n+1} - 1$.
(ii) $f_1 + f_3 + \cdots + f_{2n-1} = f_{2n}$.
(iii) $f_1 + f_2 + \cdots + f_n = f_{n+2} - 1$.
(iv) $f_n < 2^n$.

Example 5.2.5. Every integer greater than 1 has a prime divisor.

First, we look at the base case $n = 2$. Since 2 is a prime and is divisible by itself, the base case is true.

Suppose $k \geq 2$ and assume that each of the integers $2, 3, \ldots, k$ has a prime divisor. We show that $k + 1$ has a prime divisor.

If $k + 1$ is a prime number then $k + 1$ has a prime divisor, namely itself. Otherwise, there exist integers a and b such that $2 \leq a, b \leq k+1$ and $k+1 = a \cdot b$. By induction hypothesis, a has a prime divisor, say p. Clearly, p is a prime divisor of $k + 1$.

Hence, by the strong induction principle, the result follows. □

Exercise 5.2.6. Use the strong induction principle to prove that every integer $n \geq 2$ is a product of primes. (In fact, up to an order, one can write n exactly one way as a product of primes. This result is known as the *Fundamental Theorem of Arithmetic*.)

Exercise 5.2.7. (a) Prove that every amount of postage that is at least 12 rupees can be made from 4-rupee and 5-rupee stamps.
(b) Formulate the problem replacing 4, 5 and 12 by 5, 7 and some suitable base case and solve it.
(c) Can you formulate the problem for any given pair a and b of positive integers which are relatively prime and solve it?
Hint: (a) The answer amounts to showing that for $m \geq 12$, the equation $4x + 5y = m$ has a solution in nonnegative integers. Prove directly for $12 \leq m \leq 15$, and then for $k \geq 15$ show that $4x + 5y = k + 1$ has a solution in nonnegative integers assuming solutions for $12, 13, \ldots, k$. For (b) and (c) Corollary 5.3.6 may be helpful.

Exercise 5.2.8. If A is the matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, then show that $A^n = \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix}$, for all $n \in \mathbb{N}$. Here f_n stands for the n-th Fibonacci number.

5.3 The Well-ordering Principle

In this section, we introduce the well-ordering principle which is used frequently in mathematics, especially in number theory.

Definition 5.3.1. Suppose $A \subseteq \mathbb{N}$. Then a natural number a is called the *least element* of A if (i) $a \in A$ and (ii) $\forall x \in A$, we have $a \leq x$.

For the set of even natural numbers, 2 is the least element and for the set of odd natural numbers, 1 is the least element. Suppose $A = \{n \in \mathbb{N} : n^2 \geq 5\}$. Then 3 is the least element of A.

The Well-Ordering Principle

Let $C \subseteq \mathbb{N}$ be nonempty. Then C has a least element.

Remark 5.3.2. The well-ordering principle states that for $C \subseteq \mathbb{N}$, if C is nonempty, then C has a least element. This is equivalent to its contrapositive, that is, "If $C \subseteq \mathbb{N}$ does not have a least element, then C is empty".

Theorem 5.3.3 (Division Algorithm)**.** *Let $a, b \in \mathbb{Z}$ with $a \in \mathbb{N}$. Then there exist unique integers q and r such that (i) $b = aq + r$ and (ii) $0 \leq r < a$.*

Let us understand the statement. It is a rigorous version of what is called the 'long division'. Let $b = 25$ and $a = 4$, then $q = 6$ and $r = 1$. If $b = -25$ and $a = 4$, then $-25 = (-7) \times 4 + 3$ so that $q = -7$ and $r = 3$.

Proof. Consider the set

$$S := \{x \in \mathbb{Z} : \exists k \in \mathbb{Z} \text{ such that } x = b - ak\}.$$

The set S is nonempty, since $b = b - a \times 0 \in S$. Let $S^+ := \{x \in S : x \geq 0\}$. If $b \geq 0$, then $b \in S^+$. If $b < 0$, then we take $k = -b$ so that $b - (-b)a = b(1-a) \geq 0$. Thus S^+ is nonempty.

If $0 \in S^+$, then we have $0 = b - aq$ for some $q \in \mathbb{Z}$. Thus, $b = aq + 0$ as required.

If $0 \notin S^+$, then $\emptyset \neq S^+ \subseteq \mathbb{N}$. Hence by the well-ordering principle, S^+ has a least element, say $r \in S^+$. It follows that $r = b - aq$ for some $q \in \mathbb{Z}$, that is, $b = aq + r$. To complete the proof, we need to show that $r < a$. If $r \geq a$, then $0 \leq r - a = b - qa - a = b - (q+1)a$. We therefore conclude that $r - a \in S^+$ and since $a > 0$, $r - a < r$. This contradicts the fact that r is the least element of S^+. Thus we have proved $0 < r < a$.

Now we shall prove the uniqueness part. What is to be proved? If we have two pairs (q_1, r_1) and (q_2, r_2) satisfying the two conditions, namely (i) $b = aq_1 + r_1$ and $b = aq_2 + r_2$ and (ii) $0 \leq r_1 < a$ and $0 \leq r_2 < a$, then we are required to prove that $q_1 = q_2$ and $r_1 = r_2$.

We prove this by contradiction. Assume, without loss of generality, that $r_1 < r_2$. Then from $b = aq_1 + r_1 = aq_2 + r_2$ we deduce that $a(q_1 - q_2) = r_2 - r_1$. Since $a > 0$ and the RHS is positive we infer that $q_1 - q_2 > 0$, that is, $q_1 - q_2 \in \mathbb{N}$. Hence the LHS, namely $a(q_1 - q_2)$ is a positive integral multiple of a and hence greater than or equal to a: $a(q_1 - q_2) \geq a$. But on the left side we have $r_2 - r_1 < a - r_1 \leq a$ and hence the LHS is strictly less than a. This contradiction proves that $r_1 = r_2$. From $b = aq_1 + r_1 = aq_2 + r_1$, it follows that $a(q_1 - q_2) = 0$. Since $a \neq 0$, we see that $q_1 = q_2$. This proves the uniqueness part of the theorem. □

Definition 5.3.4. Let a, b be integers, not both zero. The *greatest common divisor* (denoted by $\gcd(a, b)$) of a and b is the positive integer d such that
(i) d is a common divisor of a and b, and
(ii) if c is a common divisor of a and b, then c divides d.

In the following theorem, we will see a beautiful and useful result on greatest common divisors of integers.

Theorem 5.3.5. *Let a, b be integers, not both zero, and d be the greatest common divisor of a and b. Then there exist integers x, y such that $d = ax + by$.*

Proof. We use the well-ordering principle to prove the result. Consider the set

$$S := \{k \in \mathbb{N} : \text{there exist } m, n \in \mathbb{Z} \text{ such that } k = am + bn\}.$$

Is S nonempty? Well, take $m = a, n = b$. By the well-ordering principle, S has a least element d. We show that $d = \gcd(m, n)$. Since $d \in S$, we have $d = ax + by$ for some $x, y \in \mathbb{Z}$.

(i) First, we claim that d divides a and b. Let r be the remainder of a when divided by d, that is, $0 \leq r < d$ and $a = dq + r$ for some $q \in \mathbb{Z}$. Note that d divides a iff $r = 0$. Can $r > 0$ be true? In that case $r \in S$, since $r = a - dq = a(1 - x) + b(-y) \in S$. That is, S would have an element, namely r, smaller that d! This cannot be true, since d is the least element of S. Thus, $r = 0$, that is, d divides a. You can similarly argue that d divides b.
(ii) Next, suppose c is a common divisor of a and b. Since $d = ax + by$, c divides

d. (Why?)
This establishes that $d = \gcd(m, n)$. □

From the above result you can easily derive the following.

Corollary 5.3.6. *Two integers m and n are relatively prime if and only if there exist integers a and b such that $am + bn = 1$.* □

We leave it as an exercise. *Hint:* If $am + bn = 1$, can there be a positive integer $d < 1$ such that $am + bn = d$?

Example 5.3.7. Let us apply the well-ordering principle to prove the statement in Example 5.2.5, that is,

Every integer greater than 1 has a prime divisor.

We prove this by contradiction. Assume that there exists an integer $n > 1$ such that n is not divisible by any prime. Consider the set

$$S := \{k \in \mathbb{N} : \ k > 1 \text{ and } k \text{ is not divisible by any prime number}\}.$$

Since $n \in S$, S is a nonempty subset of $\mathbb{N}$. Therefore, by the well-ordering principle, S has a least element, say m.

Since $m \in S$, we have $m > 1$ and m is not a prime. (Why?) Therefore, there exist integers $1 < r, s < m$ such that $m = r \cdot s$. Since $r < m$ we have $r \notin S$. (Why?) Because $r > 1$ and $r \notin S$, there exists a prime p which divides r. This implies p also divides m, because r divides m. That is, m has a prime divisor, a contradiction. □

Exercise 5.3.8. Use the well-ordering principle to prove the assertions in Example 5.1.1, and Exercise 5.1.3.

Exercise 5.3.9. Use the well-ordering principle to prove that every integer $n \geq 2$ is a product of primes. *Hint:* Use contradiction.

5.4 Equivalence of the three principles

In this section, we prove that the three principles discussed earlier are equivalent. What exactly does it mean? If one of the principles holds true for $\mathbb{N}$, then the other two are also true. Note that we have not proved any of these principles. Rather, we assumed them to be true and applied them for proving some other results.

Theorem 5.4.1. *The following are equivalent:*

(1) **The Induction Principle:** *Let $A \subseteq \mathbb{N}$. Assume that (i) $1 \in A$ and that (ii) $k \in A$ implies $k + 1 \in A$. Then $A = \mathbb{N}$.*

(2) **The Strong Induction Principle:** *Let $B \subseteq \mathbb{N}$. Assume that (i) $1 \in B$ and that (ii) $\{1, 2, \ldots, k\} \subseteq B$ implies $k + 1 \in B$. Then $B = \mathbb{N}$.*

(3) **The Well-Ordering Principle:** *Let $C \subseteq \mathbb{N}$ be nonempty. Then C has a least element.*

Proof. (1) $\Rightarrow$ (2): Assume the induction principle to be true.

Let B satisfy the hypothesis of the strong induction principle. We wish to show that $B = \mathbb{N}$. As we have the induction principle at our disposal, we devise a set to which the induction principle can be applied. With this in mind, let us define $A := \{n \in \mathbb{N} : \{1, 2, \ldots, n\} \subseteq B\}$ and show that $A = \mathbb{N}$. Suppose this is done. Then, for $n \in \mathbb{N}$, we have $n \in A$. This would give $\{1, \ldots, n\} \subseteq B$. In particular, we get $n \in B$, and therefore $\mathbb{N} \subseteq B$, yielding $B = \mathbb{N}$ as required.

Now, $1 \in A$, since by hypothesis (i) on B, $\{1\} \subseteq B$. Now let $k \in A$. By definition of A, this means $\{1, 2, \ldots, k\} \subseteq B$. By the hypothesis (ii) on B, it follows that $k + 1 \in B$. Thus $\{1, 2, \ldots, k + 1\} \subseteq B$, that is, $k + 1 \in A$.

Thus, $A \subseteq \mathbb{N}$ is such that (i) $1 \in A$ and (ii) $k \in \mathcal{A}$ implies $k + 1 \in A$. By the induction principle, we have $A = \mathbb{N}$, and we are done.

(2) $\Rightarrow$ (3): Assume the strong induction principle to be true. How do you prove the well-ordering principle? In fact, you can get the idea from Remark 5.3.2. Suppose $C \subseteq \mathbb{N}$ does not have a least element. We should be able to conclude that $C = \emptyset$. This will be the case if $S = \mathbb{N} \setminus C = \mathbb{N}$. Can you see that $S = \mathbb{N}$?

