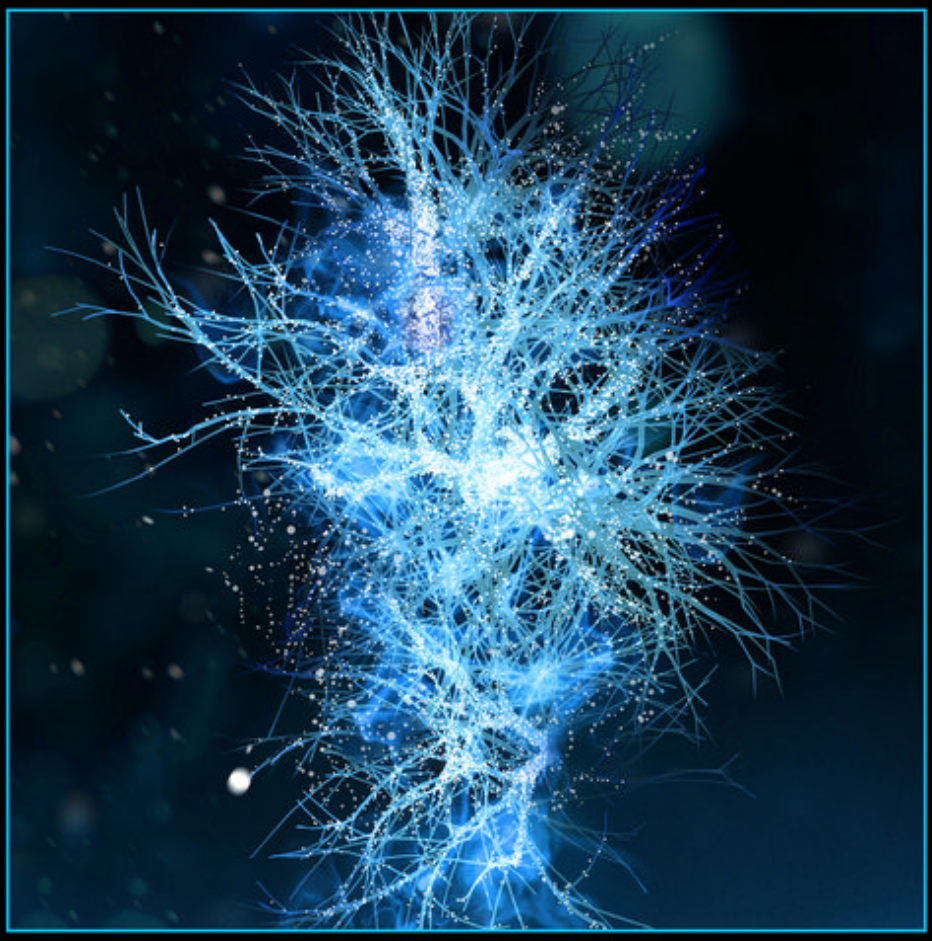
VLADIMIR LEPETIC

PRINCIPLES OF MATHEMATICS

A Primer



WILEY

PRINCIPLES OF MATHEMATICS

PRINCIPLES OF MATHEMATICS

A Primer

VLADIMIR LEPETIC
DePaul University

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Lepetic, Vladimir, 1950- author.
Principles of mathematics : a primer / Vladimir Lepetic.
pages cm
Includes index.
ISBN 978-1-119-13164-9 (cloth)
1. Mathematics–Philosophy. I. Title.
QA8.4.L447 2016
510.1–dc23
2015025151

Typeset in 10.5/12.5pt, Times by SPi Global, Chennai, India.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

1 2016

To Ivan and Marija

CONTENTS

PREFACE

I suspect that everyone working in academia remembers the days when as a student, being unhappy with the assigned textbook, they promised themselves one day to write the "right" book – one that would be easy and fun to read, have all the necessary material for successfully passing those (pesky) exams, but also be serious enough to incite the reader to dig further. It would, one hoped, open new venues to satisfy readers' curiosity provoked by frequent, but causal, remarks on the themes beyond the scope of an introductory text. Thus, the hope continues, students would be seduced by the intimation of mathematical immensity, sense the "flavor" of mathematical thinking, rather than just learning some "tricks of the trade." Of course, undertaking such an enterprise might easily be considered by many to be too ambitious at best and pure hubris at worst, for many potential readers are likely aware of the plethora of excellent books already available. Well then, why this book and not another one? First, this book is what its title says – a primer. However, unlike many other introductory texts, it is intended for a wide variety of readers; it (or parts of it) can be used as a starter for college undergraduate courses fulfilling "general education" math requirements, but also as an overture to more serious mathematics for students aspiring for careers in math and science. I am sure that rather smart high school students could also use the book to maintain and enhance their enthusiasm for mathematics and science. In any case, I am of the opinion that "introduction" and "rigor" should not exclude each other. Similarly, I don't think that avoiding discussion of weighty issues necessarily makes a text reader-friendly, in particular when mathematics and the sciences are in question. I do think, however, that many profound issues can be introduced accessibly to a beginner and, most importantly, in a way that provokes intellectual

curiosity and consequently leads to a better appreciation of the field in question. Another thing, in my opinion equally significant, that any introductory mathematics text should convey is the importance of recognizing the difference and mutual interconnectedness in "knowing," "understanding," and "explaining" (i.e., "being able to explain") something. Admittedly any book, even the most advanced one, is an introduction in such an endeavor; yet one has to start somewhere so why not with a "primer" like this. So, with such philosophy in mind – which by the way can also serve as an apologia, albeit not a very transparent one – all efforts have been made to meticulously follow well-established mathematical formalism and routines. Incidentally, contrary to some educators, I don't think that the standard mathematical routine "definition–lemma–theorem–proof–corollary" is necessarily a deterrent to learning the subject. The only way one can see the "big picture" is to acquire a unique technique that will empower one to do so. The real beauty reveals itself after years of study and practice. If you want to successfully play a musical instrument you need to learn to read the notes, learn some music theory, and then relentlessly practice until you reach a reasonable command of your instrument. After years of doing so, and if you are lucky enough, you become an artist. The payoff, however, is immense.

The book contains six chapters and literally hundreds of solved problems. In addition, readers will find that every chapter ends with a number of supplementary exercises. It is my experience that in a one-semester course with class meeting twice a week, an instructor can leisurely cover two chapters of his/her choice, and they would still have ample time to pick and choose additional topics from other chapters they deem interesting/relevant. The first chapter contains a fairly detailed introduction to Set Theory, and it may be also considered as a conceptual/"philosophical" introduction to everything that follows. Chapters 2 and 3 on Logic and Proofs follow naturally, although I would not have many arguments against those who would prefer to start with Chapters 2 and 3 and subsequently discuss Set Theory. Readers interested only in Functions and/or Group Theory can, after Chapter 1, immediately jump to Chapters 4 and/or 5. Similarly, the last chapter on Linear Algebra can be approached independently provided, of course, the reader has been at least briefly introduced to the necessary preliminaries from Chapters 1, 4, and 5. Finally, the case could be made that a text of this kind, in order to justify an implicitly hinted philosophy, should necessarily have a chapter on Topology and Category Theory. I wholeheartedly agree. However, that would require adding at least another 300 pages to this primer, and the sheer volume of such a book would likely be a deterrent rather than an enticement for a beginner and thus defeat its very purpose. Postponing topology and category theory for some later time hopefully will be just a temporary weakness. It is not unreasonable to expect that after carefully going through Sets and Functions, for instance, the reader will anticipate further subtleties in need for clarification and reach for a book on topology. Similarly, those wondering about a possible theory that would subsume all others might find Category Theory an appropriate venue

to reach that goal. In any case, provoking intellectual curiosity and imagination is the main purpose of this text, and the author wishes that the blame for any failure in this endeavor could be put, or at least partially placed, on the shoulders of the course instructor. Alas, the shortcomings are all mine.

I am aware that it is impossible, and no effort would be adequate, to express my gratitude to all of my teachers, colleagues, and students who have over decades influenced my thinking about mathematics. This impossibility notwithstanding, I need to mention Ivan Supek who, at my early age, introduced me to the unique thinking about mathematics, physics, and philosophy, coming directly from Heisenberg, whose assistant and personal friend he had been for years. Vladimir Devidé who discovered for me the world of Gödel and, many years later, my PhD adviser, Louis Kauffman, who put the final touch on those long fermenting ideas.

Lastly, I want to thank Ivan Lepetic, who painstakingly read the whole manuscript, made many corrections and improvements, and drew the illustrations.

VLADIMIR LEPETIC

1

SET THEORY

> "The question for the ultimate foundations and the ultimate meaning of mathematics remains open; we do not know in which direction it will find its final solution nor even whether a final objective answer can be expected at all.
>
> "Mathematizing" may well be a creative activity of man, like language or music, of primary originality, whose historical decisions defy complete objective rationalization."
>
> H. Weyl[1]

1.1 INTRODUCTION

The fact that you chose to read this book makes it likely that you might have heard of Kurt Gödel,[2] the greatest logician since Aristotle,[3] whose arguably revolutionary discoveries influenced our views on mathematics, physics, and philosophy,

[1] Hermann Klaus Hugo Weyl (1885–1955), German mathematician, Yearbook of the American Philosophical Society, 1943 (copyright 1944).
[2] Kurt Gödel (1906–1978), Austrian–American logician, mathematician, and philosopher.
[3] J.A. Wheeler said that "if you called him the greatest logician since Aristotle you'd be downgrading him" (quoted in Bernstein, J., *Quantum Profiles*, Princeton University Press, 1991. Also in Wang, H. *A Logical Journey,* MIT Press, 1996).

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

comparable only to the discoveries of quantum mechanics. Well, even if you have not heard of him I want to start by rephrasing his famous theorem:

Mathematics is inexhaustible!

Notwithstanding the lack of a definition of what mathematics is, that still sounds wonderful, doesn't it? At this point, you may not fully understand the meaning of this "theorem" or appreciate its significance for mathematics and philosophy. You may even disagree with it, but I suppose you would agree with me that mathematics is the study of abstract structures with enormous applications to the "real world." Also, wouldn't you agree that the most impressive features of mathematics are its certainty, abstractness, and precision? That has always been the case, and mathematics continues to be a vibrant, constantly growing, and definitely different discipline from what it used to be. I hope you would also agree (at least after reading this book) that it possesses a unique beauty and elegance recognized from ancient times, and yet revealing its beauty more and more with/to every new generation of mathematicians. Where does it come from? Even if you accept the premise that it is a construct of our mind, you need to wonder how come it represents/reflects reality so faithfully, and in such a precise and elegant way. How come its formalism matches our intuition so neatly? Is that why we "trust" mathematics (not mathematicians) more than any other science; indeed, very often we define truth as a "mathematical truth" without asking for experimental verification of its claims? So, it is definitely reasonable to ask at the very beginning of our journey (and we will ask this question frequently as we go along): Does the world of mathematics exist outside of, and independently of, the physical world and the actions of the human mind? Gödel thought so. In any case, keep this question in mind as you go along – it has been in the minds of mathematicians and philosophers for centuries.

The set theory that we start with comes as a culmination of 2000 years of mathematics, with the work of the German mathematician George Cantor[4] in the 1890s. As much as the inception of set theory might have had (apparently) modest beginnings, there is virtually no mathematical field in which set theory doesn't enter as the very foundation of it. And it does it so flawlessly, so naturally, and in such a "how-could-it-be-otherwise" way, that one wonders why it took us so long to discover it. And arguably, there is no concept more fundamental than the concept of the set. (Indeed, try to answer the question: What is a real number without referring to set theory?) Be it as it may, now we have it. We start our journey through the "Principles," with the basic formalism of set theory.

No one shall be able to drive us from the paradise that Cantor created for us.[5]
D. Hilbert

[4]Georg Ferdinand Ludwig Philipp Cantor (1845–1918), German mathematician, the "father" of Set Theory.
[5]David Hilbert (1862–1943), German mathematician.

1.2 SET THEORY – DEFINITIONS, NOTATION, AND TERMINOLOGY – WHAT IS A SET?

You are probably familiar with the notion in mathematics of a set as a collection, an aggregate or a "group"[6] of certain "(some)things," or a collection of certain "objects"[7] that form a whole. We assume the existence of some domain of those "objects," out of which our mind will build a "whole." Cantor suggested that one should *imagine a set as a collection into a whole A of definite and separate objects of our intuition or our thought.* These objects are called members or elements of a set. For example, we can consider the set of all planets in the solar system,[8] or the set of all living people on Earth. Or, we can consider the set of all living females on this planet. Those would be well-defined sets, and by the very "definition," that is, the description of the set, our mind effortlessly constructs the concept of a "whole." On the other hand, calling for a set of all tall men, or a set of all big planets, triggers a similar concern. What is "a tall man" or "a big planet?" Obviously, describing a set of real objects by means of their characteristics can be problematic due to the imprecision of everyday language. So, it is fair to say that once the nature of objects defining a set is unambiguously stated, the whole entity, and not the individual elements, becomes the object of our study. Consequently, what we care about is the relationship between different sets as well as the very consistency of the "set" concept.

As you can see, at the very beginning of our discussion, we are introducing a concept that looks, to say the least, pretty vague, especially since we are doing mathematics, which we expect to be the epitome of precision. So, at this point in the process of devising our theory – *The Naïve Set Theory* – we will use the words "*set*" and "*is an element of*" without properly defining them. We will simply assume that we know exactly what they mean and hope that we won't run into any inconsistencies and paradoxes. In addition, we need the basic logical vocabulary consisting of "*not,*" "*and,*" "*or,*" and "*if … then …*." That's it! With so little, how can one satisfy the credo of modern mathematics – a "philosophy" by the name of Cantorism – that *everything (mathematical) is a set*? This idea is not as outlandish as you may think, so I suggest you wait for a while before deciding whether to accept this doctrine or not. Remember the Pythagoreans[9] who thought that everything is a natural number. You can imagine their dismay upon

[6]To be precise, we want to make sure that here by the "group" we do not mean the mathematical term "group" as in Group Theory, but simply a group of certain objects or elements.

[7]The term "object" could be misleading too, for sometimes by the "object" people instinctively think of "(some)thing" that is, a "thing" that can be touched, seen, and so on. Since objects of a set theory can be ordinary things, like pencils, chairs, people, or animals, and they can also be very abstract in nature, like numbers, functions, and ideas, maybe the term "entity" instead of the "object" would be more appropriate.

[8]Of course, "all" in this case, by mathematical standards, might be a somewhat imprecise quantifier, but let's assume at this point that there will be no surprises of striping off a "planetary status" of an object in our solar system, as we have recently witnessed in the case of Pluto.

[9]Religious sect founded by Pythagoras of Samos (ca. 570–495) Ionian–Greek philosopher.

learning of the incommensurability of the side and the diagonal of a square. The discovery of $\sqrt{2}$ must have been a catastrophe for this secluded sect, let alone the pain of disclosing the findings to the uninitiated. Legend has it that for his unfortunate discovery Hippasus[10] was drowned by the members of this mystic brotherhood. Later, we learned about certain other sets of numbers – the set of real numbers, for instance – which is fundamentally more "infinite" than anything we knew before. To understand those we definitely need sets.[11] We may continue on this rather vague path and also say that a set is a "thing" that is a collection of other things (which themselves could be sets) called the elements of the set. These hazy definitions by synonym suffice for most purposes, for our mind is able to grasp (the essence of) the concept regardless of the abstractness of the definition. Indeed, we want these concepts to be sufficiently abstract in order to avoid contradictions, especially when dealing with the foundation of mathematics. At the same time, very few so "simple" ideas in mathematics proved to be so fecund with the repercussions to almost all fields of mathematics. Not surprisingly, Mathematical Logic and Philosophy of Mathematics in particular became exceptionally interesting and rich fields notwithstanding the paradoxes spurred by much ingenious work on the foundations of mathematics and set theory.

So, before we start with the formalism of set theory, I want to tell you something rather funny and interesting, something that will keep showing up over and over again in the foundation of mathematics. This will certainly provoke some curiosity in you and at the same time show you the richness of ideas that set theory contains, and how our mind detects paradoxes in apparently simple concepts – concepts that this very mind came up with. The following is known as the Russell[12] Paradox. (Remember, the notion of "*elementhood*" or "*membership*" does not prevent us from thinking of sets as being elements of (i.e., belonging to) other sets.) So, let's follow Cantor and imagine all the *definite distinguishable concepts of your/our intellect.* One of them could be the idea of unicorns – it doesn't matter that you/we know they don't "exist." (They do exist in your mind, right?) Well, let's think about the collection of definite concepts of our intellect that doesn't contain itself. Let me explain. It is easy to think of, say, a set of all horses (or unicorns if you wish) on Earth. This set obviously represents a set that does not contain itself as a member. A set of horses is not a horse, of course. Now, can you think of a set that would be a member of itself? How about a set of all ideas? It is an idea, right? So is it a member of itself? Or, how about a set of all sets? It is a reasonable idea too. But, it is again also a set, hence a member of itself. Well, let's think about it. Let's call any set that doesn't contain itself

[10]Hippasus of Metapontum (ca. fifth century BC), Pythagorean philosopher.

[11]Could it be that even sets are not "everything"? Well, yes! It is possible that we may need an even more fundamental structure to address, among other things, the even "greater," Absolute Infinities. The discussion of those we leave for some other time.

[12]Bertrand Arthur William Russell (1872–1970), British philosopher, logician, and mathematician.

as one of its elements an ordinary set, say, $\mathcal{O}$ and the one that does – an extraordinary, $\mathcal{E}$. Now, here is what Russell said: Consider a set of all ordinary sets $\mathcal{O}$. It exists – Cantor said so – since it is a distinguishable concept of one's intuition or one's thought. So we could safely claim:

1. $\mathcal{O}$ is an ordinary set!
   Suppose not. Suppose it is extraordinary and thus contains itself as one of its elements. But every set in $\mathcal{O}$ is ordinary. Thus $\mathcal{O}$ is ordinary. But this is a contradiction! Therefore, our assumption was wrong; $\mathcal{O}$ is definitely ordinary. Well, is it? No!? What if we say:
2. $\mathcal{O}$ is an extraordinary set!
   Suppose not. Suppose $\mathcal{O}$ is ordinary. Since $\mathcal{O}$ contains all ordinary sets, it has to contain itself as one of its members. But that makes it extraordinary. This is a contradiction. Our assumption that $\mathcal{O}$ is ordinary was wrong. Therefore, $\mathcal{O}$ is extraordinary.

Obviously (1) and (2) are contradictory.
Here is another well-known example of a finite set, which we cannot properly make out[13]:

Consider two sets of adjectives: set $\mathcal{A}$ of self-descriptive adjectives we call *autologous* (*autological)* and set $\mathcal{H}$ of nonself-descriptive adjectives, called *heterologous* (*heterological*), that is, the set of all adjectives not belonging to $\mathcal{A}$. For example, set $\mathcal{A}$ contains adjectives such as *English*, *finite*, *derived*, and *pentasyllabic*. That is, they do have the properties they describe. On the other hand, set $\mathcal{H}$ contains adjectives such as *German*, *French*, *black, white*, and *monosyllabic*, that is, obviously none of them belongs to $\mathcal{A}$. Now, what about "*heterologous*"? Which set does it belong to? What I am asking is this: Is "*heterologous*" heterologous?

If this sounds confusing to you, and it's perfectly all right if it does, for it is confusing indeed. Here is Russell again with an analogous "story" (and I assure you this is not some silly game of words) to help us out:

> *There is a small town with only one (strange) barber. The strange thing about him is that he shaves all men in town that do not shave themselves. Now, does he shave himself or not?*

So, what are we to make of it? At the very beginning, we are dealing only with two concepts, "*set*" and "*an element of*," and we are faced with a fundamental problem of definitions that seems irresolvable. We cannot allow a seed of contradictions sitting at the very concept we want as our foundation. How do we start?

[13]Due to Kurt Grelling (1886–1942) and Leonard Nelson (1882–1927), German mathematicians and philosophers.

How do we build a fundamental structure of mathematics, a structure precise enough and rich enough, to encapsulate "all of mathematics" and all the rules of inference, without contradictions and without any ambiguities? Mathematicians and philosophers have been thinking about these questions for thousands of years, going back to Euclid's[14] axiomatic treatment of geometry, to Leibnitz's[15] ideas of mathematical logic, to Hilbert's dream of unifying all of mathematics under the umbrella of a formal axiomatic system, to the works of Cantor, Russell, and Whitehead,[16] and many others. In any case, the theory that Cantor developed, indeed a mathematical theory unlike any before, proved to be the best candidate to fulfill that. Mathematics arose on a system of axioms and precise formalism, which we want to be

1. consistent;
2. complete; and
3. decidable.

That a formal system is "consistent" means that we should not be able to prove, in finitely many steps, an assertion and its negation at the same time. *A* and *not-A* cannot (should not) be true at the same time. By "complete" we mean a system that is rich enough to allow us to determine whether *A* or *not-A* is a theorem, that is, a true statement. And finally, "decidable" refers to what is known as "the decision problem" (the famous "Entscheidungsproblem" in German), that is, a procedure, an algorithm by which we can (always) determine, in a finite number of steps, whether something is a theorem or not. That's what we want. Not much to ask for, wouldn't you say? After all, consistent and complete should imply that a decision procedure is at hand. Well, it's not. It can't be done! Mr Gödel said so.[17]

Here is how Hilary Putnam[18] "illustrates" Gödel's theorem:

(i) *That, even if some arithmetical (or set-theoretical) statements have no truth value, still, to say of any arithmetical (or set-theoretical) statement that it has (or lacks) a truth value is itself always either true or false (i.e. the statement either has a truth value or it doesn't).*

[14]Euclid (of Alexandria) (ca. 325–270 BC), Greek mathematician/geometer.
[15]Gottfried Wilhelm Leibniz (1646–1716), German mathematician and philosopher.
[16]Alfred North Whitehead (1861–1947), British mathematician, logician, and philosopher.
[17]"The human mind is incapable of formulating all its mathematical intuitions, that is, if it has succeeded in formulating some of them, this very fact yields new intuitive knowledge, for example, the consistency of this formalism. This may be called the 'incompletability' of mathematics." Kurt Gödel, *Collected Works,* Oxford University Press, 2001.
[18]Putnam, H., *Mathematics Without Foundation,* in Philosophy of Mathematics, 2nd ed., Cambridge University Press, 1983.

(ii) *All and only decidable statements have a truth value.*

For a statement that a mathematical statement $\mathcal{S}$ is decidable may itself be undecidable. Then, by (ii), it has no truth value to say "$\mathcal{S}$ has a truth value" (in fact falsity; since if $\mathcal{S}$ has a truth value, then $\mathcal{S}$ is decidable, by (ii), and if $\mathcal{S}$ is decidable, then "$\mathcal{S}$ is decidable" is also decidable). Since it is false (by the previous parenthetical remark) to say "$\mathcal{S}$ has a truth value" and since we accept the equivalence of "$\mathcal{S}$ has a truth value" and "$\mathcal{S}$ is decidable", then it must also be false to say "$\mathcal{S}$ is decidable". But it has no truth value to say "$\mathcal{S}$ is decidable". Contradiction.

Did you get it? Think about it. It literally grows on you. The whole point of all of "this" is that you start getting a "feel" for what mathematics really is and where we are actually "going." Anyway, after this "warm-up," let's start slowly and from the beginning.

First, we assume that there is a *domain,* or *universe* $\mathcal{U}$, of objects, some of which are sets.

Next, we need the formalism in which all our statements about sets can be precisely written – let's call it ***the language of set theory*** $\mathcal{L}$. This formal language contains a specific *alphabet,* that is, a list of symbols that we judiciously use and a number of specific statements that are called *axioms.* What are they? Well, in order to start somewhere and in order to avoid an infinite regress, we choose (there has to be (?)) a set of propositions that are not proved (not provable) but can be used in sound construction of our formalism. In addition, we create a basis for (all?) mathematics, which is inherently beautiful, and thus we can use it as an aesthetical criterion that all other sciences can measure up to. Similarly, there exists a collection of (mathematical) words or symbols that we do not define in terms of others – undefined does not mean meaningless – but simply take as given. Those we call primitives. This idea is as old as mathematics itself. Remember Euclid? The first lines of his *Elements* read as follows:

1. A point is that which has no parts.
2. A curve is length without width.
3. The extremity of a curve is a point.
4. A surface is that which has only a length and a width.
5. The extremity of a surface is a curve, and so on.

Surely, you feel some uneasiness about these statements. Still, the whole gigantic structure of Euclidean geometry, unquestioned for 2000 years, is based on these axioms. Putting aside the controversy among mathematicians on how fundamental these axioms are in general, as well as the question of their effectiveness, these axioms are needed and they are here to stay.

We also need the *formal rules of inference* so that the *language* we use is precise enough to derive all the theorems of our theory.

In addition to the aforementioned four basic symbols, we will soon need some more. So, we list the somewhat extensive alphabet of the language we are going to use throughout the book:

$\in$: element; a member; $x \in A$: x is an element of set A
$\notin$: not an element; not a member; $x \notin A$: x is not an element of set A
$\ni$: such that; sometimes "s.t."
^{c} : complement; A^c: complement of set A
$\setminus$: difference; $A \setminus B$: A difference B; sometimes just: A "minus" B
Δ : symmetric difference: $A \Delta B$: symmetric difference of A and B
$\subseteq$: subset; $A \subseteq B$: A is a subset of B
$\subset$: proper subset: $A \subset B$
$\cap$: intersection: $A \cap B$
$\cup$: union: $A \cup B$
$\varnothing$: the empty set
$\times$: Cartesian product; $A \times B$: Cartesian product of sets A and B
$\mathbf{N}$: the natural numbers
$\mathbf{Z}$: the integers
$\mathbf{Q}$: the rational numbers
$\mathbf{R}$: the real numbers
$\mathbf{Z}^+$: the nonnegative integers
$\mathbf{Q}^+$: the nonnegative rational numbers
$\mathbf{R}^+$: the nonnegative real numbers
$|A|$: the cardinal number (cardinality) of A
$\forall$: for all; for every; for any; $\forall x \in A$: for every x from A
$\exists$: there exists
$\exists!$: there exists a unique …
$\nexists$: (same as $\sim \exists$) does not exist
$\wedge$: and; sometimes also "&"
$\vee$: or
$\rightarrow$: "conditional"; "implication"; $a \rightarrow b$ if a then b. Sometimes same as "$\Rightarrow$"
$\leftrightarrow$: "biconditional"; $a \leftrightarrow b$ a if and only if b; "iff"; Sometimes same as "$\Leftrightarrow$"
$\sim$: "negation"; "it is not the case that"; "opposite of"
$=$: equal
$\equiv$: equivalent
iff: "if and only if"; "$\Leftrightarrow$"; "$\leftrightarrow$"

Definition 1.1 A set is said to be a well-defined set iff there is a method of determining whether a particular object is an element of that set.

The precise "description" of a set and its elements is based on the following axioms.

Axiom 0 (Set Existence)[19] There exists a *set*, that is, $\exists A\ (A = A)$. In other words, we postulate that there exists something, a "thing," an entity, we call a *set.*

Once a set A is given, we say that "*a is an element of A*" or that "*a is a member of A,*" and we write $a \in A$. Similarly, if a is not a member of A, we simply write $a \notin A$.

It is worth mentioning again that the expression "*an element of,*" that is, an elementhood relation, is also the elemental concept for which it is difficult to find a suitable alternative, so we also take it as an undefined predicate.

Example 1.1

$$A = \{a, b, c, d, e, f\}$$

is a set whose elements are $a, b, c, d, e, f,$ that is, $a \in A, b \in A, c \in A,$ and so on. This is nicely illustrated by the Venn diagram (Figure 1.1).

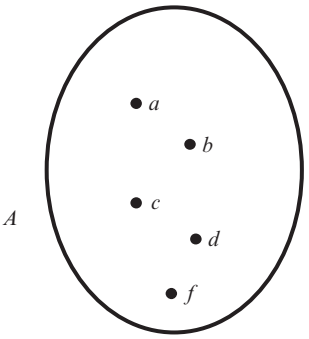


Figure 1.1 Venn diagram

Often it is convenient, especially when it is impossible to list all the elements of a set, to introduce a set using the so-called set-builder notation. We write

$$A = \{x | P(x)\}$$

and we read: A is a set of all x, such that $P(x)$, where $P(x)$ designates some property that all x'*s* possess, or P is a condition that specifies some property of all objects x.

[19] We will have more to say about these axioms later.

For instance, if we want A to be a set of all natural numbers greater or equal to 5 we write:

$$A = \{x |\, x \geq 5, \quad x \in \mathbf{N}\}$$

Certainly nothing prevents us from considering a set whose elements are also sets. In other words, we can have a set $X = \{x, y, w, z\}$, where x, y, w, and z are sets themselves. ■

Example 1.2 Suppose we consider

$$X = \{\text{Alice}, \text{Bob}\}$$

as a set whose two elements are persons Alice and Bob. *Set X* is definitely different from set, say,

$$Y = \{\text{Alice}, \{\text{Bob}\}\}$$

which also has two elements, but this time the elements are: Alice and $\{Bob\}$, that is, the element $\{Bob\}$ is itself a set containing one element – *Bob*.

Formally, we write:

$$\text{Alice} \in Y, \text{Bob} \notin Y, \text{but} \{\text{Bob}\} \in Y$$

Of course, we could have constructed a set

$$Z = \{\{\text{Alice}, \{\text{Bob}\}\}\}$$

which has only one element, namely, Y. Do you see why? It may help if we represent sets by Venn diagrams, where $X, Y,$ and Z (Figure 1.2) look as follows:

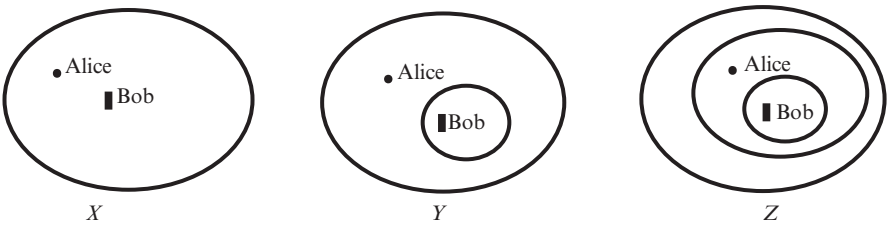


Figure 1.2 Sets *X, Y,* and *Z*

■

Axiom 1 (Axiom of extensionality) A set is uniquely determined by the elements it contains, that is, two sets are considered equal if they have the same elements. Less clearly but often said: a set is determined by its extension.

Example 1.3 Sets $A = \{a, b, c, d\}$ and $B = \{d, a, a, a, b, c, c, d\}$ are considered the same, that is, we say that $A = B$. ■

So, we have

Definition 1.2 Given sets A and B, we say that A equals B, and we write $A = B$ if and only if every element of A is an element of B and every element of B is an element of A. For the sake of completeness and more precision (at this point maybe prematurely[20]), using formal logic notation, we express this as follows:

$$A = B \leftrightarrow (\forall x)(x \in A \leftrightarrow x \in B)$$

Definition 1.3 Given two sets A and B, we say that A is a subset of B, and we write $A \subseteq B$ if and only if every element of A is also an element of B (Figure 1.3), that is,

$$A \subseteq B \leftrightarrow (\forall \ x \in A, x \in B)$$

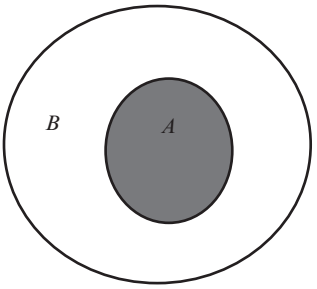


Figure 1.3 Subset $A \subseteq B$

Note that B could be "larger" than A, that is, that all elements of A are elements of B, but not all elements of B are elements of A. To distinguish between these subtleties, we state the following

Definition 1.4 Given two sets A and B, we say that A is a proper subset of B, $A \subseteq B$, if and only if every element of A is an element of B, but not all elements of B are elements of A.

Equality of sets can now be restated as

$$A = B \leftrightarrow A \subseteq B \& B \subseteq A$$

[20]This formalism will become more clear after you have studied Chapter 2.

What we are saying here is that two sets are considered equal solely on the basis of their elements (i.e., what's in the sets and how many) and not on the "arrangement" or a repeat of some of the elements in the respective sets.

Example 1.4 Show that, if a set A is a set of all integers n, where every n is expressible as $n = 2p$, with $p \in \mathbf{Z}$, that is,

$$A = \{n \in \mathbf{Z} | n = 2p, \quad p \in \mathbf{Z}\}$$

and B analogously described as

$$B = \{m \in \mathbf{Z} \, | m = 2q - 2, \quad q \in \mathbf{Z}\}$$

then $A = B$.

Solution Set A is the set of all even integers. We would like to see whether any integer of the form $2p$, for some $p \in \mathbf{Z}$, can also be written in the form $2q–2$, for some $q \in \mathbf{Z}$. Suppose there is an $n \in \mathbf{Z}$, such that $n = 2p$, for some integer p we want to find an integer q, so that $n = 2q - 2$. Thus,

$$\begin{aligned} 2q - 2 &= 2p \\ 2q &= 2p + 2 = 2(p + 1) \\ q &= p + 1 \end{aligned}$$

Therefore, for $n = 2p$, and $p \in \mathbf{Z}$, $q = p + 1$. It follows that

$$2q - 2 = 2(p + 1) - 2 = 2p + 2 - 2 = 2p$$

Hence, $A \subseteq B$.

Let's now assume that an integer can be expressed as $m = 2q - 2$, for some $q \in \mathbf{Z}$. Suppose, furthermore, that

$$2p = 2q - 2 = 2(q - 1)$$

that is,

$$p = q - 1$$

So, if $m = 2q - 2$, with $q \in \mathbf{Z}$, we write

$$2p = 2(q - 1) = 2q - 2$$

We conclude that $B \subseteq A$. Since $A \subseteq B$ and $B \subseteq A$, it follows that $A = B$ by definition of set equality. ■

Example 1.5 Let A be a set of all solutions of the equation $x^2 = 2x$, and let B be a set of all solutions of the equation $(x-1)^2 = 1$. Then, it is easy to see that $A = B$. ■

Axiom 2 (Comprehension axiom)[21]

(i) For any *reasonable*[22] property P, there exists a set containing exactly those elements that are defined by that property; In particular, mathematical entities that have a certain property in common constitute a set.

Certainly nothing prevents us from considering a set whose elements are also sets. In other words, we can have a set $X = \{x, y, w, z\}$, where x, y, w, and z are sets themselves. So we postulate:

(ii) Sets are mathematical entities, and, hence, they may in turn appear as elements of a set.

This is one of the reasons why one should not restrict oneself on a style of letters that represent sets. Thus, although we will most frequently use capitals to designate sets, occasionally it will be more convenient to use lowercase letters.