Can the integer 1 be in C? No, otherwise 1 is the least element of C. Therefore, $1 \in S$. Suppose $k \geq 1$ and $\{1, \ldots, k\} \subseteq S$. Then none of the integers $1, 2, \ldots, k$ is in C. Can $k + 1$ be in C? No, otherwise $k + 1$ is the least element of C. Therefore, $k + 1 \in S$. Hence, by the strong induction principle, $S = \mathbb{N}$ and therefore $C = \emptyset$. This proves (2) $\Rightarrow$ (3).

(3) $\Rightarrow$ (1): Assume the well-ordering principle to be true. Let $A \subseteq \mathbb{N}$ satisfy the hypotheses of the induction principle. We wish to show that $A = \mathbb{N}$.

Suppose this is false. Then the set $C := \mathbb{N} \setminus A$ is nonempty. By the well-ordering principle, C has a least element, say ℓ. Since $\ell \in C$, ℓ is not in A.

Now, ℓ cannot be 1, since $1 \in A$. Thus, $\ell \geq 2$, which gives $\ell - 1 \in \mathbb{N}$. Since ℓ is least in C, $\ell - 1 \notin C$, that is, $\ell - 1 \in A$. By the hypothesis (ii) on A, this implies $\ell \in A$, that is, $\ell \notin C$. This is impossible, and therefore $A = \mathbb{N}$. $\square$

Remark 5.4.2. We have proved that the three principles are equivalent. In practice, one may have an easy proof for a result using one of the principles, whereas proving the same using other principles may be difficult, though not impossible.

In the proofs of the existence of prime divisors of a positive integers $n > 1$, we used the strong induction principle as well as the well-ordering principle. Similarly, to prove division algorithm in integers, one usually uses the well-ordering principle. Readers may try to prove these results using the other principles.

Chapter 6

Countability of Sets

The primitive idea of "counting" a set is to set up a bijection with a known set. The words 'calculus' and 'calculation' have their origin with such a correspondence with a pile of stones!

6.1 Sets with same cardinality

Definition 6.1.1. We say that two sets X and Y have the *same cardinality* if either both are empty or there is a bijection from one onto the other. (Intuitively, this means that X and Y "have the same number of elements." Because of this we may even say that X and Y are *equinumerous*.) Note that "having the same cardinality" is an equivalence relation.

Example 6.1.2. (i) $\mathbb{N}$ and $2\mathbb{N}$, the set of even positive integers, have the same cardinality. See Example 3.2.19.
(ii) Any two closed intervals $[a, b]$ and $[c, d]$ have the same cardinality. See the discussion on bijection between intervals on page 67.
(iii) Any two open intervals (a, b) and (c, d) have the same cardinality. See Ex. 3.4.16 on page 68.
(iv) $(0, 1)$ and $[0, 1]$ have the same cardinality. See Example 3.4.20.
(v) $(-1, 1)$ and $\mathbb{R}$ have the same cardinality. See Ex. 3.4.17 and Ex. 3.4.18.
(vi) $\mathbb{Z}$ and $\mathbb{N}$ have the same cardinality. See Ex. 3.2.31.
(vii) $(0, 1)$ and $(0, \infty)$ have the same cardinality. (Hint: Consider $f(x) = \frac{x}{1-x}$ for $x \in (0, 1)$).
(viii) $(0, \infty)$ and $\mathbb{R}$ have the same cardinality. (Hint: Consider $f(x) = \log x$.)

Often setting up a bijection between two sets is not easy. However, the next theorem say that it is enough to find a one-one map from each one to the other.

Theorem 6.1.3 (Schröeder-Bernstein)**.** *Let X and Y be sets. Assume that there exist $f\colon X \to Y$ and $g\colon Y \to Y$ which are one-one. Then there exists a bijection $h\colon X \to Y$.*

We provide here a simple proof of the theorem by breaking X into disjoint parts A and B and Y into disjoint parts C and D such that f when restricted to

A produces a bijection between A and C, and g when restricted to D produces a bijection between D and B. Producing a bijection is then straight forward. For doing so, we will use the following useful result by Knaster-Tarski.

Lemma 6.1.4 (Knaster-Tarski)**.** *Let $F\colon P(X) \to P(X)$ be a map. Assume that it is increasing in the sense that if $A \subseteq B$, then $F(A) \subseteq F(B)$. Then F has a fixed point, that is, there exists $S \subseteq X$ such that $F(S) = S$.*

Proof. Consider the set $\mathcal{C} := \{C \subseteq X : C \subseteq F(C)\}$. Note that $\emptyset \in \mathcal{C}$ and therefore $\mathcal{C}$ is nonempty. Let S be the union of all members of $\mathcal{C}$. We claim that $F(S) = S$. For any $C \in \mathcal{C}$, since F is increasing, we have $C \subseteq F(C) \subseteq F(S)$, yielding $S \subseteq F(S)$. This further gives $F(S) \subseteq F(F(S))$, that is, $F(S) \in \mathcal{C}$ and therefore $F(S) \subseteq S$. □

Proof of Schröeder-Bernstein Theorem. Consider $F\colon P(X) \to P(X)$ given by $F(S) := X \setminus g(Y \setminus f(S))$. Look at the Figure 6.1.

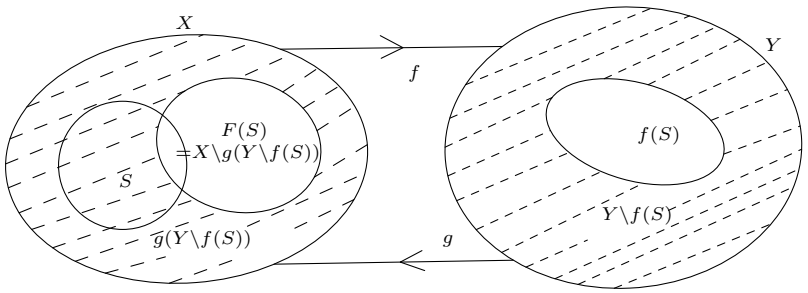


Figure 6.1: Schröder-Bernstein Theorem (Figure 1)

Then F is increasing: for $A \subseteq B \subseteq X$, we have $f(A) \subseteq f(B) \subseteq Y$, that is, $Y \setminus f(B) \subseteq Y \setminus f(A)$. This yields, $g(Y \setminus f(B)) \subseteq g(Y \setminus f(A))$, and therefore

$$F(A) = \Big(X \setminus g(Y \setminus f(A))\Big) \subseteq \Big(X \setminus g(Y \setminus f(B))\Big) = F(B).$$

By Lemma 6.1.4, there is $B \subseteq X$ such that $F(B) = B$, that is, $X \setminus g(Y \setminus f(B)) = B$, i.e., $g(Y \setminus f(B)) = X \setminus B$.

Set $A = X \setminus B$, $D = f(B)$ and $C = Y \setminus D$ (see Figure 6.2). Then, $g(C) = A$, and $f\colon B \to D$ and $g\colon C \to A$ are bijections. Clearly, $h\colon X \to Y$ defined by

$$h(x) = \begin{cases} f(x), & \text{if } x \in B, \\ g^{-1}(x), & \text{if } x \in A, \end{cases}$$

is a bijection. □

Another proof of Schröder-Bernstein Theorem

Lemma 6.1.5. *Let φ be a one-one map of A into itself. If $\varphi(A) \subseteq C \subseteq A$, then A is bijective with C.*

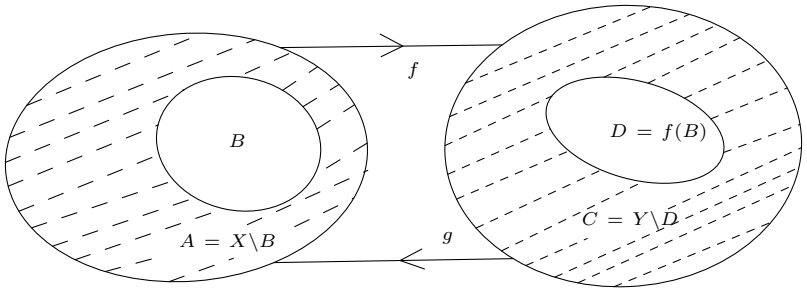


Figure 6.2: Schröder-Bernstein Theorem (Figure 2)

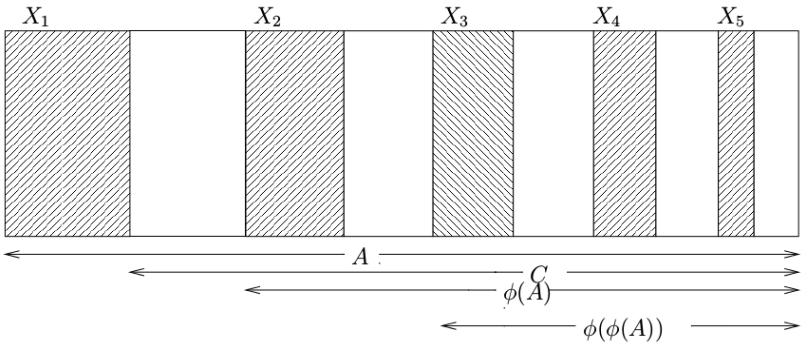


Figure 6.3: Figure for Lemma 6.1.5

Idea behind the proof. Look at the Figure 6.3. Since φ is one-one, it carries $\varphi(A) \subseteq C \subseteq A$ into a smaller version of the same thing. If we iterate φ then successive images of A and C alternate:

$$A \supseteq C \supseteq \varphi(A) \supseteq \varphi(C) \supseteq \varphi(\varphi(A)) \supseteq \varphi(\varphi(C)) \supseteq \dots$$

Thus φ maps each of the vertical strips at the left bijectively onto its second neighbour on the right (like C with $\varphi(C)$, $\varphi(A)$ with $\varphi(\varphi(A)$ and so on). Hence a bijection is obtained if we allow φ to act on the shaded strips and leave everything else fixed. The proof below formalizes this idea.

Proof. Let $X_1 := A \setminus C$. Define inductively $X_{n+1} := \varphi(X_n)$. Let $X := \cup_{n\in\mathbb{N}} X_n$. Define $\psi\colon A \to A$ by setting

$$\psi(a) = \begin{cases} \varphi(a), & \text{if } a \in X, \\ a, & \text{if } a \notin X. \end{cases}$$

We claim that ψ is a bijection from A to C.

Note that ψ maps X to X and $A \setminus X$ to $A \setminus X$. It follows that ψ is one-one.

If $a \in A$, either $\psi(a) = a \in A \setminus X \subseteq C$ or $\psi(a) = \varphi(a) \in \varphi(A) \subseteq C$. Thus ψ maps A into C.

Let $b \in C$. If $b \in A \setminus X$, then $\psi(b) = b$. If $b \in X$, then choose n so that $b \in X_n$. This integer $n \neq 1$, for, $X_1 \cap C = \emptyset$. Hence $b = \varphi(x)$ for some $x \in X_{n-1}$. Now, $x \in X$ so that $\psi(x) = b$. Thus the range of ψ is C. □

2nd proof of Schröder-Bernstein Theorem. Let $\varphi := g \circ f$. Then φ is one-one and $\varphi(A) \subseteq C := g(B) \subseteq A$. By the lemma there is a bijection ψ from A to $g(B)$. Then $g^{-1} \circ \psi$ is a bijection from A to B. □

Example 6.1.6. (i) $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$ have the same cardinality. *Hint:* The map $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ given by $f(m,n) := 2^m 3^n$ is one-one. For an explicit bijection, see Example 6.3.5.
(ii) The set $\mathbb{Q}$ of rational numbers and $\mathbb{N}$ have the same cardinality. Look at $\mathbb{Q} \to \mathbb{Z} \times \mathbb{N} \to \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.
(iii) A closed interval $[a,b]$ and an open interval (c,d) have the same cardinality. *Hint:* Produce one-one map from each to a subinterval of the other. See Example 6.1.2.
(iv) The sets $A := (0,1)$ and $B := A \times A$ have the same cardinality. *Hint:* Use non-recurring decimal expansion to get a one-map of B into A. For example, consider $g(0.x_1x_2\ldots, 0.y_1y_2\ldots) := 0.x_1y_1x_2y_2\ldots$.
(v) The sets $(0,1)$ and $P(\mathbb{N})$, the power set of $\mathbb{N}$, have the same cardinality. *Hint:* For a one-one map $f : P(\mathbb{N}) \to (0,1)$ define $f(A) = 0.x_1x_2\ldots$ by $x_i = 1$, if $i \in A$, and 2, otherwise. Use non-recurring binary expansion to define similarly one-one maps of $(0,1)$ into $P(\mathbb{N})$.