Example 1.6 Let $x_1, x_2, \ldots, x_n$ be a collection of n sets, then

$$X = \{x_1, x_2, \ldots, x_n\}$$

is also a set (Figure 1.4).

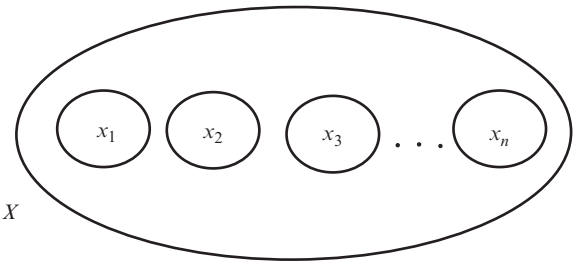


Figure 1.4

[21] This is sometimes called the Comprehensive principle.

[22] What is "*reasonable*" is debatable, and in any case a rather vague concept. We won't be discussing these subtleties here.

Having elements of a set being sets themselves gives us more flexibility in dealing with only one kind of object. Thus, we don't need to postulate the existence of every possible element of the various structures we intend to study. It follows, let's emphasize this again, that every set x is a unique element of another set, namely, $\{x\}$.

After accepting the fact that the elements of a set are sets, let's take a closer look at Axiom 2: Let X be a set, and let Y be a set whose elements are exactly those elements $x \in X$ with a property P, that is,

$$Y = \{x \in X | P(x)\}$$

So, let the particular property be $x \notin x$. (Remember, x is a set.) In other words, whatever set X may be, if

$$Y = \{x \in X | x \notin x\}$$

then for every y,

$$y \in Y \quad \text{iff} \quad y \in X \quad \text{and} \quad y \notin y \tag{*}$$

Is it possible that $Y \in X$? Let's see. If $Y \in X$, we have two possibilities: either $Y \in Y$ or $Y \notin Y$. Suppose $Y \in Y$. Then, from $Y \in X$ and (*) it follows that $Y \notin Y$ – obviously a contradiction. Suppose that $Y \notin Y$. Then, again, from $Y \in X$ and (*) it follows that $Y \in Y$ – a contraction again. We conclude that it is impossible that $Y \in X$. (You may remember this argument from before.)

Now, let me digress a bit and say something about two very important concepts that will be discussed in much more detail in Chapter 4. Many readers are familiar with the concepts of *relations* and *function*: For the time being, let's just say that:

A *relation* R is uniquely determined by pairs of elements x and y that are somehow related.

A *function* $f : X \to Y$ is uniquely determined by the pairs of two objects, an argument $x \in X$ and a functional value $f(x) \in Y$.

Now let's look at these via Axioms 1 and 2: For instance, the usual relation $\leq$ on the set of natural numbers describes a particular property, so we can construct a set R consisting of pairs of natural numbers (a, b) where $a \leq b$, that is

$$R = \{(a, b) | a \leq b, a, b \in \mathbf{N}\}$$

Similarly, we think of a function f as the following set of pairs:

$$f = \{(x, f(x)) | x \in X, f(x) \in Y\}$$

■

Axiom 3 There exists a set $\mathcal{U}$, called the universal set, such that for all sets A, if x is an element of A, then x is an element of $\mathcal{U}$ (Figure 1.5). Symbolically,

$$\forall A \ (x \in A \rightarrow \ x \in \mathcal{U})$$

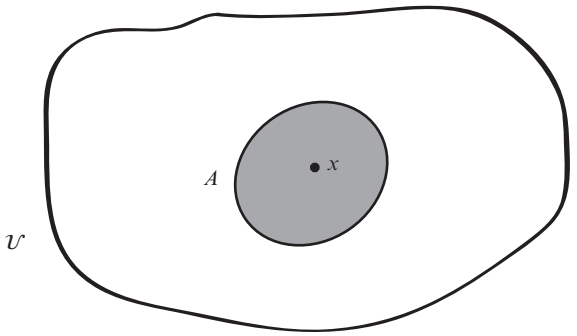


Figure 1.5

Axiom 4 If $x \in \mathcal{U}$ and A is a set, the statement $x \in A$ is a proposition that can either be true or false, but not both.

Saying something so "obvious" is not that trivial, as will become evident shortly.

Example/Exercise 1.7

(i) Is $a = \{a\}$?
(ii) Is $a \in \{a\}$?
(iii) Is $a \subseteq \{\{a\}\}$?
(iv) Is $a \in \{\ a, \{a\}\}$?

1.3 SETS GIVEN BY A DEFINING PROPERTY

As we have seen in the previous section, we often describe sets the following way:

$$A = \{x | P(x)\}$$

and we say: A is a set of all x such that $P(x)$, where $P(x)$ designates some property that all x's possess, or P is a condition that specifies some property of all objects x. In other words, $x \in A \leftrightarrow P(x)$. (see Axiom 2).

Note: Some sets have a universally accepted notation, so let's just agree at this point, without further explanation, to denote the set of natural numbers[23]

$$\mathbf{N} = \{\ 0, 1, 2, 3,\ \ldots\ \}$$

the set of integers

$$\mathbf{Z} = \{\ \ldots\ -3, -2,\ -1, 0, 1, 2, 3,\ \ldots\ \}$$

the set of rational numbers (which we will define later) **Q**, and the set of real numbers (also to be defined later), **R**.

Example 1.8 If, for instance, we say

$$P(x)\ :\ x \in \mathbf{N} \quad \text{and } x \text{ is even}$$

then, in set-builder notation, we write

$$\{x | P(x)\}$$

by which we mean the set of all natural even numbers. ■

Example 1.9 $A = \{x \in \mathbf{N} | 10 \leq x \leq 25\} = \{10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25\}$, that is, the set A is a set of all natural numbers greater than or equal to 10, and less than or equal to 25. ■

Definition 1.5 (The empty set) A set with no elements is called the **empty set**, denoted by the symbol $\varnothing = \{\}$.

Definition 1.5′ A set $\varnothing$ is said to be an empty set if

$$\varnothing = \{x | x \neq x\}$$

Equivalently, we can argue as follows: let X be a set and let there be a set $A = \{X | X \neq X\}$. Then, $X \in A \Rightarrow X \neq X$, which is a contradiction. Thus, A is empty.

The "existence" of the empty set is postulated by

Axiom 5 (Empty set (null set) axiom) There is a set with no elements.

[23]Many authors do not include 0 in **N** (in particular, for historical reasons) and, indeed sometimes that may be more convenient, and they reserve the following notation for nonnegative integers: $\mathbf{Z}^+ = \{0, 1, 2, 3, \ldots\} = \mathbf{N} \cup \{0\}$.

Example 1.10 Here are some examples of empty sets:

(i) $\{ n \in \mathbf{N} | n < 0\} = \varnothing$
(ii) $\{x \in \mathbf{Q} | x^2 = 2\} = \varnothing$
(iii) $\{x \in \mathbf{R} | x = x + 1\} = \varnothing$
(iv) $\{x \in \mathbf{R} | x^2 < 0\} = \varnothing$ ■

Example/Exercise 1.11 Determine whether or not, and why, are any of the following sets empty:

(i) $A = \{\{\varnothing\}, \{\{\varnothing\}\}\}$
(ii) $B = \{\{\{\varnothing\}\}, \{\varnothing\}, \{\{\{\varnothing\}\}\}\}$
(iii) $C = \{\{\{\{\varnothing\}\}\}\}$

A remarkable property of the empty set is given by the following:

Theorem 1.1 A set with no elements is a subset of any set, that is, if A is any set, and $\varnothing$ is the empty set, then

$$\varnothing \subseteq A$$

Proof Suppose that is not true, that is, suppose that there exists a set $\varnothing = \{\ \}$ (with no elements), and a set A such that $\varnothing \nsubseteq A$. That would mean, by definition of a subset, that there would be an element of $\varnothing$, which is not an element of A. But there can be no such element, since $\varnothing$ has no elements by definition. This contradiction leads us to conclude that the assumption $\varnothing \nsubseteq A$ was wrong; therefore, the theorem is true. ■

Example/Exercise 1.12 Show that $\{\varnothing\} \subseteq A$ for every set A.

You can think of this yet another way. *Any* set X is defined by a property P, possessed by all of its members, that is, if $x \in X$, then x has a property P. In particular, all elements of $\varnothing$ have to be defined by a certain property P, that is, if $x \in \varnothing$, then x has a property P. But, it is false to say that x is an element of $\varnothing$ (since $\varnothing$ has no elements), and since a false statement implies any proposition, it is true that if $x \in \varnothing$, then P holds for all the elements of $\varnothing$. Now, since P is a property defining a set X, it follows that $\varnothing \subseteq X$. All of this, as much as it may sound confusing to you now, will become more clear after you have studied Chapter 2.

Now you can try to prove the following:

Corollary 1.1 The empty set is unique, that is, there is only one set with no elements.

If you accept Axioms 3 and 4, then it is fun to contemplate the next claim, which might otherwise sound as an outrageous doctrine of set theory.

Claim Everything (mathematical?) is a set.

The "Proof" would go something like this:

Suppose there is a (mathematical) object X that is not a set. Then, X has no elements hence, X is equal to an empty set by Axiom 3, which contradicts the assumption that it is not a set. ■

This is pretty cute, don't you think? And, as a very fundamental concept, it will prove to be very useful. However, as a little exercise, try to think how you would dispute the aforementioned proof.

Axiom 6 (Pairing axiom) For any two sets X and Y, there is a set whose elements are these two sets, namely $\{X, Y\}$. We call the set $\{X, Y\}$ the **unordered pair** or **doubleton** of X and Y, that is

$$\{X, Y\} = \{Z | Z = X \text{ or } Z = Y\}$$

Theorem 1.2 Given two sets X and Y there is a *unique* set Z whose elements are X and Y.

Proof Since Axiom 6 established the existence of at least one set Z, whose elements are X and Y, the only thing we need to show is its uniqueness. Suppose then that there is another set Z' whose elements are also X and Y. But if X and Y are the elements of both Z and Z', by the axiom of extensionality, we have that $Z = Z'$. ■

The axiom of pairing gave us enough means to construct more sets, starting from just the empty set.

Example 1.13 One way of constructing many simple sets, each having at most two elements, is as follows:

$$\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\{\varnothing\}, \{\{\varnothing\}\}\}, \ldots$$ ■

From Axiom 6, it also follows that if $X = Y$, then $\{X, X\} = \{X\}$. For obvious reasons, we call this set the **singleton** $\{X\}$, or singleton of X. This is formalized by

Theorem 1.3 For any set X, there is a set whose only element is X.

As you can see, the key feature of set theory is that following Axioms 1–6 we can, in principle, construct a set from any object, or collection of objects, satisfying a certain property P and consider that as a mathematical object in its own right. In other words, we could consider a set X, which is a set of all sets x with a property P, that is

$$X = \{x | x \text{ is a set with property } P\}$$

As much as this principle is powerful, it has some fatal flaws. Consider this:

Let ***One*** be a set of all one-element sets, that is

$$\mathbf{One} = \{x | x \text{ is a one-element set}\}$$

Then nothing prevents us from forming the one-element set $\{\mathbf{One}\}$ whose only element is **One**. Immediately you recognize a Russell-like paradox:

$$\mathbf{One} \in \{\mathbf{One}\} \in \mathbf{One}$$

This can get even more intriguing. By Axiom 1, we can construct a set of (all) sets

$$\mathbf{U} = \{x | x \text{ is a set}\}$$

Since **U** is a set, it follows that $\mathbf{U} \in \mathbf{U}$. Obviously, in order to avoid circularities such as this one, we cannot treat **U** as any other "normal" set. We will have to say more about this later.

For now, let me incite your curiosity a bit more, especially in case you still have some doubts about the existence of the empty set. Let's assume the existence of the so-called **pure sets**, that is, sets that would exist even if there was nothing else but sets – no you and me, no people, no stars and planets, and so on, and simply refer to them as *Sets* (with a capital "*S*"). While the existence of the empty set $\varnothing$ becomes evident right away, we can immediately conceive the set whose only member is the empty set, that is, $\{\varnothing\}$ and, unsurprisingly, the next would be $\{\{\varnothing\}\}$, followed by $\{\varnothing, \{\varnothing\}\}$, and so on and so forth. So, we recognize the collection of sets mentioned in the previous example as pure sets – *Sets*. Observe that their "nature" is rather unique. That is, all *Sets* are sets but sets are not *Sets.* (The set of horses is not a *Set*.) After inaugurating the concept of *Sets* why not construct additional (particular) ones, respectively, assign familiar names to them, and thus obtain "something" from "nothing." One way to do it would be:

$$\varnothing = \mathbf{0}$$

$$\{\varnothing\} = \mathbf{1}$$

$$\{\{\varnothing\}\} = \mathbf{2}$$

$$\{\{\{\varnothing\}\}\} = \mathbf{3}$$

and so on.

The other way would be:

$$\mathbf{0} = \varnothing$$
$$\mathbf{1} = \{\varnothing\}$$
$$\mathbf{2} = \{\varnothing, \{\varnothing\}\}$$
$$\mathbf{3} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$$

and so on.

Certainly, you can think of some other way to construct "something" from "nothing."

However, before continuing, do

Example/Exercise 1.14 Prove that $\varnothing \neq \{\varnothing\}$.

Definition 1.6 Given sets X and Y, we say that

$$\langle X, Y\rangle = \{\{X\}, \{X, Y\}\}$$

is an ordered pair.

Analogously, we define an ordered n-tuple:

Definition 1.7 Let $X_1, X_2, X_3, \ldots, X_n$ be sets where $n \in \mathbf{N}, n \geq 3$. We define an ordered n-tuple recursively as follows:

$$\langle X_1, X_2, X_3, \ldots, X_n\rangle = \langle X_1, \langle X_2, X_3, \ldots, X_n\rangle\rangle$$

Theorem 1.4 For any sets $X, Y, U, V, \langle X, Y\rangle = \langle U, V\rangle$ iff $X = U$ and $Y = V$.

Proof That $X = U$ and $Y = V$ implies $\langle X, Y\rangle = \langle U, V\rangle$ is trivial, so we need to examine only that $\langle X, Y\rangle = \langle U, V\rangle$ implies $X = U$ and $Y = V$.

Suppose that $\langle X, Y\rangle = \langle U, V\rangle$ which by definition means that

$$\{\{X\}, \{X, Y\}\} = \{\{U\}, \{U, V\}\} \qquad (*)$$

We should consider two cases: (i) $X = Y$ and (ii) $X \neq Y$.

(i) If $X = Y$ then

$$\langle X, Y\rangle = \{\{X\}, \{X, Y\}\} = \{\{X\}, \{X, X\}\} = \{\{X\}, \{X\}\} = \{\{X\}\}$$

is a singleton, so $\langle U, V\rangle$ has to be a singleton too. Thus, $U = V$. But that means that

$$\{\{U\}, \{U, V\}\} = \{\{U\}, \{U, U\}\} = \{\{U\}, \{U\}\} = \{\{U\}\}$$

With the assumption that equality in $(*)$ holds, we have that $\{\{X\}\} = \{\{U\}\}$, that is, $X = U$ hence,

$$X = Y = U = V$$

(ii) If $X \neq Y$ then from (*) it follows that the singleton $\{X\}$ must correspond to the singleton $\{U\}$ and, likewise, the doubleton $\{X, Y\}$ corresponds to the doubleton $\{U, V\}$. We conclude that

$$X = U \quad \text{and} \quad Y = V$$ ■

Definition 1.8 (Cardinal number) Let A be a set. If there are exactly n *distinct* elements in A, where n is a finite natural number, we say that the set A is a **finite set** and that n is the **cardinality** of A, or that n is the **cardinal number** of A, and we denote cardinality by $|A|$, (Figure 1.6).

You may have an uneasy feeling about this definition. Considering the fact that A was said to be a finite set, the definition seems to be too restrictive. Everything is fine if a set has, say, 3175 elements – the cardinal number is 3175. Naturally, one would ask: what about sets that have infinitely many elements? How would we characterize the "number" of elements of an infinite set? After all, the issue of infinities (as we will see shortly) is *the* issue of set theory. It turns out that this is one of the most intuitively difficult mathematical concepts of the theory. Can we "enumerate" a set with infinitely many members regardless of their "nature?" Cantor used the symbol $\overline{\overline{A}}$ to indicate the cardinal number of set A, emphasizing double abstraction: first from the nature of elements and second from their order, and he said:

> *Every set A has a definite* ***"power" which we will call its "cardinal number."*** *We will call by this name the general concept, which by means of our active faculty of*

thought arises from the set A when we make abstraction of its various elements x and of the order in which they are given.

… This number has an existence in our minds as an intellectual image or projection of the given set.

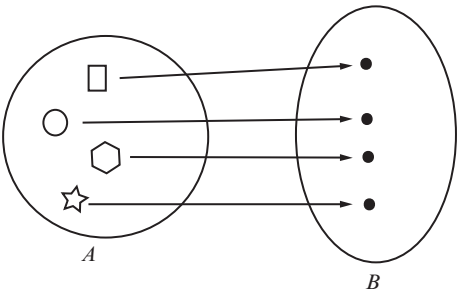


Figure 1.6 The concept of cardinality for a four-element set A à la Cantor

With all that said, and with tongue-in-cheek, let's say for the time being that the *cardinality* of a set means the "number" of the elements of a set or, even better, the "size" of a set.

Example 1.15 Let A be the set from Example 1.9, then $|A| = 16$. ■

Example 1.16 What is $|\varnothing|$?
Well, since the empty set $\varnothing$ has no elements, it follows that $|\varnothing| = 0$. ■

Definition 1.9 A set is said to be **infinite** if it is not finite.

The existence of the "infinite" set is provided by

Axiom 7 (Axiom of infinity) There exists a set I that contains the empty set $\varnothing$ and the singleton of each of its members, that is

$$\varnothing \in I \;\&\forall x \in I, \{x\} \in I$$

Example 1.17 Let I be a set defined in Axiom 7. Observe that $\varnothing \in I$, but also, $\{\varnothing\} \in I$, $\{\{\varnothing\}\} \in I$, …. So, with this family of complex singletons, we have indeed obtained an infinite set of more abstract nature:

$$I = \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \ldots\}$$ ■

Example 1.18 Here is how Dedekind[24] argued that at least one infinite set exists: *Given some arbitrary thought τ_1, there is a separate thought τ_2, namely that τ_1 is an object of thought. But there also exists a thought τ_3, that is, a thought of τ_1 and τ_2. And so on ad infinitum. Thus, set of thoughts is infinite.* ■

The claims of the next two examples are usually accepted as obvious.

Example 1.19 A set of all natural numbers **N** is infinite. ■

Example 1.20 A set of all integers **Z** is infinite. ■

We will discuss the intricacies of infinite sets in a little while.

Definition 1.10 We say that two sets A and B are **equivalent** (or **equinumerous**) or that they have the same cardinality, and we write

$$A \sim B \quad \text{iff} \quad |A| = |B|$$

Following Cantor, we say that cardinal number of a set A is what A has in common with all sets equivalent to A.

Example 1.21 Given sets $A = \{1, 2, 3\}, B = \{a, b, c\}$, and $C = \{b, c, a\}$, we say that $A \sim B$, and $A \sim C$, but only $B = C$. ■

Theorem 1.5 Given three sets A, B, C, such that $A \sim B$, and $B \sim C$, then $A \sim C$.

Proof Easy. You should do it! ■

Now that we have a rudimentary knowledge of sets, in order to finish this section and have some fun, I have to tell you something else. Something about those strange sets I have mentioned in the introduction. In a sense, you may think of what follows as a "historical" progress toward the paradoxes Russell pointed out to us.

With the concept of a set handy, and assuming also that the attributes that apply to a set are not mutually contradictive, then, by an *extension* of such an idea, we can easily contemplate a set that contains sets as its elements. Why not, right? For example, the concept of a finite set F is easily conceivable. Its extension $\mathcal{F}$ would be a set of all sets with finitely many elements, that is

$$\mathcal{F} = \{F | F \text{ is a finite set}\} \tag{1.1}$$

[24] Julius Wilhelm Richard Dedekind (1831–1916), German mathematician.

Similarly, with the idea of an infinite set I (say, a set $\mathbf{N}$) let's define its extension as a set

$$\mathcal{I} = \{ I \,|I \text{ is an infinite set}\} \tag{1.2}$$

Observe that, while all the elements of $\mathcal{F}$ are finite sets, $\mathcal{F}$ itself is an infinite set. That makes $\mathcal{F}$ not a member of itself, but a member of $\mathcal{I}$. Symbolically,

$$\mathcal{F} \notin \mathcal{F}, \text{ and } \mathcal{F} \in \mathcal{I}$$

On the other hand, it is clear that

$$\mathcal{I} \in \mathcal{I}$$

Again, do you see where we are going? Let's call on Russell again. Consider the concept of a "*set that is not a member of itself,*" and let's call its extension

$$\mathcal{R} = \{X|X \text{ is a set } \& \, X \in X\} \tag{1.3}$$

From (1.1) – (1.3), we see that $\mathcal{F} \in \mathcal{R}$ and $\mathcal{I} \notin \mathcal{R}$. But how about $\mathcal{R}$? Is it a member of itself or not? From the aforementioned discussion, it follows that

$$\mathcal{R} \in \mathcal{R} \text{ iff } \mathcal{R} \notin \mathcal{R} \tag{1.4}$$

But this is impossible! Either $\mathcal{R}$ is a member of itself or not. Claim (4) is a contradiction par excellence. Thus, we state (we are forced to state):

Theorem 1.6 There is no set $\mathcal{R}$ such that

$$\mathcal{R} = \{X \,|X \text{ is a set } \& \, X \notin X\}$$

The reason I keep on mentioning this quintessential paradox is because of its profound mathematical/philosophical importance. I'll stop here abruptly, again quoting Russell: "*Whatever involves all of a collection must not be one of the collection.*" What he actually said was: just forget about those "crazy" sets, consider only those sets that are ordinary.[25] Can you do that? Can you just forget about the "crazy" sets? I could never do that. They keep coming up in many different branches of mathematics, physics, and philosophy. It seems our mind, once having become aware of them, simply cannot let go. In any case, we continue our discussion of sets by introducing the formalism that will enable us to "calculate" and discover even more interesting "stuff."

[25] At this point, you may want to revisit the discussion on ordinary and extraordinary sets on pages 4 and 5.

1.4 THE ALGEBRA OF SETS

In order to reasonably carry on a mathematical discussion in the context (a set) of specific elements, we can often visualize the entity whose existence we postulated in Section 1.2 by Axiom 1. For example, we may consider a set of all students at the university, or we may consider a set of all books in your school library, a set of all animals in the zoo, or a set of all real numbers, and so on. In each of these cases, we call this a universe of discourse, or the universal set of the given discourse. So, we formally state

Definition 1.11 By universal set $\mathcal{U}$, we mean the set of all the elements *under discussion* (all the objects under consideration).

Note the important qualification "*under discussion*" in the aforementioned definition. Without it, the concept of a universal set would create a rather difficult problem. Namely, one could be tempted to consider the universal set $\mathcal{U}$ as a set of all "objects," that is, a set of everything. Why not, right? But then, in particular, $\mathcal{U}$ would contain itself as a member, and that would be a problem indeed as we have indicated at the beginning of this chapter.

Definition 1.12 Given a universal set $\mathcal{U}$, and A and B the two subsets of $\mathcal{U}$, we define the **union** of A and B, denoted $A \cup B$, as a set of all the elements $x \in \mathcal{U}$, such that x is an element of A or x is an element of B (Figure 1.7), that is

$$A \cup B = \{x \in \mathcal{U} | x \in A \text{ or } x \in B\}$$

Note that in this definition "or" is the inclusive "or" (as opposed to "either–or").

Example 1.22 Let A be a set of all even whole numbers, that is, all even integers, and let B be a set of all odd whole numbers, that is, all odd integers. Then, $A \cup B$ represents the set of all whole numbers, that is, the set of all integers. Recall, we denoted that set by the symbol $\mathbf{Z}$. ■

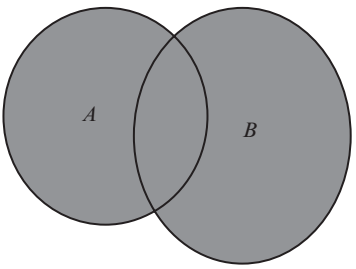


Figure 1.7 $A \cup B$

Example 1.23 Given two sets $X = \{a, b, c, d, e\}$ and $Y = \{@, \#, \$, \&\}$, then

$$X \cup Y = \{a, b, c, d, e, @, \#, \$, \&\}$$ ■

The union of an infinite sequence of sets is defined in the same way.

Definition 1.12′

$$\cup A = \bigcup_{n=0}^{\infty} A_n = A_0 \cup A_1 \cup A_2 \ \ldots \ = \ \{x | x \in \ A_n, n \in \mathbf{N}\}$$

In general, considering the abstract nature of a set, the existence of the union as a set is postulated by

Axiom 8 (Union axiom) For any set X, there is a set that is the union of all the elements of X.

As much as the concept of the union of two sets is easy to understand, Axiom 8 might take some time to absorb, so you can skip it until you have studied Chapter 2. For now let's just say that one can think of the expression in Definition 1.12′ as $\cup\{A_n | n \in \mathbf{N}\}$.

Example 1.24 Let's take just two sets, A_1 and A_2, and consider $\{A_1, A_2\}$. Suppose $x \in \cup\{A_1, A_2\}$. That is true iff $x \in X$ for some $X \in \{A_1, A_2\}$. But the only X's in $\{A_1, A_2\}$ are A_1 and A_2. Thus $x \in \cup\{A_1, A_2\}$ iff $x \in A_1$ or $x \in A_2$. But that's exactly what we are saying with $x \in A_1 \cup A_2$. ■

Example 1.25 Suppose we have three sets A, B, C. Then, there is a set with these sets as its elements:

$$\{A\} \cup \{B\} \cup \{C\} = \{A, \ B\} \cup \{C\} = \{A, B, C\}$$ ■

Example 1.26 The next simple fact is that

$$\cup\{X | X \in \{A\}\} = A$$ ■

Following the aforementioned three examples, it should not be difficult to work out:

Example/Exercise 1.27 Determine whether the following is true:

(i) $\cup\{X\} = X$

(ii) $\cup\varnothing = \cup\{\varnothing\} = \varnothing$
(iii) $\{\varnothing\} \cup \varnothing = \{\varnothing\}$

Definition 1.13 Let A and B be sets. The **intersection** of A and B, denoted $A \cap B$, is the set of all elements $x \in \mathcal{U}$, such that x is an element of A and x is an element of B (Figure 1.8), that is

$$A \cap B = \{x \in \mathcal{U} | x \in A \ \& \ x \in B\}$$

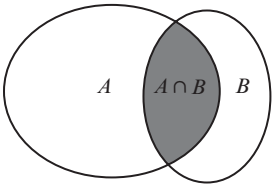


Figure 1.8 $A \cap B$ intersection

Example 1.28 Given two sets $A = \{1, 2, 3, a, b, c\}$ and $B = \{3, b, x, y\}$, then

$$A \cap B = \{3, b\}$$ ■

Example/Exercise 1.29 Show that for all sets A, B, and C

(i) $A \subseteq A \cup B$ and $B \subseteq A \cup B$
(ii) If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$

The intersection of an infinite sequence of sets is defined analogously

Definition 1.13′

$$\cap_{n=0}^{\infty} A_n = A_0 \cap A_1 \cap A_2 \cdots = \{x | (\forall n \in \mathbf{N})x \in A_n\}$$

Example/Exercise 1.30 Show that

$$A \cap B \subseteq A \quad \text{and} \quad A \cap B \subseteq B$$

Example/Exercise 1.31 Convince yourself that

(i) $A \subseteq B$ iff $A \cup B = B$
(ii) $A \subseteq B$ iff $A \cap B = A$
(iii) $\{\varnothing\} \cap \varnothing = \varnothing$

Example/Exercise 1.32 The union of empty sets is clearly an empty set. You may be wondering now: what about $\cap\varnothing$? This is much trickier. Can you see why?

Definition 1.14 Let A and B be two sets. We say that A and B are disjoint, if $A \cap B = \varnothing$ (Figure 1.9).

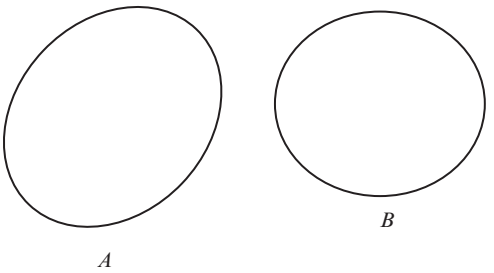


Figure 1.9 $A \cap B = \varnothing$.

Example 1.33 Consider the following:

Let $A_1 = \{0\}, A_2 = \{0, 1\}, A_3 = \{0, 1, 2\}, \ldots, A_{i+1} = \{0, 1, 2, \ldots, i\}, \ldots$. So we have an infinite collection of $A's$, such that for every $n \in \mathbf{N}^+$, $n \in A_{n+1}$.[26] Thus, $\mathbf{N}^+ = A_1 \cup A_2 \cup \cdots$ and $A_1 \cap A_2 \cap \cdots = 0$. ■

Example 1.34 Consider a set $\mathbf{R}$. Let set A be the interval $(-3, 5)$, and set B the interval $(3, 8)$.

Find:

(i) $A \cap B$

(ii) $A \cup B$

Solution First, recall the definition of intervals on the set of real numbers $\mathbf{R}$:

An open interval

$$O = (a, b) = \{x | a < x < b\}$$

A closed interval

$$C = [a, b] = \{x | a \leq x \leq b\}$$

Of course, we can have a half-open–half-closed interval, such as

$$O_C = (a, b] = \{x | a < x \leq b\}$$

or

$$C_O = [a, b) = \{x | a \leq x < b\}$$

[26] In order to avoid confusion, when starting with zero in our collection of A's, for the time being, we put $\mathbf{N} \cup \{0\} = \mathbf{N}^+$, which is also designated by $\mathbf{Z}^+$.

Now, observe that set

$$A = (-3, 5) = \{x | -3 < x < 5\}$$

and set

$$B = [3, 8) = \{x | 3 \leq x \leq 8\}$$

Hence,

(i) $A \cap B = [3, 5) = \{x | 3 \leq x < 5\}$
(ii) $A \cup B = (-3, 8] = \{x | -3 < x \leq 8\}$ ■

Example/Exercise 1.35 Let the universal set be a set of all integers, that is, $\mathcal{U} = \mathbf{Z}$, and let $A = \{x \in \mathbf{Z} | x = 2n, n \in \mathbf{Z}\}$, and $B = \{y \in \mathbf{Z} | 2m + 1, m \in \mathbf{Z}\}$, then

$$A \cap B = \varnothing$$

Convince yourself that this is indeed so.

Example/Exercise 1.36[27] Consider the oldest mathematician among chess players and the oldest chess player among mathematicians. Could they be two different persons?

Definition 1.15 We say that a collection $A_1, A_2, A_3, \ldots, A_n$ is a **partition** $P(A)$ (Figure 1.10) of a set A iff

(i) $A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n = A$ and
(ii) $A_1 \cap A_2 \cap A_3 \cap \cdots A_n = \varnothing$

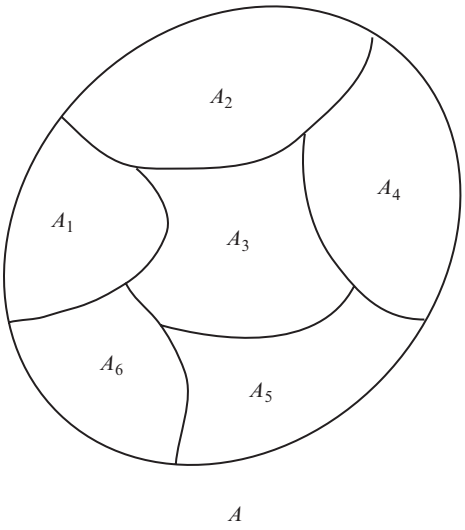


Figure 1.10 Partition $P(A)$

[27] Shen, S., Vereshchagin, N. K., *Naïve Set Theory*, American Mathematical Society, 2002.

Definition 1.16 Given two sets A and B, we say that the **difference** of A and B, denoted $A\backslash B$, and read "A minus B," is the set of all elements x from $\mathcal{U}$, such that x is in A and x is not in B (Figure 1.11). We write

$$A \setminus B = \{x \in \mathcal{U} | x \in A \ \& \ x \notin B\}$$

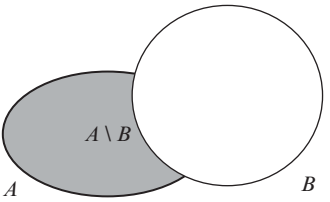


Figure 1.11

Example 1.37 Let $A = \{a, b, c, d, e, f, g\}$ and $B = \{c, e, g, h, i, k\}$, then

$$A \setminus B = \{a, b, d, f\}$$ ■

Example/Exercise 1.38 Prove the following:

(i) $A\backslash\varnothing = A$
(ii) $A\backslash A = \varnothing$
(iii) $A \cap (B \setminus A) = \varnothing$

Definition 1.17 Let A and B be sets. The **symmetric difference** of A and B, denoted $A\Delta B$ (Figure 1.12), is defined

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

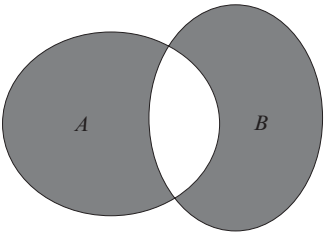


Figure 1.12 $A \setminus B$

Example/Exercise 1.39 Convince yourself that

(i) $A\Delta B = 0$ iff $A = B$
(ii) $A\Delta\varnothing = A$

Example/Exercise 1.40 Show that

$$A \,\Delta\, B = (A \cup B) \setminus (A \cap B)$$

Definition 1.18 Let A be a subset of the universal set $\mathcal{U}$. We define the **complement** of A, denoted A^c, as the set of all elements x from $\mathcal{U}$ (Figure 1.13), such that x is not in A:

$$A^c = \{x \in \mathcal{U} | x \notin A \}$$

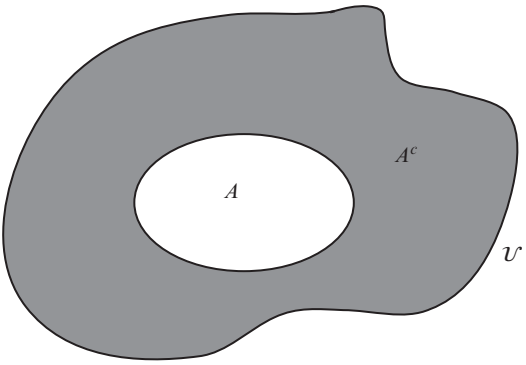


Figure 1.13 A^c

Example 1.41 Prove that $A \setminus B = A \cap B^c$. ■

Proof The proof is easy. We need to show that $\forall x$ if $x \in A \setminus B$ then $x \in A \cap B^c$, and also that $\forall x$ if $x \in A \cap B^c$ then $x \in A \setminus B$.