Example 6.1.7. Recall Cantor's Theorem (see Theorem 3.4.31) which states that for any set X there is no map from X onto $P(X)$. Can there be a one-one map from $P(X)$ into X? Suppose there is one. Note that there is a one-one map $f\colon X \to P(X)$ namely $f(x) = \{x\}$. Therefore, by Schröder-Bernstein theorem, there is a bijection $g\colon X \to P(X)$. However, this would contradict Cantor's theorem, since g would be onto. We conclude that *for any set X there exists no one-one map from $P(X)$ to X.*

6.2 Finite sets

In this section, we discuss finite sets, and their properties that give rise to the notion of number of elements in finite sets.

For any $n \in \mathbb{N}$, let I_n denote the subset $\{k : 1 \leq k \leq n\}$ of $\mathbb{N}$.

Definition 6.2.1. A set A is said to be *finite* if either $A = \emptyset$ or there is a bijection $f\colon A \to I_n$ for some $n \in \mathbb{N}$.

A set which is not finite is said to be *infinite*.

Lemma 6.2.2. *If $m < n$, there is no one-one map of I_n into I_m.*

Proof. We prove the result by induction on m. Let $P(m)$ be the statement: *Given $n > m$, no map $f\colon I_n \to I_m$ is one-one.*

Suppose $m = 1$ and $n > 1$. A map $f\colon I_n \to I_1 = \{1\}$ is not one-one, since $f(1) = f(n) = 1$ and $n \neq 1$. Thus, P(1) is true.

Assume $P(m)$ to be true for some $m \geq 1$.

Let $n > m+1$. Let, if possible, $f\colon I_n \to I_{m+1}$ be one-one. There are two possibilities for $f(n)$.

Case 1: Let $f(n) = m+1$. Look at the Figure 6.4. Consider the map $g\colon I_{n-1} \to I_m$ given by $g(j) = f(j)$. (Note that g is the restriction of f on I_{n-1}. It is usually denoted by $f|_{I_{n-1}}$.) Then g is one-one. This contradicts P_m.

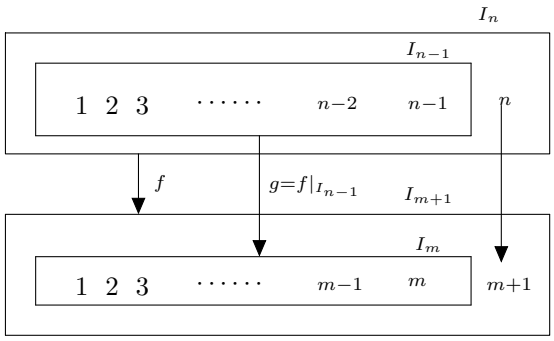


Figure 6.4: Case 1 of Lemma 6.2.2

Case 2: Let $f(n) = r < m+1$. Look at the Figure 6.5. Since f is one-one, there is at most one $1 \leq k < n$ such that $f(k) = m+1$. We define $g\colon I_{n-1} \to I_m$ by setting $g(j) = f(j)$ for $j \neq k$ and $g(k) = r = f(n)$, if there is such k. Then g is one-one, which contradicts $P(m)$.

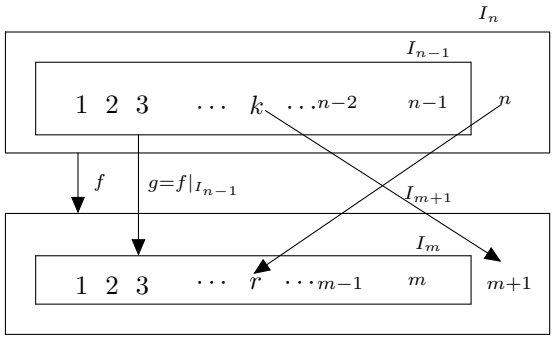


Figure 6.5: Case 2 of Lemma 6.2.2.

Thus we conclude that such an f cannot exist. In other words, $P(m+1)$ is also true. By the principle of induction, $P(m)$ is true for all m. □

Lemma 6.2.3. *If $m < n$, then there is no onto map $f\colon I_m \to I_n$.*

Proof. Let, if possible, $f\colon I_m \to I_n$ be onto, where $m < n$. We define $g\colon I_n \to I_m$ as follows: for $r \in I_n$ put $g(r) = \min f^{-1}(r)$. Then g is one-one: for if $g(r) = g(s)$, then there exists $k \in I_m$ such that $k \in f^{-1}(r) \cap f^{-1}(s)$, which

gives $r = f(k) = s$. This is a contradiction, because there is no one-one map from I_n to I_m, by the previous lemma. □

The following result is now immediate.

Theorem 6.2.4. *If $f\colon I_m \to I_n$ is a bijection, then $m = n$.* □

Why is this result significant? If we have a large pile of coins/stones, when each of us counts, we may arrive at different numbers. But we know for sure, if we count correctly, all of us should arrive at the same number. This result is the rigorous version of our intuitive assertion! Thus if we have a bijection $f\colon A \to I_m$ and another bijection $g\colon A \to I_n$, we arrive at a bijection $g \circ f^{-1} : I_m \to I_n$. By the above result we know that $m = n$. That is, we may "count" by different methods, but A will have "the same number of elements"!

Definition 6.2.5. A nonempty finite set A is said to have n elements, if there is a bijection $f\colon A \to I_n$. Note that in view of the above theorem this is well-defined. We denote the *number of elements* in A by $|A|$. Moreover, we write $|\emptyset| = 0$.

Note that for $n \in \mathbb{N}$, we have $|I_n| = n$.

Lemma 6.2.6. *Let $f\colon A \to I_n$ be one-one. Then A is finite and $|A| \leq n$.*

Proof. We use the well-ordering property of $\mathbb{N}$, that is, every nonempty subset of $\mathbb{N}$ has a least element. Let $r_1 = \min\{f(a) : a \in A\} \subseteq \mathbb{N}$, $r_2 = \min\{f(A) \setminus \{r_1\}\}$. Note that $r_1 \geq 1$ and $r_2 > r_1$ so that $r_2 \geq 2$. We proceed by induction to construct $r_1 < r_2 < \cdots < r_k$ where $r_k \geq k$. This process will stop at some stage in the sense that $f(A) \setminus \{r_j : 1 \leq j \leq k\} = \emptyset$ for some $k \leq n$. For, otherwise, if $k > n$, then $r_k \geq k > n$. This contradicts the fact that $r_k \in I_n$.

Clearly, $f(a) \in \{r_1, r_2, \ldots, r_k\}$ for each $a \in A$. Moreover, for $1 \leq j \leq k$, $r_j = f(a)$ for a unique $a \in A$, since f is one-one. Define $g\colon A \to I_k$ by $g(a) = i$, if $f(a) = r_i$. It is evident that g is a bijection.

Therefore, A is finite and $|A| = k \leq n$. □

Corollary 6.2.7. *Let $f\colon I_n \to A$ be onto. Then A is finite and $|A| \leq n$.*

Proof. Define $g\colon A \to I_n$ by setting

$$g(a) = \min f^{-1}(a) = \min\{k \in I_n : f(k) = a\}.$$

Then g is one-one and the result follows from Lemma 6.2.6. □

Proposition 6.2.8. *If A is finite and $B \subseteq A$, then B is finite and $|B| \leq |A|$.*

Proof. Let $f\colon A \to I_n$ be a bijection. Let $g\colon B \hookrightarrow A$ be the inclusion map (that is, $x \mapsto x$, the restriction of the identity map A to B) of B in A. Then the composition $f \circ g$ is a one-one map of B into I_n. By Lemma 6.2.6, the result follows. □

Proposition 6.2.9. *If A is finite and B is a proper subset of A, then $|B| < |A|$.*

Proof. Let us prove the result when $C \subset A$ misses exactly one element of A, that is, $C = A\backslash\{a\}$, $a \in A$, say. Let $f\colon A \to I_n$ be a bijection. Let $f(a) = k$. We define $g\colon C \to I_{n-1}$ as follows. Let $g(x) = f(x)$ if $f(x) < k$ and $g(x) = f(x) - 1$ if $f(x) > k$. Observe that g maps C into I_{n-1}.

We claim that $g\colon C \to I_{n-1}$ is one-one. Let $x_1 \neq x_2 \in C$.

Case 1. $g(x_1) = f(x_1)$ and $g(x_2) = f(x_2)$. Since f is one-one, it follows that $g(x_1) \neq g(x_2)$.

Case 2. $g(x_1) = f(x_1)$ and $g(x_2) = f(x_2) - 1$ (or the other way round). In this case, $g(x_1) < k$ and $g(x_2) \geq k$, since $f(x_2) > k$. Hence $g(x_1) \neq g(x_2)$.

Case 3. Let $g(x_1) = f(x_1) - 1$ and $g(x_2) = f(x_2) - 1$. Since f is one-one, it follows that $g(x_1) \neq g(x_2)$.

We now show that g is onto I_{n-1}. If $r \in I_{n-1}$ is such that $r < k$, since f is onto, there exists $x \in A$ such that $f(x) = r$ and hence $g(x) = r$. Let $r \in I_{n-1}$ be such that $r \geq k$. Since f is onto I_n, there exists $x \in A$ such that $f(x) = r + 1$. Since $r + 1 > k$, $x \neq a$ so that $x \in C$ and we have $g(x) = f(x) - 1 = r$. Therefore, $g\colon C \to I_{n-1}$ is a bijection. We have thus proved that if C is of the form $A \setminus \{a\}$, then $|C| = n - 1$.

Now let B be any proper subset of A. Then there exists $a \in A$ such that $a \notin B$. Hence $B \subseteq C = A \setminus \{a\}$. It follows from the last Proposition 6.2.8 that $|B| \leq |C| = n - 1$. □

The following well-known principle is a restatement of Lemma 6.2.2.

Proposition 6.2.10 (Pigeonhole Principle)**.** *Let $m, n \in \mathbb{N}$ be such that $m < n$. If $f\colon I_n \to I_m$ is a map, then there exists $i, j \in I_n$ such that $i \neq j$ and $f(i) = f(j)$.*

In other words, if A, B are finite sets with $|A| > |B|$ and $f\colon A \to B$ is a map, then f is not one-one. □

Why is this called Pigeonhole principle? If there are n pigeons and there are m pigeonholes with $m < n$ and if all n pigeons stay in these m pigeonholes, we 'expect' that at least two pigeons stay in the same pigeonhole. The proposition is a rigorous version of our intuition. Pigeonhole principle is one of the most basic and powerful tools in Combinatorics.

Exercise 6.2.11. A typical application of pigeonhole principle. Let $k \geq 2$. Let $a_1, \ldots, a_{k+1} \in \mathbb{Z}$. Prove that there exist $1 \leq i \neq j \leq k$ such that $a_i - a_j$ is divisible by k.

Exercise 6.2.12. Let A and B be finite sets with $A \cap B = \emptyset$. Show that $A \cup B$ is finite. What is the number of elements in $A \cup B$?

Exercise 6.2.13. Let A and B be finite sets. Show that $A \cap B$ and $A \cup B$ are finite, and

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Hint: If $|A| = m, |B| = n$, and $|A \cap B| = k$, construct a bijection between $A \cup B$ and I_{m+n-k}.

Exercise 6.2.14. Let X be a finite set and $f\colon X \to X$ be a map. Show that the following are equivalent:

(a) f is a bijection.
(b) f is one-one.
(c) f is onto.

Exercise 6.2.15. Let A and B be finite sets and $f\colon A \to B$ be a map. Prove the following:
(a) If f is one-one, than $|A| \leq |B|$.
(b) If f is onto, than $|A| \geq |B|$.
(c) If $f\colon A \to B$ and $g\colon B \to A$ are one-one, then $|A| = |B|$, and f and g are bijections.

Exercise 6.2.16. Let A be an infinite set and $B \subseteq A$ a finite subset. Show that $A \setminus B$ is infinite.

Exercise 6.2.17. Let A be such that there is a bijection $f\colon A \to \mathbb{N}$. Can A be finite?

6.3 Countable sets

Definition 6.3.1. A set A is said to be *countable* if either A is finite or there is a bijection $f\colon A \to \mathbb{N}$. A set of the latter type is said to be *countably infinite*. Note that in view of Ex. 6.2.17, a set cannot both be finite and countably infinite.

A set which is not countable is said to be *uncountable*.

Let A be a countably infinite set and $f\colon \mathbb{N} \to A$, a bijection. Then for each $i \in \mathbb{N}$, we can denote $f(i) \in A$ by a_i. In particular, elements of A can be enumerated as $a_1, a_2, \ldots$, where a_i is the image of i under f. In other words, $A = \{a_i : i \in \mathbb{N}\}$.

Example 6.3.2. $\mathbb{Z}_+$, $\mathbb{Z}$ are countably infinite. (See Ex. 3.2.31.)

Example 6.3.3. Any infinite subset of $\mathbb{N}$ is countably infinite. (Hint: Use the well-ordering principle.)

Exercise 6.3.4. Suppose A and B are disjoint countably infinite sets. Show that $A \cup B$ is countably infinite.

Example 6.3.5. The set $\mathbb{N} \times \mathbb{N}$ is countably infinite.

Look at the grid in Figure 6.6 in which the nodes are the points of $\mathbb{N} \times \mathbb{N}$. The arrows explain how to map points in $\mathbb{N} \times \mathbb{N}$ to $1, 2, 3, 4$, etc.

Thus, we start counting from $(1, 1)$, then start counting the points of $\mathbb{N} \times \mathbb{N}$ in the line joining $(2, 1)$ to $(1, 2)$ (along the direction of south-east to north west), and continuing this process, at the m-th stage we continue counting the points on the line joining $(m, 1)$ to $(1, m)$ along the specified direction.

The figure suggests how to set up a bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. But how to find such a bijection explicitly?

Let us call the slanting line joining $(m, 1)$ and $(1, m)$ as L_m. Observe that there are m points (of $\mathbb{N} \times \mathbb{N}$) on L_m, and any point on this line is of the form (r, s) with $r + s = m + 1$, that is, of the form $(m - k, k + 1)$, $0 \leq k \leq m - 1$.

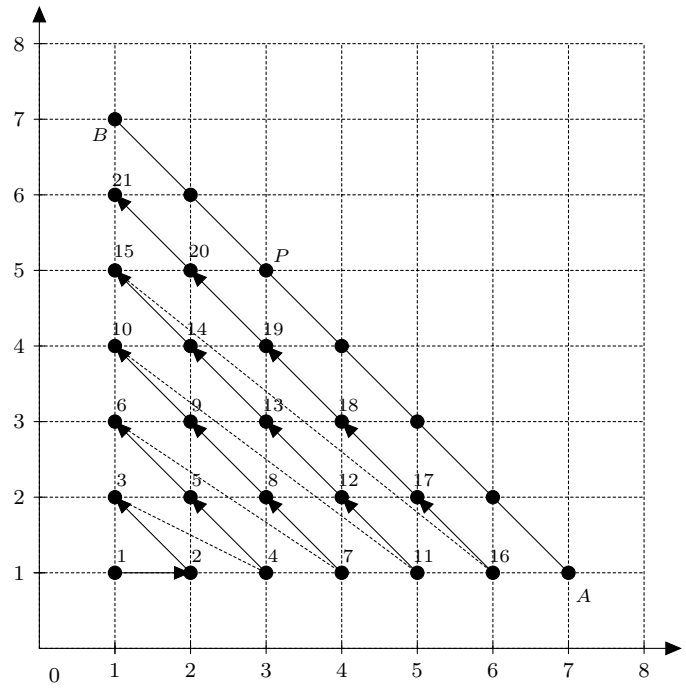


Figure 6.6: Bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$

Note that $(m-k, k+1)$ will be the $(k+1)$-th point on the line L_m (as we move from south-east to north-west direction). Now, the number of points on all the lines preceding this line, namely on the lines L_j with $1 \leq j < m$, is $\frac{(m-1)m}{2}$. Therefore, the point $(m-k, k+1)$ will be the $\left(\frac{(m-1)m}{2} + k + 1\right)$-th point in our listing. If $(r, s) \in \mathbb{N} \times \mathbb{N}$, then we set $r + s = m + 1$ and $(r, s) = (m - k, k + 1)$, where $0 \leq k \leq m - 1$. Then, we define $f(r, s) = \frac{(m-1)m}{2} + k + 1$. It is easy to see that this map is one-one.

Why is it onto? Given any $n \in \mathbb{N}$, we choose the largest m such that $\frac{(m-1)m}{2} < n$. Then we have $\frac{(m-1)m}{2} < n \leq \frac{m(m+1)}{2}$. Thus we expect n to correspond to a point on L_m. Note that there are m natural numbers j which satisfy the above inequalities. Let $k+1 := n - \frac{(m-1)m}{2}$. Then $f(m-k, k+1) = n$.

The following graded exercise gives an explicit bijection between $\mathbb{N}$ and $\mathbb{Q}^+$.

Exercise 6.3.6. (i) Define $\varphi \colon \mathbb{N} \to \mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$ by

$$\varphi(n) := \begin{cases} \frac{n}{2}, & \text{if } n \text{ is even} \\ -\frac{n+1}{2}, & \text{if } n \text{ is odd.} \end{cases}$$

Show that φ is a bijection.
(ii) Define $f \colon \mathbb{N} \to \mathbb{Q}^+ = \{x \in \mathbb{Q} : x > 0\}$ by $f(1) = 1$ and for $n > 1$

$$f(n) = p_1^{\varphi(n_1)} p_2^{\varphi(n_2)} \cdots p_k^{\varphi(n_k)},$$

where $n = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$ is the prime factorization of n. Show that f is a bijection, and therefore $\mathbb{Q}^+$ is countably infinite.
(iii) Deduce that $\mathbb{Q}$ is countably infinite. (See Ex. 6.3.4.)

Proposition 6.3.7. *For a nonempty set A the following are equivalent.*
(i) *A is countable.*
(ii) *There is a one-one map of A into $\mathbb{N}$.*
(iii) *There is an onto map from $\mathbb{N}$ onto A.*

Proof. We shall sketch a proof asking the reader to supply the details.

(i) $\implies$ (ii): If A is finite then there exists a bijection $f\colon A \to I_n$. Define $g : A \to \mathbb{N}$ by setting $g(a) = f(a)$. Then g is as required. If A is countably infinite, then there exists a bijection $f\colon A \to \mathbb{N}$ and hence it is one-one.

(ii) $\implies$ (iii): Let $f\colon A \to \mathbb{N}$ be a one-one map. Define $g\colon \mathbb{N} \to A$ as follows. Let $B = f(A)$. Then f is a bijection of A to B. Fix $c \in A$. Define $g\colon \mathbb{N} \to A$ as follows:

$$g(n) = \begin{cases} f^{-1}(n) & \text{if } n \in B \\ c & \text{if } n \notin B \end{cases}.$$

Then g is onto, since $g(B) = A$.

(iii) $\implies$ (i): Let $f\colon \mathbb{N} \to A$ be onto. For $a \in A$, define $E_a := f^{-1}(a)$. Then we have $\emptyset \neq E_a \subseteq \mathbb{N}$. Let n_a be the least element of E_a. Then the collection $\{n_a : a \in A\}$ is either finite or infinite. In either case, using well-ordering principle, arrange them either as a finite sequence $n_1 < n_2 < \cdots < n_k$ or $n_1 < n_2 < \cdots < n_k < n_{k+1} < \cdots$. Define $g\colon A \to \mathbb{N}$ by setting $g(a) = j$ if $f(n_j) = a$. You are asked to check that g is a bijection of A either with I_k or with $\mathbb{N}$. □

Corollary 6.3.8. (i) *If $f\colon A \to B$ is one-one and B is countable, then A is countable.*
(ii) *If $f\colon A \to B$ is onto and A is countable, then B is countable.*
(iii) *A subset of a countable set is countable.*

Proof. (i) Since B is countable, there is a one-one map $g\colon B \to \mathbb{N}$. Then $g \circ f\colon A \to \mathbb{N}$ is one-one, and so A is countable.
(ii) Similar to (i).
(iii) Let $A \subseteq B$ and B is countable. Since the inclusion map $i\colon A \hookrightarrow B$ is one-one, the result follows from (i). □

Corollary 6.3.9. (i) *If I is a countable set and A_i is a countable set for each $i \in I$, then $A := \cup_{i\in I} A_i$ is countable. That is, a countable union of countable sets is countable.*
(ii) *A finite Cartesian product of countable sets is countable.*

Proof. We shall sketch a proof asking the reader to supply the details.

(i) Without loss of generality, assume that $A_i \neq \emptyset$. Take one-one maps $f\colon I \to \mathbb{N}$ and $g_i\colon A_i \to \mathbb{N}$ for $i \in I$. Define $g\colon A \to \mathbb{N} \times \mathbb{N}$ as follows: for $a \in A$ choose $n := \min\{f(i) : a \in A_i\}$ and put $g(a) := (n, g_n(a))$. Then g is one-one. Since $\mathbb{N} \times \mathbb{N}$ is countable, in view of Corollary 6.3.8 (i), A is countable.

(ii) Suppose A and B are nonempty and countable, and $f\colon A \to \mathbb{N}$ and $g\colon B \to \mathbb{N}$ are one-one maps. Then $h\colon A \times B \to \mathbb{N} \times \mathbb{N}$ defined by $h(a, b) =$

$(f(a), g(b))$ is one-one. Since $\mathbb{N} \times \mathbb{N}$ is countable, so is $A \times B$. Now, we can use induction on the number of sets in the Cartesian product. □

Exercise 6.3.10. Use Corollary 6.3.9 to give another proof of countability of $\mathbb{Q}$. *Hint:* Let $q \in \mathbb{N}$. Let A_q be the set of rational numbers whose denominator is q. Then A_q is countable and $\mathbb{Q}$ is the union of A_q's.

Exercise 6.3.11. Show that the set $F(\mathbb{N}) = \{A \subseteq \mathbb{N} : A \text{ is finite}\}$, that is, the collection of all finite subsets of $\mathbb{N}$, is countable.

Exercise 6.3.12. A complex number is said to be an *algebraic number* if it is a root of a polynomial with integer coefficients. Show that the set of algebraic numbers is countable. *Hint:* Show that the set of polynomials with integer coefficient is countable. Use (ii) and (iii) of Corollary 6.3.9.

Exercise 6.3.13. Show that $P(\mathbb{N})$ is not countable. (*Hint:* Use Cantor's Theorem 3.4.31.)

Exercise 6.3.14. Show that the set $2^{\mathbb{N}}$ of all functions from $\mathbb{N}$ to $\{0, 1\}$ is not countable. *Hint:* The set under question is bijective with $P(\mathbb{N})$.