So first, suppose we take any $x \in A \setminus B$. That means that $x \in A$ and $x \notin B$, which in turn implies $x \in B^c$. So, $x \in A$ and $x \in B^c$, and therefore $x \in A \cap B^c$.

Conversely, if $x \in A \cap B^c$ then $x \in A$ and $x \in B^c$, that is, $x \in A$ and $x \notin B$ and thus $x \in A \setminus B$. ■

Example/Exercise 1.42 Let $A, B \subseteq \mathcal{U}$ be any two subsets of the universal set. Show that

$$A \subseteq B \text{ iff } B^c \subseteq A^c$$

Example 1.43 Let $A = [0, 1)$, $B = (-1, 1)$, and $C = (-2, 1]$.
Find

(i) $A^c \cap B^c \cap C^c$
(ii) $(A \cap B) \cup C \cup B^c$

Solution

(i) The complements of $A, B,$ and C are as follows:

$$A^c = ([0, 10))^c = (-\infty, 0) \cup [1, \infty)$$
$$B^c = ((-1, 1))^c = (-\infty, -1] \cup [1, \infty)$$
$$C^c = ((-2, 1])^c = (-\infty, -2] \cup (1, \infty)$$

Then

$$A^c \cap B^c \cap C^c = ((-\infty, 0) \cup [1, \infty)) \cap ((-\infty, -1] \cup [1, \infty)) \cap ((-\infty, -2] \cup (1, \infty))$$
$$= (-\infty, -2] \cup (1, \infty)$$

(ii)

$$(A \cap B) \cup C \cup B^c = ([0, 1) \cap (-1, 1)) \cup (-2, 1] \cup (-\infty, -1] \cup [1, \infty)$$
$$= [0, 1) \cup (-2, 1] \cup (-\infty, -1] \cup [1, \infty)$$
$$= (-\infty, -1] \cup [0, \infty)$$ ■

Theorem 1.7

(i) $A \cap \varnothing = \varnothing$
(ii) $A \cup \varnothing = A$
(iii) $A \cap A^c = \varnothing$
(iv) $A \cup A^c = \mathcal{U}$
(v) $\mathcal{U}^c = \varnothing$
(vi) $\varnothing^c = \mathcal{U}$

Proof (i): Let A be any set. Suppose $\cap \varnothing \neq \varnothing$, that is, suppose there exists an $x \in A \cap \varnothing$. By the definition of intersection, $x \in A$, and $x \in \varnothing$. But this is impossible since $\varnothing$ has no elements by definition. Thus,

$$A \cap \varnothing = \varnothing$$

Now you should try to prove parts (ii)–(vi) of the theorem. ■

Example/Exercise 1.44 Consider three sets A, B, and C. Is it possible that

$$A \cap B \neq \varnothing, \ A \cap C = \varnothing \text{ and } (A \cap B) \setminus C = \varnothing?$$

Theorem 1.8 (Set identities) For all sets A, B and C.

1.8.1

(i) $A \cap B = B \cap A$

(ii) $A \cup B = B \cup A$ (Commutative Laws for intersection and union)

1.8.2

(i) $A \cap A = A$

(ii) $A \cup A = A$ (Idempotent Laws)

1.8.3

(i) $A \cap (B \cap C) = (A \cap B) \cap C$

(ii) $A \cup (B \cup C) = (A \cup B) \cup C$ (Associative Laws for intersection and union)

1.8.4

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

(ii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (Distributive Laws)

1.8.5

$A \cap \mathcal{U} = A$

1.8.6

$A \cup \mathcal{U} = \mathcal{U}$

1.8.7

(i) $A \cup (A \cap B) = A$

(ii) $A \cap (A \cup B) = A$ (Absorption Laws)

1.8.8

(i) $(A \cap B)^c = A^c \cup B^c$

(ii) $(A \cup B)^c = A^c \cap B^c$ (DeMorgan's Laws)

Example/Exercise 1.45 Prove 1.8.1–1.8.8 of Theorem 1.8.

Proof Remember, two sets A and B are equal iff $A \subseteq B$ and $B \subseteq A$. Thus, in each case, we need to show that any x, being an element of the set on the left-hand side (LHS) of our equation is also an element of the set on the right-hand side (RHS) of our equation, and vice versa. So,

1.8.4 (ii):

Suppose $x \in A \cap (B \cup C)$. By the definition of intersection that means that $x \in A$ and $x \in (B \cup C)$. That gives us two possible cases.

Case 1: $x \in A$ and $x \in B$, by the definition of union. Hence, $x \in A$ and $x \in B$ implies that $x \in A \cap B$, therefore,

$$x \in (A \cap B) \cup (A \cap C)$$

Case 2: $x \in A$ and $x \in C$, again by the definition of union. Hence, $x \in A$ and $x \in C$ implies that $x \in A \cap B$, therefore, $x \in (A \cap C) \cup (A \cap B)$. From Theorem 1.8.1 (ii), it follows that

$$x \in (A \cap B) \cup (A \cap C)$$

In both cases, $x \in (A \cap B) \cup (A \cap C)$.
Hence, we have proved that

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C) \qquad (*)$$

Suppose now that $x \in (A \cap B) \cup (A \cap C)$. By the definition of union that means that either

$$x \in (A \cap B) \quad \text{or} \quad x \in (A \cap C)$$

So, again, we have two possibilities.

Case 1: $x \in (A \cap B)$. By the definition of intersection, this implies that

$$x \in A \quad \text{and} \quad x \in B$$

Well, x being an element of B, means that x is also an element of $(B \cup C)$. We have that $x \in A$, and

$$x \in A \quad \text{and} \quad x \in (B \cup C)$$

Therefore, by the definition of intersection,

$$x \in A \cap (B \cup C)$$

Now consider
Case 2: $x \in (A \cap C)$. By the definition of intersection, this implies that

$$x \in A \quad \text{and} \quad x \in C$$

Since x is an element of C, it also has to be an element of $(B \cup C)$. So, again, we have that

$$x \in A \text{ and } x \in (B \cup C)$$

Therefore, by the definition of intersection,

$$x \in A \cap (B \cup C)$$

In both cases, $x \in A \cap (B \cup C)$. Hence, we proved that

$$(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C) \qquad (**)$$

Since both subset relations (*) and (**) have been proved, it follows by definition of set equality that

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

as stated in Theorem 1.8.3 (ii).

Now you can prove Theorem 1.8.3 (i).

Proof 1.8.8 (i)

We need to prove that for every x

$$\text{if } x \in (A \cap B)^c \text{ then } x \in A^c \cup B^c$$

Suppose $x \in (A \cap B)^c$. By the definition of complement, $x \notin A \cap B$. But this implies that $x \notin A$ or $x \notin B$. Saying that $x \notin A$ means that $x \in A^c$. Similarly, if $x \notin B$, then $x \in B^c$. Hence, $x \in A^c$ or $x \in B^c$ and by the definition of union this implies that

$$x \in A^c \cup B^c$$

So, we have proved that

$$(A \cap B)^c \subseteq A^c \cup B^c \qquad (*)$$

Let's now consider the converse, that is, let's show that for every x

$$\text{if } x \in A^c \cup B^c \text{ then } x \in (A \cap B)^c$$

Suppose that $\in A^c \cup B^c$. By definition of union, it follows that $x \in A^c$ or $x \in B^c$. So we have to consider two cases.

Case 1: $x \in A^c$. Being an element of A^c means that $x \notin A$, and therefore x cannot be in $A \cap B$ either, that is

$$x \notin A \cap B$$

Well, since $x \notin A \cap B$, it is definitely true that

$$x \in (A \cap B)^c \qquad (**)$$

Case 2: $x \in B^c$ would lead us, by the similar arguments, to the same conclusion (**):

$$x \in (A \cap B)^c$$

Thus, we have also proved that

$$A^c \cup B^c \subseteq (A \cap B)^c$$

By the definition of equality of sets, (*) and (**) imply that $(A \cap B)^c = A^c \cup B^c$, as was to be shown. ■

Now you can prove Theorem 1.8.8(ii).

Example 1.46 (Generalized distributive property) Let $A_i \in \mathcal{U}$, $i \in \mathbf{N}$, and let $B \in \mathcal{U}$.[28] Show that

$$B \cup (\cap_{i=1}^{n} A_i) = \cap_{i=1}^{n} (B \cup A_i), \quad \forall n \in \mathbf{N}$$

Solution We will do the proof by the *Method of Mathematical Induction*:

First, we note that the statement is trivially true when $n = 1$. Theorem 1.8.4(b) assures us that the claim is true for $n = 2$. We will assume that it is also true for $n = k$. If we could prove that it is also true for $n = k + 1$, then the claim is true for any $n \in \mathbf{N}$. Consider

$$\begin{aligned} B \cup (\bigcap_{i=1}^{k+1} A_i) &= B \cup (\bigcap_{i=1}^{k} A_i \cap A_{k+1}) \\ &= (B \cup (\bigcap_{i=1}^{k} A_i) \cap A_{k+1}) \end{aligned}$$

(Since we assumed that the claim is valid for $n = k$)

$$\begin{aligned} &= \bigcap_{i=1}^{k} (B \cup A_i) \cap (B \cap A_{k+1}) \\ &= \bigcap_{i=1}^{k+1} (B \cup A_i) \end{aligned}$$

So, our proposition is true for $n = k + 1$ and thus,

$$B \cup (\bigcap_{i=1}^{n} A_i) = \bigcap_{i=1}^{n} (B \cup A_i) \quad \forall n \in \mathbf{N}$$ ■

[28] If you are unfamiliar with the "Proof by induction" method, you can skip this example until you have studied Chapter 4.

Example/Exercise 1.47 Let $A_i \in \mathcal{U}$, $i \in \mathbf{N}$, and let $B \in \mathcal{U}$. Show that

$$B \cap (\bigcup_{i=1}^{n} A_i) = \bigcap_{i=1}^{n} (B \cap A_i), \quad \forall n \in \mathbf{N}$$

Example 1.48 Let $A, B \subseteq \mathcal{U}$ be any two sets. Show that

$$(A \cap B) \cup (A \cap B^c) = A$$

Solution

$$\begin{aligned}(A \cap B) \cup (A \cap B^c) &= A \cap (B \cup B^c) \\ &= A \cap \mathcal{U} \\ &= A\end{aligned}$$

On the other hand,

$$\begin{aligned}A &= A \cap \mathcal{U} \\ &= A \cap (B \cup B^c) \\ &= (A \cap B) \cup (A \cap B^c)\end{aligned}$$

■

Example/Exercise 1.49 Let $\mathcal{U} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ be the universal set, and let $A = \{1, 3, 5, 8\}$, $B = \{2, 3, 4, 5\}$, and $C = \{3, 4, 6, 7, 8\}$. Using these sets, convince yourself that Theorems 1.8.4 and 1.8.8 are indeed true.

Example 1.50 Prove that the following statements are equivalent:

(i) $A \subseteq B$
(ii) $A \cap B = A$
(iii) $A \cup \mathrm{B} = B$

■

Proof To prove that (i) implies (ii), let's assume that $A \subseteq B$. We need to establish that $A \cap B \subseteq A$ and that $A \subseteq A \cap B$. But, since $A \cap B \subseteq A$ for all A and B, it is sufficient to prove that $A \subseteq A \cap B$. So, if $x \in A$, it follows from (i) that $x \in B$ and therefore $x \in A \cap B$. Hence, $A \subseteq A \cap B$.

To prove that (ii) implies (iii), let's assume that $A \cap B = A$ holds. Then,

$$\begin{aligned} A \cup B &= (A \cap B) \cup B = (A \cup B) \cap (B \cup B) \\ &= (A \cup B) \cap B = B \end{aligned}$$

Finally, to prove that (iii) implies (i), we assume that $A \cup B = B$ holds. Then, since $A \subseteq A \cup B$ for all A and B, it follows that $A \subseteq B$. ■

Example 1.51 Let $\mathcal{U} = \mathbf{R}$, $A = [0, 1)$, $B = (-1, 1)$ and $C = (-2, 1]$.[29] Determine

$$(A \cap B) \cup (A^c \cap C^c)$$

Solution

If $A = [0, 1)$, then $A^c = (-\infty, 0) \cup [1, \infty)$.
If $B = (-1, 1)$, then $B^c = (-\infty, -1] \cup [1, \infty)$.
If $C = (-2, 1]$, then $C^c = (-\infty, -2] \cup (1, \infty)$.
So, we have

$$\begin{aligned} (A \cap B) \cup (A^c \cap C^c) &= ([0, 1) \cap (-1, 1)) \cup (((-\infty, 0) \\ &\quad \cup [1, \infty)) \cap ((-\infty, -2] \cup (1, \infty))) \\ &= ([0, 1) \cap (-1, 1)) \cup ((-\infty, -2) \cup (1, \infty)) \\ &= [0, 1) \cup (-\infty, -2] \cup (1, \infty) \\ &= (-\infty, -2] \cup [0, 1) \cup (1, \infty) \end{aligned}$$

■

Example 1.52 (Generalized DeMorgan's Law) Prove that for all $n \in \mathbf{N}$, if $A_1, A_2, A_3, \ldots, A_n$ are sets, then

$$(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} (A_i)^c$$

■

***Proof*[30]** The formula is obviously true for $n = 1$. (Why?) Suppose it is also true for $n = k$, that is, suppose

$$(\bigcup_{i=1}^{k} A_i)^c = \bigcap_{i}^{k+1} (A_i)^c$$

[29] We assume here that the reader is at least vaguely familiar with the properties of real numbers and she/he won't mind that we have not yet precisely defined the set **R.**

[30] Here, again, if you are not familiar with mathematical induction, you may skip this proof until you have learned it in later chapters.

We have to prove that it is also valid for $k+1$, that is

$$(\bigcup_{i=1}^{k+1} A_i)^c = \bigcap_{i}^{k+1} (A_i)^c$$

which would imply that our formula is valid for all n.

Recalling the properties of union and DeMorgan's law for two sets, we get

$$\begin{aligned}(\bigcup_{i=1}^{k+1} A_i)^c &= (\bigcup_{i=1}^{k} A_i \cup A_n)^c \\ &= (\bigcup_{i=1}^{k} A_i)^c \cap (A_n)^c \\ &= (\bigcup_{i=1}^{k} (A_i)^c) \cap (A_n) \\ &= \bigcup_{i=1}^{k+1} (A_i)^c\end{aligned}$$

Since the formula holds for $k+1$, it holds for every $n \in \mathbf{N}$. ■

In a similar way, you can work out

Example/Exercise 1.53 Prove that for all $n \in \mathbf{N}$, if $A_1, A_2, A_3, \ldots, A_n$ are sets, then

$$(\bigcap_{i=1}^{n} A_i)^c = \bigcup_{i=1}^{n} (A_i)^c$$

Example 1.54 Let $A, B, C \in \mathcal{U}$ be any three sets. Prove that

$$(A \cup B) \backslash C = (A \backslash C) \cup (B \backslash C)$$

Solution

$$\begin{aligned}(A \cup B) \setminus C &= (A \cup B) \cap C^c \\ &= C^c \cap (A \cup B) \\ &= (C^c \cap A) \cup (C^c \cap B) \\ &= (A \cap C^c) \cup (B \cap C^c) \\ &= (A \setminus C) \cup (B \setminus C)\end{aligned}$$

On the other hand, we could have said:

$$\begin{aligned}(A \setminus C) \cup (B \setminus C) &= (A \cap C^c) \cup (B \cap C^c) \\ &= (A \cup B) \cap C^c \\ &= (A \cup B) \setminus C\end{aligned}$$

Thus, we have our proof. ■

Example 1.55 Let $A,\ B,\ C \in \mathcal{U}$ be any three sets. Prove that

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

Solution

$$\begin{aligned}C \setminus (A \cap B) &= C \cap (A \cap B)^c \\ &= C \cap (A^c \cup B^c) \\ &= (C \cap A^c) \cup (C \cap B^c) \\ &= (C \setminus A) \cup (C \setminus B)\end{aligned}$$

Similarly,

$$\begin{aligned}(C \setminus A) \cup (C \setminus B) &= (C \cap A^c) \cup (C \cap B^c) \\ &= C \cap (A^c \cup B^c) \\ &= C \cap (A \cap B)^c \\ &= C \setminus (A \cap B)\end{aligned}$$

which completes our proof. ■

Example 1.56 Show that for any two sets A and B

$$A^c \setminus B^c = B \setminus A$$

Solution

$$\begin{aligned}A^c \setminus B^c &= A^c \cap (B^c)^c \\ &= B \cap A^c = B \setminus A\end{aligned}$$

■

Example/Exercise 1.57 Prove that for any sets $A, B, C \in \mathcal{U}$

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

Example/Exercise 1.58 Show that for any sets A_i and C, the following is true.

(i) $C \setminus (\underset{i=1}{\overset{n}{\cup}} A_i) = \underset{i=1}{\overset{n}{\cap}} (C \setminus A_i)$

(ii) $C \setminus (\underset{i=1}{\overset{n}{\cap}} A_i) = \underset{i=1}{\overset{n}{\cup}} (C \setminus A_i)$

Theorem 1.9 Let $A, B \subseteq \mathcal{U}$ be any two sets. Then,

(i) $A \subseteq B$ iff $\mathcal{U} \setminus B \subseteq \mathcal{U} \setminus A$

(ii) $A \subseteq B$ iff $A \cap (\mathcal{U} \setminus B) = \varnothing$

Proof

(i) First, we prove that $A \subseteq B$ implies that $\forall x \in \mathcal{U} \setminus B$, $x \in \mathcal{U} \setminus A$. Let's see:
Suppose $x \in \mathcal{U} \setminus B$ then $x \notin B$. On the other hand, since $A \subseteq B$ if $y \in A$, then $y \in B$ too, which implies that for any $y \notin B$, $y \notin A$. Thus, $x \notin B$ implies that $x \notin A$ and therefore $x \in U \setminus A$. Hence, $\mathcal{U} \setminus B \subseteq \mathcal{U} \setminus A$.
Suppose $\mathcal{U} \setminus B \subseteq \mathcal{U} \setminus A$. We need to prove that it implies $A \subseteq B$. Well, if $x \in \mathcal{U} \setminus B$ then $x \in \mathcal{U} \setminus A$, which furthermore implies that *if* $x \notin B$ then $x \notin A$, and since $\mathcal{U} \setminus B \subseteq \mathcal{U} \setminus A$, it follows that $A \subseteq B$ as claimed.

(ii) First, we prove that $A \subseteq B$ implies $A \cap (\mathcal{U} \setminus B) = \varnothing$:
Suppose $A \subseteq B$, then for any $x \in A$ is true that $x \in B$. Therefore, $x \notin \mathcal{U} \setminus B$, and thus
$A \cap (\mathcal{U} \setminus B) = \varnothing$.
Next, let $A \cap (\mathcal{U} \setminus B) = \varnothing$. We need to prove that it implies that $A \subseteq B$. Consider

$$\begin{aligned} A \cap (\mathcal{U} \setminus B) &= A \cap (\mathcal{U} \cap B^c) \\ &= (A \cap \mathcal{U}) \cap B^c \\ &= (A \cap B^c) \\ &= A \setminus B \\ &= \varnothing \end{aligned}$$

Thus, $A \subseteq B$, as claimed. ■

1.5 THE POWER SET

Definition 1.19 (Power set) Given a set X, the set of all subsets of the set X, is called the **power set of** X, that is

$$\mathcal{P}(X) = \{A | A \subseteq X\}$$

The existence of a power set is postulated by

Axiom 9 For any set X, there is a set consisting of all the subsets of X.

It is easy to convince yourself that the following theorem is true.

Theorem 1.10 For any set X, $\varnothing,\ X \in \mathcal{P}(X)$.

Example 1.59 Let $X = \{a, b, c\}$, then $\mathcal{P}(X) = \{\ \varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$.

Note that the empty set and the set itself are considered members of this set of sets. ■

Example 1.60

(i) What is the power set of the empty set?
(ii) What is the power set of $\{\varnothing\}$?

Solution

(i) Since $\varnothing$ is a subset of any set, set $\varnothing$ has only one subset, namely itself. Therefore,

$$\mathcal{P}(\varnothing) = \{\varnothing\}$$

(ii) By the definition of the power set, the set $\{\varnothing\}$ has exactly two subsets: $\varnothing$, and the set $\{\varnothing\}$ itself, that is

$$\mathcal{P}(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}$$

■

Example 1.61 Let $S_{n+1} = S_n \cup \mathcal{P}(S_n)$, with $S_0 = \varnothing$. Then, we can recursively construct the sequence of sets as follows:

$$\begin{aligned}
S_0 &= \varnothing \\
S_1 &= S_0 \cup \mathcal{P}(S_0) = \varnothing \cup \mathcal{P}(\varnothing) = \varnothing \cup \{\varnothing\} = \{\varnothing\} \\
S_2 &= S_1 \cup \mathcal{P}(S_1) = \{\varnothing\} \cup \mathcal{P}(\{\varnothing\}) \\
&= \{\varnothing\} \cup \{\varnothing, \{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} \\
S_3 &= S_2 \cup \mathcal{P}(S_2) = \{\varnothing, \{\varnothing\}\} \cup \mathcal{P}(\{\varnothing, \{\varnothing\}\}) \\
&= \{\varnothing, \{\varnothing\}\} \cup \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\} \\
&= \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}
\end{aligned}$$

and so on. ■

Theorem 1.11 Let A and B be any two sets. If $A \subseteq B$, then $\mathcal{P}(A) \subseteq \mathcal{P}(B)$.

Proof We have to show that $\forall X \in \mathcal{P}(A), X \in \mathcal{P}(B)$.

First note that if $A \subseteq B$, then every subset of A is also a subset of B. Suppose that $X \in \mathcal{P}(A)$. Now recall that, by the definition of power set, X being an element of $\mathcal{P}(A)$, has to be a subset of A. But $A \subseteq B$, so $X \subseteq B$ too. It follows immediately that $X \in \mathcal{P}(B)$, which was to be shown. Thus,

$$\mathcal{P}(A) \subseteq \mathcal{P}(B)$$ ■

The next question we may ask is: Given a set X with n elements, how do we find the number of subsets of X? The following theorem answers this question.

Theorem 1.12 Let X be any set such that $|X| = n$, $n \in \mathbf{N}$. Then, $|\mathcal{P}(X)| = 2^{|X|} = 2^n$.

For the proof, we need the following lemma.

Lemma 1.1 Let X be any set, and let $x_0 \in X$ be *any* element of X, then there are as many subsets of X that contain x_0 as there are subsets of X that do not contain x_0.

We reason as follows: suppose we take a set $X_n = \{x_1, x_2, x_3, \ldots, x_n\}$ and $X_{n-1} = \{x_1, x_2, x_3, \ldots, x_{n-1}\}$, that is, a set with one, say x_0, fewer elements than X_n. Evidently, $X_{n-1} \subseteq X_n$. Then, we argue, if we collect all the subsets of X_{n-1} together with those same subsets, where each one of them is adjoined with $x_0 \in X_n$, we will get twice as many subsets of X_n than of X_{n-1}. Formally, and more precisely, the proof of the lemma goes as follows:

Proof of Lemma Let's express the set X as a union of two subsets $A = \cup_i A_i$ and $B = \cup_j B_j$, that is, the union of collections of subsets A_i and B_j.

$$\begin{aligned} X &= A \cup B \\ &= (\cup_i A_i) \cup (\cup_j B_j) \end{aligned}$$

such that $x_0 \in A_i$, $\forall\, i$, and $x_0 \notin B_j$, $\forall j$. In other words, every A_i is a subset of X, and every B_j is a subset of $X \setminus \{x_0\}$.

Observe that the number of subsets in collection A is the same as the number of subsets in collection B. Indeed, every B_j subsets of X that do not contain x_0 can be matched up with $B_j \cup \{x_0\} = A_j$. Thus, there are as many subsets of X that contain x_0 as there are those that do not. ■

Now, we proceed with the proof of Theorem 1.12.

Proof [31] First note that Examples 1.59 and 1.60 in particular are in accord with the theorem. We need to prove that the theorem holds for any set X.

Consider the statement of the theorem when $n = 0$. We ask if a set with zero elements, that is, the empty set, has $2^0 = 1$ subset? The answer is yes, as we have shown in Example 1.60. So our theorem is true in the case $n = 0$. Let's assume that it is also true for $n = k$, that is, we assume that any set with k elements has 2^k subsets. If we could show that the theorem is also true for $n = k + 1$, then it is true for any n.

Let X be a set with $k + 1$ elements, and let $x_0 \in X$. From the previous lemma, we have learned that there is an equal number of subsets of X that contain x_0, and those that do not. What does that mean? Well – and this is the crux of the matter – that tells us there are twice as many subsets of X as there are subsets of $X \setminus \{x_0\}$. But $|X \setminus \{x_0\}| = k$, that is, $X \setminus \{x_0\}$ has k elements by our assumption, hence the number of subsets of $X \setminus \{x_0\} = 2^k$, that is

$$|\mathcal{P}(X \setminus \{x_0\})| = 2^k$$

as our inductive hypothesis required.

It follows that the number of subsets of X equals twice the number of subsets of $X \setminus \{x_0\}$, that is

$$|\mathcal{P}(X)| = 2 \cdot 2^k = 2^{2k+1}$$

as was to be shown.

In other words, the important conclusion is

$$|\mathcal{P}(X)| = 2^{|X|}$$ ■

At this point, it may be intuitively clear to everyone that the power set of any finite set, regardless of its size, is again a finite set. For infinite sets, of course, power sets are infinite.

1.6 THE CARTESIAN PRODUCT

Definition 1.20 Let $n \in \mathbf{N}$, and let $x_1, x_2, \ldots, x_n$ be a collection of n, not necessarily distinct, elements. We say that $(x_1, x_2, \ldots, x_n)$ is an **ordered *n*-tuple** of n elements, in which we distinguish the first, the second, and so on elements.

Definition 1.21 Two ordered n-tuples $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ are said to be **equal** iff

$$x_1 = y_1, x_2 = y_2, \ldots, x_n = y_n$$

[31] If you are not familiar with the technique of mathematical induction, you can skip this proof in the first reading. After mathematical induction is introduced in the following chapters, you can come back to the proof.

Definition 1.22 (Cartesian product) Let A and B be two sets. The **Cartesian product** of A and B, denoted $A \times B$, is the set

$$A \times B = \{(a, b) |\ a \in A, b \in B\}$$

Given n sets $A_1, A_2, \ldots, A_n$, then the n-fold Cartesian product of $A_1, A_2, \ldots, A_n$ is

$$\begin{aligned} A_1 \times A_2 \times \cdots \times A_n &= \prod_i^n A_i \\ &= \{(a_1, a_2, \ldots, a_n) | a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n\} \end{aligned}$$

Example 1.62 Let $A = \{a, b, c\}$ and $B = \{1, 2, 3, \}$ then

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

■

Theorem 1.13 Let $A_1, A_2, \ldots, A_n$ be sets, where $n \in \mathbf{N}$, and $n \geq 2$, then the Cartesian product $A_1 \times A_2 \times \cdots \times A_n$ is a set defined recursively by

$$A_1 \times A_2 \times \cdots \times A_n = A_1 \times (A_2 \times A_3 \cdots \times A_n)$$

Example/Exercise 1.63 Convince yourself that

$$A \times B \neq B \times A$$

Example/Exercise 1.64 Prove that $A \times \varnothing = \varnothing \times A = \varnothing$.

From the aforementioned discussion, we conclude that if A and B are (finite) sets, and if one of them is empty, then the Cartesian product $A \times B$ is empty. In other words, if neither A nor B is empty, then there is $a \in A$ and $b \in B$ so that $(a, b) \in A \times B$. The rather difficult question is : Can we generalize this to infinite sets, that is, can we say that *the Cartesian product of a nonempty family of nonempty sets is nonempty*?

Example/Exercise 1.65 Suppose that $A \neq \varnothing$, and that $B \neq \varnothing$. Show that

$$A \times B = B \times A, \quad \text{iff } A = B$$

Theorem 1.14 If $A, B,$ and C are sets, then

(i) $(A \cup B) \times C = (A \times C) \cup (B \times C)$

(ii) $(A \cap B) \times C = (A \times C) \cap (B \times C)$
(iii) $(A \setminus B) \times C = (A \times C) \setminus (B \times C)$

Definition 1.23 If $A = B$, then we write $A \times A = A^2$.

Example 1.66 $\mathbf{R} \times \mathbf{R} = \mathbf{R}^2$ is our familiar 2-dimensional Euclidean plane. ■

1.7 THE SETS N, Z, AND Q

The sets of numbers **N**, **Z**, and **Q** have been mentioned several times already, but now we want to address some more interesting things about them.

The set of **natural numbers** is a collection

$$\mathbf{N} = \{1, 2, 3, \ldots\}$$

As was mentioned before, one will often find that some authors, especially those working in mathematical logic and computer science, prefer to include "0" (zero) in the set **N**, which is mostly for convenience. Consider this:

Suppose we came up with numerals such as this:

$$\mathit{I}, \mathit{II}, \mathit{III}, \mathit{IIII}, \ldots$$

Such a sequence can be considered a counterpart of natural numbers $1, 2, 3, 4, \ldots$ constructed with only one object "I." On the other hand, if we wanted to begin with zero, construction of our sequence would require two objects "**0**" and "I" and we would have

$$\mathit{0}, \mathit{0I}, \mathit{0II}, \mathit{0III}, \ldots$$

representing $0, 1, 2, 3, \ldots$. So, it is debatable whether it is advantageous to consider zero as a natural number. I hope the reader won't find this confusing, since it will be evident from the very context of every argument what is meant by the set **N**.

Also, you will often hear that the set **N** is called the set of **counting numbers**, or even the set of **nonnegative integers.** However, **natural numbers** is the name most commonly used, and it is historically the most appropriate one.

More importantly, note that whether you are expressing the set of natural numbers with zero or as $\mathbf{N} = \{1, 2, 3, \ldots\}$, the amount of information contained in this notation is astounding. Namely, just a few elements of this set, that is, "1," "2," "3," with the ellipsis " … " following them, suffice to "completely describe" the whole (infinite) set. In other words, our mind is able to grasp the enormous amount of information contained in **N** by recognizing just a few "examples" and that very significant "dot, dot, dot." We feel that we know exactly what kind of

numbers we are talking about when discussing the set **N**. But – what is a (natural) number? Well, assuming that we know what we mean by the terms "number," "1," and "successor," we can formally define a set of natural numbers by using axioms due to the Italian mathematician Giuseppe Peano.[32]

P1.: 1 is a natural number, that is, $1 \in \mathbf{N}$ (i.e., $\mathbf{N} \neq \emptyset$).[33]

P2.: $\forall\, n \in \mathbf{N},\ \exists\, n' = S(n) \in \mathbf{N}$, called the successor of n.

P3.: $n' \neq 1$, that is, there exists no number whose successor is 1.

P4.: If $n' = m'$ then $n = m$, that is, there is no number or there is exactly one number whose successor is the given number.

More generally, we can state Peano's axioms, and this time including "zero," as follows:

Let X be a set such that:

(P1′): There is a special element $0_X \in X$.

(P2′): There is a function $S : X \to X$ such that the following holds: For every $x, y \in X$, if $x' = S(x) = S(y) = y'$then $x = y$.

(P3′): For every $x \in X$, $0_X \neq S(x)$.

(P4′): For every $A \subseteq X$, if $0_X \in A$ and $S(x) \in A$ whenever $x \in A$, then $A = X$.

If we take X to be the set **N** with $0_X = 0$, that is, $X = \mathbf{N} = \{\, 0, 1, 2, 3, \ldots \}$ and defining the function S by $n \mapsto n + 1$, we see that **N** satisfies axioms $P1' - P4'$.

Theorem 1.15 The set **N** with a special element 0 and the successor function S defined by $n \mapsto n + 1$ satisfies Peano's axioms.

Definition 1.24 (Russell's hereditary principle) A property is said to be "hereditary" in the natural number series if, whenever it belongs to a number n it also belongs to $n + 1$. Similarly, a set is said to be "hereditary" if, whenever n is an element of a set, so is $n + 1$.[34]

Speaking of sets in everyday parlance, we usually think of them as a collection of objects, whatever the "objects" are. In mathematics, however, we can equally well speak (and we often do) of "pure sets" – sets whose members are other pure sets, like the empty set itself. Can we use those to construct other familiar sets?

At this point, you may recall Axioms 0–6 and revisit our discussion on pages 19 and 20, and in particular Example 1.13, where we listed a sequence of sets:

[32] Giuseppe Peano (1858–1932).

[33] One can equally well take zero to be the element of **N** and start with it as the first natural number.

[34] Bertrand Russell, *Introduction to Mathematical Philosophy.*

$\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\{\varnothing\}, \{\{\varnothing\}\}\}, \ldots$. One way to construct natural numbers could be as follows:

$$\begin{aligned}
\mathbf{0} &= \varnothing = \{\} \\
\mathbf{1} &= \{0\} = \{\{\}\} \\
\mathbf{2} &= \{0, 1\} = \{0, \{0\}\} = \{\{\}, \{\{\}\}\} \\
\mathbf{3} &= \{0, 1, 2\} = \{0, \{0\}, \{0, \{0\}\}\} = \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}\} \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
\mathbf{n} &= \{0, 1, 2, \ldots, n-2, n-1\} = \{0, 1, 2, \ldots, n-2\} \cup \{n-1\} \\
&= (n-1) \cup \{n-1\}
\end{aligned}$$

On the other hand, with the empty set $\varnothing = \{\ \}$ and a successor function[35] defined by

$$S(x) = \{x\}$$

we can have

$$\begin{aligned}
\mathbf{0} &= \varnothing = \{\} \\
\mathbf{1} &= S(0) = \{\varnothing\} = \{\{\}\} \\
\mathbf{2} &= S(1) = \{1\} = \{\{\{\}\}\} \\
\mathbf{3} &= S(2) = \{2\} = \{\{\{\{\}\}\}\}
\end{aligned}$$

and so on.

We can say that each natural number n is equal to the set of the natural number preceding it, $1, 2, 3, \ldots, n-1$.

Alternatively, defining zero as

$$\mathbf{0} = \{\{\}\}$$

and the successor of x as

$$S(x) = x \cup \{x\}$$

we have

$$\begin{aligned}
S(\varnothing) &= \varnothing \cup \{\varnothing\} \\
S(S(\varnothing)) &= S(\varnothing) \cup \{S(\varnothing)\} \\
&= \varnothing \cup \{\varnothing\} \cup \{\varnothing \cup \{\varnothing\}\}
\end{aligned}$$

[35]Here, for the sake of simplicity, we will designate a generic set by a lowercase x.

$$S(S(S(\varnothing)) = S(\varnothing) \cup \{S(\varnothing)\} \cup \{S(\varnothing) \cup \{S(\varnothing)\}\}$$
$$= \varnothing \cup \{\varnothing\} \cup \{\varnothing \cup \{\varnothing\}\} \cup \{\varnothing \cup \{\varnothing\} \cup \{\varnothing \cup \{\varnothing\}\}\}$$

and so on (cf. Example 1.61).