Example 6.3.15. $\mathbb{R}$ is uncountable.

There are many interesting proofs of this result. Here we mention a few of them.
(i) Note that $\mathbb{R}$, $(0, 1)$ and $P(\mathbb{N})$ have the same cardinality. However, $P(\mathbb{N})$ is uncountable, so is $\mathbb{R}$.
(ii) *Diagonal Trick:* If $\mathbb{R}$ is countable then $(0, 1)$ is countable. Let $f : \mathbb{N} \to (0, 1)$ be a bijection. Since $f(i) \in (0, 1)$, it has a non recurring decimal expansion as $f(i) = 0.a_{i1}a_{i2}\ldots$.

Define $b_n = a_{nn} + 1$ if $a_{nn} < 9$ and $b_n = 8$ if $a_{nn} = 9$. Now consider the number $b = 0.b_1b_2\ldots$. Clearly, $b \in (0, 1)$ and is different from $f(n)$ for any $n \in \mathbb{N}$.
(iii) This result can also be proved using the least upper bound (LUB) property of $\mathbb{R}$ and Nested Interval theorem. See the expository article "Uncountability of $\mathbb{R}$" by S Kumaresan in the collection [8].

Exercise 6.3.16. Show that the set of irrational numbers is uncountable.

Exercise 6.3.17. Prove that the set $\mathbb{C}$ of complex numbers and $\mathbb{R}$ have the same cardinality. (*Hint:* Use item (iv) of Example 6.1.6.)

A complex number is *transcendental* if it is not algebraic.

Corollary 6.3.18. *The set of transcendental numbers is uncountable.* □

Exercise 6.3.19. Show that the set $F(\mathbb{N}, \mathbb{N}) := \{f : \mathbb{N} \to \mathbb{N}\}$ is uncountable. *Hint:* Diagonal trick.

Theorem 6.3.20. *The following are equivalent for a set X.*
(i) *The set X is infinite.*
(ii) *There exists a countably infinite subset S of X.*
(iii) *There exists a proper subset Y of X such that X and Y have the same cardinality.*

We sketch a proof of this theorem.

Proof. (i) $\Rightarrow$ (ii). Use induction and Ex. 6.2.16 to construct a one-one map from $\mathbb{N} \to X$.

(ii) $\Rightarrow$ (iii). Let $S = \{s_i : i \in \mathbb{N}\}$ be a countably infinite subset of X. Define $Y = X \setminus \{s_1\}$. Consider

$$f(x) := \begin{cases} s_{i+1} & \text{if } x = s_i \text{ for some } i \in \mathbb{N} \\ x & \text{if } x \notin S. \end{cases}$$

Check that f is a bijection.

(iii) $\Rightarrow$ (i). Suppose X is finite and Y a proper subset of X. Then by Proposition 6.2.9, $|Y| < |X|$. This shows that if X is finite then (iii) cannot hold. □

Exercise 6.3.21. Let X be uncountable and A a countable subset of X. Show that $X \setminus A$ is uncountable. More precisely, show that X and $X \setminus A$ have the same cardinality.
Hint: Let B be a countably infinite subset of $X \setminus A$. Use a bijection from $A \cup B$ to B to construct a bijection from X and $X \setminus A$.

6.4 Comparing cardinality

Recall that for a finite set X, we denote by $|X|$, the number of elements in X. If Y is another finite set, then $|X| \leq |Y|$ iff either $X = \emptyset$ or there is an one-one map from X into Y.

Definition 6.4.1. For sets X and Y, we write
(i) $|X| = |Y|$, if X and Y are of same cardinality, and $|X| \neq |Y|$, otherwise.
(ii) $|X| \leq |Y|$ (read as "the cardinality of X is *less than or equal to* the cardinality of Y") if either $X = \emptyset$ or there is an one-one map from X into Y.
(iii) $|X| < |Y|$, (read as "the cardinality of X is *less than* the cardinality of Y") if $|X| \leq |Y|$ and $|X| \neq |Y|$.

Note that if $|X| \leq |Y|$ and $|Y| \leq |X|$, then by Schröder-Bernstein theorem there exists a bijection between them. Hence we conclude that $|X| = |Y|$.

Example 6.4.2. (i) If X is finite, then $|X| < |\mathbb{N}|$.
(ii) If X is infinite, then $|\mathbb{N}| \leq |X|$.
(iii) If X is an uncountable set, then $|\mathbb{N}| < |X|$.
(iv) For any set X, $|X| < |P(X)|$. (Follows from Cantor's Theorem 3.4.31)

Example 6.4.3. It follows from the Cantor's Theorem that there are infinitely many infinite sets with strictly increasing cardinalities. More precisely,

$$|\mathbb{N}| < |P(\mathbb{N})| < |P(P(\mathbb{N}))| < |P(P(P(\mathbb{N})))| < \cdots .$$

Chapter 7

Order Relations

In this chapter, we deal with relations which exhibit some special properties. Our main aim is to introduce partial and total orders on a set. We also give a brief account of axiom of choice, Zorn's lemma and well-ordering principle.

7.1 Partial and Total Orders

In Chapter 4, we discussed different types of relations on sets, namely, reflexive, symmetric, anti-symmetric and transitive relations, with a number of examples. We urge the reader to go through the first two sections of Chapter 4 for reviewing the basic concepts and examples.

Definition 7.1.1. We say that a relation R on X is a *partial order* if it is reflexive, anti-symmetric and transitive. If R is a partial order, it is customary to denote it by $\leq$ (or by $\preceq$, if there is a possibility of confusion) and write $x \leq y$ in place of xRy. We read this as "x is less than or equal to y". The ordered pair $(X, \leq)$ is called a *partially ordered set* (sometime, a *poset*, in short). If $x \leq y$ holds and if $x \neq y$, we say "x is less than y" and write $x < y$.

We write the negation of $x \leq y$ as $x \not\leq y$ (read as "x is not less than or equal to y). A partial order is a *total order* if for each $x, y \in X$ either $x \leq y$ or $y \leq x$. In that case, for each $x, y \in X$, either x and y are equal, or x is less than y, or y is less than x. It is the analogue of Law of Trichotomy. A set X with a total order $\leq$ (or more precisely, the ordered pair $(X, \leq)$) is called a totally ordered set.

Note that $x \not\leq y$ implies $y < x$ only if X is totally ordered, and not in general.

We now revisit some of the examples of relations defined in Chapter 4 and check which of them are partial orders and if true which of them are total orders.

Example 7.1.2. (1) $(\mathbb{R}, \leq)$ is a partially ordered set with usual ordering on $\mathbb{R}$, "less than or equal to". Is this a total order?

(2) For any set X, $(P(X), \subseteq)$ is a partially ordered set. Note that it is not a total order if X has more than one element.

(3) Consider the relation $\preceq$ on $\mathbb{N}$ defined as follows: $m \preceq n$ iff n is a multiple

of m. Then $(\mathbb{N}, \preceq)$ is a partially ordered set which is not totally ordered.

(4) On any set X, define $\preceq$ by $x \preceq y$ iff $x = y$. Then $(X, \preceq)$ is a partially ordered set. When is it a totally ordered set?

Remark 7.1.3. In our opinion, Examples 7.1.2 (2) and (3) are the prototype examples of partial order.

Exercise 7.1.4. Can an equivalence relation on a set be a partial order?

Example 7.1.5. (1) "Strictly less than ($<$)" on $\mathbb{N}$ is not a partial order. It is not reflexive.

(2) On the set $\mathbb{Z}^*$ of nonzero integers, the relation defined by xRy iff y is a multiple of x is not a partial order. The relation is not anti-symmetric. Look at 1 and -1.

(3) On $\mathbb{R}$ the relation defined by xRy iff $xy > 0$ is not a partial order. This relation is not reflexive and not anti-symmetric.

Exercise 7.1.6. Consider the standard total order $\leq$ on $\mathbb{R}$. For (x_1, y_1), $(x_2, y_2) \in \mathbb{R}^2$, define $(x_1, y_1) \leq (x_2, y_2)$ iff $x_1 \leq x_2$ and $y_1 \leq y_2$. Show that this relation on $\mathbb{R}^2$ is a partial order which is not a total order.

Example 7.1.7. Let $(X, \leq)$ and $(Y, \leq)$ be partially ordered sets. On $X \times Y$, define

$$(x_1, y_1) \preceq (x_2, y_2) \text{ iff either } x_1 < x_2 \text{ or } (x_1 = x_2 \text{ and } y_1 \leq y_2).$$

Note that if $(x_1, y_1) \preceq (x_2, y_2)$, then we have $x_1 \leq x_2$.

We claim that $(X \times Y, \preceq)$ is a partially ordered set. Further if X and Y are totally ordered, then $(X \times Y, \preceq)$ is totally ordered.

This order is known as the *lexicographic* or *dictionary order* on $X \times Y$. Refer to Examples 4.2.8 and 4.2.10.

We show that $\preceq$ is reflexive, anti-symmetric and transitive.

Let $(x, y) \in X \times Y$. Then $x = x$ and $y \leq y$. This implies $(x, y) \preceq (x, y)$. Hence $\preceq$ is reflexive.

Let $(x_1, y_1) \preceq (x_2, y_2)$ and $(x_2, y_2) \preceq (x_1, y_1)$. Then we have $x_1 \leq x_2$ and $x_2 \leq x_1$. This implies $x_1 = x_2$. Thus we have $(x_1, y_1) \preceq (x_1, y_2)$ and $(x_1, y_2) \preceq (x_1, y_1)$. This gives $y_1 \leq y_2$ and $y_2 \leq y_2$. Hence $y_1 = y_2$, and therefore $(x_1, y_1) = (x_2, y_2)$. This shows that $\preceq$ is anti-symmetric.

Next suppose $(x_1, y_1) \preceq (x_2, y_2)$ and $(x_2, y_2) \preceq (x_3, y_3)$. We claim that $(x_1, y_1) \preceq (x_3, y_3)$.

Now $(x_1, y_1) \preceq (x_2, y_2)$ implies $x_1 \leq x_2$ and $(x_2, y_2) \preceq (x_3, y_3)$ implies $x_2 \leq x_3$. Hence $x_1 \leq x_3$. We have two cases.

Case 1: $x_1 < x_3$. Then by definition $(x_1, y_1) \preceq (x_3, y_3)$.

Case 2: $x_1 = x_3$. In this case, we have $x_1 \leq x_2$ and $x_2 \leq x_3 = x_1$. This implies $x_1 = x_2 = x_3$. Therefore, $(x_1, y_1) \preceq (x_2, y_2)$ implies $y_1 \leq y_2$ and $(x_2, y_2) \preceq (x_3, y_3)$ implies $y_2 \leq y_3$. Consequently, $y_1 \leq y_3$. Thus we have $(x_1, y_2) \preceq (x_3, y_3)$ and hence $\preceq$ is transitive.

This proves that $\preceq$ is a partial order.

Next, assume that X and Y are totally ordered. We show that $\preceq$ is a total order.

Let $(x_1, y_1), (x_2, y_2) \in X \times Y$. Then either $x_1 \leq x_2$ or $x_2 \leq x_1$. Similarly, either $y_1 \leq y_2$ or $y_2 \leq y_1$. We have four cases.
(i) $x_1 < x_2$. Then $(x_1, y_1) \preceq (x_2, y_2)$.
(ii) $x_2 < x_1$. Then $(x_2, y_2) \preceq (x_1, y_1)$.
(iii) $x_1 = x_2$ and $y_1 \leq y_2$. Then $(x_1, y_1) \preceq (x_2, y_2)$.
(iv) $x_1 = x_2$ and $y_2 \leq y_1$. Then $(x_2, y_1) \preceq (x_1, y_1)$.
This proves that $\preceq$ is a total order.