Thus, our newly designed natural numbers look like this:

$$\mathbf{0} = \{\{\}\}$$
$$\mathbf{1} = \{\{\}, \mathbf{0}\} = \{\{\}, \{\{\}\}\}$$
$$\mathbf{2} = \{\{\}, \mathbf{0}, \mathbf{1}\}$$

and so on. (In those examples, I purposely wrote natural numbers bold-faced to emphasize their "set-theoretical nature.")

Now, let me show you two things that can cause you some headache.

First, suppose we ask: Is it true that $S(x)$ has one element more than the set x? (One would expect that this is indeed true. After all, that's exactly how we constructed $S(x)$.) Well, let's see. Since $S(x) = x \cup \{x\}$, certainly $x \subseteq S(x)$. Now, $S(x)$ obviously contains x, which is also an element of $\{x\}$. But –and now comes the caveat – in order for this element (i.e., $\{x\}$) to be an *extra* element, we need $x \notin x$ (?!). On the other hand, if $x \in x$ then $\{x\}$ is a subset of x, and then $x \cup \{x\} = x$.

Second, as you might have anticipated, the three different ways (defined earlier) of identifying natural numbers with pure sets are not the only ones – there are infinitely many. What one would expect though is that they are all equivalent. Well, let's see. Consider only two versions:

(i)

$$\mathbf{0} = \varnothing$$
$$\mathbf{1} = \{\varnothing\}$$
$$\mathbf{2} = \{\{\varnothing\}\}$$
$$\mathbf{3} = \{\{\{\varnothing\}\}\}$$
$$\vdots$$

and

(ii)

$$\mathbf{0} = \varnothing$$
$$\mathbf{1} = \{\varnothing\}$$
$$\mathbf{2} = \{\varnothing, \{\varnothing\}\}$$
$$\mathbf{3} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$$
$$\vdots$$

Obviously, the 3 from (i) and the 3 from (ii) are not the same. From the set-theoretic standpoint $\{\{\{\varnothing\}\}\} \neq \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$. So, the question "What is a number?" is not as trivial as some might have thought.

Let's introduce another concept, which will prove to be very useful later.

Definition 1.25 We say that a set I is *inductive* if $\varnothing \in I$, and if for all $x \in I$ the successor $S(x) \in I$.

Do inductive sets exist? We will assume that there exists at least one inductive set.

Theorem 1.16 If two sets I and J are the inductive sets, then $I \cap J$ is also inductive.

Proof Following Definition 1.24, we need to show that

(i) $\varnothing \in I \cap J$, and

(ii) whenever $x \in I \cap J$, then $S(x) \in I \cap J$ too

For (i): Since both I and J are inductive, $\varnothing \in I$ and $\varnothing \in$ J, thus $\varnothing \in I \cap J$.

For (ii): If $x \in I \cap J$, then $x \in I$ and $x \in J$. But since I and J are inductive, $S(x) \in I$ and $S(x) \in J$. Hence, $S(x) \in I \cap J$. ■

As a simple exercise, you can now prove.

Theorem 1.17 The set **N** is inductive.

In more general terms, we state

Principle of Induction:

Let X be some set with $0_X \in X$ such that for all properties P, if 0_X has property P, and the successor function $S(x)$ has the same property P whenever $x \in X$ has it, then every element of X has property P.

This becomes "obvious" if we take $X = \mathbf{N}$ and $0_X = 0$. We will have to say more about the principle of induction later but for now let's illustrate it with

Theorem 1.18 Let X be the set that satisfies Peano's axioms. Then, for every $x \in X$ different from 0_X there exists $y \in X$ such that $x = S(y)$.

Proof Let $A = \{x \in X | x = 0_X \text{ or } x = S(y), y \in X\}$.

By definition, $0_X \in A$. On the other hand, if x is an element of A then, again by definition, there has to be a $y \in X$, such that $x = S(y) \in A$ and therefore $S(x) = S(S(y)) \in A$. Thus, $A = X$. In other words, for every $x \neq 0_X$ there exists $y \in X$, such that $x = S(y)$.

We continue by "describing" a set of **integers** as a collection

$$\mathbf{Z} = \{ \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots \}$$

where, as in the case of natural numbers, a few elements of the set, together with "$\ldots$," capture much of the relevant information about the set **Z**. (The symbol **Z**, which Cantor used to denote integers, comes from the German word *die Zahl = a number, Zahlen = to number*.)

At this point, we want to list the rules of arithmetic, the "**axioms of the set Z**," which are generally well known but rarely justified in introductory textbooks. Also, these rules, as well as many of the "everybody-knows-it" facts, point to some more advanced algebraic structures that will be studied later.

Consider the set **Z** with two operations defined on it: addition "+," and multiplication "·," so that from now on we will be working with the structure $(\mathbf{Z}; +, \cdot)$. Hence our **rules of arithmetic** are as follows:

1. $a + b \in \mathbf{Z}, \ \forall\, a, b \in \mathbf{Z}$
2. $a \cdot b \in \mathbf{Z}, \ \forall\, a, b \in \mathbf{Z}$
3. $a + (b + c) = (a + b) + c, \ \forall\, a, b, c \in \mathbf{Z}$
4. $a \cdot (b \cdot c) = (a \cdot b) \cdot c, \ \forall\, a, b, c \in \mathbf{Z}$
5. $a + b = b + a, \ \forall\, a, b \in \mathbf{Z}$
6. $a \cdot b = b \cdot a, \ \forall\, a, b \in \mathbf{Z}$
7. $a \cdot (b + c) = a \cdot b + a \cdot c, \ \forall\, a, b, c \in \mathbf{Z}$
8. $\exists\, 0 \in \mathbf{Z}$, s.t. $0 + a = a + 0 = a, \forall\, a \in \mathbf{Z}$
9. $\exists\, 1 \in \mathbf{Z}$, s.t. $1 \cdot a = a \cdot 1 = a, \ \forall\, a \in \mathbf{Z}$
10. $\exists\, (-a) \in \mathbf{Z}$, s.t. $a + (-a) = (-a) + a = 0, \ \forall\, a \in \mathbf{Z}$ ■

Example 1.67 Prove that (i) the additive and (ii) the multiplicative identities are unique. ■

Proof (i) Suppose there are two additive identity, 0 and $0'$, then according to rule (7)

$$0 + 0' = 0' \quad \text{since 0 is an additive identity. But}$$
$$0 + 0' = 0 \quad \text{since } 0' \text{ is an additive identity too. Therefore}$$
$$0 = 0 + 0' = 0'$$

Now you should be able to prove part (ii). ■

Example 1.68 Prove that

$$a \cdot 0 = 0, \quad \forall a \in \mathbf{Z}$$ ■

Proof

$$a \cdot 0 = a \cdot (b + (-b)) = a \cdot b - a \cdot b = 0$$ ■

Example/Exercise 1.69 Show that if $a \in \mathbf{Z}$, then

$$(-1)\, a = -a$$

Example/Exercise 1.70 Show that if $a,\ b \in \mathbf{Z}$, then

(i) $(-a)b = a(-b) = -ab$
(ii) $(-a)(-b) = ab$

Example/Exercise 1.71 Show that $\forall a, b, c \in \mathbf{Z}$, and $a \neq 0$, if $ab = bc$, then

$$a = c$$

Example/Exercise 1.72 Show that if $a, b \in \mathbf{Z}$, and $a \cdot b = 0$, then either $a = 0$ or $b = 0$.

Here is another property of the set $\mathbf{Z}$ by the name of the

Trichotomy Law

$\forall a, b \in \mathbf{Z}$, only one of the following holds

(i) $a < b$
(ii) $a = b$
(iii) $a > b$

Example 1.73 Prove that for any $a \in \mathbf{Z}$, $a > 0$ iff $-a < 0$. ■

Proof Suppose $a > 0$, then

$$a\ + (-a) > 0 + (-a) > (-a)$$

which implies

$$0 > (-a)$$ ■

Example/Exercise 1.74 Prove that for any $a, b \in \mathbf{Z}$, such that $a > 0$ and $b < 0$,

$$a \cdot b < 0$$

Example 1.75 Prove that for any $a, b \in \mathbf{Z}$, such that $a < 0$ and $b < 0$,

$$a \cdot b > 0$$ ■

Proof Suppose $a < 0$ and $b < 0$, then $-a > 0$ and $-b > 0$. Hence

$$(-a) \cdot (-b) = a \cdot b > 0$$ ■

Example 1.76 Let $a, b \in \mathbf{Z}$, and $a > 0, b > 0$. Prove that $a < b$ iff $a^2 < b^2$. ■

Proof Suppose $a > 0$ and $b > 0$, and, furthermore, suppose that $a < b$, then, since $a < b$,

$$a^2 < a \cdot b < b^2$$

as was to be shown.

Now suppose that $a^2 < b^2$. Then,

$$a \cdot a < a \cdot b < b \cdot b$$

Therefore, $a < b$, as claimed. ■

Example/Exercise 1.77 Let $a, b \in \mathbf{Z}$, and let $a < 0$ and $b < 0$. Show that $a < b$ iff $b^2 < a^2$.

Theorem 1.19 There are no integers between 0 and 1.

Proof Suppose there is a set

$$A = \{a \in \mathbf{Z} | 0 < a < 1\}$$

Suppose, furthermore, that $A \neq \emptyset$. Then there is a least element $a_0 \in A$. Now, a_0 being an element of A means that $0 < a_0 < 1$, which implies that $0 < a_0^2 < a_0$. But then it follows that $a_0^2 \in A$, and therefore a_0 is not the least element of A. Hence $A = \emptyset$, that is, there are no elements between 0 and 1 in **Z**. ■

So far, we haven't discussed the numbers of the form a/b, where $a, b \in \mathbf{Z}$ and $b \neq 0$. Those are ostensibly fully "legitimate" numbers and we have to include them in our family of numbers.

In order to describe those numbers, called **rational numbers Q**, we cannot proceed in the same way as before, that is, we cannot give a few examples that would be sufficient to encapsulate all properties of the set **Q**. We need to refer to set theory. So we define the set of all rational numbers **Q** as follows.

Definition 1.26 We say that the set

$$\mathbf{Q} = \left\{ x \,\middle|\, x = \frac{p}{q}\ ,\ \ p, q \in \mathbf{Z},\ \ q \neq 0 \right\}$$

is the set of rational numbers. In other words, we say a number x is rational if and only if $x = p/q$ for some integers p and q, with $q \neq 0$. In addition, to make things simpler, occasionally we request that p and q be relatively prime, that is, that there is no number that divides p and q at the same time (except, of course the number 1). By doing this, we are simply collecting all the numbers expressible as a quotient of two integers reduced to simplest form.

Observe that, based on everything we have discussed so far,

$$\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q}$$

which makes the following theorem obvious.

Theorem 1.20 Every integer is a rational number.

Proof It's easy – you should do it! ■

Theorem 1.21 The sum of two rational numbers is rational.

Proof Suppose $x, y \in \mathbf{Q}$. Then, by Definition 1.26, we know that $x = a/b$ and $y = c/d$ for some $a, b, c, d \in \mathbf{Z}$, with $b \neq 0, d \neq 0$. Then,

$$x + y = \frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd}$$

$ad + cb$ is the sum of two integers, therefore an integer, say, p, and bd as the product of two integers is also an integer, say, q. So we have a quotient of two integers p/q, with $q \neq 0$. Hence, $x + y$ is a rational number. ■

Theorem 1.22 The set $\mathbf{Q}$ is *dense*, that is, between any two rational numbers there is at least another one, that is

$$\forall a, b \in \mathbf{Q}, \quad (a < b), \exists\, c \in \mathbf{Q}, \text{ such that } a < c < b$$

Thus, there are infinitely many.

Proof If $a, b \in \mathbf{Q}$, then $a = m/n$ and $b = p/q$. Consider

$$c = \frac{a + b}{2} = \frac{mq + mp}{2nq}$$

c, itself a rational number, is obviously an arithmetic mean of two rational numbers a and b, that is

$$a < c < b$$

as was to be shown. ■

I want to show you some less obvious, and rather intriguing, properties of sets **N**, **Z**, and **Q.** First, recall that

(i) Set **N** has a least element but not a greatest.
(ii) Set **Z** has neither a least nor a greatest element.
(iii) Both sets **N** and **Z** are infinite.

Also, $\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q}$, and since **N** and **Z** are infinite sets, **Q** has to be infinite too.

But how "big" are the infinities of **N** and **Z** and **Q**? In other words, if $\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q}$, how do we compare those "three infinities?" Recall, in Definition 1.8, we said that two sets A and B are equivalent if and only if their cardinal numbers are the same, that is, they have the same number of elements. We need to examine the "number" of elements in infinite sets.

We will follow Cantor and call the cardinal number of **N**, $\aleph_0$ (aleph zero), that is, we say

$$|\mathbf{N}| = \aleph_0$$

Now, what about $|\mathbf{Z}|$ and $|\mathbf{Q}|$, and what about $\aleph_0$ itself? To address those questions, and some others pertaining to set **R**, we need to introduce briefly one of the most important concepts in the whole of mathematics – the concept of a function. We will devote much more time to functions later (see Chapter 4), but for now we will just state two (equivalent) definitions.

Definition 1.27 Given two sets X and Y, we say that a **function f from set X to set Y** is a map that assigns to every element of X a unique element of Y (Figure 1.14). We write this as follows:

$$f : X \to Y$$

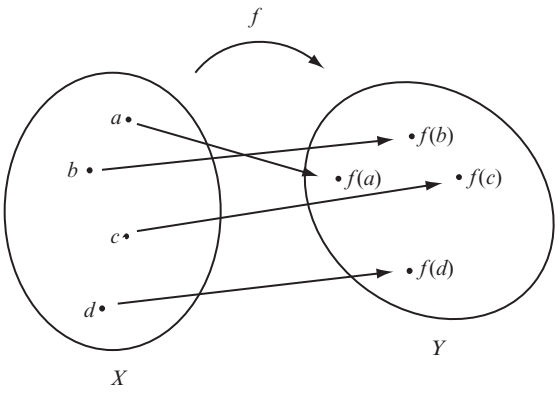


Figure 1.14 Function $f : X \to Y$.

Set X is called the **domain** of f and Y the **codomain** of f.[36]

[36]Some finesses in the definitions of domain, codomain, and range will be addressed in Chapter 4.

Sometimes, it is convenient to simply write

$$X \xrightarrow{f} Y$$

If there is no need to explicitly name the function, we abbreviate the notation by writing

$$x \mapsto f(x)$$

For example, if $x \in \mathbf{R}$, $x \mapsto x^2$ would indicate the function $f : \mathbf{R} \to \mathbf{R}$ that maps every real number to its square (Figure 1.15).

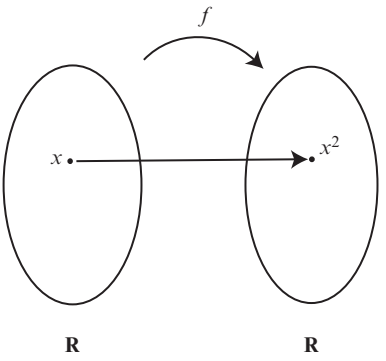


Figure 1.15 Function $x \mapsto x^2$

Definition 1.28 A function f from set X into set Y, is a set of all ordered pairs (x, y), where for all $x \in X$ there exists a unique $y \in Y$, such that $(x, y) \in f$, that is

$$f = \{(x, y) | x \in X, y \in Y\}$$

$f(x) \in Y$ is said to be an image of $x \in X$. We say that set X is the **domain** of f, set Y is the **codomain**, and the set of all images of elements of X is the **range** of f.

Definition 1.29 Given a function $f : X \to Y$, and $A \subseteq X$, we say that the set

$$f[A] = \{f(x) | x \in A\}$$

is the *image* of A under action of f.

Consequently, if $B \subseteq Y$ we call

$$f^{-1}[B] = \{x \in X | f(x) \in B\}$$

the *preimage* of B under action of f.

Definition 1.30 Two functions are said to be *equal* if and only if they have the same domain and assign the same value to every member of their common

domain. Symbolically,

$$f = g \leftrightarrow (\forall x \in X, \quad f(x) = g(x))$$

Definition 1.31 The function $f : X \to Y$ is said to be **one-to-one**, (**1 − 1**, or an **injection**) if and only if

$$\forall x_1, x_2 \in X, \text{ if } f(x_1) = f(x_2) \text{ then } x_1 = x_2$$

or equivalently

$$\forall\, x_1,\ x_2 \in X, \text{ if } x_1 \neq x_2 \text{ then } f(x_1) \neq f(x_2)$$

Often injections are designated with the special arrow "$\rightarrowtail$" (Figure 1.16), so for an injection we write

$$f : X \rightarrowtail Y$$

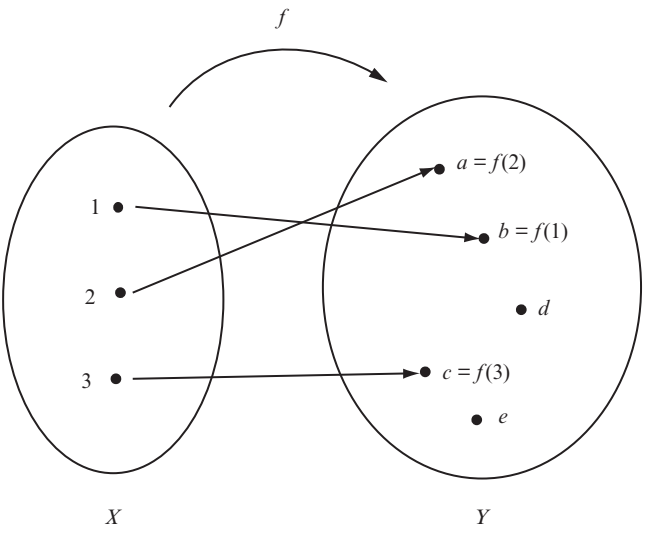


Figure 1.16 One-to-one function

Example/Exercise 1.78[37] Let $f : X \rightarrowtail Y$ be an injection, and let $A, B \subseteq X$. Show that

(i) $f[A \cap B] = f[A] \cap f[B]$
(ii) $f[A \cup B] = f[A] \cup f[B]$

[37] For this example and the others involving functions, you may want to consult Chapter 4.

Definition 1.32 We say that a function $f : X \to Y$ is **onto** (or a **surjection**, Figure 1.17) iff

$$\forall y \in Y, \exists\, x \in X, \text{ such that } f(x) = y$$

(Sometimes, we use "$\twoheadrightarrow$" to indicate surjection.)

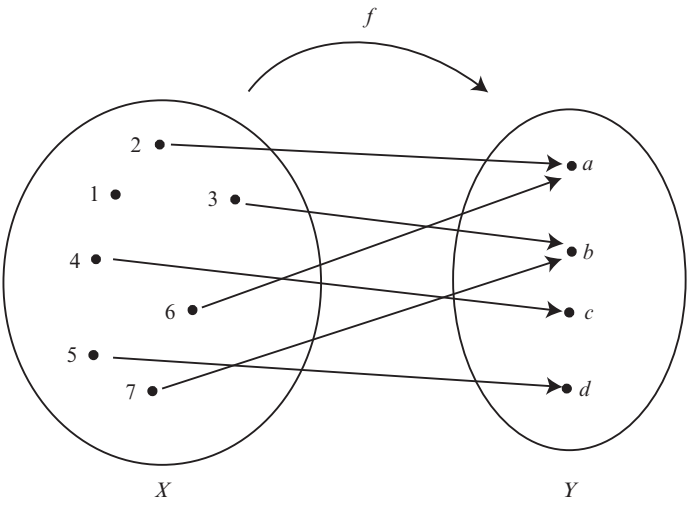


Figure 1.17 Surjection

Definition 1.33 A function $f : X \to Y$, that is both *one-to-one and onto*, we call a **bijection** or a **one-to-one correspondence** (Figure 1.18) between sets X and Y (sometimes, we use "$\rightarrowtail$" to indicate bijection).

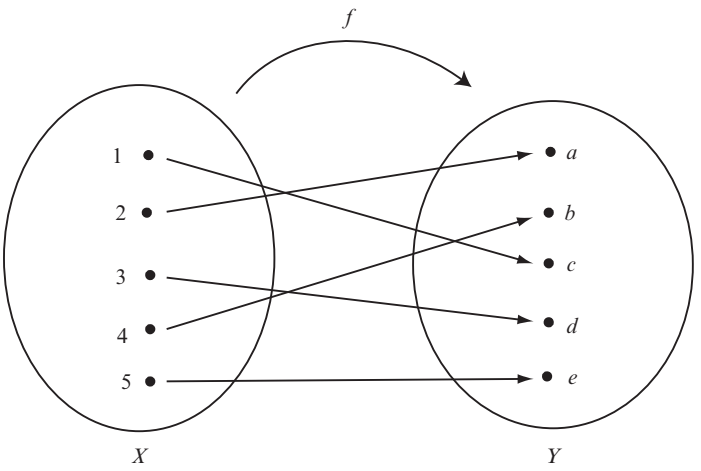


Figure 1.18 Bijection (one-to-one correspondence)

Example 1.79 Now we can restate Definition 1.6 and say: The cardinal number of a finite set A is a natural number n if there exists a bijection between A and the set $\{x \in \mathbf{N} \mid 1 \leq x \leq n\}$, that is

$$f : A \rightarrowtail\!\!\!\!\!\rightarrow \{x \in \mathbf{N} \mid 1 \leq x \leq n\}$$ ■

Similarly, the concept of the equivalence of sets can be restated more precisely:

Definition 1.34 Given two sets X and Y, we say that they have the *same cardinality* and we write $|X| = |Y|$, iff there is a one-to-one correspondence between X and Y, that is, there exists a function $f : X \to Y$, which is one-to-one and onto. Recall, in Definition 1.10, we call sets with the same cardinality *equivalent* and we write $X \sim Y$.

Definition 1.35 We say that set X has more elements than set Y, if there exists a function $f : X \to Y$ which is onto, but no function $g : X \to Y$ which is one-to-one.

Theorem 1.23 If X and Y are any two sets, such that there exist one-to-one mappings $X \xrightarrow{f} Y$ and $Y \xrightarrow{g} X$, then $|X| = |Y|$.

Definition 1.36 Let $f : X \to Y$ be a bijection. We say that

$$f^{-1} : Y \to X$$

is the **inverse function** of f, if the following is true:

$$f^{-1}(y) = x \quad \text{iff} \quad f(x) = y$$

Definition 1.37 We say that $h = g \circ f : X \to Z$ is a **composition** of functions f and g, that is

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

if

$$h(x) = g(f(x))$$

Theorem 1.24 Let $f : X \to Y$ and $g : Y \to Z$ be two bijections. Then, $g \circ f$ is also a bijection.

Proof See Chapter 4. ■

Definition 1.38

A set A is said to be **finite** iff there is a bijection from a set $S = \{1, 2, 3, \ldots, n\}$ to A (see Example 1.79).

A set A is said to be infinite if there is no such bijection.

Example 1.80 Prove that if two **finite** sets X and Y have the same number of elements, that is $|X| = |Y| = n$, then there exists a function $h : X \to Y$, which is one-to-one and onto.

Solution According to Definition 1.38, a set A is said to be finite if there exists $n \in \mathbf{N}$, such that, given a set $S = \{1, 2, 3, \ldots, n\}$, there exists some function

$$f : S \to A$$

which is one-to-one and onto. Accordingly, for our sets X and Y, there exist functions

$$f : S \to X \text{ and } g : S \to Y$$

both one-to-one and onto. Since f is a bijection, by Definition 1.34, it follows that $f^{-1} : X \to S$ is a bijection too. Hence,

$$g \circ f^{-1} : X \to Y$$

is a bijection too. If we take $h = g \circ f^{-1}$ we have our proof. ■

The proofs of the following two, very important, theorems we leave for Chapter 4.

Theorem 1.25 (Schröder–Bernstein) If A and B are any two sets, and if there exist injections $f : A \to B$ and $g : B \to A$, then there exists a bijection between A and B, and thus $|A| = |B|$.[38]

Theorem 1.26 If A and B are any two sets, then exactly one of the following is true:

(i) $|A| = |B|$
(ii) $|A| < |B|$
(iii) $|A| > |B|$

[38] Ernst Schröder (1841–1902), German mathematician. Felix Bernstein (1878–1956), German mathematician.

Theorem 1.27 Let A, B and C be any three finite sets. Then,

(i) $|A \cup B| = |A| + |B| - |A \cap B|$

(ii) $|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$

(iii) $|A \times B| = |A| \cdot |B|$

Proof

(i) Note that

$$|A \cup B| = |A \setminus B| + |A \cap B| + |B \setminus A| \tag{1.5}$$

On the other hand, observe that

$$|A| = |A \setminus B| + |A \cap B| \tag{1.6}$$

and

$$|B| = |B \setminus A| + |A \cap B| \tag{1.7}$$

Combining (1.5)–(1.7), we get the desired result.

(ii) For this proof, we will use (i) and the following identities:

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \text{ and } (A \cap B) \cap (B \cap C) = A \cap B \cap C$$

So we have

$$\begin{aligned} |A \cup B \cup C| &= |A \cup B| + |C| - |(A \cap C) \cup (B \cap C)| \\ &= |A| + |B| - |A \cap B| + |C| - |A \cap C| \\ &\quad - |B \cap C| + |A \cap B \cap C| \\ &= |A| + |B| + |C| - |A \cap B| - |A \cap C| + |A \cap B \cap C| \end{aligned}$$

(iii) From the definition of the Cartesian product of two finite sets A and B, for any $(a, b) \in A \times B$, there are $|A|$ possibilities for a, and $|B|$ possibilities for b, and therefore $|A \times B| = |A| \cdot |B|$. ■

Theorem 1.28 If A is a *finite* set with cardinality k, and $x \notin A$, then $|A \cup \{x\}| = k + 1$.

Proof First note, that if $A = \varnothing$ then $|A| = |\varnothing| = 0$, therefore $|A \cup \{x\}| = 1 = 0 + 1$.

If $A \neq \varnothing$ then $A \sim \mathbf{N}_k$, where $\mathbf{N}_k = \{1, 2, \ldots, k\}$. It follows that

$$A \cup \{x\} \sim \mathbf{N}_k \cup \{k+1\} = N_{k+1}. \text{ Thus } |A \cup \{x\}| = k+1$$ ■

Things are quite different for infinite sets, as the following example illustrates.

Example 1.81 With $\mathbf{N} = \{1, 2, 3, \ldots\}$ show that $|\mathbf{N} \cup \{0\}| = \aleph_0$

Solution Consider a function $f : (\mathbf{N} \cup \{0\}) \to \mathbf{N}$ defined by

$$f(x) = x + 1$$

It is easy to see that f is a bijection.

Consequently, $|\mathbf{N} \cup \{0\}| = \aleph_0$. ■

The next theorem is almost trivial now. However, for its proof we need to invoke the technique of mathematical induction (see Chapter 3).[39]

Theorem 1.29 For every $k \in \mathbf{N}$, every subset A of $\mathbf{N}_k$ is finite.

Proof Let $k \in \mathbf{N}$ be any natural number and let $A \subseteq \mathbf{N}_k$. If $k = 1$ then $A = \varnothing$ or $A = \mathbf{N}_k$ and thus A is finite. Suppose that all the subsets of $\mathbf{N}_k$ are finite for some number k. Now, let $A \subseteq \mathbf{N}_{k+1}$, then $A \setminus \{k+1\} \subseteq \mathbf{N}_k$ which, by our induction hypothesis, is finite. Thus, A is finite. Suppose not. Then, we could write

$$A = (A \setminus \{k+1\}) \cup \{k+1\}$$

which is finite by the previous theorem. We conclude that for every $k \in \mathbf{N}$, every subset of $\mathbf{N}_k$ is finite. ■

Definition 1.39 We say that A is *less than* or *equinumerous* with B, if there is a one-to-one function $f : A \to B$, and we write $A \preccurlyeq B$.

Definition 1.40 A set A is *less than or equal* to B in "size" if it is equinumerous with at least one subset of B, that is

$$|A| \leq |B| \leftrightarrow (\exists C)(C \subseteq B \ \& \ |A| = |C|)$$

Theorem 1.30 For any sets A and B if $A \subseteq B$, then $A \preccurlyeq B$.

[39] You can skip this proof until you have read Section 3.4.

Proof Let $C \subseteq B$ be a set, such that $|A| = |C|$. Then, there has to be a bijection $f : A \to C$, which means that f is an injection from A to B. On the other hand, if there exists an injection $f : A \to B$, then the image $f[A]$ is a subset of B, that is, $f[A] \subseteq B$. However, $|A| = |f[A]|$, and thus, $A \preccurlyeq B$. ■

Now we return to our sets **N**, **Z**, and **Q**.

Definition 1.41 A set is called **countably infinite** or **denumerable** (sometimes, just **countable**)[40] iff it has the same cardinality as the set of natural numbers **N**. If that is not the case, we say that a set is **uncountable**.

Example 1.82 Let $\mathbf{N}_{\text{even}}$ be the set of all even natural numbers

$$\mathbf{N}_{\text{even}} = \{2, 4, 6, 8, \dots\}$$ ■

Obviously, $\mathbf{N}_{\text{even}} \subseteq \mathbf{N}$, so what is the cardinal number of $\mathbf{N}_{\text{even}}$? In order to make the answer rather obvious let's take $\mathbf{N} = \{1, 2, 3, \dots\}$ and then establish a one-to-one correspondence between **N** and $\mathbf{N}_{\text{even}}$ in the following way:

$$\begin{aligned} \mathbf{N} &= \{1, 2, 3, 4, \dots\} \\ &\updownarrow\ \updownarrow\ \updownarrow\ \updownarrow\ \dots \\ \mathbf{N}_{\text{even}} &= \{2, 4, 6, 8, \dots\} \end{aligned}$$

that is, we have obtained the following correspondence: $1 \leftrightarrow 2, 2 \leftrightarrow 4, 3 \leftrightarrow 6, 4 \leftrightarrow 8$, and so on.

In other words, we are considering a function $f : \mathbf{N} \to \mathbf{N}_{\text{even}}$ defined by

$$f(n) = 2n, \quad \forall n \in \mathbf{N}$$

Obviously, the described function is one-to-one and onto, therefore,

$$|\mathbf{N}| = |\mathbf{N}_{\text{even}}|$$

We came to the surprising and unexpected conclusion that, regardless of the fact that $\mathbf{N}_{\text{even}}$ is a proper subset of **N**, indeed just a "half" of **N**, they still have the same cardinality, that is, they are equivalent.

This shocking result, discovered by Cantor, disputed one of Euclid's famous axioms that seemed so "self-evident" for centuries: "*The whole is greater than its part.*" Euclid also stated "[*t*]*hings that coincide with one another are equal to*

[40] Sometimes, it is said that a set is *countable*, if it is either finite or denumerable.

one another." Obviously, "things" had to be reconsidered.

Reminding yourself of Definition 1.39, you may now try

Example/Exercise 1.83 Decide whether the following statements are true:

(i) $2\mathbf{Z} \preccurlyeq \mathbf{Z}$
(ii) $2\mathbf{Z} \sim \mathbf{Z}$

With the concept of equivalence, and Cantor's aforementioned discovery, we are now in a position to define an infinite set yet another way:

Definition 1.42 A set X is infinite if there exists at least one proper subset of X with the same cardinality as X, that is, a set is infinite if it is equivalent to at least one of its proper subsets.

Example/Exercise 1.84 Convince yourself that a set of all natural numbers and a set of all squares of natural numbers have the same cardinality.

Theorem 1.31 (Cantor) Let X be any set. Then, $|X| < |\mathcal{P}(x)|$.

Proof Let $f : X \to \mathcal{P}(X)$ be a function defined by $f(x) = \{x\}$, that is, to every $x \in X$ we associate a singleton $\{x\}$. It is easy to see that f is an injection. Indeed, if $f(x_1) = f(x_2)$ that is, $\{x_1\} = \{x_2\}$, then $x_1 = x_2$. Suppose that there also exists a bijection $g : X \to \mathcal{P}(X)$. Define

$$Y = \{x \in X | x \notin g(x)\}$$

Since g is a bijection, there exists a unique $x \in X$, such that $Y = g(x)$, and we ask: Is $x \in g(x)$ or not? Suppose $x \in g(x)$, then by definition of Y, $x \notin Y$; conversely, if $x \notin g(x)$ then $x \in Y$. But that contradicts our request that $Y = g(x)$. Thus the proof. Needless to say, the similarity with Russell's paradox is evident. You may want to compare this theorem with Theorem 1.12. ■

Corollary Set $\mathcal{P}(\mathbf{N})$ is uncountable.

Example/Exercise 1.85 Convince yourself that

(i) $\aleph_0 + \aleph_0 = \aleph_0$
(ii) $2\aleph_0 = \aleph_0$

Example 1.86 Show that **Z** is countable.

Solution As the definition of countability requires, we need to find a function $f : \mathbf{N} \to \mathbf{Z}$ that is one-to-one and onto. Let's rearrange the elements of the set **Z** this way:

$$\mathbf{Z} = \{0, 1, -1, 2, -2, 3, -3, 4, -4, \ldots\}$$

The pattern is self-evident, and we are sure that all integers have been "collected." Now, as before, we establish a correspondence

$$\begin{array}{rccccccccc} \mathbf{N} = \{1, & 2, & 3, & 4, & 5, & 6, & 7, & \ldots\} \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \cdots \\ \mathbf{Z} = \{0, & 1, -1, & 2, -2, & 3, -3, & \ldots\} \end{array}$$

It is clear from the aforementioned scheme that no integer has been missed or counted twice in the process of matching it with a corresponding natural number. The "function" defined by the aforementioned pattern is obviously the function $f : \mathbf{N} \to \mathbf{Z}$ given by

$$f(n) = \begin{cases} \frac{n}{2} & \text{if } n \text{ is an even natural number} \\ -\frac{n-1}{2} & \text{if } n \text{ is an odd natural number} \end{cases}$$

The function is one-to-one and onto, telling us that $|\mathbf{Z}| = |\mathbf{N}| = \aleph_0$. ■

Example/Exercise 1.87 Find another rearrangement of the elements of **Z** and **N** to establish a bijection between these two sets and prove that $|\mathbf{Z}| = \aleph_0$

Example/Exercise 1.88 Try to find another explicit formula for the function $f : \mathbf{N} \to \mathbf{Z}$ that would produce the result obtained in the previous example.

For the next exercise, we will need the following:

Theorem 1.32 Given three sets A, B, and C, such that $|A| = |B|$ and $|B| = |C|$, then $|A| = |C|$.

Proof Suppose that there exist two bijections

$$f : A \to B \quad \text{and} \quad g : B \to C$$

telling us that $|A| = |B|$ and $|B| = |C|$. Consider the composition of f and g

$$g \circ f : A \to C$$

By Theorem 1.24, $g \circ f$ is a bijection too, thus $|A| = |C|$. ■

Exercise 1.89 Suppose that $A \sim B$ and $C \sim D$. Show that $(A \times C) \sim (B \times D)$.

Solution Since $A \sim B$ and $C \sim D$, there exist respective bijections $f : A \to B$ and $g : C \to D$. Define the function $h : (A \times C) \to (B \times D)$ by

$$h(a, c) = (f(a), g(c)), \quad a \in A, \quad c \in C$$

As constructed above, h is evidently a bijection, hence $(A \times C) \sim (B \times D)$.

Example/Exercise 1.90 Show that $|2\mathbf{Z}| = \aleph_0$.

From all that has been said so far, it becomes evident why we define a set as infinite if it could be made equivalent to a proper subset of itself. And vice versa: a set is said to be finite, if it could not be made equivalent to at least one of its subsets.

A natural question one could ask at this point is: What about the cardinality of the set of rational numbers? Recall that $\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q}$. It is conceivable, then, that $\mathbf{Q}$ is much "bigger" than $\mathbf{N}$. Also, remember we proved in Theorem 1.22 that the set $\mathbf{Q}$ is dense. We can rephrase this by saying that in the ordering of rational numbers in terms of size, there is no next-larger rational number to any given number. So, it is definitely nontrivial to ask about larger infinities. In other words, all the sets we discussed so far have been countably infinite, that is, all of them have been of the "size" of $\aleph_0$. Now, considering the density of set $\mathbf{Q}$, one could expect the "size" of $\mathbf{Q}$ to be much larger than $\aleph_0$. However, the next theorem points to a different conclusion.

Theorem 1.33 The set of all positive rational numbers $\mathbf{Q}^+$ is countable.

Proof We would like to construct a "reasonable" function from $\mathbf{N}$ to $\mathbf{Q}^+$, and, hopefully, make it a bijection. Consider the following diagram suggested by Cantor:

1 / 1	**1 / 2**	**1 / 3**	**1 / 4**	**1 / 5**	**1 / 6**	**1 / 7**	…
2 / 1	2/2	**2 / 3**	2/4	**2 / 5**	2/6	2/7	…
3 / 1	**3 / 2**	3/3	**3 / 4**	**3 / 5**	3/6	**3 / 7**	…
4 / 1	4/2	**4 / 3**	4/4	**4 / 5**	4/6	**4 / 7**	…
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	

If we reduce each fraction to its lowest form and remove any repetition (i.e., we keep only the bold-faced numbers), we indeed obtain the set of all positive

rational numbers $\mathbf{Q}^+$. We can now establish a one-to-one correspondence between the natural numbers and the positive rational numbers in the following way:

$$1 \leftrightarrow 1/1, \quad 2 \leftrightarrow 1/2, \quad 3 \leftrightarrow 2/1, \quad 4 \leftrightarrow 3/1, \quad 5 \leftrightarrow 1/3, \quad 6 \leftrightarrow 1/4,$$
$$7 \leftrightarrow 2/3, \quad 8 \leftrightarrow 3/2, \quad 9 \leftrightarrow 4/1, \quad 10 \leftrightarrow 5/1, \ \ldots$$

Note that all the elements of $\mathbf{Q}^+$ have been "accounted for" exactly once, and each one of them has been matched to one and only one natural number. Such a one-to-one correspondence is indeed well defined, and, consequently, we conclude that $|\mathbf{N}| = |\mathbf{Q}^+|$. ■

Now try to prove

Theorem 1.34 The set of all rational numbers is countable, that is, $|\mathbf{Q}| = |\mathbf{N}|$.
Hint: Consider the fact that $|\mathbf{Q}^+| = |\mathbf{Q}^-|$, and that for any set A, $A \cup A = A$.

More generally, we have

Theorem 1.35 Any subset of a countable set is countable.

Proof Let X be a set such that $|X| = |\mathbf{N}|$, that is, X is countable. Let $Y \subseteq X$. Y could be finite or infinite. If it is finite, there is nothing to prove – Y is countable by definition. So, let Y be an infinite set. We would like to find a one-to-one correspondence between $\mathbf{N}$ and Y, that is, we are looking for a function $f : N \to Y$, such that f is one-to-one and onto.

Now, considering that X is countable, we can arrange the elements of X as a sequence

$$x_1, x_2, x_3, \ \ldots$$

Since $Y \subseteq X$, this sequence must contain all the elements of Y. We search among all the x_i's for the elements of Y, and arrange them in the order of occurrence as

$$f(1), f(2), f(3), \ \ldots$$

In other words, $\forall\, x_i \in X$, $\exists\, f(i) \in Y$. Since all the elements $x_1, x_2, x_3, \ \ldots$ are distinct, the function f is one-to-one. Now, since every $f(i)$ is found by sequentially searching through all of $x_1, x_2, x_3, \ \ldots$, and is constructed as an image of a natural number, f is also onto. Therefore, f is a bijection from $\mathbf{N}$ to Y, which proves that Y is countable. ■

Theorem 1.36 Let $A_0, A_1, A_2, \ \ldots$ be a sequence of countable sets. Then the union

$$A = \bigcup_{i=1}^{\infty} A_i = A_0 \cup A_1 \cup A_2 \cup \cdots$$

is also a countable set.

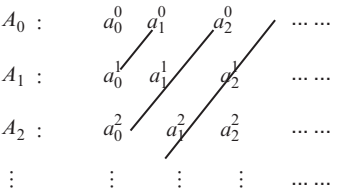
Proof Assuming that every $A_i = \{a_0^i, a_1^i, a_2^i, \ldots, a_n^i, \ldots\} \neq \varnothing$, we can find an onto function

$$\pi^i : \mathbf{N} \to A_i$$

such that for every i

$$\pi^i(n) = a_n^i$$

Again, following Cantor, we can construct a table:



Collecting the elements on the diagonals, we obtain the elements of the union fully enumerated:

$$A = \{a_0^0, a_0^1, a_1^0, a_0^2, a_1^1, a_2^0, a_0^2, \ldots\}$$

and the one-to-one correspondence $\pi^i(n)$ between A and $\mathbf{N}$ is evident:

$$\begin{array}{lccccccccc} A = \{ & a_0^0, & a_0^1, & a_1^0, & a_0^2, & a_1^1, & a_2^0, & \ldots \} \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \ldots \\ \mathbf{N} = \{ & 1, & 2, & 3, & 4, & 5, & 6, & \ldots \} \end{array}$$

■

Corollary The set of all finite subsets of a countable set is countable.

Example 1.91 Let A and B be two countable sets. Show that $A \cup B$ is countable.

Solution Since A and B are countable, we can express them as

$$A = \{a_1, a_2, a_3, \ldots\} \quad \text{and} \quad B = \{b_1, b_2, b_3, \ldots\}$$

Let's now define a function $f : \mathbf{N} \to A \cup B$ by the following diagram:

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & & \\ \downarrow & \downarrow & \downarrow & \downarrow & \ldots & \\ a_1 & b_1 & a_2 & b_2 & & \end{array}$$

Obviously, the function is one-to-one and onto, thus, $A \cup B$ is countable. ■

Example 1.92 Another way to prove that a set of rational numbers is countable goes as follows: Let

$$A_i = \left\{ q \,\middle|\, q = \frac{z}{i},\ z \in \mathbf{Z},\ i \in \mathbf{N} \right\}$$

Since A_i is obviously countable and $\mathbf{Q} = \cup_i^{\infty} A_i$, by Theorems 1.35 and 1.36, the set $\mathbf{Q}$ is countable too. ■

As a good exercise, you should now prove Theorems 1.37–1.39.

Theorem 1.37 If Y is a countable set, and if there exists an injection $f : X \to Y$, then X is also countable.

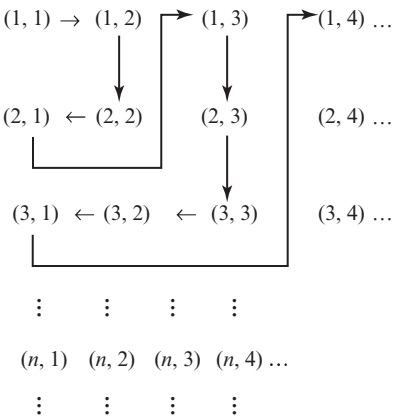
Theorem 1.38 If X is a countable set, and there exists a surjection $f : X \to Y$, then Y is also countable.

Theorem 1.39 Every two countably infinite sets are equivalent.

Theorem 1.40 $\mathbf{N} \times \mathbf{N}$ is countable.

Proof It suffices to show that $|\mathbf{N} \times \mathbf{N}| = |\mathbf{N}|$, that is, to construct a bijection $(\mathbf{N} \times \mathbf{N}) \to \mathbf{N}$.

One way to visualize it would be to count by following the arrows:



The other way to do it would be to arrange $\mathbf{N} \times \mathbf{N}$ as an infinite rectangular array of ordered pairs of natural numbers:

$$\begin{array}{cccc} (1,1) & (1,2) & (1,3) & (1,4) & \dots \\ (2,1) & (2,2) & (2,3) & (2,4) & \dots \\ (3,1) & (3,2) & (3,3) & (3,4) & \dots \\ \vdots & \vdots & \vdots & \vdots & \\ (n,1) & (n,2) & (n,3) & (n,4) & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

Thus, we have constructed a countable set of countable sets, that is

$$\begin{aligned} A_1 &= \{(1,1)\ (1,2)\ (1,3)\ (1,4)\ \dots\} \\ A_2 &= \{(2,1)\ (2,2)\ \ (2,3)\ (2,4)\ \dots\} \\ A_3 &= \{(3,1)\ (3,2)\ \ (3,3)\ (3,4)\ \dots\} \\ &\vdots \\ A_n &= \{(n,1)\ (n,2)\ \ (n,3)\ (n,4)\ \dots\} \\ &\vdots \end{aligned}$$

This is exactly the structure we have encountered in Theorem 1.36, so we conclude that $\mathbf{N} \times \mathbf{N}$ is a countable set.

Alternatively, we could have said: since (n, m) from our aforementioned list is clearly the mth element of A_n, why not consider a function

$$f : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$$

defined by $f(n, m) = 2^n\ 3^m \cdot f$ is obviously an injection, hence, by Theorem 1.37, we conclude again that $\mathbf{N} \times \mathbf{N}$ is countable. ■

Consequently, the following theorems hold:

Theorem 1.41 If A and B are countable sets, then $A \times B$ is countable.

Theorem 1.42 $\mathbf{Q} \times \mathbf{Q}$ is a countable set.

Theorem 1.43 $\mathbf{N} \times \mathbf{N} \times \mathbf{N}$ is a countable set.

More generally,

Theorem 1.44 The Cartesian product of a finite number of countable sets is countable.

1.8 THE SET R – REAL NUMBERS I

A natural question one may ask at this point is: Does there exist any "larger" set of numbers after the set **Q**? After all, we did say that $\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q}$, and we did talk about how "big" and dense the set **Q** is. But again, how big is "big?" If a set is infinite, can one construct a set with greater cardinality? We start this section, predictably, with Cantor's answer:

Now that we have established the fact that, regardless of how "big" **Q** is (and, remember, still equinumerous to **N**), we can always find a bigger one. The inevitable question is whether **Q** is axiomatically rich enough to accommodate everything we want to do mathematically. The answer, of course, is no, it is not. As many a reader may well remember the classic example from high school algebra courses, the solutions of a simple quadratic equation $x^2 = 2, x_{1,2} = \pm\sqrt{2}$ cannot be found in **Q**, that is, $\sqrt{2}$ is not a rational number; it is irrational. So in order to adhere to the spirit of Plato[41] and elude his harsh judgment: *He is unworthy of the name of man who does not know that the diagonal of a square is incommensurable with its sides* – we prove the following, well-known

Theorem 1.45 $\sqrt{2}$ is not a rational number.

Proof First, recall Definition 1.26, where we defined rational numbers as

$$\mathbf{Q} = \left\{ x \;\middle|\; x = \frac{p}{q}, \quad p, q \in \mathbf{Z}, \quad q \neq 0 \right\}$$

Without loss of generality, let's take p and q to be *relatively prime*.

Suppose that, contrary to the statement of the theorem, $\sqrt{2}$ is a rational number, that is, suppose

$$\sqrt{2} \in \mathbf{Q}$$

Then, there exist $p, q \in \mathbf{Z}$ with $q \neq 0$, such that

$$\sqrt{2} = \frac{p}{q} \tag{1.8}$$

Squaring both sides of (1.8) we get

$$2 = \frac{p^2}{q^2} \tag{1.9}$$

Or, after multiplying both sides by q^2

$$2q^2 = p^2 \tag{1.10}$$

[41] Plato (427–347 BC).

Equation (1.10) implies that p^2 is an even integer, so p has to be even too! (Why?) Let's express this fact by writing

$$p = 2k, \quad k \in \mathbf{Z} \tag{1.11}$$

Substituting (1.11) into (1.10), we get

$$2q^2 = 4k^2 \tag{1.12}$$

or

$$q^2 = 2k^2 \tag{1.12'}$$

which tells us that q^2 is even, and therefore q is even too. As in the case of p, we express the fact that q is an even integer by writing it as

$$q = 2l, \quad l \in \mathbf{Z} \tag{1.13}$$

Substituting (1.11) and (1.13) into (1.8) gives

$$\sqrt{2} = \frac{p}{q} = \frac{2k}{2l}$$

which contradicts our assumption that p and q are relatively prime. Hence, our assumption was wrong; therefore, the theorem is true. ■

Example/Exercise 1.93 Show that $\sqrt{3} \notin \mathbf{Q}$

(Hint: Start, as usual, assuming that

$$\sqrt{3} = \frac{p}{q} \in \mathbf{Q}, \quad q, p \in \mathbf{Z}, \ \ q \neq 0$$

and from $3q^2 = p^2$, consider the cases when q is an even number and when q is an odd number.)

Evidently, in addition to the numbers that we have encountered so far (natural numbers, integers, and rational numbers), there exists another set of numbers called **irrational numbers** $\boldsymbol{I}$, that is, the numbers that cannot be found in any of the previously studied sets **N**, **Z,** and **Q**. The set that contains all of them we call the **set of real numbers R**. We have the following structure:

$$\begin{array}{ccccccc} & & & & \mathbf{I} & & \\ & & & & | & & \\ \mathbf{N} & \subseteq & \mathbf{Z} & \subseteq & \mathbf{Q} & \subseteq & \mathbf{R} \end{array}$$

or

$$\mathbf{Q} \cup \mathbf{I} = \mathbf{R}$$

An astute reader would now expect a formal definition of the set **R**, with all the axioms precisely laid down in order to fully understand all the intricacies of **R**. However, because of the complexities of such a formalism, some additional, more advanced concepts are needed for a full and rigorous definition of the structure. So, we will postpone the full formal definition for later and will now introduce a rather "heuristic" definition and some properties of **R**, which, we hope, will suffice for at least an intuitive appreciation of the richness and importance **R**.

We are about to venture deeper into the land of Cantor, "*a paradise from which no one shall drive us.*"

Without proof, we state

Theorem 1.46 There is a one-to-one correspondence between the set **R** and the points on the number line.

Theorem 1.47 The set **R** and the set of points in the open interval $(0, 1)$ are equivalent.

Proof All we need to do is to find an appropriate bijection $f : (0, 1) \to \mathbf{R}$ and we have the proof. Any bijection $(0, 1) \to \mathbf{R}$ that is not defined at 0 and 1 will do. Let's try a function defined by

$$f(x) = \frac{1 - 2x}{x^2 - x} \qquad (*)$$

f is certainly not defined at 0 and 1. Is it a bijection? Let's see. Take $x_1,\ x_2 \in (0,\ 1)$ and suppose that $f(x_1) = f(x_2)$, that is

$$\frac{1 - 2x_1}{x_1^2 - x_1} = \frac{1 - 2x_2}{x_2^2 - x_2}$$

Then,

$$(1 - 2x_1)(x_2^2 - x_2) = (1 - 2x_2)(x_1^2 - x_1)$$

or

$$(x_1 - x_2)(x_1 + x_2 - 2x_1x_2 - 1) = 0$$

If we could prove that $x_1 = x_2$, then our function is one-to-one. $x_1 = x_2$ if

$$x_1 + x_2 - 2x_1x_2 - 1 \neq 0$$

Suppose that is not the case, that is, suppose

$$x_1 + x_2 - 2x_1x_2 - 1 = 0$$

then, with little algebraic reshuffling, we get

$$-x_1 - x_2 + x_1x_2 + 1 = -x_1x_2$$

or

$$(x_1 - 1)(x_2 - 1) = -x_1x_2 \qquad (**)$$

Since $x_1, x_2 \in (0, 1)$, that is, $0 < x_1 < 1$ and $0 < x_2 < 1$, it would follow that the LHS of (**) implies

$$(x_1 - 1)(x_2 - 1) > 0$$

while the RHS implies

$$-x_1x_2 < 0$$

which, of course, is impossible. We conclude that $x_1 + x_2 - 2x_1x_2 - 1 \neq 0$ indeed. Therefore, $x_1 = x_2$ and our function f is one-to-one. Next, we need to show that f is also onto, that is, that for every $y \in \mathbf{R}$, there exists an $x \in (0, 1)$, such that $f(x) = y$. From $(*)$ it follows that one possible $x \in (0, 1)$ is

$$x = \frac{y - 2 + \sqrt{y^2 + 4}}{2y}$$

with $y \in \mathbf{R}, y \neq 0$.

Hence,

$$\begin{aligned} f(x) &= \frac{1 - 2x}{x^2 - x} \\ &= \frac{1 - 2\left(\frac{y-2+\sqrt{y^2+4}}{2y}\right)}{\left(\frac{y-2+\sqrt{y^2+4}}{2y}\right)^2 - \left(\frac{y-2+\sqrt{y^2+4}}{2y}\right)} \\ &= \frac{y(8 - 4\sqrt{y^2 + 4})}{8 - 4\sqrt{y^2 + 4}} = y \end{aligned}$$

Hence, our function is one-to-one and onto, that is, $f : (0, 1) \to \mathbf{R}$ is a bijection and therefore $|(0, 1)| = |\mathbf{R}|$. ■

Example 1.94 If you find the aforementioned theorem too technical, or too complicated to be convincing, consider the statement: "There is the same number of points on a line segment 1 cm long as on one that is 1 km long." Here is the "proof":

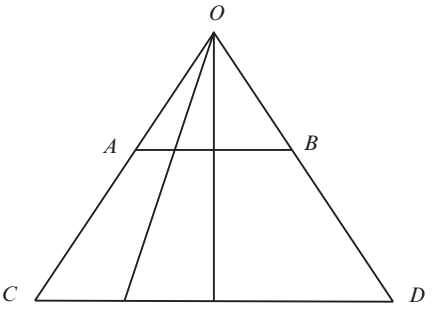


Figure 1.19

As you can see, to every point on segment AB there corresponds one and only one point on segment CD, regardless of the difference in their lengths. So, a bijection between the points of segment AB and segment CD is evident. ■

If you are at least somewhat familiar with trigonometric functions, the following example is also a good "visual" proof of the equivalence of "short" and "long" segments.

Example 1.95 The mapping $(\frac{-\pi}{2}, \frac{\pi}{2}) \mapsto \mathbf{R}$ defined by

$$f(x) = \tan x$$

is clearly a bijection (Figure 1.20).

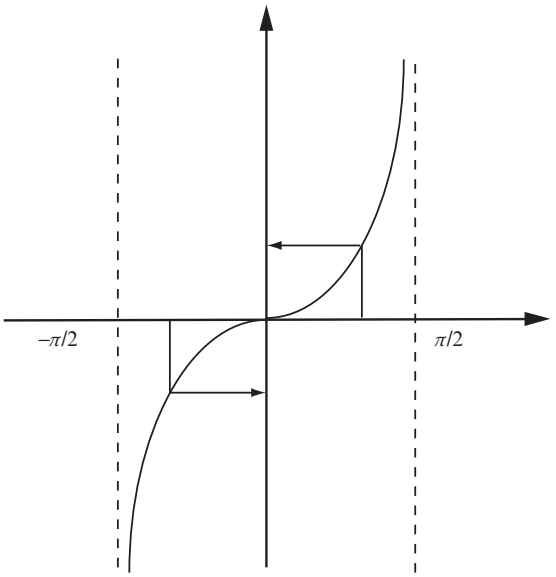


Figure 1.20

■

Example/Exercise 1.96 Prove that $|[0, 1]| = |[0, 1)|$.

Example/Exercise 1.97 Prove that $|(0, 1)| = |(a, b)|, \forall\, a, b \in \mathbf{R}$ (see Theorem 1.47).

Theorem 1.48 The set $\mathbf{R}$ is *dense*, that is, between any two real numbers there is another one; therefore, there are infinitely many. You ought to recall that a similar statement has been made regarding rational numbers (cf. Theorem 1.22). Here, however, we are talking about an even "higher" density. Nevertheless, proof of this theorem should not be difficult for you.

Theorem 1.49 (Cantor) The set $\mathbf{R}$ is uncountable.

Proof Consider a set of all real numbers between 0 and 1. Is this set countable or not? Suppose it is countable. In that case, these numbers have a decimal representation and we can list them all as follows:

$$\left.\begin{array}{l} 0.\, a_{11}a_{12}a_{13}\ \ldots\ a_{1n}\ \ldots \\ 0.\, a_{21}a_{22}a_{23}\ \ldots\ a_{2n}\ \ldots \\ 0.\, a_{31}a_{32}a_{33}\ \ldots\ a_{3n}\ \ldots \\ \qquad\vdots \\ 0.\, a_{n1}a_{n2}a_{n3}\ \ldots\ a_{nn}\ \ldots \\ \qquad\vdots \end{array}\right\} \qquad (*)$$

Note that every decimal digit in (*) has two indices; the first one indicating which member of the sequence it belongs to (i.e., which row in the aforementioned sequence), and the second indicating which decimal place the digit is in (e.g., let's say that $0.4758\ldots$ is the number in the third row, then $4 = a_{31}, 7 = a_{32}$, $5 = a_{33}, 8 = a_{34}$, and so on).

With a construction like this, we should be able to associate to every number in (*) one and only one element from $\mathbf{N}$, that is, we should be able to "count" them. For example,

$$\begin{array}{rl} 1 & \leftrightarrow\ 0.\, a_{11}a_{12}a_{13}\ \ldots\ a_{1n}\ \ldots \\ 2 & \leftrightarrow\ 0.\, a_{21}a_{22}a_{23}\ \ldots\ a_{2n}\ \ldots \\ 3 & \leftrightarrow\ 0.\, a_{31}a_{32}a_{33}\ \ldots\ a_{3n}\ \ldots \\ & \qquad\vdots \\ n & \leftrightarrow 0.\, a_{n1}a_{n2}a_{n3}\ \ldots\ a_{nn}\ \ldots \\ & \qquad\vdots \end{array}$$

Well, no. We cannot. Regardless of the construction of the sequence (*) of real numbers between 0 and 1, there are still numbers in this interval that are not contained in the list (*). Consider this: suppose we choose a decimal number

$$d = 0.\,\overline{a}_{11}\overline{a}_{22}\overline{a}_{33}\ \ldots$$

such that $\overline{a}_{ii} \neq a_{ii}$, that is, $\overline{a}_{11} \neq a_{11}$, $\overline{a}_{22} \neq a_{22}$, $\overline{a}_{33} \neq a_{33}$, and so on.

We can do this very easily. Say we first check a_{11} in (*). If it is different from 1 we put $\overline{a}_{11} = 1$, but if it is equal to 1 we put $\overline{a}_{11} = 2$. Then we check a_{22} and do the same: if $a_{22} \neq 1$, then we write $\overline{a}_{22} = 1$, and if $a_{22} = 1$, we write $\overline{a}_{22} = 2$. We continue this process for all a_{ii} in our sequence (*). Of course, this is not the only way one can construct

$$d = 0.\,\overline{a}_{11}\overline{a}_{22}\overline{a}_{33}\ \ldots$$

In general, we can reason the following way: in the case that $a_{nn} = 0$, we have nine choices to make $\overline{a}_{nn}$ different. In the case that $a_{nn} \neq 0$, we still have eight choices for $\overline{a}_{nn}$. Hence, for every decimal digit, we have at least eight choices and, therefore, we have infinitely many choices for the number d. But whatever "technique" we use, note that the number d cannot be found in our sequence (*), since it differs from the first number of (*) in the first decimal place. With the second number of (*), it differs in the second decimal place, and so on. We conclude that since the real number d is different from all the numbers in our sequence (*), the sequence does not contain all the numbers between 0 and 1, contrary to our starting assumption. Hence, the set of numbers between 0 and 1 is uncountable. Since $(0, 1) \subseteq \mathbf{R}$, it follows that the set $\mathbf{R}$ is uncountable. ■

This proof is known as Cantor's diagonal argument. It turned out to be very important in mathematics and logic. There have been various versions of diagonal arguments, and the gist of it led to a number of very important results in mathematics. We have already encountered some of them in Section 1.7. You can find references to Cantor's argument over and over again in many discourses in mathematics, physics, computer science, philosophy, and so on. The next ingenious and witty example by Smullyan and Fitting[42] goes something like this:

Example 1.98 Suppose there is a book with infinitely many pages: page 1, page 2, page 3, and so on. Obviously, the set of pages is countable. Furthermore, suppose that on each page there is a list, a set, of some natural numbers: on page 1 there is a set N_1, on page 2 there is a set N_2, on page 3 there is a set N_3, … on page n there is a set N_n, and so on. The question is : Is every natural number listed in the book? The answer is no. There must be at least one set of natural

[42]R. M. Smullyan, M. Fitting, *Set Theory and the Continuum Problem,* Clarendon Press – Oxford, 1996.

numbers that is not listed in the book. In other words, there exists a set N, which is different from every one of the sets: $N_1, \ldots, N_n, \ldots$. Let's see why.

First consider the number 1 – either 1 belongs to set N_1 or it doesn't. We include it in N, **only if it does not** belong to N_1. Thus, whatever future decisions we make concerning the numbers $2, 3, \ldots, n, \ldots$, we know that $N \neq N_1$ because, **only one** of the two sets N and N_1, contains 1 and the other doesn't. Next, we consider the number 2. We put it into N only if it does not belong to N_2, and that makes $N \neq N_2$ (since one of them contains 2 and the other doesn't). We continue the process for every natural number n. This way, we constructed N such that for every n, $N \neq N_n$. What we have shown is that, given any countably infinite sequence $N_1, N_2, \ldots, N_n, \ldots$ of sets of natural numbers, there exists a set N of natural numbers (namely, the set of all n such that n doesn't belong to N_n) such that N is different from each of the sets $N_1, N_2, \ldots, N_n, \ldots$. This means that no countable set of sets of natural numbers contains every set of natural numbers, that is, the set of all sets of natural numbers is uncountable. ■

You may find the following two examples also engaging.

Example 1.99 Let $r_1, r_2, r_3, \ldots$ be any sequence of real numbers, and let $[a_1, b_1], [a_2, b_2], [a_3, b_3], \ldots$ be a sequence of closed intervals where $a_i, b_i \in \mathbf{R}$, satisfying the following:

(i) $a_i < b_i, \quad \forall i$
(ii) $[a_i, b_i] \subseteq [a_j, b_j], \quad \forall i > j$
(iii) $r_i \notin [a_i, b_i]$

Because of (ii), $[a_1, b_1] \cap [a_2, b_2] \cap [a_3, b_3] \cap \cdots \neq \emptyset$. So suppose

$$r \in [a_1, b_1] \cap [a_2, b_2] \cap [a_3, b_3] \cap \cdots$$

Now, r cannot be one of $r_1, r_2, r_3, \ldots$ because of (iii) and therefore

$$r \notin [a_1, b_1] \cap [a_2, b_2] \cap [a_3, b_3] \cap \cdots$$

Since the sequence $r_1, r_2, r_3, \ldots$ was arbitrarily chosen, it follows that no countable set of real numbers contains all real numbers. ■

Example/Exercise 1.100 Give another example of a real number not in the list (*) on page 77, that is, construct another proof of the uncountability of real numbers.

Example/Exercise 1.101 Prove that there are infinitely many possibilities to choose from along the original diagonal in (*) to construct another real number.

Although, generally, it is not easy to prove that a set is uncountable, fortunately, the following theorem is not difficult.

Theorem 1.50 The set of all irrational numbers is uncountable.

Proof Recall the set $\mathbf{R} \setminus \mathbf{Q} = \mathbf{I}$ is the set of irrational numbers.

Suppose now that the set $\mathbf{I}$ is countable. In that case, we could list all irrational numbers in a sequence $i_1, i_2, i_3, \ldots$. On the other hand, since rational numbers *are* countable, we can certainly list them as $q_1, q_2, q_3, \ldots$. Consequently, we could construct the following list:

$$i_1, q_1, i_2, q_2, i_3, q_3, \ldots \qquad (*)$$

Since $\mathbf{R} = \mathbf{Q} \cup \mathbf{I}$, the list (*) by construction should contain all real numbers and be countable. But that is impossible since $\mathbf{R}$ is uncountable. On the other hand, as we have established before, $\mathbf{Q}$ is countable. Hence, contrary to our supposition, the set $\mathbf{I}$ of all irrational numbers must be uncountable. ■

Theorem 1.51 Let A and B be two sets such that $A \subseteq B$. If A is uncountable, then B is uncountable too.

Proof Suppose not, that is, suppose B is countable. A being uncountable, and also a subset of the countable B, contradicts Theorem 1.35 thus B has to be uncountable too. ■

Example 1.102 If a set A is uncountable, is it equivalent to $\mathbf{R}$?

Solution Of course not! Suppose we take $A = \mathcal{P}(\mathbf{R})$. A is definitely uncountable, but at the same time $|\mathcal{P}(\mathbf{R})| > |\mathbf{R}|$. ■

Theorem 1.52 Let $\mathcal{F} = \{f | f : \mathbf{N} \to \{0, 1\}\}$ be the set of all functions from $\mathbf{N}$ to $\{0, 1\}$. Then, $|\mathcal{F}| = |\mathcal{P}(N)|$

Proof Let $\Phi : \mathcal{F} \to \mathcal{P}(\mathbf{N})$ be a function defined as follows:

$$\forall f \in \mathcal{F}, \Phi(f) = \{x \in \mathbf{N} \mid f(x) = 1\}$$

We would like to show that Φ is a bijection. So, let's take $f_1, f_2 \in \mathcal{F}$ such that $f_1 \neq f_2$. It follows that there exists $n \in \mathbf{N}$ such that $f_1(n) \neq f_2(n)$. Suppose f_1, f_2 are such that $f_1(n) = 1$ and $f_2(n) = 0$. Then,

$$n \in \{x \in \mathbf{N} \,|f_1(x) = 1\} = \Phi(f_1)$$

and similarly

$$n \notin \{x \in \mathbf{N} | f_2(x) = 1\} = \Phi(f_2)$$

Thus, $\Phi(f_1) \neq \Phi(f_2)$, that is, Φ is a one-to-one function. Is it onto? Well, consider a set $A \in \mathcal{P}(\mathbf{N})$. Then $A \subseteq \mathbf{N}$, and the characteristic function (see Chapter 4)

$$\chi_A : \mathbf{N} \to \{0, 1\}$$

is obviously an element of $\mathcal{F}$. Furthermore,

$$\Phi(\chi_A) = \{x \in \mathbf{N} \mid \chi_A(x) = 1\} = A$$

Thus, Φ is onto. Consequently, $|\mathcal{F}| = |\mathcal{P}(N)|$. ■

The following is also true.

Theorem 1.53 A set $F = \{f | f : \mathbf{N} \to \mathbf{N}\}$ of all functions from $\mathbf{N}$ to $\mathbf{N}$ is uncountable.

The mind of thee upon these lines of ours,
Till thou see through the nature of all things,
And how exists the interwoven frame

It has no bounds, no end, no limit,
And it matters not what part of the universe you are in;
Wherever you are, from the spot you take up,
It stretches to infinity in all directions. …

Titus Lucretius Carus[43]

What is that thing which does not give itself, and which if it were to give itself would not exist? It is infinite!

Leonardo da Vinci[44]

1.9 A SHORT MUSING ON TRANSFINITE ARITHMETIC

The Hilbert Hotel

Let's imagine an Infinity Hotel, also (appropriately) known as the Hilbert Hotel, with infinitely many rooms (numbered 1, 2, 3, … and so on forever). As an infinite

[43]Titus Lucretius Carus (ca. 99–55 BC), *De Rerum Natura.*
[44]Leonardo da Vinci (1452–1519), *Notebooks.*

number of guests (mathematicians [sic!] attending a mathematics convention) occupy all rooms, the receptionist is convinced there are no vacancies and all latecomers should be turned away. "Not so," the manager exclaimed. When the next VIP arrives, move the person from room 1 to room 2, the person from room 2 to room 3, the person from room 3 to room 4, etc. This leaves room 1 vacant while everyone else is properly accommodated. In case more latecomers arrive the manager repeats the process. So infinitely many newcomers are accommodated. It turned out that infinitely many physicists came to the conference too, but the manager is not worried at all. He keeps guest from room 1 in room 1 but moves the guest from room 2 to room 4, the guest from room 3 to room 9, … , the guest from room n to room n^2 and so on forever. Obviously infinitely many rooms are now ready to accommodate all the physicists. And, as you might have anticipated, when in addition to all previous guests, infinitely many philosophers and infinitely many rock concert fans arrive, all of them are accommodated by similar methods. But this is not the end of the story. As is often in life, things turn unexpectedly odd. The Infinity Hotel became so profitable and soon enough infinitely many infinity hotels opened up: Hotel 1, Hotel 2, Hotel 3, … and so on forever. However, one day all the guests from those hotels, for some strange reason decided they wanted to move to the original Infinity Hotel. Our ingenious manager now has to accommodate infinitely many guests from each of infinitely many hotels. Here is what he does. Consider all prime numbers (there are infinitely many of them): $2, 3, 5, 7, 11, 13, \ldots$, *and then do the following: put infinitely many guests from Hotel 1 into rooms 2, 4, 8, 16,* … $2, 4, 8, 16, \ldots$ (*i.e.*, $2^1, 2^2, 2^3, 2^4, \ldots$)*; those from Hotel* 2 *into rooms* $3, 9, 27, 81, \ldots$ (*i.e.*, $3^1, 3^2, 3^3, \ldots$ *etc.*); *those arriving from Hotel* 3 *into rooms* $5, 25, 125, 625, \ldots$ (*i.e.*, $5^1, 5^2, 5^3, \ldots$ *etc.*). *Continuing this process manager is sure that while accommodating all the guests from all the hotels no two persons will occupy the same room.*

At the beginning of Section 1.8, we asked: How big is "big." How do we determine whether one set is "larger" than the other? Let's think about this for a moment. We have already established that any infinite subset of a set of natural numbers is countable. Can we prove that any infinite set contains a countable subset? Let's see. Take any infinite set X. We can always pick a nonempty (infinite) countable subset $A \subseteq X$ the following way: remove one element, say, a_1 from X. Certainly, $X \setminus \{a_1\}$ is still an infinite set. Let's remove another element, a_2, which keeps $X \setminus \{a_1, a_2\}$ still infinite. We continue this process choosing $a_3, a_4, a_4, \ldots$ to be removed from X. Thus, we have extracted from X a countable set $A = \{a_1, a_2, a_3, \ldots, a_n, \ldots\}$ and X nevertheless remains to be an infinite set. As you might have anticipated by now, we could continue with these arguments further and, for instance, remove from A the set of all elements with even indices, $B = \{a_2, a_4, a_6, \ldots\}$ and $A \setminus B$ remains countably infinite. We conclude that the cardinality of an infinite set does not change if we adjoin a countable set to it. And certainly cardinality of an uncountable set won't change if we extract a countable subset from it. It is reasonable to wonder how come that all infinities

are not the same? Also, recall that the sets we may want to study could have as elements whole families of countable or uncountable elements. The following few examples will make this more transparent.