Example 7.1.8. Consider the dictionary order on $\mathbb{R} \times \mathbb{R}$. Then $(1, 5) \preceq (5, -1)$ and $(1, -1) \preceq (1, 5)$. However, $(1, 1) \not\preceq (1, 0)$.

Exercise 7.1.9. Consider $\mathbb{Z} \times \mathbb{Z}$ with dictionary order. What is the relation between the elements (i) $(-1, 0)$ and $(2, -20)$ and (ii) $(1, 10)$ and $(1, -10)$?

Exercise 7.1.10. Can you distinguish the dictionary order on $\mathbb{R}^2$ from the order defined in Ex. 7.1.6? If $(x_1, y_1) \leq (x_2, y_2)$, will $(x_1, y_1) \preceq (x_2, y_2)$ hold? What about the converse?

Exercise 7.1.11. With $\preceq$ as the dictionary order on $\mathbb{R}^2$, can you identify the following subsets of $\mathbb{R}^2$?
(i) $\{(x, y) \in \mathbb{R}^2 : (0, 0) \preceq (x, y)\}$.
(ii) $\{(x, y) \in \mathbb{R}^2 : (1, 0) \preceq (x, y) \preceq (2, 1/2)\}$,
(Never mind, if you find it difficult; we will take up this in Example 7.2.45.)

Example 7.1.12. Let $\mathbb{C}$ be the set of complex numbers. If $z = x + iy$ and $w = u + iv$ are complex numbers, we say that zRw if either $x < u$ or $x = u$ and $y \leq v$. Note that this is 'essentially' the same as the dictionary order on $\mathbb{R}^2$. This gives a total order on $\mathbb{C}$.

Exercise 7.1.13. On $\mathbb{C}$, define $z_1 R z_2$ iff $|z_1| \leq |z_2|$. Is this a partial order on $\mathbb{C}$?

Exercise 7.1.14. We define a relation R on $\mathbb{N}$ as follows. We say that aRb iff $a \leq 2b$. Is it a partial order?

Exercise 7.1.15. Recall inverse of a relation defined in Definition 4.2.13. If R is a partial order (respectively, total order) on X, is the inverse relation, R^{-1} a partial order (respectively, total order) on X? Justify.

Exercise 7.1.16. Suppose Y is a nonempty set and $(Y, \leq)$ is a partially ordered set. Let $f \colon X \to Y$ be a map. Define a relation R on X by $x_1 R x_2$ iff $f(x_1) \leq f(x_2)$.
(i) Show that R is a partial order on X iff f is one-one.
(ii) Suppose $(Y, \leq)$ is totally ordered. When is R a total order on X?

7.2 Chains, bounds and maximal elements

Definition 7.2.1. Let R be a relation on X and $Y \subseteq X$. The relation R' on Y, defined by $y_1 R' y_2$ in Y iff $y_1 R y_2$ in X, is called the *restriction* of R to Y.

If $(X, \leq)$ be a partially ordered set and $Y \subseteq X$, then it is easy to check that the restriction of $\leq$ to Y is a partial order on Y. Following the standard practice, we denote both the partial orders by the same symbol $\leq$.

Definition 7.2.2. A nonempty subset Y of a partially ordered set $(X, \leq)$ is called a *chain* in X, if $(Y, \leq)$ is a totally ordered set.

Example 7.2.3. Let $X = \{1, 2, 3, 4, 5, 6, 9, 12, 18, 24\}$ with partial order defined by $x \preceq y$ iff x divides y. Then $(\{1, 2, 4, 12, 24\}, \preceq)$ is a chain in X.

Example 7.2.4. Consider the subset $A := \{2^n : n \in \mathbb{N}\}$ in the partially ordered set $(\mathbb{N}, \preceq)$ in Example 7.1.2 (3). Then A is a chain in $\mathbb{N}$.

Example 7.2.5. A chain need not be a countable set. For example, consider $X = [0, 1]$ and $P(X)$ with inclusion as the partial order. Then the family $\{[0, t] : 0 < t \leq 1\}$ is a chain in $P(X)$.

Definition 7.2.6. Let $(X, \leq)$ be a partially ordered set. Let $A \subseteq X$ be nonempty. We say that an element $u \in X$ is an *upper bound* of A if for all $a \in A$, we have $a \leq u$. Note that if u is an upper bound of A and if $u \leq v$, then v is also an upper bound of A.

Note that an upper bound of a set A need not be an element of A. See Example 7.2.11.

Example 7.2.7. Consider $X = \{a, b, c, d\}$ and the partially ordered set $P(X)$ with inclusion as the partial order. Let $A = \{\{a, b\}, \{b, c\}, \{a, b, c\}\}$. Then $\{a, b, c\}$ and X are upper bounds of A.

Example 7.2.8. Consider the partially ordered set $(\mathbb{N}, \preceq)$ in Example 7.1.2 (3). Let $A = \{3, 4, 6, 9\}$. Then, 36 is an upper bound of A. In fact, the multiples of 36 are the upper bounds of A. Suppose A is a finite subset of $\mathbb{N}$. Can you identify the upper bounds of A?

Example 7.2.9. Consider the partially ordered set $\mathbb{N}$ with standard order of real numbers. Then there exist no natural number ℓ, which is an upper bound of $\mathbb{E}$, the set of even natural numbers.

Definition 7.2.10. Let $(X, \leq)$ be a partially ordered set and $A \subseteq X$ be a nonempty subset of X. We say that an element $\ell \in X$ is a *lower bound* of A if for all $a \in A$, we have $\ell \leq a$.

Note that if ℓ is a lower bound of A and if $u \leq \ell$, then u is also a lower bound of A.

Example 7.2.11. Consider $(\mathbb{R}, \leq)$ with the standard order "$\leq$".

(i) For $A = [0, 1)$, 0 is a lower bound and 1 is an upper bound.

(ii) For $B = \{x \in \mathbb{R} : x > -1\}$, -1 is a lower bound. However, B does not have an upper bound.

(iii) For $C = (0, 1)$, 1 is an upper bound and 0 is a lower bound. Note that they are not elements of A.

Example 7.2.12. Consider $X = \{a, b, c\}$ and the partially ordered set $P(X)$ with inclusion as the partial order. Let $A = \{\{a, b\}, \{b, c\}\}$. Then $\emptyset$ and $\{b\}$ are lower bounds of A. Is $\{a\}$ a lower bound of A? Can you find upper bounds of A? How many can you find?

Exercise 7.2.13. Find all lower bounds of the sets in Examples 7.2.7, 7.2.8 and 7.2.9, whenever they exist.

Example 7.2.14. Consider $(\mathbb{Z}, \leq)$ with the standard partial order. Then 1 is a lower bound of $A = \mathbb{N}$. There is no upper bound for $\mathbb{N}$.

Consider $A = -\mathbb{N} \subset \mathbb{Z}$, the set of negative integers. Then -1 is an upper bound of A, but A does not have a lower bound.

Remark 7.2.15. A set A in a partially ordered set may have an upper bound (or a lower bound), but may not have any inside A. For example, see Example 7.2.11 (iii).

Definition 7.2.16. Let $(X, \leq)$ be a partially ordered set and A be a nonempty subset of X. An element $a \in A$ is called a *maximum* of A if for all $x \in A$, we have $x \leq a$.

Similarly, an element $b \in A$ is called a *minimum* of A if for all $x \in A$, we have $b \leq x$.

Note that a maximum (respectively, a minimum) of a set A is an upper bound (respectively, a lower bound) of A which lies in A.

Example 7.2.17. Let $X = P(\mathbb{R})$ with inclusion as the partial order. Then $\mathbb{R}$ and $\emptyset$ are the maximum and minimum of X, respectively.

Let $X = P(\mathbb{R}) \setminus \{\emptyset, \mathbb{R}\}$ with inclusion as the partial order. That is, X is the set of all nonempty proper subsets of $\mathbb{R}$ ordered by inclusion. Then there is no maximum or minimum of X. Supply a proof of this claim.

Example 7.2.18. Let $(\mathbb{N}, \preceq)$ be a partially ordered set with $x \preceq y$ iff x divides y.

(i) Let $A = \{2, 3, 4, 6, 12, 18\}$. Then A has no maximum and no minimum elements. Note, however, there exist upper bounds and lower bounds of A in $\mathbb{N}$.

(ii) The set $B = \{2, 3, 4, 6, 12\}$ has 12 as a maximum element. However, B does not have a minimum element.

Exercise 7.2.19. Let A be a subset of a partially ordered set $(X, \leq)$. Assume that $a, b \in A$ are two maximums of A. Prove that $a = b$. In other words, the maximum of a set, if exists, is unique.

State and prove an analogue for minimum.

Exercise 7.2.20. If X is a partially ordered set in which the maximum is also a minimum, what can you conclude about X?

Definition 7.2.21. Let $(X, \leq)$ be a partially ordered set and A be a nonempty subset of X. An element $\alpha \in X$ is said to be a *least upper bound* (LUB, in short) or *supremum* of A if (i) α is an upper bound of A and (ii) for each upper bound β of A, we have $\alpha \leq \beta$.

A *greatest lower bound* (GLB, in short) or *infimum* of a set is defined analogously.

Exercise 7.2.22. Prove that the LUB of a set, if exists, is unique. State and prove the analogue for GLB of a set.

While these concepts are important, for example in Real Analysis, what is more important for applications of partially ordered sets are the objects defined next.

Definition 7.2.23. Let $(X, \leq)$ be a partially ordered set. We say that $x_0 \in X$ is a *maximal element* if $a \in X$ and $x_0 \leq a$ implies $a = x_0$.

Let us understand what we mean by a maximal element x_0 in X. It says that there is no element in X which is greater than x_0. That is, only element in X which is greater than or equal to x_0 is x_0 itself.

Contrast this with the maximum of a set. If a is the maximum of A, then all elements $x \in A$ satisfy $x \leq a$.

Definition 7.2.24. Let $(X, \leq)$ be a partially ordered set. An element $y_0 \in X$ is called a *minimal element* if $b \in X$ and $y_0 \geq b$ implies $y_0 = b$.

The couple of examples below will make the concepts clear.

Example 7.2.25. This is perhaps the best example to understand the concept of maximal elements. Let S be any set with at least two elements and $X = P(S) \setminus \{\emptyset, S\}$ with inclusion as the partial order on X. Then for any $t \in S$, $A = S \setminus \{t\}$ is a maximal element in X. To prove this, suppose $B \in X$ is such that $A \subseteq B$. We show that $B = A$. If $A \neq B$, then B has at least one element which is not in A. But the only element of S which is not in A is t. Hence $S = A \cup \{t\} \subseteq B$, which gives $B = S$. Since $S \notin X$, this cannot happen, and we get $B = A$.

Note that none of these maximal elements A is the maximum in X, since $\{t\} \in X$ and $\{t\} \not\subseteq A = S \setminus \{t\}$.

Can you show that each maximal element of X is of the form $S \setminus \{t\}$ for some $t \in S$? Thus X has as many maximal elements as X has elements. In particular, for $S = \mathbb{R}$, X has uncountably many maximal elements.

Can you find the minimal elements of X?

Exercise 7.2.26. Discuss the minimal and maximal elements in (1) $P(\mathbb{R})$, (2) $P(\mathbb{R}) \setminus \{\emptyset\}$, (3) $P(\mathbb{R}) \setminus \{\mathbb{R}\}$ all being ordered by inclusion.

Example 7.2.27. For $k \in \mathbb{N}$, $k \geq 2$, let H_k be the set of all integral multiples of k. Thus, $H_k := \{mk : m \in \mathbb{Z}\}$. Let $X := \{H_k : k \in \mathbb{N}, k \geq 2\}$. The inclusion relation is a partial order on X. What are the maximal elements of X?
Hint: Observe that $H_a \subseteq H_b$ iff $b|a$.