Example 1.103 Consider the following set $\mathcal{F} = \{\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}\}$. Observe that $\mathcal{F}$ is a finite (thus countable) family of sets. No matter that all of its elements are infinite sets themselves. ■

Example 1.104 For each $i \in \mathbf{N}$, let's construct a family of sets $\mathbf{N}_i$, where each $\mathbf{N}_i$ is the set of all natural numbers divisible by i, that is, starting with $\mathbf{N}_1 = \mathbf{N}$ we have

$$\begin{aligned}
\mathbf{N}_2 &= \{2, 4, 6, \ldots, 2n, \ldots\} \\
\mathbf{N}_3 &= \{3, 6, 9, \ldots, 3n, \ldots\} \\
&\vdots \\
\mathbf{N}_i &= \{i, 2i, 3i, \ldots, ni, \ldots\} \\
&\vdots
\end{aligned}$$

Thus, we have obtained an infinite family of infinite countable sets

$$\mathcal{F} = \{\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}_3, \ldots, \mathbf{N}_i, \ldots\}$$ ■

Next is the example that we have encountered in a slightly different context earlier. Recall that the set $\boldsymbol{Q}^+$ is a countable set (Theorem 1.33), and this time let's look at Cantor's proof from the perspective of an infinite family of sets:

Example 1.105 Define

$$\begin{aligned}
\mathbf{Q}_1^+ &= \left\{\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right\} \\
\mathbf{Q}_2^+ &= \left\{\frac{2}{1}, \frac{2}{2}, \frac{2}{3}, \ldots, \frac{2}{n}, \ldots\right\} \\
\mathbf{Q}_3^+ &= \left\{\frac{3}{1}, \frac{3}{2}, \frac{3}{3}, \ldots, \frac{3}{n}, \ldots\right\} \\
&\vdots \\
\mathbf{Q}_m^+ &= \left\{\frac{m}{1}, \frac{m}{2}, \frac{m}{3}, \ldots, \frac{m}{n}, \ldots\right\} \\
&\vdots
\end{aligned}$$

So we have obtained an infinite, countable family

$$\mathcal{F} = \{\mathbf{Q}_1^+, \mathbf{Q}_2^+, \mathbf{Q}_3^+, \ldots, \mathbf{Q}_m^+, \ldots\}$$

whose elements are exactly the rows in Cantor's diagram. ■

Thus, questions about the "nature" of infinities, and indeed, how many "infinities" there are, are inevitable.

Here is Cantor again:

Definition 1.43 (Cantor) A set A is greater than a set B if and only if B is equivalent to some subset of A, but A is not equivalent to any subset of B.

Cantor also showed that $\aleph_0$ is the smallest infinite cardinal number, and following Cantor we have shown that $\aleph > \aleph_0$. We have established the fact that there are at least two different sorts of infinite sets, two different "kinds" of infinities, that is, two different kinds of cardinal numbers. Let's remind ourselves what Cantor meant by the cardinal number of a set X:

> *… the general concept which by means of our active faculty of thought, arises from the aggregate X when we make abstraction of the nature of its various elements x and of the order in which they are given.*

Now we ask: Is there a cardinal number greater than $\aleph$? Cantor's answer is this: For any set X, there exist sets larger than X, in particular $\mathcal{P}(X)$. So, for instance $|\mathbf{N}| < |\mathcal{P}(\mathbf{N})|$, thus we are prompted to consider the following:

$$\begin{aligned} \aleph_0 &= |\mathbf{N}| \\ 2^{\aleph_0} &= |\mathcal{P}(\mathbf{N})| \\ 2^{2^{\aleph_0}} &= |\mathcal{P}(\mathcal{P}(\mathbf{N}))| \\ &\vdots \end{aligned}$$

Consequently, we can naturally proceed and construct a hierarchy of **transfinite** cardinals:

$$\aleph_0, 2^{\aleph_0}, 2^{2^{\aleph_0}}, 2^{2^{2^{\aleph_0}}}, \ldots \qquad (*)$$

Hence

Theorem 1.54 There is an infinite sequence of infinite cardinals

$$\aleph_0 < \aleph_1 < \aleph_2 \cdots$$

where $\aleph_0 = |\mathbf{N}|$, $\aleph_1 = |\mathcal{P}(\mathbf{N})|$, $\aleph_2 = |\mathcal{P}(\mathcal{P}(\mathbf{N}))|$, and so on.

Proof We have learned from Theorem 1.8.1 that $\aleph_0 < \aleph_1$. Next, consider $\aleph_2$. It is obviously a cardinal number of the power set of the set $\mathcal{P}(\mathbf{N})$. Thus, according to Theorem 1.31, $\aleph_1 < \aleph_2$. So, we have established that

$$\aleph_0 < \aleph_1 < \aleph_2 \cdots$$

There is no reason to stop at $\aleph_2$, so in general for any n we have

$$\aleph_{n-1} < \aleph_n = |\mathcal{P}(\mathcal{P}(\mathcal{P}(\ldots \ (\mathcal{P}(\mathbf{N}))) \ \ldots)|$$

Hence, there is indeed a sequence of infinite cardinals

$$\aleph_0 < \aleph_1 < \aleph_2 < \cdots < \aleph_n < \cdots$$

which we recognize as our sequence (*). If you wish to "visualize" the aforementioned sequence, you may consider the following picture, but keep in mind, the line pictured is not a real line (Figure 1.21).

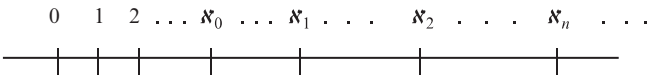


Figure 1.21

■

Where is the cardinal number of **R** in this sequence? It can be shown (cf. Theorem 1.12) that the cardinal number of the reals

$$|\mathbf{R}| = |\mathcal{P}(\mathbf{N})| = 2^{\aleph_0}$$

We conclude that the set of all real numbers **R** is equivalent to the set of all subsets of natural numbers $\mathcal{P}(\mathbf{N})$.

So, our sequence (*) is as expected

$$\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ \ldots \qquad (**)$$

assuming that there is no cardinal number between $\aleph_0$ and $2^{\aleph_0}$, that is, no cardinal number greater than $\aleph_0$ and less than $2^{\aleph_0}$. Well, can we assume this? And why? Cantor said yes, $2^{\aleph_0} = \aleph_1 = c$,[45] that is, we take $|\mathbf{R}| = \aleph_1$. However, he was unsuccessful in proving it. This is what is known as the **Continuum Hypothesis (CH)**.

[45] Remember, we are assuming the Continuum Hypothesis, that is, $c = \aleph_1$

Let's have some fun and reflect on the issue some more. It is important to keep in mind that in the following lines, as in this whole section, we are just musing, and by no means do we expect to give the proof of the hypothesis, or anything even close to a definite answer. Some would suggest that the question of CH is subjective, and the whole issue has to do with how strong a Platonist you (the mathematician) are.

First, recall that the sets $\mathbf{N}$ and $\mathbf{Q}$ are countable and the set $\mathbf{R}$ is not. Also, remember $\mathbf{N} \subseteq \mathbf{Q} \subseteq \mathbf{R}$. Cantor conjectured that there is no set X such that it has more elements than $\mathbf{N}$ and fewer elements than $\mathbf{R}$, that is

$$\nexists X \text{ s.t. } |\mathbf{N}| < X < |\mathbf{R}|$$

Remember (cf. Chapter 1.5), given a set A such that $|A| = n$, then $|\mathcal{P}(A)| = 2^{|A|} = 2^n$. Hence, there are more elements in $\mathcal{P}(A)$ than in A. Consequently, for given $n = |A|$ there are $2^n - (n + 1)$ sets having the number of elements greater than n and less than 2^n. In general, for any sets A and B if

$$|B| \leq |A| \leq |\mathcal{P}(B)|$$

then either $A \sim B$ or $A \sim \mathcal{P}(B)$. Translating this to alephs, we get to the **generalized continuum hypothesis:**

$$2^{\aleph_n} = \aleph_{n+1}$$

Once we have "convinced" ourselves of this fact, it is natural to contemplate the extension of this to sets $\mathbf{N}$ and $\mathbf{R}$. Suppose there is a set X with more elements than $\mathbf{N}$ and fewer than $\mathbf{R}$. Then X should be such that

$$|\mathbf{N}| < |X| < 2^{|\mathbf{N}|}$$

Now, $2^{|\mathbf{N}|} > |\mathbf{N}|$ and (cf. Theorems 1.12, 1.31, and Definition 1.43) $2^{|\mathbf{N}|}$ is the number of elements of $\mathbf{R}$. Hence, $2^{|\mathbf{N}|}$ cannot be the cardinality of any set between $\mathbf{N}$ and $\mathbf{R}$. Everything said earlier is correct (except the last sentence – "Hence … " is kind of a stretch), but somehow your instinct might be telling you that something is still missing – our "proof" is not satisfying. No wonder many mathematicians have unsuccessfully struggled with the problem for years. In 1931, Kurt Gödel used the techniques of mathematical logic to show that Continuum Hypothesis could not be disproved on the basis of available axioms.[46] That, of course, does not mean that it could be proved either. In 1963, Paul Cohen took it one step further and showed that it was also impossible to prove the Continuum Hypothesis. All efforts were unfruitful, because the assumptions of set theory,

[46]Zermelo–Fraenkel axioms.

which Cantor and others used, were independent of the Hypothesis. Being aware of Gödel's work, Cohen concluded:

> *Set Theory with the assumption of the Continuum Hypothesis is consistent; Set Theory with the denial of Continuum Hypothesis is consistent.*

So, what are we to make of all this? Obviously, we are venturing into the territories beyond everyday experiences. In standard mathematics, saying that you *cannot* prove that $A = B$ but that you *can* prove that $A \neq B$ would sound pretty odd. Here, however, we are talking "the other" mathematics. Simply put, alephs are definitely different kinds of numbers (for lack of a better word), or at least "numbers" that many of us have never thought about before, and consequently every statement regarding them has to be pondered over with special care. Before we continue with our "regular" mathematics, I cannot resist the temptation of showing you something I find extremely fascinating. What follows is again mostly due to Cantor. The concepts that we will briefly touch upon are generally uncontroversial nowadays. However, although the logical consistency of the theory is indisputable, one might occasionally hear some dissonant voices regarding the existence and the "reality" of infinities. I'll let you make up your own mind.

> Alfred North Whitehead, however, would say: "*Our minds are finite, and yet even in the circumstances of finitude we are surrounded by possibilities that are infinite, and the purpose of life is to grasp as much as we can of that infinitude.*"

With this encouragement in mind, we may continue a bit further. Mathematics – any mathematics – is about thinking, wouldn't you agree? And thinking is "due" to our mind (whatever that may be). So, let's also agree, for starters at least, that admitting that mathematics (as well as science and philosophy) has its limitations does not imply that there are limitations of the universality of reason. (All right, I concede that this is a rather big assumption, but let's not dwell on it for the time being.) Accordingly, here are some new realms that our mind can explore.

Recall what Cantor meant by the cardinal number: the cardinal number of a set X is what X has in common with all the sets equivalent to X. We get cardinal number(s) by simply counting: $1, 2, 3, \ldots, \aleph_0$. In other words, the cardinal number indicates how many members there are in a given set. Nothing is said as to how they are *ordered.* That's why, you may recall, he denoted the cardinal number of a set X as $\overline{\overline{X}}$.[47] It is worth repeating that the double bar indicates double abstraction, first from the nature of the elements and second from their order. Now, consider the sequence (**) on page 84. We start with 1 and then 2, and so

[47] This should not cause confusion with our notation $|X|$. See Definition 1.9.

on until we reach the first transfinite cardinal, the second transfinite cardinal, and so on. In other words, we have the sequence

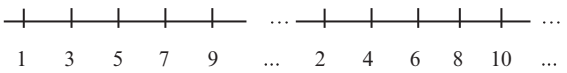
$$1,\ 2,\ 3,\ \ldots\ ,\ \aleph_0, \aleph_1, \aleph_2, \aleph_3,\ \ldots \qquad (***)$$

in which subscripts indicate the ordering by the size of transfinite cardinals.

We can lump them all in some set, the set of cardinals, but now we also distinguish which one is the first element, which one is the second, and so on. With such ordering, we obtained the set of *ordinal numbers.* What are those? Strictly speaking, every time when we count and use the expressions "first," "second," "third," and so on, we talk about ordering the elements in a set. Think about it this way: We can use natural numbers to count (and that's why some call them "counting numbers"), and so on to answer the question of "how many" of a certain object we have: one, two, three, and so on, and in this case, we call them *cardinals.* But if we want to answer the question "in which order" the objects are arranged, and so on, which object is first, second, third, and so on, we call them *ordinals.* Now comes the important part. Suppose we list the elements of the set **N** in the following way:

$$1, 3, 5, 7,\ \ldots\ , 2, 4, 6,\ \ldots$$

that is, we first list all odd natural numbers and then we list all even natural numbers. We could picture this as in Figure 1.22.

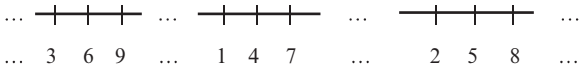


(1)

Figure 1.22

Suppose now that we want to enumerate them. How could we do that? As you already know, we would "exhaust" the whole of set **N** just to enumerate the odds, and we would still be left with infinity many evens (i.e., $\aleph_0$ of them) without any means of counting them. The same problem would arise if we wanted to first list all numbers, say, divisible by 3, and then those which leave remainder 1 after division by 3, and those that leave remainder 2 after division by 3, that is, $3, 6, 9,\ \ldots\ , 1, 4, 7,\ \ldots\ , 2, 5, 8,\ \ldots$

This would look something like in Figure 1.23.



(2)

Figure 1.23

Obviously, we need some other symbol to take care of this "problem." Cantor introduced the symbol ω to account for the problem of emerging *transfinites*. Let's superimpose a "picture" of $\{1, 2, 3, 4, 5, \ldots, n, \ldots\}$ on (1) in Figure 1.22:

The first transfinite ordinal ω corresponds to 2, $\omega + 1$ to 4, and so on. So, how many transfinite numbers are there? Let's look at (2) again, but this time taking into account the just acquired concept of the transfinite ω. We get something like Figure 1.24.

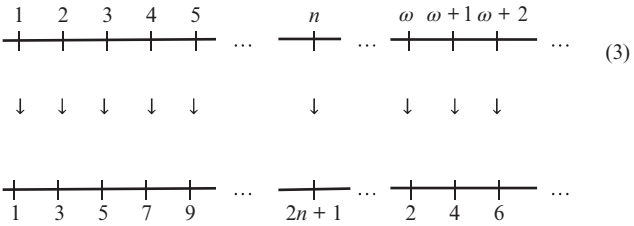


Figure 1.24

Yet another way to look at this.

Consider "Zeno-like" running on the real line such that every "step" (every number) is at half the distance of the previous one. That will look something like this:

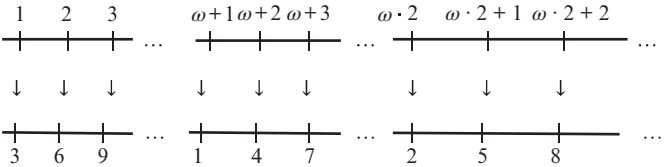


Figure 1.25

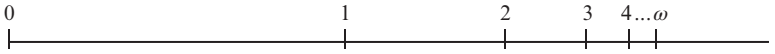


Figure 1.26

Now, "superimpose" a copy of Figure 1.25 onto each of the spaces between the points 0 and 1, 1 and 2, 2, and 3, and so on. That will look something like Figure 1.27.

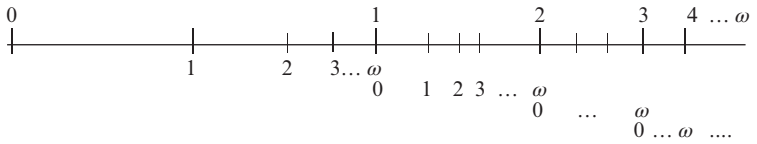


Figure 1.27

Repeating the process again we come to:

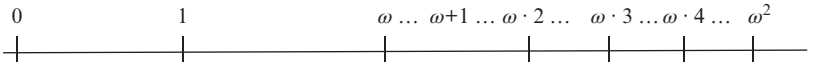


Figure 1.28

Doing it one more time gives: Figure 1.29

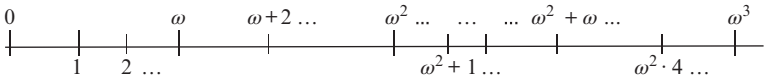


Figure 1.29

Repeating the process over and over again, we get the "final result" that would look something like Figure 1.30:

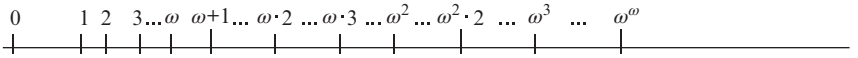


Figure 1.30

Of course, there is no reason to stop there, so the answer to the question of how many ordinal numbers there are is: There are infinitely many of them! Well, said Cantor, let's collect them "all," according to the following recipe: If α is an ordinal number, then we can always find the next ordinal $\alpha + 1$, and once we obtain a definite sequence of increasing ordinals, then we can find the last ordinal, called $\lim(\alpha)$, which is greater than all the α's. Thus, the following series of ordinals (which we tried to "visualize" earlier) is

$$1, 2, \ldots, \omega$$

$$\omega + 1,\ \omega + 2,\ \ldots$$
$$\omega \cdot 2,\ \omega \cdot 2 + 1,\ \omega \cdot 2 + 2,\ \ldots$$
$$\omega \cdot 3,\ \omega \cdot 3 + 1,\ \omega \cdot 2 + 2,\ \ldots$$
$$\vdots$$
$$\omega^2,\ \omega^2 + 1,\ \omega^2 + 2,\ \omega^2,\ \ldots$$
$$\omega^2 + \omega,\ \omega^2 + \ (\omega + \ 1),\ \omega^2 + (\omega + 2),\ \ldots$$
$$\vdots$$
$$\omega^3,\ \omega^3 + 1,\ \omega^3 + 2,\ \ldots$$
$$\vdots$$

An important note is in order. Unlike finite ordinals, the infinite ordinals demand a particular "order of operation," namely, commutativity no longer holds. Observe that $1 + \omega = \omega$, but $\omega + 1$ is the next "number" after ω. In other words,

$$1 + \omega = \omega \neq \omega + 1$$

Similarly, $2 \cdot \omega = \omega$, but $\omega \cdot 2 = \omega + \omega$.

We continue this way until we reach

$$\omega^\omega,\ \omega^\omega + 1,\ \omega^\omega + 2,\ \ldots$$
$$\vdots$$

And on and on until we reach

$$\omega^{\omega^\omega},\ \omega^{\omega^\omega} + 1,\ \omega^{\omega^\omega} + 2,\ \ldots$$
$$\vdots$$

Continuing this way (and now it really gets complicated), we come to a new sequence

$$\varepsilon_0, \varepsilon_1, \varepsilon_2,\ \ldots$$

where $\varepsilon_0 = \omega^{\omega^{\omega^{\cdot^{\cdot^{\cdot}}}}}$.

We can go on like this forever, right? Why not? Well, we "have been going" forever already while "counting" to ω, so "going" to ε_0 means – what?

Do you see where "*this*" is going? Do you see where our mind is taking us? Do you feel the richness of the underlining theory? Talking about "big," really, really "big," infinitely big, infinitely, infinitely big. The Absolute???

We know that the infinite exists without knowing its nature, just as we know that it is untrue that numbers are finite. Thus it is true that there is an infinite number, but we don't know what it is.[48]

So Cantor said:

The Absolute can only be acknowledged and admitted, never known, not even approximately.

Before I offer you another paradox, let's sum up what we know about ordinals:

(i) There is a first ordinal.
(ii) For each ordinal, there is an immediate successor ordinal.
(iii) For each set of ordinals, there is an ordinal which is the first succeeding them all.

So, we get the familiar sequence (cf. Example 1.61)

$$\begin{aligned}
0 &= \varnothing \\
1 &= \{\varnothing\} \\
2 &= \{0,1\} = \{\varnothing, \{\varnothing\}\} \\
3 &= \{0,1,2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \\
&\vdots \\
\omega &= \{0,1,2,3,\ldots\} \\
&= \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \ldots\} \\
\omega+1 &= \{0,1,2,3,\ldots,\omega\} \\
\omega+2 &= \{0,1,2,3,\ldots,\omega,\omega+1\} \\
&\vdots \\
&\vdots \\
\omega\cdot 2 &= \{0,1,2,3,\ldots,\omega,\omega+1,\ldots\} \\
\omega\cdot 2+1 &= \{0,1,2,3,\ldots,\omega,\omega+1,\ldots,\omega\cdot 2\} \\
&\vdots \\
&\vdots
\end{aligned}$$

Now consider

[48] Blaise Pascal, *Penseés.*

The Burali-Forti[49] Paradox

Suppose we are tempted to form a set Ω of *all* (infinitely many) ordinals. Can we do that? After all, we have a set of all natural numbers, rational numbers, real numbers, and so on, so why not do the same with ordinals? Well, if the set Ω exists, then it is the set (of ordinals) like any other. But then, by condition (iii), there must be *another ordinal*, $\Omega + 1$, the first to succeed it, that is, the first to succeed all the members of Ω. In other words, $\Omega < \Omega + 1$. But that contradicts the assumption that Ω contains *all* ordinals. We conclude: *The ordinal numbers do not form a set.*

In the same way by which we can always find more ordinals, we can always "find" more cardinals. Consider this sequence:

$$\aleph_0, \aleph_1, \aleph_2, \ \dots\ , \aleph_\omega, \aleph_{\omega+1}, \ \dots\ \aleph^{\omega^\omega}, \ \dots$$

As you might have anticipated by now, we do not stop here. A whole new universe of more and more complex structures opens up and the mathematics of transfinite turns out to be an exceptionally rich and philosophically exciting theory.

The assiduous reader may feel a little uneasy by now. After all the discussion of the transfinite, the author, with all of his fascination with alephs, so far has still not satisfactorily defined the very culprit of all of this – the real number(s). We will do that in a moment, but let's see briefly some of the remarkable features of ***arithmetic of transfinites***. We start with

Theorem 1.55 Let $\aleph$ be any infinite cardinal then

(i) $0 + \aleph = \aleph$
(ii) $n + \aleph = \aleph, \ \ \forall n \in \mathbf{N}(n \text{ is finite})$

Proof

(i) Let X be a set s.t. $|X| = \aleph$. We know that $0 = |\varnothing|$, thus we have

$$0 + \aleph = |\varnothing \cup X| = |X| = \aleph$$

(ii) Since X is an infinite set, there exists a set $A = \{a_1, \ a_2, \ \dots\ , \ a_n\ \} \subseteq X$, then

$$n + \aleph = |A \cup X| = |X| = \aleph \qquad \blacksquare$$

[49] Cesare Burali-Forti (1861–1931), Italian mathematician.

Example 1.106 Let n be a finite cardinal number. Then

$$n + \aleph_0 = |\{1, 2, 3, \ldots, n\} \cup \{n+1, n+2, \ldots\}| = \aleph_0$$ ■

Example 1.107 From Theorem 1.55, it follows that

$$\aleph_0 + 0 = \aleph_0 + 1 = \aleph_0 + 10^{10^{10^{10}}} \qquad (*)$$

You might immediately object that equation (*) cannot be true since, by elementary school algebra, (*) implies that $1 = 0$, which is obviously an absurdity. But remember, we are not doing ordinary algebra! Although the addition of transfinites is (well?) defined, interestingly enough, subtraction is not. Why not? Well, consider

$$\aleph_0 + 1 = \aleph_0$$

This one we can believe (we have proved even more: $\aleph_0 + \aleph_0 = \aleph_0$). Using ordinary algebra, we could go a step further and argue that

$$\begin{aligned} 1 + 0 &= 1 + \aleph_0 - \aleph_0 \\ &= (1 + \aleph_0) - \aleph_0 \\ &= \aleph_0 - \aleph_0 \\ &= 0 \end{aligned}$$

concluding that

$$1 = 0$$

which, of course, is nonsense. Therefore, we are forced to accept that $\aleph_0 - \aleph_0$ simply is not defined. ■

However, things are quite different for addition. Consider the following:

Example 1.108

$$\aleph_0 + \aleph_0 = |\{2, 4, 6, \ldots\} \cup \{1, 3, 5, \ldots\}| = |\{1, 2, 3, 4, 5, \ldots\}| = \aleph_0$$ ■

Actually, this can be generalized even further.

Theorem 1.56 Let $\aleph_\alpha$ and $\aleph_\beta$ be two infinite cardinals such that $\aleph_\alpha < \aleph_\beta$. Then,

$$\aleph_\alpha + \aleph_\beta = \aleph_\beta$$

Proof Let A and B be two sets s.t. $|A| = \aleph_\alpha$ and $|B| = \aleph_\beta$. If $\aleph_\alpha < \aleph_\beta$, then there exists a one-to-one function

$$f : A \to B$$

We need to show that A and $f(A)$ are equivalent sets. Consider the function

$$g : A \to f(A)$$

defined by the same rule as f, except that it is restricted to map A into $f(A) \subseteq B$. Since we choose f as one-to-one, g is also one-to-one by construction. Note that although f may not be onto, we would like g to be onto. To see that g is onto, let's take some $y \in f(A)$. By definition of $f(A)$ there has to be an $x \in A$, such that

$$y = f(x) = g(x)$$

We conclude that g is onto, and therefore a bijection. Thus A and $f(A)$ are equivalent, that is, $|A| = |f(A)|$.

Since $f(A) \subseteq B, f(A) \cup B = B$, we get the following:

$$\begin{aligned} \aleph_\alpha + \aleph_\beta &= |f(A)| + |B| \\ &= |f(A) \cup B| \\ &= |B| \\ &= \aleph_\beta \end{aligned}$$

which was to be proved. ■

Example 1.109

$$\aleph_0 + 2^{\aleph_0} = \aleph_0 + \aleph_1 = \aleph_0 + c = c$$ ■

Example 1.110 Let's examine the "sum" $c + c$ on the interval $[0, 1]$:

$$c + c = \left| \left[0, \frac{1}{2}\right] \cup \left(\frac{1}{2}, 1\right] \right| = c$$ ■

Example 1.111

$$\aleph + 2^{\aleph} = 2^{\aleph}$$

Since $\aleph_0$ is the first infinite cardinal, it follows that $\aleph_0 \leq \aleph$ for any other infinite cardinal. Hence, as a consequence of Theorem 1.56, $\aleph_0$ behaves as a neutral element with respect to addition of infinite cardinals, that is, it always holds that

$$\aleph + \aleph_0 = \aleph_0 + \aleph = \aleph$$

Since cardinals can be added, we conclude that

$$\aleph + \aleph + \cdots + \aleph = n \cdot \aleph = \aleph, \quad \forall\ \in \mathbf{N} \tag{*}$$

■

However, we also have

Theorem 1.57 Let $\aleph$ be an infinite cardinal. Then,

$$0 \cdot \aleph = 0$$

Proof In Example/Exercise 1.64, you were asked to prove that for any set A, $A \times \varnothing = \varnothing \times A = \varnothing$. If you haven't done it, let's do it now so we can use that to prove our theorem.

Suppose $A \times \varnothing \neq \varnothing$. Then, there exists an $n \in A \times \varnothing$ such that $n = (x, y)$, with $x \in A$ and $y \in \varnothing$. But this contradicts the fact that $\varnothing$ has no elements. Thus, our supposition was wrong and we conclude that

$$A \times \varnothing = \varnothing \times A = \varnothing$$

as claimed. Since we didn't specify A to be any particular set, we take that our assertion also holds for any set; therefore, $\mathbf{N} \times \varnothing = \varnothing$ as well as $\mathbf{R} \times \varnothing = \varnothing$. Now the proof of the theorem follows immediately: Consider a set A such that $|A| = \aleph$.

$$0 \cdot \aleph = |\varnothing \times A| = |\varnothing| = 0$$

■

So far so good. But now the next natural question arises: if we accept the statement (*) from Example 1.111, how far can we push the multiplication of alephs? In other words, what is $\aleph_0 \cdot \aleph_0$? Or, in general, $\aleph \cdot \aleph_0$?

Recalling our discussion of the Cartesian product from Chapter 1.6, you can easily convince yourself that if we are given k finite sets $A_1, A_2, \ldots, A_k$ such that

$$A_1 = n_1, |A_2| = n_2, \ldots, |A_k| = n_k$$

then

$$|A_1 \times A_2 \times \ldots \times A_k| = n_1 \cdot n_2 \cdot \ldots \cdot n_k$$

Indeed, each element in $A_1 \times A_2 \times \ldots \times A_k$ is a k-tuple of the form $(a_1, a_2, \ldots, a_k)$, where $a_i \in A_i$. Thus, there are n_1 ways to choose the first element in a k-tuple, n_2 ways to choose the second one, and so on. Therefore,

there are $n_1 \cdot n_2 \cdot \ldots \cdot n_k$ elements in $A_1 \times A_2 \times \ldots \times A_k$. We extend this formalism to calculate the product of infinite cardinals by

Definition 1.44 Let A and B be two sets such that $|A| = \aleph_\alpha$ and $|B| = \aleph_\beta$ are respective infinite cardinals. Then, we define

$$\aleph_\alpha \cdot \aleph_\beta = |A \times B|$$

Example 1.112

$$\begin{aligned} \aleph_0 \cdot \aleph_0 &= |\mathbf{N} \times \mathbf{N}| \\ &= |\mathbf{N}| \\ &= \aleph_0 \end{aligned}$$

Note that we utilize Theorem 1.37 in the second step. ■

This rule is valid for any other aleph, that is

$$\aleph \cdot \aleph = \aleph$$

Here is another good example:

Theorem 1.58 $\mathbf{R} \sim \mathbf{R} \times \mathbf{R}$. That is to say $|\mathbf{R} \times \mathbf{R}| = |\mathbf{R}|$.

Proof Consider a function

$$f : \mathbf{R} \to \mathbf{R} \times \mathbf{R}$$

defined by $f(x) = (x, 0), \ \forall x \in \mathbf{R}$. f is clearly a one-to-one function. In order to complete the proof, we also need another one-to-one function

$$g : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$$

Since the cardinality of the interval $(0, 1)$ is the same as the cardinality of $\mathbf{R}$, rather than working with the whole $\mathbf{R}$ we prefer to work with

$$g : (0, 1) \times (0, 1) \to (0, 1)$$

defined by

$$g(0.\, a_1 a_2 a_3 \ \ldots\ , 0.\, b_1 b_2 b_3 \ \ldots) = 0.\, a_1 b_1 a_2 b_2 a_3 b_3 \ \ldots$$

with the only restriction that the a_i' s and $b_i's$ not be repeating nines. Thus, g is a well-defined function and it is clearly one-to-one. The Schröder–Bernstein theorem immediately leads to the desired proof. ■

Example/Exercise 1.113 Show that $\mathbf{R} \times \mathbf{R} \times \mathbf{R} \sim \mathbf{R}$.

Theorem 1.59 Let A and B be two sets such that $|A| = \aleph_\alpha$ and $|B| = \aleph_\beta$ are respective infinite cardinals. Then,

$$\aleph_\alpha \cdot \aleph_\beta = \aleph_\beta \cdot \aleph_\alpha$$

Proof Consider a function

$$f : A \times B \to B \times A$$

defined by

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

The function f is obviously a bijection. That implies

$$\aleph_\alpha \cdot \aleph_\beta = |A \times B| = |B \times A| = \aleph_\beta \cdot \aleph_\alpha$$

which was to be proved. ■

Without proof, we state

Theorem 1.60 If $\aleph_\alpha$ and $\aleph_\beta$ are two infinite cardinals such that $\aleph_\alpha \leq \aleph_\beta$, then

$$\aleph_\alpha \cdot \aleph_\beta = \aleph_\beta$$

that is, the larger of the two cardinals.

The following examples illustrate another "unusual" consequence of multiplication of alephs.

Example 1.114

$$\aleph \cdot \aleph = \aleph^2$$

but also

$$\aleph \cdot \aleph = \aleph$$

Thus,

$$\sqrt{\aleph} = \aleph$$ ■

A few more peculiar properties of infinite cardinals are listed without proof in the following theorem.

Theorem 1.61 If $n \in \mathbf{N}$ is any finite cardinal, then

(i) $n^{\aleph_0} = \aleph_0^{\aleph_0} = \mathfrak{c}^{\aleph_0} = \mathfrak{c}$
(ii) $2^{\mathfrak{c}} = n^c = \aleph_0^{\mathfrak{c}} = \mathfrak{c}^{\mathfrak{c}}$

Example 1.115

(i) $c \cdot c = c$[50]
(ii) $\aleph_0 \cdot c = c$
(iii) $\aleph_0 \cdot \aleph = \aleph$
(iv) $\aleph \cdot \aleph \cdot \aleph = \aleph$

■

Example 1.116

(i) $c^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \aleph_0} = 2^{\aleph_0} = c$
(ii) $c^c = (2^{\aleph_0})^c = 2^{\aleph_0 c} = 2^c = c$ ■

Example/Exercise 1.117 Prove that $c^{\aleph_0} = c$.

It is very important to stress again that, regardless of the fact that we "know" how to multiply cardinals, the division is not defined. Our inherent intuition is worthless when dealing with alephs. Here is a simple example: Suppose we can divide cardinals. Then, it would be natural to infer the following:

$$\aleph_0 \cdot \frac{1}{\aleph_0} = 1$$

From expression (*) on page 95, it would in particular follow that $2\aleph_0 = \aleph_0$, so by ordinary algebra, we could write

$$2 \cdot \aleph_0 \frac{1}{\aleph_0} = \aleph_0 \cdot \frac{1}{\aleph_0}$$

which would entail

$$2 = 1$$

[50]Remember, we are assuming the Continuum Hypothesis, that is, $c = \aleph_1$

Obviously, we cannot accept something this absurd. Thus, our assumption of the possibility of division was wrong. We conclude that, the same as with subtraction, division is also not defined! Again, we realize that mathematics has a life of its own – we just have to discover the beautiful new world hidden under the surface.

Before concluding this section, I need to tell you about an axiom that may seem fairly obvious to you. Indeed, we have done much of our set theory tacitly assuming its validity. As a matter of fact, it has not been recognized by mathematicians for a long time. And even today, regardless of the many beautiful results one can prove with it, many mathematicians are rather skirmish about it. The discomfort that they feel is mostly due to its nonconstructive nature and some very unexpected and counterintuitive implications that follow. Let's devote a short subsection to the (in)famous Axiom of Choice.

Axiom of Choice

In mathematics, there are arguably very few so "simple" and "self-evident" and still so controversial axioms as The Axiom of Choice (AC). As B. Russell said: *At first it seems obvious, but the more you think about it the stranger the deductions from this axiom seem to become; in the end you cease to understand what is meant by it.* Many crucial concepts in different branches of mathematics, as well as the (proofs of) theorems therein, are based on it. However, do note that I put "simple" and "self-evident" in quotation marks. Being "simple" and "self-evident" can be misleading indeed! For instance, on page 46, after introducing the Cartesian product, we asked whether one could extend the conclusion (evidently valid for finite sets) to infinite ones (see Example 1.64) as well. Similarly, when discussing the Continuum Hypothesis, we have encountered the sequence of alephs

$$\aleph_0, \aleph_1, \aleph_2, \aleph_3, \ \dots \ , \aleph_\omega, \ \dots$$

and we ask: Suppose an infinite set of infinite sets is given, *is it possible to choose one element from each set without giving a* ***rule of choice*** *in advance?*

It turns out that the issues involved are very profound and we will end this chapter with just a rudimentary exposition of the subject.

A very well-known and witty formulation of this question, which I like to call "*On Socks and Shoes,*" is due to (who else but) Bertrand Russell who said: Suppose there are infinitely many pairs of socks and shoes. To choose one sock from each pair of identical socks requires the Axiom of Choice, but for shoes the Axiom is not needed, it suffices to simply impose a rule "always chose the left shoe" and we are done. Once again, the phrase "*infinitely many*" is crucial, for with finite sets of socks we wouldn't have the problem. (Can you figure out why?) Let's start with a few simpler examples.

Example 1.118 Consider a set $S = \{A, B, C\}$, where A, B, and C are disjoint sets such that

$$A = \{a_1, a_2\}, B = \{b_1, b_2\}, \text{ and } C = \{c_1, c_2\}$$

Suppose we want to construct a set $\mathcal{A}$ by choosing for its elements one and only one element from each of the sets A, B, C. For instance, one possible "*choice set*" could be

$$\mathcal{A} = \{a_1, b_1, c_1\}$$

where the "*choice function*" was: "*take the 'first' element from each set*" (whatever "*the first*" means – in this case, obviously, an element with index 1). ■

However, how would you do

Example/Exercise 1.119

(i) If $\varnothing \in S$ what would be the choice set?
(ii) If $S = \varnothing$ what would be the choice set?

Example 1.120 Let

$$S = \{A | A \subseteq \mathbf{N}, A \neq \varnothing\}$$

be a collection of all nonempty subsets of $\mathbf{N}$, then we can simply define the "choice function" by saying $f(A) =$ smallest member of A. ■

Example 1.121 Let

$$S = \{I = [a, b] | a, b \in \mathbf{R}, d(a, b) < \infty\}$$

that is, a collection of all intervals of real numbers with finite length. Then, we can define $f(I)$ to be the midpoint of the interval I. ■

Now comes a problem: Consider Example 1.118, again assuming this time that the sets $A, B, C \in S$ are open intervals, that is

$$A = (a, b), B = (c, d) \quad \text{and} \quad C = (e, f); \quad a, b, c, d, e, f \in \mathbf{R}$$

How would you choose an element from each of the sets to construct the set $\mathcal{A}$? (Say, you first consider our familiar interval $(0, 1)$, how would you take the least element from it?) To make it even more intriguing, take the set S to be the set of all nonempty subsets of $\mathbf{R}$. How would we find a suitable function f to collect an element from all of those subsets? So we ask: If an infinite set of infinite sets is given, is it possible to choose one element from each set without giving the rule of choice in advance? Yes, it is possible, said Zermelo.[51] In 1904, he introduced the

[51] Ernst Zermelo (1871–1953), German mathematician.

Axiom of Choice (AC)

Let S be a collection of mutually disjoint nonempty sets; then, there exists a set $\mathcal{A}$ consisting of exactly one member chosen from each set in the collection S. In other words, given any family of nonempty sets,

$$S = \{A_i | i \in I, I = \text{index set}\}$$

there exists a function – the choice function

$$f : I \to \underset{i \in I}{\cup} A_i$$

such that $f(i) = a_i \in A_i$.

Equivalently, we can approach AC as follows:

Let $S = \cup_{i \in I} A_i$ be a nonempty family of nonempty sets. Then, the Cartesian product $\prod_{i \in I} A_i$ of the sets A_i is the set of all choice functions $f : I \to \cup_{i \in I} A_i$ where $f(i) = a_i \in A_i$, for all $i \in I$. In other words, for every $i \in I$, f chooses a point a_i from each set A_i.

Hence, we can restate the Axiom of Choice as follows: The Cartesian product of a nonempty family of nonempty sets is nonempty.

Note that the axiom only claims the existence of the choice function. It doesn't say anything about its construction.

As an example, let's prove

Theorem 1.62 Every infinite set has a countably infinite subset.

Proof Let S be an infinite set. Consider a set $A_1 = S \backslash \{a_1\}$, where $a_1 \in S$. A_1 is certainly not empty since S is not empty. Furthermore, A_1 is infinite, for if A_1 were finite S would be finite too, contradicting our original assumption. Next, we can consider a set.

$A_2 = A_1 \backslash \{a_2\} = S \backslash \{a_1, a_2\}$, $a_2 \in S$. A_2 is also infinite, and in particular it contains an element a_3. Can we continue these arguments ad infinitum? Well, to continue with this argumentation we need AC, and we construct A_i for any $1 \leq i \leq n, i \in \mathbf{N}$ according to the aforementioned prescription. We claim that A_i is infinite. But then, there is $a_{n+1} \in A_n$, such that $A_{n+1} = A_n \backslash \{a_{n+1}\}$ is also infinite. Note that if $i < j$, then $a_i \in A_{i+1}$, but $a_j \notin A_{i+1}$. Now, if we let $B = \{a_i | i \in \mathbf{N}\}$, then B is infinite and $|B| = |\mathbf{N}|$. ■

Going back to our list of cardinals, it is reasonable to ask: Can we form a set C of all cardinal numbers? Well, let's try that. Suppose C is a set of all cardinals. Then, for every $c \in C$, there exists a set A_c such that $c = |A_c|$. Furthermore, let

$$A = \underset{c \in G}{\cup} A_c$$

Consider now $\mathcal{P}(A)$ and let $|\mathcal{P}(A)| = \alpha$. Then, since

$$|\mathcal{P}(A)| = \alpha$$

we have

$$|\mathcal{P}(A)| \leq |A|$$

On the other hand, by Cantor's theorem

$$|\mathcal{P}(A)| > |A|$$

so we have a contradiction. We see that, the same as with the ordinals, the axioms of set theory fail to accommodate the cardinals also.

Finally, without proof, we list three crucial theorems of mathematics:

Theorem 1.63 (Zorn lemma) Let X be a nonempty partially ordered set, whose every linearly ordered subset has an upper bound in X. Then X contains at least one maximal element.

Theorem 1.64 (Zermelo's well-ordering theorem) Every nonempty set X can be well ordered.

Theorem 1.65 The following are equivalent:

(i) Axiom of choice
(ii) Zorn lemma
(iii) Well-ordering theorem.

It might be appropriate to conclude this subsection with a quote you may philosophically disagree with but, nevertheless, you have to admit it is rather captivating:

> *… For me, and I suppose for most mathematicians, there is another reality, which I will call "mathematical reality" … I believe that mathematical reality lies outside us, that our function is to discover or observe it, and that the theorems which we prove, and which we describe grandiloquently as our "creations" are simply our notes of our observations.*[52]

1.10 THE SET R – REAL NUMBERS II

So far, we have dealt with real numbers more or less heuristically. We assumed their existence for a simple reason: set **Q** obviously was not sufficiently rich

[52]Hardy, G. H., *A Mathematician's Apology,* Cambridge University Press, 1967.

enough to accommodate everything we wanted to do mathematically. Also, as the reader is obviously aware of by now, the importance of set **R**, and therefore its proper definition, can hardly be overstated. We need to introduce a few more concepts in order to adequately address real numbers.

Definition 1.45 (A ring) A **ring** is a set R with two binary operations on it: "+" and "$\cdot$" called "addition" and "multiplication,"[53] respectively, such that

1. addition is commutative: $a + b = b + a, \ \ \forall a, b \in R$;
2. addition is associative: $a + (b + c) = (a + b) + c, \ \ \forall a, b, c \in R$;
3. addition has a neutral element with respect to addition: $\exists\, 0 \in X, \ \ \text{s.t.} \ \ a + 0 = 0 + a = a$,
4. addition has an inverse: $\forall a \in R, \exists(-a) \in R, \ \ \text{s.t.} \ \ a + (-a) = (-a) + a = 0$;
5. multiplication is associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c, \ \ \forall a, b, c \in R$;
6. multiplication is distributive with respect to addition:

$$a \cdot (b + c) = a \cdot b + a \cdot c, \ \ \forall a, b, c \in R$$

Definition 1.46 (A field) A **field** Φ is a set with two binary operations on it, such that Φ ***is a commutative ring with the identity with respect to "multiplication,"*** that is, in addition to (1)–(6) from Definition 1.44, there are three more properties that have to be satisfied:

1. $a \cdot b = b \cdot a, \ \ \forall a, b \in \Phi$
2. There exists a unique element $1 \in \Phi$, which we call the **identity** (sometimes unity) with respect to multiplication, s.t. $1 \cdot a = a \cdot 1 = a, \ \ \forall a \in \Phi$, and
3. for every element $a \in \Phi$, there exists a **multiplicative inverse** $a^{-1} \in \Phi$, s.t.

$$a \cdot a^{-1} = a^{-1} \cdot a = 1$$

Example 1.122 The sets Z, Q, and R, with the usual addition and multiplication, are rings. Q and R are also fields. ■

Example/Exercise 1.123 Convince yourself that the set **R** is a field.

Example 1.124 The set $2\mathbf{Z}$ of even integers with the usual addition and multiplication is a ring. Note that it doesn't have an identity with respect to multiplication. ■

Example 1.125 I hope that you are familiar with the concept of a polynomial of n th degree in one variable, that is, a function of the form

[53] "Addition" and "Multiplication" are names that we conveniently associate with "+" and "$\cdot$" These operations do not necessarily have to be our ordinary addition and multiplication.

$$f(x) = P_n(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0$$

where $n, n-1, n-2, \ldots \in \mathbf{Z}^+$, that is, nonnegative integers,[54] and $a_n, a_{n-1}, \ldots a_0 \in R$. You can easily convince yourself that a set of all polynomials is a ring. On the other hand, a set of polynomials is not a field for obvious reasons: there is no multiplicative inverse that is also a polynomial. ■

Example 1.126 If you are familiar with matrices, you can immediately recognize that the set of all square $(n \times n)$ – matrices form a noncommutative ring. ■

Definition 1.47 We say that a field Φ is an **ordered field** if the following is satisfied:

1. If $a, b \in \Phi$, then one and only one of the following holds:

$$a < b, \quad a = b, \text{ or } a > b$$

2. If $a, b, c \in \Phi$, s.t. $a > b$, and $b > c$, then $a > c$
3. If $a, b, c \in \Phi$, and if $a > b$, then $a + c > b + c$.
4. If $a, b, c \in \Phi$ and if $a > b$, with $c > 0$, then $ac > bc$.

Example 1.127 Prove that $a > 0$ iff $-a < 0$. ■

Proof

(i) If $a > 0$, then $a + (-a) > 0 + (-a) \Rightarrow 0 > -a$
(ii) If $-a < 0$, then $-a + a < 0 + a \Rightarrow 0 < a$ ■

Example 1.128 Prove that if $a > 0$ and $b < 0$, then $a \cdot b < 0$. ■

Proof Suppose $a > 0$ and $b < 0$. Then,

$$-b > 0$$

therefore,

$$a \cdot (-b) = -(a \cdot b) > 0$$

Hence,

$$a \cdot b < 0$$ ■

[54] $Z^+ = N \cup \{0\} = N^+$.

Example/Exercise 1.129 Prove that if $a > 0$ and $b > 0$, then $a \cdot b > 0$.

Example/Exercise 1.130 Prove that if $a \neq 0$, then $a^2 > 0$.

Example/Exercise 1.131 Prove that if $a < 0$ and $b < 0$, then $a \cdot b > 0$.

Now, we are ready for some important definitions that will safely lead us to a better insight into real numbers.

Definition 1.48 Let Φ be an ordered field, and let A be a nonempty subset of Φ. We say that A is **bounded above**, if there exists an element $a \in \Phi$, such that $x \leq a, \ \ \forall x \in A$. We call a an **upper bound** of A.

Similarly, we say that A is **bounded below**, if there exists a $b \in \Phi$, such that $x \geq b, \ \ \forall x \in A$. We call b a **lower bound** of A.

We say that A is **bounded** if it is bounded above and below.

Definition 1.49 Let A be a nonempty subset of Φ. We say that $a \in \Phi$ is the **least upper bound or a *supremum*** of A iff a is an upper bound of A, and for every other upper bound x of A, $a \leq x$. We write $a = \sup A$.

Definition 1.50 Let A be a nonempty subset of Φ. We say that $b \in \Phi$ is the **greatest lower bound or *infimum*** of A iff b is a lower bound of A, and for every other lower bound x of A, $b \geq x$. We write $b = \inf A$.

Definition 1.51 A field Φ is said to be **completely ordered** if the **completeness property** is satisfied, that is, if every nonempty bounded set $S \subseteq \Phi$ has a supremum in the field.

Theorem 1.66 If a nonempty set A has a supremum, then $\sup A$ is unique.

Proof Suppose there are two elements x_1 and x_2, both supremums of a set A. By definition, both x_1 and x_2 are upper bounds of A, and since x_1 is a supremum, x_1 is less or equal to any other upper bound, in particular, $x_1 \leq x_2$. On the other hand, x_2, being a supremum, is less or equal to any other upper bound, in particular, $x_2 \leq x_1$. Hence $x_1 = x_2$. ■

Now you should be able to prove

Theorem 1.67 If a nonempty set A has an infimum, then $\inf A$ is unique.

Finally, we have

Definition 1.52 The set of real numbers **R** is a completely ordered field.

As an additional exercise you may want to revisit Theorems 1.66 and 1.67 and simply replace the words "nonempty set A" by "nonempty subset A of R."

Let's pause for a moment and reflect on all of this. Suppose we are familiar only with rational numbers and take a subset of all rational numbers such that $(p/q)^2 < 2$. This subset does not have a supremum, because if it did have a supremum, say a, we could eventually get $a^2 = 2$. But we have proved (see Theorem 1.45) that this is impossible. So, indeed, we want a set of numbers, call it $\mathbf{R}$, with a property that any nonempty subset $A \subseteq R$, which is bounded above, has a supremum. Well, said Dedekind,[55] suppose we knew only the infinite set $\mathbf{Q}$. Here is what we could do. Let's partition – *cut* – set $\mathbf{Q}$ into two subsets L and R, such that (1) every element of L is smaller than every element of R, and (2) R has no least element. The idea being that every rational number is either an element of L or an element of R. Thus, we have

$$L = \{x \in \mathbf{Q} | x < r\} \quad \text{and} \quad R = \{x \in \mathbf{Q} | x > r\}$$

For instance, our (in)famous $\sqrt{2}$ would be represented by the cut $[L, R]$ such that

$$L = \left\{ \frac{p}{q} \middle| \left(\frac{p}{q}\right)^2 < 2, \;\; p, q \in \mathbf{Z} \right\}$$

and

$$R = \left\{ \frac{p}{q} \middle| \left(\frac{p}{q}\right)^2 > 2, p, q \in \mathbf{Z} \right\}$$

It may be worth mentioning at this point that in 1872, when Dedekind introduced his "cut," topology did not exist. Today's treatment of the "Dedekind cut," as a topological space in open interval topology, had to wait for better times. Dedekind's idea still holds today: real numbers cannot be represented in terms of discrete mathematical objects. The only way to consistently represent arbitrary real numbers is by infinite sets. (Remember the statement of Cantorism at the beginning of this chapter: *Everything is a set*.)

We conclude this discussion with

Definition 1.53 A real number is a pair $[R, L]$ of infinite sets.

For the sake of completeness, let's put together everything we have said about the field of real numbers and state it explicitly: Let $\mathbf{R}$ be a set with two binary operations on it, called **addition, "+,"** and **multiplication, "·."** These operations satisfy the following properties:

[55] Richard Dedekind (1831–1916).

1. $\forall a, b \in \mathbf{R},\ \ a + b = b + a \in \mathbf{R}$
2. $\forall a, b \in \mathbf{R},\ \ a \cdot b = b \cdot a \in \mathbf{R}$
3. $\forall a, b, c \in \mathbf{R},\ \ a + (b + c) = (a + b) + c$
4. $\forall a, b, c \in \mathbf{R},\ \ a \cdot (b \cdot c) = (a \cdot b) \cdot c$
5. $\forall a, b, c \in \mathbf{R},\ \ a \cdot (b + c) = a \cdot b + a \cdot c$
6. $\exists\, 0 \in \mathbf{R},$ s.t. $0 + a = a + 0 = a,\ \ \forall a \in \mathbf{R}$
7. $\exists\, 1 \in \mathbf{R},$ s.t. $1 \cdot a = a \cdot 1 = a,\ \ \forall a \in \mathbf{R}$
8. $\forall a \in \mathbf{R},\ \ \exists(-a) \in \mathbf{R},$ s.t. $a + (-a) = (-a) + a = 0$
9. $\forall a \in \mathbf{R},\ \ \exists a^{-1} \in \mathbf{R},$ s.t. $a \cdot a^{-1} = a^{-1} \cdot a = 1$

We call the field **R** the field of real numbers. Consequently, we have a theorem that summarizes the most important algebraic properties of the field **R**.

Theorem 1.68 For any real numbers $a, b, c, d \in \mathbf{R}$, the following holds:

(i) *Cancellation Law for Addition*: If $a + b = a + c$, then $b = c$

(ii) *Possibility of Subtraction*: Given a and b, there is exactly one x such that

$$a + x = b$$

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

(iv) $a \cdot (b - c) = a \cdot b - a \cdot c$

(v) $0 \cdot a = a \cdot 0 = 0$

(vi) *Cancellation Law for Multiplication*: If $ab = ac$ and $a \neq 0$, then $b = c$

(vii) If $b \neq 0$, then $a/b = ab^{-1}$

(viii) If $a \neq 0$, then $(a^{-1})^{-1} = a$

(ix) *Zero Product Property:* If $a \cdot b = 0$, then $a = 0$ or $b = 0$

(x) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$

(xi) $(-a) \cdot (-b) = a \cdot b$

(xii) *Rule of addition of fractions:*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \quad b \neq 0,\ \ d \neq 0$$

(xiii) *Rule of multiplication of fractions:*

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}, \quad b \neq 0,\ \ d \neq 0$$

(xiv) *Rule of division of Fractions:*

$$\frac{a/b}{c/d} = \frac{a \cdot d}{b \cdot c}, \quad b \neq 0,\ c \neq 0,\ d \neq 0$$

(xv) *Trichotomy Law:* Given any two real numbers a, b, only one of the three relations holds:

$$a < b, b < a, \quad \text{or} \quad a = b$$

(xvi) *Transitive Law*: If $a < b$ and $b < c$, then $a < c$

(xvii) If $a < b$, then $a + b < b + c$

(xviii) If $a < b$ and $c > 0$, then $ac < bc$

(xix) If $a \neq 0$, then $a^2 > 0$

We have already proved many of the statements in the aforementioned theorem in a different context. The reader shouldn't have any problems proving the remaining parts.

Definition 1.54 Suppose $a \in \mathbf{R}$, we define the **absolute value of** a by

$$|a| = \begin{cases} a & \text{if} \quad a \geq 0 \\ -a & \text{if} \quad a < 0 \end{cases}$$

Theorem 1.69

(i) For any $a, b \in \mathbf{R}$, $|ab| = |a||b|$.

(ii) For any $a, b \in \mathbf{R}$, $|a + b| \leq |a| + |b|$ (*Triangle inequality*).

Proof

(i) Suppose $a > 0$ and $b > 0$. Then, by definition, $|a| = a$ and $|b| = b$. Thus, $|a||b| = ab$. On the other hand, $|ab| = ab$. We conclude that $|ab| = |a||b|$. If $a < 0$ and $b < 0$, then $|a| = -a$ and $|b| = -b$, so we again have

$$|ab| = ab = (-a)(-b) = |a||b|$$

(ii) Consider the following obvious inequalities:

$$-|a| \leq a \leq |a| \tag{1.14}$$

$$-|b| \leq b \leq |b| \tag{1.15}$$

Adding (1.14) and (1.15), we get

$$-(|a| + |b|) \leq a + b \leq |a| + |b|$$

which implies

$$|a + b| \leq |a| + |b|$$

■

Example/Exercise 1.132 Prove that

$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}, \quad b \neq 0$$

1.11 SUPPLEMENTARY PROBLEMS

1. Given four sets $A = \{a, b, c, d\}, B = \{a, a, c, b, d, d\}, C = \{d, b, a, c, 0\}$, and $D = \{d, b, a, c\}$, determine which sets are equal.

2. Is $a = \{a\}$? Why? Why not?

3. Is $0 = \{\}$? Why? Why not?

4. Let $A = \{a, b, c\{a\}, \{\{a\}\}, \{a, d\}, d\}$
(i) Is $a \in A$?
(ii) Is $\{a\} \subseteq A$?
(iii) Is $\{\{a\}\} \in A$?
(iv) Is $\{\{a\}\} \subseteq A$?
(v) Is $\{a, b, c\} \subseteq A$?

5. Let $\mathcal{U} = \{a, b, c, d, e, f, g\}$ be a universal set, and let $A = \{b, c, d, f\}$, $B = \{a, b, c\}$, and $C = \{d, e, f, g\}$. Find
(i) $A \cap B$
(ii) $A \cup B$
(iii) $A \cap C$
(iv) $B \backslash A$
(v) $A \backslash (B \cap C)$

6. Let A be a set. Show that
(i) $A \cup \varnothing = A$
(ii) $A \cap \varnothing = \varnothing$
(iii) $A \cup A = A$
(iv) $A \cap A = A$
(v) $A \backslash \varnothing = A$

7. Let the universal set be the set of all natural numbers, that is, let $\mathcal{U} = N$ and $A = \{x | x = 2n, n \in \mathbf{N}\}$, find
(i) $A \cap \mathbf{N}$
(ii) $A \cup \mathbf{N}$
(iii) A^c

8. Let the universal set $\mathcal{U}$ be the set of all real numbers $\mathbf{R}$, and let $A = \{x \in \mathbf{R} | 0 \leq x \leq 1\}$, and $B = \{x \in \mathbf{R} | -3 < x \leq 3\}$. **Find**

(i) $A \cup B$

(ii) $A \cap B$

(iii) A^c

(iv) B^c

(v) $(A \cap B)^c$

9. What is the cardinality of each of the following sets?

(i) $\{a\}$

(ii) $\{\{a\}\}$

(iii) $\{\varnothing\}$

(iv) $\{\varnothing, \{\varnothing\}\}$

(v) $\{a, \{\{\varnothing\}\}\}$

10. Show that for all sets $A, B,$ and C

(i) If $A \subseteq B$ and $A \subseteq C$ then $A \subseteq B \cap C$

(ii) If $A \subseteq C$ and $B \subseteq C$ then $A \cup B \subseteq C$

11. Show that if $A \subseteq B$, then $B = A \cup (B \backslash A)$.

12. Show that for all sets $A, B,$ and C

$$(A \cup B) \backslash C = (A \backslash C) \cup (B \backslash C)$$

13. Show that for any sets A and B

$$A \backslash B = A \backslash (A \cap B)$$

14. Let $A \subseteq C$ and $B \subseteq C$. Prove the following assertions:

(i) $C \backslash (C \backslash A) = A$

(ii) $C \backslash (A \cap B) = (C \backslash A) \cup (C \backslash B)$

(iii) $C \backslash (A \cup B) = (C \backslash A) \cap (C \backslash B)$

15. Show that for all sets $A, B,$ and C

$$(A \backslash B) \backslash C = (A \backslash C) \backslash (B \backslash C)$$

16. Prove: $(A \cup B) \cap B^c = A$ iff $A \cap B = \varnothing$.

17. Let A and B be subsets of X. Prove that

$$(A \subseteq B) \Leftrightarrow [(x \backslash B) \subseteq (x \backslash A)]$$

18. Show that (i) and (ii) are logically equivalent:
(i) A and B are disjoint sets.
(ii) $A \subseteq \mathcal{U} \backslash B, \quad B \subseteq \mathcal{U} \backslash A$.

19. Show that for all $A \neq B \neq \emptyset, \quad A \times B \neq B \times A$.

20. Prove that for all sets $A, B, C,$ and D

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

21. Suppose $A = \{a, b\}$ and $B = \{c, b\}$. Find
(i) $\mathcal{P}(A \cap B)$
(ii) $\mathcal{P}(A \cup B)$

22. Let $A, B \subseteq \mathcal{U}$. Show that

$$(A \cup B) \cap (A^c \cup B^c) = A \Delta B$$

23. Let $A, B \subseteq \mathcal{U}$. Show that
(i) $A \Delta B = (A \cup B) \backslash (A \cap B)$
(ii) $A \Delta (B \Delta C) = (A \Delta B) \Delta C$

24. Show that for all sets $A, B,$ and C
(i) $(A \cap B) \times C = (A \times C) \cap (B \times C)$
(ii) $(A \cup B) \times C = (A \times C) \cup (B \times C)$

25. Show that for all sets, B and C

$$(A \backslash B) \times C = (A \times C) \backslash (B \times C)$$

26. Verify that Definition 1.23 is a good definition, that is, prove that if X is an infinitely countable set, then it has a proper subset with the same cardinality.

27. Which of the following is true:
(i) $\mathbf{N} \subseteq \mathbf{Z}$
(ii) $\mathbf{Q} \subseteq \mathbf{Z}$

(iii) $\mathbf{R} \cap \mathbf{Q} = \mathbf{Q}$
(iv) $\mathbf{Z} \cup \mathbf{Q} = \mathbf{Q}$
(v) $\mathbf{Q} \cap \mathbf{Z} \cap \mathbf{N} = \mathbf{N}$

28. Prove that if $X \subseteq \mathbf{N}$, then X is either countably infinite or finite.

29. Prove that if
(a) X is countable and $Y \subseteq X$ is finite, then $X \backslash Y$ is countable.
(b) X is uncountable and $Y \subseteq X$ is countable, then $X \backslash Y$ is uncountable.

30. Prove that $A = [0, 1]$ and $B = [0, 2]$ have the same number of elements.

31. Prove that $A = (0, 1)$ and $B = (0, 2)$ have the same number of elements.

32. Determine the cardinality of the following sets:
(i) $\mathbf{N} \cap [1, \pi]$
(ii) $\mathbf{N} \cup [1, \pi]$

33. Determine the cardinality of the following sets:
(i) $\mathbf{Q}^3$
(ii) $\mathbf{Q}^{\mathbf{R}}$

34. Define $\mathbf{N}^k = \underbrace{\mathbf{N} \cdot \mathbf{N} \cdot \cdots \cdot \mathbf{N}}_{k \text{ times}}$. Prove that $\mathbf{N}^3 \sim \mathbf{N}$ that is $|\mathbf{N}^3| = |\mathbf{N}|$

35. Let S_n be the set of all subsets of $\mathbf{N}$ whose size is n. Prove that S_n is countable for all $n \in \mathbf{N}$.

36. Show that $Q \cap [0, 1]$ is countable.

37. Show that $(\mathbf{R} \backslash \mathbf{Q}) \sim \mathbf{R}$.

38. Show that for all $n = \mathbf{N}$, $|\mathbf{R}^n| = c$.

39. Determine the cardinality of the following sets:
(i) $\mathcal{P}(\mathbf{Z}) \times \mathcal{P}(\mathbf{Z})$
(ii) $\mathcal{P}(\mathcal{P}(\mathbf{Z}))$

40. Prove that if $A \neq \varnothing$ is a finite set and if $B = \{f | f : \mathbf{N} \to A\}$, then B is uncountable. (Hint: recall Cantor's proof for the uncountability of the set $(0, 1)$.)

41. Let $(A \to B) = \{f | f : A \to B\}$
Show that if $|A_1| = |A_2|$ and $|B_1| = |B_2|$, then

$$|(A_1 \to B_1)| = |(A_2 \to B_2)|$$

42. Prove that the countable union of sets of cardinality $\mathbf{c} = 2^{\aleph_0}$ (continuum) again has cardinality $\mathbf{c}$.

43. Here is Cantor's Paradox: Consider the set of all sets. The set of all its subsets, according to Cantor's own theorem, has a cardinal number larger than the cardinal number of the original set. Yet our original set by definition includes **all** sets. Thus, we constructed a set larger than the set of all sets. Can you resolve this paradox?

44. Here again is Russell's famous paradox: Let S be the set that contains a set X. If the set X doesn't belong to itself, so $S = \{X | X \notin X\}$,
Show that

(i) the assumption that S is a member of S leads to a contradiction;

(ii) the assumption that S is not a member of S leads to a contradiction too.

45. Explain why there are no "holes" in **R.**

2

LOGIC

How wonderful that we have met with a paradox. Now we have some hope of making progress.

N. Bohr[1]

Logic is the daughter of Greece, as are democracy, tragedy, rhetoric, history, philosophy, and mathematics. … Like diamond, logic is pure, transparent, and also most impenetrable, capable of leaving its mark on everything.

R. Omnès[2]

I seem to have been only like a boy playing on the sea shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.

I. Newton[3]

[1]Niels Henrik David Bohr (1885–1962), Danish physicist as quoted in S. Hawking (ed.) *On the Shoulders of Giants*, Running Press, Philadelphia, 2002.
[2]Roland Omnès (1931–), French theoretical physicist and philosopher.
[3]Brewster, D., *Life of Sir Isaac Newton,* Nabu Press, 2010.

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

2.1 INTRODUCTION

After our short tour through Set Theory, it seems natural to ask why logic, which we are about to embark onto, comes *after* sets. Wouldn't be more "natural" to start with logic – often described as the analysis of reasoning – and then study set theory? After all, good logical/analytical thinking is a prerequisite for doing mathematics or any other science for that matter. Nobody doubts that we need logic, and we need language to think or do mathematics. But an equally good argument could be made: If "sets are everything," and if set theory is the underlining fabric of any language or any abstract structure, hence logic too, then logic should come *after* sets. As you can see this is a question of "the chicken and the egg." I was slightly more inclined to start with set theory, but I leave it to you to decide which is the chicken and which is the egg, as well as what comes first. One thing is sure, set theory is important for mathematical logic. It made us think twice before deciding to trust our intuition in deciding what is true or false. What could be more self-evident than that the whole is greater than its parts? Yet, as you recall, we found that this is not always the case. Another thing that we have learned from set theory was that even when we think "correctly" we can run into logical contradictions. In any case, and whatever your philosophical inclinations may be, the important thing is to accept the premise of both Plato and Aristotle that all philosophy begins in wonder. So wonder we will. What I want to do in this chapter is to introduce you to *Logic,* a branch of mathematics originated most notably by Aristotle[4] almost 2400 years ago. He was by no means the first mathematician, and certainly not the only one considering the ways to scientifically explain "rational thought," but he was the first to write a treatise on logic as a set of rules for deductive reasoning – rules that should be used in all serious discourses in philosophy and science. He famously started his *Metaphysics* with the words: *All men by nature desire to know.* (Talking about wonder!) Neglecting for the time being the nagging question whether this is indeed true (and we want it to be true), it is still one of my favorite quotes. Yes, *we* want to know, and we want to know how we know. How do we know that what we (think we) know is really true? Let me try to make it simpler. We study mathematics (among other reasons) to apply it to different disciplines of science, engineering, and philosophy, to correctly design, correctly and functionally construct, to *understand,* to determine what works and what doesn't, and to decide what is true and what is false. But equally importantly we want to be able to *explain.* Indeed, our best understandings of "things," the best theories of the nature of things, are those that embody explanatory features. In particular, those that are superior, more powerful, more general than those we previously had. It is those that make us eventually

[4]Aristotle's logic (384 BC to 322 BC) was not the only one in ancient Greece. Stoics developed their own logic following the logic of Magarians, which itself was developed from the philosophy of Parmenides (515–450(?) BC), and Zeno (ca. 490 BC to 430 BC). Chrysippus (280 BC to 207 BC), for instance, came to be renowned as one of the foremost logicians of ancient Greece, superior to Aristotle himself.

understand our understanding. So, one hopes that the safest avenue to achieve this goal is

$$\text{mathematics/logic} \rightarrow \text{science} \rightarrow \text{philosophy}$$

which doesn't exclude a possible two-way approach

$$\text{mathematics/logic} \rightleftharpoons \text{science} \rightleftharpoons \text{philosophy}$$

If it is true that, according to Immanuel Kant, mathematics is determined by the form of our pure intuition so that it is impossible to imagine anything violating mathematics, we want to make sure that our mathematics is also correct. But how do we decide that the mathematics we are using is "right?" How do we know it is true? Where does mathematical certainty come from? What instruments, what "mathematics" should we use to "examine" mathematics? How do we know that the language we use to express our thoughts we use properly? Consider the following: Suppose I say:

(a) *Socrates is mortal*,
or I say:
(b) *Socrates is not mortal.*

Without knowing who or what "*Socrates*" is, logic tells us that only one of the statements (a) or (b) can be true. This is what Aristotle deemed the *Principle* (or *Law*) *of Excluded Middle.* Moreover, keeping the same principle in mind, instead of "*Socrates is mortal*," I could have said "*Xanthippe's husband will one day cease to exist*," or we could have said "*Lamprocles*' *Sophroniscus*' *and Menexenus*' *father will one day cease to exist*." Not a single word of which appears in the first sentence, but all three sentences mean the same thing. It seems that language is inseparable of logic, but logic permeates language at a more fundamental structural level.