If you know of group theory or ring theory, the maximal elements of X are known as *maximal subgroups* (respectively, maximal ideals) of $\mathbb{Z}$, considered as a group (respectively, as a ring).

Exercise 7.2.28. Are there any minimal elements in Example 7.2.27? What are the maximal and minimal elements if we allow $k = 0, 1$ in Example 7.2.27? Are there maximum and minimum elements now?

Exercise 7.2.29. Consider the relation $a \preceq b$ iff a divides b on $\mathbb{N} \setminus \{1\}$. What are the minimal elements here? Does it have a maximal element?

Have you realized that a maximal (respectively, a minimal) element in X need not be a maximum (respectively, a minimum) from the previous examples?

Exercise 7.2.30. Show that the maximum of a partially ordered set X, if exists, is a maximal element of X. State and prove the analogous statement for the minimum and minimal elements.

Exercise 7.2.31. Show that in a totally ordered set, a maximal (respectively, minimal) element is a maximum (respectively, a minimum).

Exercise 7.2.32. Show that any finite partially ordered set has a maximal element. How about the existence of minimal elements? *Hint:* Use induction. (A finite partially ordered set may not have a maximum! See Example 7.2.35.)

Exercise 7.2.33. Show that any finite totally ordered set has a maximum. How about minimum? *Hint:* Use induction.

Definition 7.2.34. The Hasse diagram for a partially ordered set $(X, \leq)$ is drawn according to the following recipe:

(i) There is a vertex for each of the points of X.

(ii) If $x \leq y$, the y is positioned above x.

(iii) If $x < y$ and if there is no intermediary z such that $x < z < y$, then a line is drawn from x to y.

Example 7.2.35. Consider $X := \{2^m 3^n : m, n \in \mathbb{Z}_+,\ 1 \leq m+n \leq 4\}$ with the partial order as in 7.1.2 (3). The Hasse diagram for X is given in Figure 7.1.

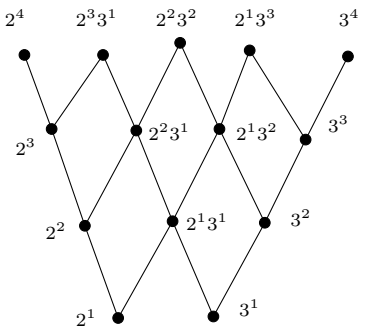


Figure 7.1: Hasse diagram for Example 7.2.35

It is clear from the Hasse diagram in Figure 7.1 that X has no maximum and minimum elements. However, it has four maximal elements and two minimal elements.

Exercise 7.2.36. Let $X = \{\{0\}, \{x\}, \{y\}, \{z\}, \{x, y\}, \{x, z\}, \{y, z\}, \mathbb{R}^3\}$ where $\{x\}$ denotes the x-axis, $\{x, y\}$ denote the xy-plane, etc. Define a partial order $\leq$ on X by inclusion. Look at the Hasse diagram for X in Figure 7.2.

Identify the maximum, minimum elements in X. If $Y := X \setminus \{\{0\}, \mathbb{R}^3\}$, what are the maximum, minimum, maximal and minimal elements in Y (if they exist)?

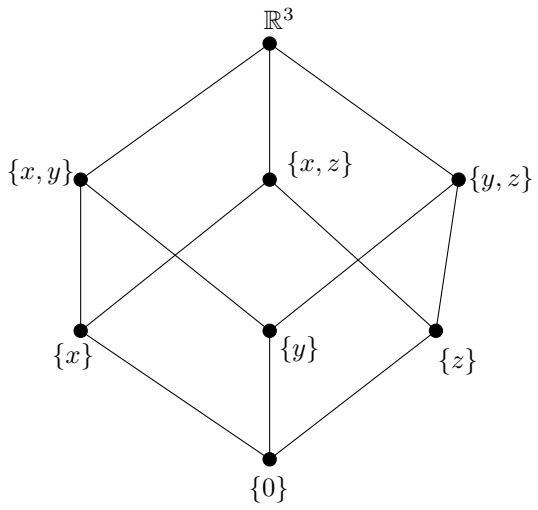


Figure 7.2: Hasse diagram for Ex.7.2.36

Exercise 7.2.37. Consider $X = P(\{a, b\})$ with partial order as inclusion and $Y = \{2, 3, 6\}$ with partial order defined by divisibility. Consider the partial order on $X \times Y$ as defined in the Ex. 7.1.6. Draw the Hasse diagram of partially ordered sets X, Y and $X \times Y$. Identify the maximum, the minimum, maximal and minimal elements (if they exist) in each of them.

Exercise 7.2.38. Let $X = \{1, 2, 3, 4, 6, 8, 12, 24\}$ be the set of positive divisors of 24. Define partial orders (i) $x \leq y$ iff x divides y, and (ii) $x \preceq y$ iff x is a multiple of y. Draw the Hasse diagrams for the two partial orders $\leq$ and $\preceq$. Can you compare the diagrams.

Exercise 7.2.39. Refer to Ex.7.2.35. Write down all the *maximal* chains in the partially ordered set. (What do you understand by a maximal chain?)

Exercise 7.2.40. The following exercises demand more background than we assume. You may skip them if you do not have the required background.

1. Discuss the minimal and maximal elements in (1) $\mathcal{H}_1$, the set of all nontrivial proper subgroups, (2) $\mathcal{H}_2$, the set of all proper subgroups and (3) $\mathcal{H}_3$, the set of all subgroups of the group $(\mathbb{Z}, +)$.

2. Find the LUB and GLB of the two element subset $\{m\mathbb{Z}, n\mathbb{Z}\}$ in the partially ordered set $\mathcal{H}$ of all subgroups of $(\mathbb{Z}, +)$.

3. Let $\mathcal{S}$ denote the set of nontrivial proper vector subspaces V of $\mathbb{R}^n$ (that is, $V \neq \{0\}$ and $V \neq \mathbb{R}^n$) where $n \geq 2$. For $V, W \in \mathcal{S}$, we define $V \leq W$ if $V \subseteq W$. Characterize the maximal and the minimal elements of $\mathcal{S}$.

4. Let V be a (not necessarily finite dimensional) vector space over $\mathbb{R}$. We say that $S \subset V$ is linearly independent if every finite subset of S is linearly independent. Let $\mathcal{S}$ be the set of all linearly independent subsets of V ordered by inclusion. Let B be a maximal element of $\mathcal{S}$. Show that the linear span of B is V.

5. Let V be a nonzero vector space over $\mathbb{R}$. Let $\mathcal{S}$ denote the set of all linearly independent subsets of V ordered by inclusion. Then the maximal elements of $\mathcal{S}$ are nothing other than the bases of V.

6. Let R be a commutative ring with identity 1. Let I be a maximal element in the set of all proper ideals of R ordered by inclusion. Show that for any $x \notin I$, there exist $r \in R$ and $z \in I$ such that $rx + z = 1$.

Exercise 7.2.41. Are there partially ordered sets in which every maximal element is a minimal element and vice-versa?

Example 7.2.42. Let X denote the set of all non-empty subsets of $\mathbb{R}$ which miss at least one integer, ordered by inclusion. Then X has countably infinite many maximal elements while it has uncountably many minimal elements.

In fact, (i) any set of the from $\mathbb{R} \setminus \{n\}$, $n \in \mathbb{Z}$, is a maximal element and (ii) any set of the form $\{x\}$, $x \in \mathbb{R}$ is a minimal element.

Exercise* 7.2.43. Let X be the set of all ordered pairs (f, S), where S is a subinterval of $[0, 1]$ such that $[1/2, 3/4] \subseteq S$ and $f\colon S \to \mathbb{R}$ is continuous such that $f(x) = 1/x$ for $x \in [1/2, 3/4]$. We say $(f, S) \leq (g, T)$ if $S \subseteq T$ and $g(x) = f(x)$ for $x \in S$ (i.e., g is an extension of f). Write down an infinite set of maximal elements of X.
Hint: Fix any t such that $0 \leq t < 1/2$. Define $g_t\colon (t, 1] \to \mathbb{R}$ by $g_t(x) = 1/x$ for $x \geq 1/2$ and $g_t(x) = \frac{1-2t}{x-t}$ for $x \in (t, 1/2)$. Then $\lim_{x\to t_+} g_t(x) = \infty$ and $\lim_{x\to 1/2} g_t(x) = 2 = g_t(1/2)$. Thus g_t is continuous and has no continuous extension. Therefore, g_t is maximal for each $t \in [0, 1/2)$.

Definition 7.2.44 (Intervals in a partially ordered set)**.** Let $(X, \leq)$ be a partially ordered set. Given $a, b \in X$ with $a \leq b$ we define the 'intervals' (mimicking Real Analysis) as follows:

$$\begin{aligned}
[a, b] &:= \{x \in X : a \leq x \text{ and } x \leq b\},\\
(a, b) &:= \{x \in X : a < x \text{ and } x < b\},\\
[a, b) &:= \{x \in X : a \leq x \text{ and } x < b\},\\
(a, b] &:= \{x \in X : a < x \text{ and } x \leq b\}.
\end{aligned}$$

Example 7.2.45. Let us give a geometric and explicit description of the interval $J = [(1, 0), (2, 1/2)]$ as a subset of $\mathbb{R}^2$ in the totally ordered set $(\mathbb{R}^2, \leq)$, where "$\leq$" is the dictionary order. We shall do this in several steps.
Step 1. When does $(1, y)$ lie in J? If $(1, y) \in J$, then $(1, 0) \leq (1, y)$, that is $y \geq 0$. Also $(1, y) \leq (2, 1/2)$ is valid for any y. Hence $(1, y) \in J$ iff $y \geq 0$. See Figure 7.3.
Step 2. Let $1 < x < 2$. When does (x, y) lie in J? Since $x > 1$, for any $y \in \mathbb{R}$, $(1, 0) < (x, y)$. Since $x < 2$ for any $y \in \mathbb{R}$, $(x, y) < (2, 1/2)$. Hence any (x, y), with $1 < x < 2$ and $y \in \mathbb{R}$ lies in J. See Figure 7.4.

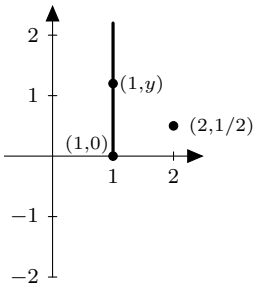


Figure 7.3: Step 1 of 7.2.45

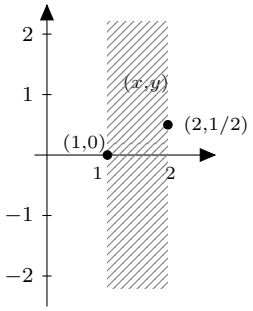


Figure 7.4: Step 2 of 7.2.45

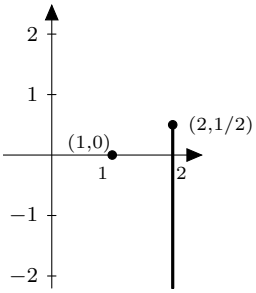


Figure 7.5: Step 3 of 7.2.45

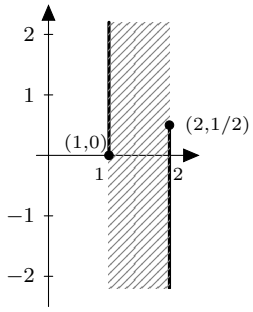


Figure 7.6: Step 4 of 7.2.45

Step 3. When does $(2, y) \in J$? Since $(1, 0) < (2, y)$, we need to ensure that $(2, y) \leq (2, 1/2)$. This is the case if $y \leq 1/2$. Hence $(2, y) \in J$ iff $y \leq 1/2$. See Figure 7.5.
Step 4. The union of sets in Steps 1-3 yields in Figure 7.6.
Step 5. If $(x, y) \in \mathbb{R}^2$ does not lie in the depicted in Figure 7.6 of Step 4, one of the following holds. (i) $x = 1, y < 0$, (ii) $x < 1, y \in \mathbb{R}$, (iii) $x = 2, y > 1/2$ and (iv) $x > 2, y \in \mathbb{R}$. It is easy to see that any such element (x, y) does not lie in J.

Hence we conclude that the picture for J is Figure 7.6.

Second approach to illustrate the interval $J = [(1, 0), (2, 1/2)]$.
Note that $J = \{(x, y) \in \mathbb{R}^2 : (1, 0) \leq (x, y) \leq (2, 1/2)\}$ is the intersection of the two sets

$$A := \{(x, y) \in \mathbb{R}^2 : (1, 0) \leq (x, y)\} \text{ and } B := \{(x, y) \in \mathbb{R}^2 : (x, y) \leq (2, 1/2)\}.$$

Step I. Identify the set A. We have $(x, y) \in A$ iff $(1, 0) \leq (x, y)$, that is, iff $x = 1, y \geq 0$ or $x > 1, y \in \mathbb{R}$. Therefore

$$A = \{(x, y) \in \mathbb{R}^2 : x = 1, y \geq 0\} \cup \{(x, y) \in \mathbb{R}^2 : x > 1, y \in \mathbb{R}\},$$

that is, A is as depicted in Figure 7.7.
Step II. Identify the set B. We have $(x, y) \in B$ iff $(x, y) \leq (2, 1/2)$, that is, iff $x < 2, y \in \mathbb{R}$ or $x = 2, y \leq 1/2$. Therefore

$$B = \{(x, y) \in \mathbb{R}^2 : x < 2, y \in \mathbb{R}\} \cup \{(x, y) \in \mathbb{R}^2 : x = 2, y \leq 1/2\},$$

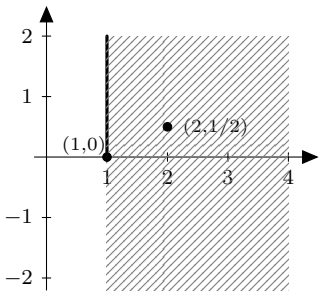


Figure 7.7: Step I of 7.2.45

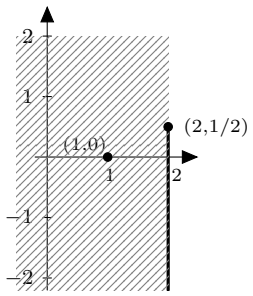


Figure 7.8: Step II of 7.2.45

that is, B is as depicted in Figure 7.8.

Step 3. The interval $J = A \cap B$ is as depicted in the Figure 7.6.

Exercise 7.2.46. Give a geometric and explicit description of the intervals (i) $[(1,0),(1,1))$ and (ii) $((1,0),(2,0)]$ in $\mathbb{R}^2$ with dictionary order.

Exercise 7.2.47. Give a geometric and explicit description of the intervals $[(a,b),(c,d)]$ in $\mathbb{R}^2$ with dictionary order.

Exercise 7.2.48. Find the following intervals in the partially ordered set $\mathbb{N}$, with divisibility as the partial order. (i) $(5,100)$, (ii) $[2,60]$, (iii) $[2,36)$ and (iv) $(3,75]$.

When is an interval $[a,b]$ in this partially ordered set a chain?

7.3 Axiom of Choice and its Equivalents

We conclude this chapter with three principles, namely, Well-ordering Principle, Zorn's Lemma and Axiom of Choice, which are widely used in mathematics assuming them to be true. We also give a few fallouts of these principles.

Definition 7.3.1. Let $(X,\leq)$ be a partially ordered set. We say that the order $\leq$ is a *well-ordering* on X if (i) it is a total order and (ii) every nonempty subset of X has a minimum. In that case, we say $(X,\leq)$ is a *well-ordered set.*

Example 7.3.2. The best example of a well-ordered set is $\mathbb{N}$ with the standard order.

Exercise 7.3.3. Show that any countable set X has a well-ordering. *Hint:* Use Ex. 7.1.16

Well-Ordering Principle

Principle 7.3.4. *On any nonempty set, there exists a well-ordering.*

Zorn's Lemma

Principle 7.3.5. *Let X be a partially ordered set. Assume further that the partial order is such that every chain in X has an upper bound in X. Then there exists a maximal element in X.*

Remark 7.3.6. The assumption on the partial order in Zorn's lemma is important. Consider $\mathcal{S} := \{S_t := (0,t) : 0 < t < 1\}$ ordered by inclusion. The order is total and hence $\mathcal{S}$ is a chain. It is easy to see that the chain has no upper bound in $\mathcal{S}$ and that $\mathcal{S}$ has no maximal element.

Exercise* 7.3.7. Show that any vector space over a field has a basis. *Hint.* Items 4, 5 of Ex. 7.2.40 and Principle 7.3.5 are relevant.

Axiom of Choice

Principle 7.3.8. *Let $\{X_i : i \in I\}$ be a nonempty family of nonempty sets. Then there exists a set A which has exactly one element from each X_i, $i \in I$.*

Remark 7.3.9. It can be proved that the axiom of choice, Zorn's lemma and the well-ordering principle are equivalent. Perhaps the most used of these is Zorn's lemma followed by the axiom of choice.

Cartesian product of a family of sets

In Chapter 2, we defined the Cartesian product a finite family of sets. Can we define the Cartesian product of an arbitrary indexed family of sets?

Definition 7.3.10. Let $\{X_i : i \in I\}$ be an indexed family of nonempty sets. Then the Cartesian product $X := \prod_{i\in I} X_i$ is defined by

$$\prod_{i\in I} X_i := \Big\{x\colon I \to \bigcup_{i\in I} X_i \mid x(i) \in X_i \text{ for each } i \in I\Big\}.$$

That is, $x \in X$ iff x is a function from I to $\bigcup_{i\in I} X_i$ such that $x(i) \in X_i$ for each $i \in I$.

It follows from Axiom of Choice that one can choose $x(i) \in X_i$ for each $i \in I$. In other words: The Cartesian product $\prod_{i\in I} X_i$ is not empty!

We usually write $x \in \prod_{i\in I} X_i$ as $x = (x_i)$, where $x_i := x(i)$. We shall call x_i as the i-th coordinate of x.

Let $\pi_j\colon \prod_{j\in I} X_i \to X_j$ be defined by $\pi_j(x) = x(j) = x_j$. This is called the j-th projection of X onto the j-th factor X_j.

As a convention, if $I = \{1, 2, \ldots, n\}$, we identify X with $X_1 \times \cdots \times X_n$, that is, with the set of ordered n-tuple $(x_1, \ldots, x_n)$. Similarly, if $I = \mathbb{N}$, we identify X with $X_1 \times X_2 \times \cdots \times X_n \times \cdots$, that is the set of ordered infinite tuples $x \mapsto (x_1, x_2, \ldots, x_n, \ldots)$.

If $V_j \subseteq X_j$, then what is $\pi_j^{-1}(V_j)$? We have $x = (x_i) \in \pi_j^{-1}(V_j)$ if and only if $\pi_j(x) = x_j \in V_j$, that is, if and only if $x = (x_j) \in X$ with $x_j \in V_j$ and $x_i \in X_i$ for $i \neq j$. Consequently, $\pi_j^{-1}(V_j) = \prod_{i\in I} U_i$ where $U_i = X_i$ for $i \neq j$ and $U_j = V_j$.

For example, if $X = X_1 \times X_2$ and $V_1 \subseteq X_1$, then $\pi_1^{-1}(V_1) = V_1 \times X_2$.

We now give a few examples to instill some confidence to work with the concept of Cartesian products.

Example 7.3.11. Let $I = \mathbb{R}$ and $X_t := \mathbb{R}$ for each $t \in I$. Then it is evident that $X := \prod_{t\in I} X_t$ is the set of all functions $f\colon \mathbb{R} \to \mathbb{R}$.

More generally, let X and Y be nonempty sets. Put $I = X$ and for each $x \in I = X$, $Y_x = Y$. Then $\prod_{x\in X} Y_x$ is the set of functions from X to Y. One usually denotes this by Y^X.

Example 7.3.12. Let $I = \mathbb{N}$ and $X_n := \mathbb{R}$. How do we visualize $\prod_{n\in I} X_n$? It is the Cartesian product of the vertical lines passing through the points $(n, 0)$ in the x-axis of $\mathbb{R}^2$.

Note that the product set is the set of all real sequences. This can be seen as in the previous example.

Example 7.3.13. Let $I := [0, \infty)$ and $X_t := [0, t]$ for $t \in I$. How do we visualize the product $\prod_{t\in I} X_t$?

Does the element $(t^2)_{t\in I}$ lie in the product? Note that if $t = 2$, then the $x(t) = 4 \notin X_t = [0, 2]$.

Note that the product is the set of functions $\{f\colon [0, \infty) \to \mathbb{R}\}$ satisfying, $0 \leq f(t) \leq t$, $t \in I$. The points $(t, f(t))$ lie in the region $\{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq x\}$.

Remark 7.3.14. Zorn's lemma is used in the following results (in a typical M.Sc. course):

1. Existence of a basis for any vector space.
2. Existence of a maximal ideal in a ring with unity.
3. Existence of algebraic closures of fields.
4. Tychonoff's theorem which asserts that the product of a family of compact spaces is compact in the product topology.
5. Hahn-Banach extension theorem in functional analysis.

Remark 7.3.15. The axiom of choice is used in the construction of a non-measurable set in the theory of Lebesgue measure.

Bibliography

[1] Ajit Kumar and S. Kumaresan, *A Basic Course in Real Analysis,* CRC Press, 2014.

[2] Robert G. Bartle and Donald R. Sherbert, *Introduction to Real Analysis*, John Wiley & Sons Inc., New York, 1972.

[3] James Munkres, *Topology,* 2nd Ed., Pearson Education (India), 2001.

[4] Paul R. Halmos, *Naive Set Theory*, Springer, 1974.

[5] A Shen and NK Vereshchagin, *Basic Set Theory*, AMS Students Mathematcal Library, 2002.

[6] Michael L. O'Leary, *A First Course in Mathematical Logic and Set Theory*, Wiley and Sons, 2016.

[7] Edwin Hewitt and Karl Stromberg, *Real and Abstract Analysis*, Springer, 1975.

[8] We also refer the reader to the expository articles written by S. Kumaresan. They are available for download from
`http://mtts.org.in/expository-articles` .

Index

About the Authors

Dr. Ajit Kumar is an associate professor at the Institute of Chemical Technology, Mumbai, India. His main interests are differential geometry, optimization and the use of technology in teaching mathematics. He received his Ph.D. from University of Mumbai. He has initiated a lot of mathematicians into the use of open source mathematics software.

Dr. Bhaba Kumar Sarma is currently a professor at Indian Indian Institute of Technology Guwahati. During his teaching of more than twenty five years he has taught at several institutions, both at undergraduate and postgraduate levels. He received his Ph.D. from the University of Delhi and works in Combinatorial Matrix Theory.

Dr. S Kumaresan is currently an NBHM visiting professor at University of Hyderabad. His initial training was at Tata Institute of Fundamental Research, Mumbai where he earned his Ph.D. He then served as a professor at University of Mumbai. His main interests are harmonic analysis, differential geometry, analytical problems in geometry, and pedagogy. He has authored six books, ranging from undergraduate level to graduate level. He was the recipient of the C.L.C Chandna award for Excellence in Teaching and Research in Mathematics in the year 1998. He was selected for the Indian National Science Academy Teacher award for 2013. He was a member of the Executive Committee of International Commission on Mathematics Instruction during the period 2007–2009.

For the last several years, authors have been associated with the Mathematics Training and Talent Search Programme, India, aimed at mathematics students at undergraduates and postgraduate levels in Universities in India.