Moving on, suppose I say:

All men are mortal.
Socrates is a man.
(Therefore) Socrates is mortal.

This syllogism[5] seems pretty reasonable, right? Can we, then, consider the following as equally valid?

All X are *Y*.
A is X.
(Therefore) A is Y

[5]This syllogism is attributed to Aristotle.

Sure, you say, it's very *logical*! How could it be otherwise? Besides, you continue, didn't we prove while discussing Set Theory in Chapter 1 that

$$\forall a \in X \text{ if } X \subseteq Y \text{ then } a \in Y.$$

In other words, if a has property X, and having property X means also having property Y, then necessarily, a has property Y too. Indeed!

Now, what about the following:

All men are mortal.
Socrates is mortal.
(Therefore) Socrates is a man.

Somehow this doesn't sound right, right?
Consider one more often quoted quip:

Nothing is better than eternal happiness.
A ham sandwich is better than nothing.
(Therefore) A ham sandwich is better than eternal happiness.

Regardless what you think of eternal happiness, the above syllogism sounds awkward to say the least. Formally though, everything is same as before.

How about this one:

$$1 + 1 \text{ contains a plus sign,}$$
$$1 + 1 = 2$$

(Therefore) 2 contains a plus sign.

Even a sentence from everyday parlance may cause consternation in a picky mathematician. Consider the following statement:

War is never the answer!

"But that's not true," exclaims a mathematician. "To the contrary, war most certainly is the answer, if the question is '*What is a three-letter word for organized armed conflict?*'"

So, as you can see from these simple examples, expressing one's thoughts correctly and soundly is not such a trivial matter. So, unlike "traditional mathematics" (for lack of better term), which does not make its method of reasoning or its language an object of study, mathematical logic considers language and the method of reasoning as a main object of study. This immediately leads us to the next thing: How do we *prove* that what we are saying is indeed so? How can we distinguish true proofs from the "proofs" that are false? What exactly are the rules

of a proof? (This is not some game of words! Think about it! This is another one of those things that literally grows on you. It is the essence of all philosophy.) It is a product of a million years of evolution, and as the reader of these words you are a member of the *Homo sapiens* species, the one who thinks about his thinking. In a word, what is the nature of mathematics? Thus, do we need some kind of mathematics of mathematics – metamathematics? As with metaphysics,[6] people sometimes associate with metamathematics a rather pejorative connotation, but it is a very serious (if seriously done) science, a very useful science, and definitely a very interesting one. Bertrand Russell and Alfred North Whitehead thought so too, and they set to write a book *Principia Mathematica* (note the (intentional?) similarity in title with Newton's *Philosophiæ Naturalis Principia Mathematica*).[7] The goal was to construct a rigorous logical structure, a formalism, on which all mathematics could be built – consistent, complete, and decidable. If after reading the first chapter, you are ready to accept the claim that mathematics *is* based on axiomatic set theory, then Russell's *Principia* could be considered a counterpart reducing all of mathematics to logic. In either case, the claim that mathematics is the study of abstract structures still holds. You are bound to ask now (unless you already wondered about this while reading Chapter 1), what do we mean by "axiom" or a "formal axiomatic method?" Well, there are essentially two methods that transformed logic into an exciting, mathematically and philosophically interesting discipline – the *algebraic method* and the *formal axiomatic method*. Suppose we want to study some (deductive) science S from a logical point of view, we need to consider two things: (i) The set of axioms[8] for the theory S and (ii) the procedures of logical inference by which the theorems of S can be derived. On the other hand, with the "algebraic method" we treat mathematical entities and the relations between them as mathematical symbols following the rules of "mathematics," that is, we manipulate the symbols according to the rules we have previously established. Of course, in the process of these "manipulations" new rules can be discovered and lead us to new discoveries.[9] Furthermore, "operations" themselves can be then treated as objects to be manipulated with/operated on, and we get more and more intricate structure of the whole new mathematics universe opened up for us.

Axioms and the "axiomatic method" are somewhat similar to the method Euclid[10] used. The paradigm of such an approach is to discover a finite set

[6] As Kant lamented: "*There was a time when metaphysics was called the queen of all sciences … Now … the queen proves despised on all sides*."

[7] Would you believe me if I told you that the title of *this* book is purely coincidental?

[8] By axioms we mean the set of propositions whose truth is assumed self-evident from the beginning. (Postulates, on the other hand, we assume to be true by convention.)

[9] All of this may look confusing to you at this point – don't worry, in due time you will develop a "sense" for it, and things will become very "natural" to you. One more comment I want to make at this point to the reader who is familiar with the following popular adage: All mathematicians are closeted Platonists. Well, that's not exactly true, I for one, am not – closeted.

[10] Euclid of Alexandria, not Euclid of Megara.

of fundamental principles (axioms) from which all infinitely many truths of mathematics can be derived. Of course, over more than 2300 years after Euclid, much of the refinements have been done by the luminaries on whose shoulders we are standing today. The formal axiomatic method freed us from logical inconsistencies when dealing with entities difficult or even impossible to define (recall, for instance, the "definition" of a set in Chapter 1). The beauty of axioms and the axiomatic method, one may argue, is that in one sense they are "free of content," and, on the other hand, they are about any set of entities that make the axioms true. Put differently, axioms determine to which set of entities the undefined terms apply. As you might have anticipated, all of this elevates the whole discussion on the nature of mathematics on a higher plane. So, do you see where we are going? Can you intuit the "big" picture?

In order to raise everything to a highly abstract level and with constant concern for precision and rigor, we need a language – a formal language – into which we can translate English (French, German, Chinese, or any other natural language) sentences. This language may not be as expressive as a natural language but that's the price we have to pay for the precision we desire. We also need a precise formation rules, a precise "grammar," so that no ambiguities can occur in our discourses. Although we can, in principle, have an infinite sequence of objects, which we will call symbols, there is only a finite number of logical symbols. Many of them we have already met, but nevertheless let's list some of them again:

The five symbols we use in propositional logic $\sim, \wedge, \vee, \rightarrow, \leftrightarrow$ (***not, and, or, if … then, if and only if***, respectively[11]) are usually called *sentential connective symbols*, which together with the usual parentheses, (a) we call *logical symbols.* In translating to and from English their role never changes. Sentences, on the other hand, will be designated by various symbols: $a, b, c, \ldots, p, q, r, \ldots$ or $A, B, C, \ldots, P, Q, R, \ldots$ or $\alpha, \beta, \gamma, \ldots$. They are *nonlogical symbols* and we call them *parameters* or *variables.* So our formal language $\mathcal{L}$ consists of the following:

(i) **Alphabet**, that is, the set of logical and nonlogical symbols

$$A = \{\sim, \wedge, \vee, \rightarrow, \leftrightarrow (,), p, q, r, \ldots\}$$

(ii) A set of **formulae**, that is, strings of symbols from A

(iii) A "**grammar,**" the set of rules of inference, that is, the rules of how to manipulate symbols in order to obtain "grammatically correct" statements, that is **well-formed-formulas (wffs)**

(iv) A set of **axioms.**

[11] These connectives are most likely due to Chrysippus.

The rules for wffs are as follows:

1. Any sentence symbol is a wff.
2. If ϕ is a well-formed formula, then so is $\sim \phi$.
3. If ϕ and ψ are well-formed formula, then so are $(\phi \wedge \psi), (\phi \vee \psi), (\phi \rightarrow \psi)$, and $(\phi \leftrightarrow \psi)$.

Example 2.1

(i) $\sim\sim\sim p$ is a wff since p is a wff by rule (1) to which we apply rule (2) three times.

(ii) $((\rightarrow p$ *is not* a wff.

(iii) $\sim (p \rightarrow q)$ *is* a wff.

(iv) *pq is not* a wff. Two sentence symbols can produce a wff only if rule (3) is satisfied. ■

Example 2.2 Suppose we are given a sequence

$$((\sim p) \rightarrow (p \wedge q)) \vee (p \wedge (q \leftrightarrow r)) \qquad (*)$$

Is it a wff?

Solution Let us check our sequence against rules (1)–(3) of well-formed formulas. If we consider $((\sim p) \rightarrow (p \wedge q))$ as ϕ and $(p \wedge (q \leftrightarrow r))$ as ψ we see that our whole expression is of the form $(\phi \vee \psi)$. On the other hand, observe that ϕ has the form of $(\alpha \rightarrow \beta)$, where α is $(\sim p)$ and β is $(p \wedge q)$ Similarly, ψ is of the form $(p \wedge \gamma)$, where γ is $(q \leftrightarrow r)$. Thus, our sequence (*) is wff indeed. ■

> "Contrariwise," continued Tweedledee, "if it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's logic."
>
> Lewis Caroll, *Alice in Wonderland*

2.2 PROPOSITIONAL CALCULUS[12]

Now we are ready to start and, as always, we would like to start with some basic definitions. But to be consistent and maintain the necessary mathematical rigor rightfully expected of logic, it is high time to precisely state what we mean by a "definition."

Definition[13] A *definition* is a statement neither true nor false. The only requirement for a definition is that it needs to be consistent in itself. A definition could

[12]Propositional calculus is also called "propositional logic," "Boolean logic" or "sentential calculus."

[13]This definition is purposely un-numerated – it has a place in the previous chapter as well as in all the chapters that follow.

also be considered as a rule for identifying the entity defined regardless of its abstract appearance.

Next we ask: what is logic? As with sets we find it almost impossible to define logic. Following Russell we will think of the *logic of an argument as that which is left over when the meaning of the argument has been removed.* Formally,

Definition 2.1 Logic (λόγος (logos)) is the study of the principles, method, and validity in deductive reasoning, especially of the structure of propositions as distinguished from their content. Since logical analysis doesn't determine the intrinsic merit of an argument but only evaluates the argument's form in order to determine whether the truth of the conclusion follows necessarily from the truth of premises, logic is often referred to as a science of reasoning. We also have

Definition 2.1′ By the **logic** of an argument, we mean the abstracted form of that argument as a "thing" independent of the content of the argument. Or simply

Definition 2.1″ Logic is a science/a study that deals with the principles and criteria of the validity of inference and demonstration.

What Frege, Russell, Whitehead, and others wanted to do was to show that, once properly established, logic could be applied to statements about anything at all, mathematics in particular (in case, you consider mathematics different from logic). So when we follow the rules of mathematics/logic, the results obtained, to paraphrase G. H. Hardy,[14] *are there not because we think so, or because our minds are shaped in one way rather than the other, but because it is so, because mathematical reality is built that way*.

We ended the last chapter with a couple of classical paradoxes wishing to incite thinking about connections between abstract formalism of set theory and logic. So before further introducing the formalism of logic, let me start with another ancient, "difficult," and still mathematically relevant paradox – the **liar paradox,**[15] known to the ancients as the **pseudomenon**. Generally attributed to Epimenides,[16] the original version of the Liar's Paradox was actually devised in the fourth century BC by the Greek philosopher Eubulides.[17] It goes like this: Epimenides, a Cretan, reportedly stated:

The Cretans always lie.

Is this true or not? *A man says that he is lying. Is what he says true or false?* (Remember Russell's barber? Well, does he shave himself or not?) Or, how about

[14]Hardy, G. H., *A Mathematician's Apology,* Cambridge Univ. Press, 1967.
[15]St. Paul also mentions it: "One of themselves, a prophet of their own, said: 'Cretans are always liars, wily beasts, lazy gluttons.' This testimony is true." [Titus I:12-13].
[16]Epimenides of Knossos, (ca. 600 BC), Cretan philosopher.
[17]Eubulides of Miletus (ca. 400 BC), Greek philosopher.

this: "*This sentence is false*." Now, compare those with the following sentence due to N. Chomsky[18]:

Colorless green ideas sleep furiously.

It looks like a nice legitimate sentence, right? It has all the attributes of a "legitimate" sentence, subject, verb, and so on. Still, you immediately recognize its meaninglessness and reject it as it doesn't tell you anything. The "liar" sentence, however, is completely different. *All Cretans lie* – there is nothing meaningless here. We know exactly what Epimenides is saying, don't we? Well, the difficulty with the liar paradox is that irrespective of whether one considers the proposition true or false, it always leads to a contradiction. Thus, what is its meaning? What are we to conclude from it? Consider again "*This sentence is false*." Suppose it is true, that is, it says of itself that it is not true. But we assumed that what it claims is true, hence it cannot be true. Suppose, now, that it is not true. Then, it says of itself that it is not true, hence by the assumption it is true.

As you can see, our mind can come up with the statements, that regardless how simple or naïve they look at the first sight, they are definitely not trivial. How, then, we are to decide their truth value? Another example could go like this: Suppose I say:

Vladimir Lepetic cannot consistently prove that this statement is true.

I can certainly make the above statement – as I just did – but regardless how hard I try I cannot prove it. You, the reader, on the other hand, can immediately show that it is true. So how am I to consider a discipline as scientific if I cannot prove a proposition that is so obvious to everyone else? You can imagine now how complicated some issues in philosophy, mathematics, and the sciences can be. Of course, if and when the propositions like the ones above occur, we want to be able to decide on their validity. In mathematics, we want the rules of logic to give precise meaning to mathematical statements, so we can distinguish with no ambiguity between valid and invalid ones. Again, the best is to start simple by introducing the basic ingredients of our discourse. So, let $S = \{p, q, r, s, \dots\}$ be a (countable) set with p, q, r, and so on representing statements or propositional variables, or simply, propositions.

Definition 2.2 A **statement** or a **proposition** is a **declarative** sentence that is either true or false but not both. A statement can be expressed in words or in mathematical symbols.

NOTE: Some authors distinguish between a *statement* and a *proposition*. They consider the former as a declarative sentence for which a truth value cannot be

[18] Avram Noam Chomsky (1928–), American linguist, philosopher, cognitive scientist.

determined and the latter as one for which that can be done. For instance, "*x is an even integer*" would be a statement and not a proposition since we cannot assign a truth value to it without knowing what x is. In this book, however, we will use "*statement*" and "*proposition*" interchangeably and consider them as the same.

Example 2.3 Consider the following:

(i) *Washington, D.C. is the capital of the USA*
(ii) $2 \times 2 = 22$
(iii) *How are you?*
(iv) *x is an even number*
(v) $\pi = 3$

Note that

(i) is a proposition, which is true.
(ii) is a proposition, which is obviously false.
(iii) is not a proposition – there is no true–false value that we can associate with it.
(iv) is not a proposition because we cannot decide whether it is true or false since we don't know what x is.
(v) is a statement, which is false. ■

Example 2.4 Consider the following:

(i) $x > 2$ is not a proposition. Since variables don't have any meaning by themselves, we cannot decide whether it is true or false. We don't even know whether x is a number or something else in which case $x > 2$ would be sheer nonsense.
(ii) Similarly, $x + y = 12$. Although it is usually considered a formula (i.e., even as one habitually assumes that x and y are numbers), this is not a proposition for we don't know what numbers x and y stand for. ■

Example/Exercise 2.5 The sentence "*Is 5 an integer?*" can certainly be answered affirmatively, but is it a proposition or not?

Many statements, of course, are not that simple and, in fact, could be extremely complicated. They could be built up out of many other statements using various *logical connectives.* Thus, we have

Definition 2.3 A statement is said to be ***simple*** or ***atomic*** if it cannot be broken down into other statements (e.g., $2 \times 2 = 4$). A statement is ***composite*** or ***compound*** if it contains several simple statements connected by words/symbols *and, or*, *thus, then*, and so on.

Example 2.6 If we formally denote statement (i) from Example 2.3, with the letter p, statement (ii) with the letter q, and with the already mentioned symbol " $\wedge$ " to stand for "*and*," we can build a composite statement $p \wedge q$, which reads

$$\textit{Washington, D.C. is the capital of the USA } \textbf{\textit{and}} \ 2 \times 2 = 22 \tag{2.1}$$

As you can see, by using "p" and "q" as *statement variables* and the symbol "$\wedge$" to denote "*and,*" we formally/abstractly express statement (2.1) as

$$p \wedge q \tag{2.1$'$}$$

The advantage of this is that all statements of the form $p \wedge q$ can be treated equivalently regardless of the "content" of the variables p and q. Similarly, we introduce a symbol "$\vee$" to denote "*or*." So, instead of the compound statement (2.1) in the form (2.1$'$), we can construct another compound that reads

$$\textit{Washington, D.C. is the capital of the USA } \textbf{\textit{or}} \ 2 \times 2 = 22 \tag{2.2}$$

whose abstract form is

$$p \vee q \tag{2.2$'$}$$

Finally, we will use the symbol "$\sim$" to denote "*not*," or "*it is not the case that …* " Thus, if p represents the statement:

"*Washington, D.C. is the capital of the USA,*" then

$\sim p$ represents the statement:

$$\textbf{\textit{It is not the case that}} \ \textit{Washington, D.C. is the capital of the USA}^{19} \tag{2.3}$$

Definition 2.4 A **propositional function** $p(x)$ is a declarative sentence about one or more symbols $x, y, z, \ldots$, which becomes a proposition when a particular meaning is assigned to these symbols.

Definition 2.5 Given a propositional function $p(x)$, we say that **interpretation of *x*** is a *meaning* assigned to x for which $p(x)$ becomes a proposition.

[19] Obviously, in everyday parlance statement (2.3) would simply be read: Washington D.C. **is not** the capital of the USA. The use of "**It is not the case that**" to negate the statement can be useful in more complicated compounds where simple "not" may cause some confusion.

Example 2.7
Let $p(x)$: *x is an even integer*

be a propositional function. Then $x = 4$ is an interpretation that makes $p(x)$ a true statement, while $x =$ *red apple* is an interpretation that makes $p(x)$ a false statement.

As it is often critical in everyday life to determine whether something is true or not, it is not surprising that in mathematics deciding whether a proposition is true or false becomes the crux of the matter. ■

Definition 2.6 The truth or falsity of a statement we call the **truth value** of the statement.

Often in order to simplify the process of finding the truth values of a statement we use *Truth Tables*. Of course, we take it for granted henceforth that truth and falsehood are two different "things" without precisely defining what "truth" and "falsehood" are. In fact, it will never matter exactly what kind of "things" they are[20] but in a given context we find it convenient to identify truth with the symbol T and falsehood with F and refer to them as the *truth values* of a given statement. Thus,

Definition 2.7 **Truth Tables** are tables in which we list all possible truth values of statements we are considering.

The definitions of the compounds that follow can also serve as a good example of applications of truth tables.

Definition 2.8 If p is a statement/proposition variable, the **negation** of p is "not p" or "*It is not the case that p*" and we denote it by $\sim p$. **The truth values** are given in Table 2.1, where T stands for "true" and F for "false."

TABLE 2.1 Truth Table for $\sim p$

p	$\sim p$
T	F
F	T

Definition 2.9 (Conjunction) Let p and q be propositions. The proposition "p and q," the ***conjunction*** of p and q, denoted $p \wedge q$, is the statement that is true only when both p and q are true, otherwise it is false. The truth table for conjunction is given in Table 2.2.

[20]Of course "… it will never matter …" is in a sense a misstatement. Of course, we DO care what is true and what is false. After all, one of the most important philosophical questions is the question of "Truth." However, here, in mathematical logic, we are concerned only with "formalism," that is the way we must reason in order to communicate coherently with other humans. Questions of "content" within our formal inferences we leave for later.

TABLE 2.2 Truth Table for $p \wedge q$

p	q	$p \wedge q$
T	T	T
F	T	F
T	F	F
F	F	F

Example 2.8 Let proposition p be: "*Today is Monday*" and proposition q: "*I am studying logic*." Then conjunction $p \wedge q$ reads: *Today is Monday* ***and*** *I am studying logic*. ■

Example 2.9 Let S be a statement: π lies between 3 and 4 Write the statement S in abstract form, that is, as a conjunction of appropriate statements p and q.

Solution Let $p: \pi > 3$, and $q: \pi < 4$. Then,

$$S: p \wedge q \equiv (\pi > 3) \wedge (\pi < 4)$$

■

Definition 2.10 (Disjunction) Let p and q be propositions. The proposition "***p or q***", the ***disjunction*** of p and q, denoted $p \vee q$, is a proposition that is false only when both p and q are false, otherwise it is true. The truth table for disjunction is given in Table 2.3.

TABLE 2.3 Truth Table for $p \vee q$

p	q	$p \vee q$
T	T	T
F	T	T
T	F	T
F	F	F

Example 2.10 Let statements p and q be as in Example 2.8. Then, disjunction $p \vee q$ reads: *Today is Monday* ***or*** *I am studying logic.*

Now that conjunction, disjunction, and negations have been defined, we can consider more complicated expressions as different combinations of basic compounds described above. For instance, we may be interested in the truth values of (i) $\sim p \wedge q$, (ii) $p \wedge \sim q$, (iii) $\sim p \vee q$, (iv) $p \vee \sim q$, (v) $\sim (p \vee q)$, and so on. Often, we want to know the truth tables for even more complex statements.

So, we restate Definition 2.3 in more general terms by ■

Definition 2.11 A statement/propositional form is an expression made up of statement variables $p, q, r,$ and so on and logical connectives " $\sim$," "$\wedge$," "$\vee$," that becomes a proposition when actual statements are substituted for the component statement variable.

Example 2.11 Construct the truth table for the following statement: $(p \wedge \sim q) \vee \sim r$.

Solution See Table 2.4. ■

TABLE 2.4 Truth Table for $(p \wedge \sim q) \vee \sim r$

p	q	$\sim q$	r	$\sim r$	$p \wedge \sim q$	$(p \wedge \sim q) \vee \sim r$
T	T	F	T	F	F	F
F	T	F	T	F	F	F
T	F	T	T	F	T	T
T	T	F	F	T	F	T
F	T	F	F	T	F	T
F	F	T	T	F	F	F
T	F	T	F	T	T	T
F	F	T	F	T	F	T

Definition 2.12 (Exclusive Or) Let p and q be propositions. The **exclusive or** of p and $q,$ denoted $p \oplus q$, is the statement that is true when exactly either p or q is true and is false otherwise, that is

$$p \oplus q = (p \vee q) \wedge \sim (p \wedge q)$$

Here is the truth table (Table 2.5).

TABLE 2.5 Truth Table for $p \oplus q$

p	q	$p \vee q$	$p \wedge q$	$\sim (p \wedge q)$	$p \oplus q$
T	T	T	T	F	F
T	F	T	F	T	T
F	T	T	F	T	T
F	F	F	F	T	F

Observe that unlike *inclusive or*, the *exclusive or* is false when both p and q are true.

Definition 2.13 (Conditional/Implication*) Let p and q be two statements. We say that "**if p then q**" (or that "**p implies q**") is the **conditional** of q by p, and

we write: $p \rightarrow q$. We define the conditional as a statement that is false only when p (the antecedent) is true and q (the consequent) is false. Otherwise, it is true. The truth table for the conditional is given in Table 2.6.

TABLE 2.6 Truth Table for $p \rightarrow q$

p	q	$p \rightarrow q$
T	T	T
F	T	T
T	F	F
F	F	T

(*) There is some subtlety here that has to be addressed. Strictly speaking, *conditional* **is not** the same as *implication* although it may seem so. That was known to the Stoics, and to their credit they recognized that the conditional has a clear or complete truth pattern only in two cases: when p is true and q is true, and when p is true and q is false. The rest has to be defined. And that's indeed what we do today. We say that conditional is true in all cases except when the antecedent is true and the consequent is false, which brings us to the issue of *conditional* and *implication.* Consider the following example:

If an apple *is a fruit then there are infinitely many primes.*

This conditional is true since both the antecedent and the consequent are true. However, the infinitude of primes has nothing to do with an apple being a fruit, that is, an apple being a fruit doesn't imply infinitude of primes. With the above observation in mind, we list a few more ways of stating the conditional (implication):

(i) *p implies q*
(ii) *q if p*
(iii) *p only if q*
(iv) *p is sufficient condition for q*
(v) *q is a necessary condition for p*

Example 2.12 Consider the statement

S: If 3 is prime greater than 2 then 3 is odd.

Take for *p:* "*3 is prime greater than 2*," and for *q*: "*3 is odd*." Formally, our statement *S* reads:

S: $p \to q$ (if T *then* T; this is the first row in the table above.)

Thus, the statement S is true.

If our statement is

S: *If* 9 *is prime greater than* 2, *then* 9 *is odd*

our conditional, formally written, is

S: $p \to q$ (if F *then* T; this is the second row in our table above.)

thus, the statement q is true.

Finally, consider

S: *If* 4 *is a prime greater than* 2 *then* 4 *is odd*,

that is

S: $p \to q$ (if F *then* F; this is the fourth row in our table above.)

the statement S is true again.

However, if the statement is

S: *If* 3 *is greater than* 2 *then* 4 *is prime*

that is

S: $p \to q$ (if T *then* F; this is the third row in our table.)

thus, the statement S is false. ■

Example 2.13 We learned in Chapter 1 that for any set A, $\varnothing \subseteq A$ (see Theorem 1.1). The proof of this theorem, in terms of a conditional statement, goes as follows:

Let A be any set, and let x be any object. Consider the conditional statement:

$$S\colon (x \in \varnothing) \to (x \in A)$$

By definition of the empty set, the antecedent of S is false, and, therefore, the statement S is true, that is $\varnothing \subseteq A$. ■

Example 2.14 Let S: $p \rightarrow q$ be a statement:

$$\textit{If } a = b \textit{ then } a^2 = b^2, \quad a, b \in \mathbf{R}$$

Write a few alternative versions of the statement S.

Solution

(i) $a = b$ *implies* $a^2 = b^2$
(ii) $a^2 = b^2$ *if* $a = b$ ■

Example/Exercise 2.15 Determine the truth value of the statement

$$S\text{: } \textit{If } 2 \cdot 2 = 22 \textit{ then } 4 \cdot 4 = 44$$

Definition 2.14 (Biconditional) Let p and q be propositions. The **biconditional of p and q**, "*p if and only if q,*" denoted $p \leftrightarrow q$, is the statement that is true when p and q have the same truth values and is false otherwise. The biconditional has Table 2.7.

TABLE 2.7 Truth Table for $p \leftrightarrow q$

p	p	$p \leftrightarrow q$
T	T	T
F	T	F
T	F	F
F	F	T

Definition 2.15 (Tautology) A **tautology** is a statement that is always true regardless of the truth value of propositions that occur in it.

Definition 2.16 (Contradiction) A **contradiction** is a statement that is always false regardless of the truth values of the propositions that occur in it.

Example 2.16 Table 2.8 gives simple examples of tautology (see also Theorem 2.2) and contradiction:

TABLE 2.8 Tautology and Contradiction

p	$\sim p$	$p \vee \sim p$	$p \wedge \sim p$
T	F	T	F
F	T	T	F

A tautology in the third column of the above table dates back to the time of Aristotle and was considered as very self-evident for a long time. ■

Example 2.17 Show that a statement S: $(p \wedge (p \to q)) \to q$ is a tautology.

Solution To prove this, we construct Table 2.9 for the statement S:

TABLE 2.9 Truth Table for $(p \wedge (p \to q)) \to q$

p	q	$p \to q$	$p \wedge (p \to q)$	$(p \wedge (p \to q)) \to q$
T	T	T	T	T
F	T	T	F	T
T	F	F	F	T
F	F	T	F	T

Since all entries in the last column of our table are "T", that is, the statement S is true for any combination of truth values of atomic statements p and q The statement S is a tautology. ■

Example 2.18 Show that a statement $S: \sim ((p \wedge q) \to q)$ is a contradiction.

Solution We show this again by constructing the table Table 2.10 for the statement S:

TABLE 2.10 Truth Table for $\sim ((p \wedge q) \to q)$

p	q	$(p \wedge q)$	$\sim ((p \wedge q) \to q)$
T	T	T	F
F	T	F	F
T	F	F	F
F	F	F	F

So, the statement S is a contradiction. ■

Now we can adequately define the implication.

Definition 2.17 We say that p implies q, and we write $p \Rightarrow q$, if the conditional $p \to q$ is a tautology.

The above definition is actually a consequence of

Theorem 2.1 Let P and Q be two statements (compound or atomic). We say that P implies Q, and we write $P \Rightarrow Q$, if and only if $P \to Q$ is a tautology.

TABLE 2.11 Truth Table for $((\sim p) \rightarrow (p \wedge q)) \vee (p \wedge (q \leftrightarrow r))$

p	q	r	$\sim p$	$p \wedge q$	$q \leftrightarrow r$	$\sim p \rightarrow (p \wedge q)$	$p \wedge (q \leftrightarrow r)$	$((\sim p) \rightarrow (p \wedge q)) \vee (p \wedge (q \leftrightarrow r))$
T	T	T	F	T	T	T	T	T
F	T	T	T	F	T	F	F	F
T	F	T	F	F	F	T	F	T
T	T	F	F	T	F	F	F	F
T	F	F	T	F	T	F	T	T
F	F	T	T	F	F	F	F	F
F	F	F	T	F	T	F	F	F

Proof First, assume that P logically implies Q, then every truth assignment that makes P true also makes Q true. Thus, no truth assignment makes P true and Q false and therefore no truth assignment makes $(P \rightarrow Q)$ false, that is $(P \rightarrow Q)$ is always true. In other words, $(P \rightarrow Q)$ is a tautology. Conversely, assume that q. is a tautology. Then, for every truth assignment S, it is not the case that P is true and Q is false. In other words, every truth assignment that makes P true makes Q also true, that is, P implies Q. ■

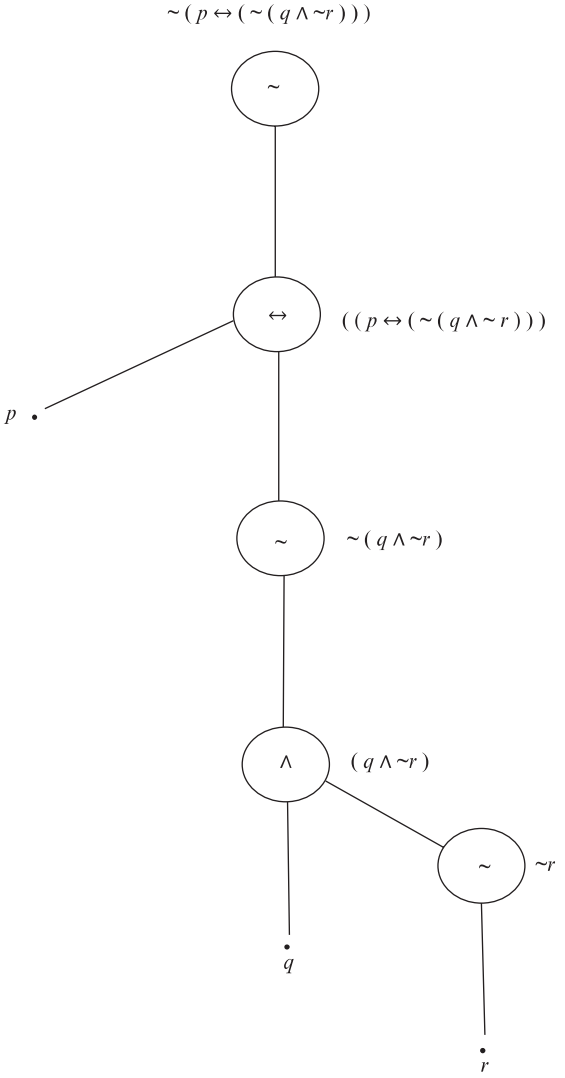
This may be the right time to revisit Example 2.2 and consider it from another angle. First, recall, we establish the fact that sequence (*) is wff. That is, wff of propositional logic is more than just a meaningless collection of symbols. It gives us a set of logical facts about the subject the atomic statements pertain to. We also know how to design a corresponding truth table. In other words, we know how to determine whether a given logical compound is true or false based on our knowledge of truth or falseness of the atomic statements. The way in which this is done defines the *semantics* of propositional calculus. The parsing of a wff can be represented by an "upside down tree," the **ancestral tree,** which we design by putting the root at the top and let it branch downward. At each branching node, we put a "descendant" of a particular "type" (i.e., a logical connective), so that a node of "∼" has one descendant and all other nodes have two. Finally, we end up with the leaves at the bottom, each labeled with a corresponding propositional (atomic) variable p_i. So, the tree of parsing the formula (*) from Example 2.2 is shown in Figure 2.1.

Since the semantics of propositional logic gives us the rules of determining the truth table of any formula, in this particular case, the truth table is Table 2.11.

Example 2.19 Let's consider a statement

$$\sim (p \leftrightarrow \sim (q \wedge \sim r))$$

It is a wff. Its tree structure is



■

And the corresponding truth table is Table 2.12.

Example 2.20 Construct the truth table and the ancestral tree for the statement

$$(p \leftrightarrow q) \rightarrow (\sim p \ \wedge \ q)$$

Solution See Table 2.13. ■

Construction of the ancestral tree is left, as a simple exercise, for the reader.

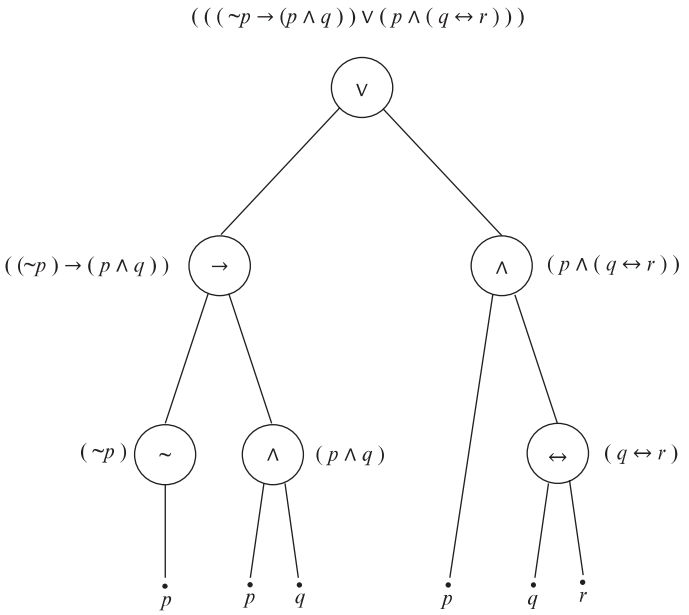


Figure 2.1 A parsing (ancestral) tree for a formula

TABLE 2.12 Truth Table for $\sim (p \leftrightarrow (\sim(q \wedge \sim r)))$

p	q	r	$\sim r$	$q \wedge \sim r$	$\sim (q \wedge \sim r)$	$(p \leftrightarrow (\sim (q \wedge \sim r))$	$\sim (p \leftrightarrow (\sim (q \wedge \sim r)))$
T	T	T	F	F	T	T	F
F	T	T	F	F	T	F	T
T	F	T	F	F	F	F	T
T	T	F	T	T	F	F	T
F	F	T	F	F	T	F	T
F	T	F	T	F	T	F	T
T	F	F	T	T	F	F	T
F	F	F	T	F	T	F	T

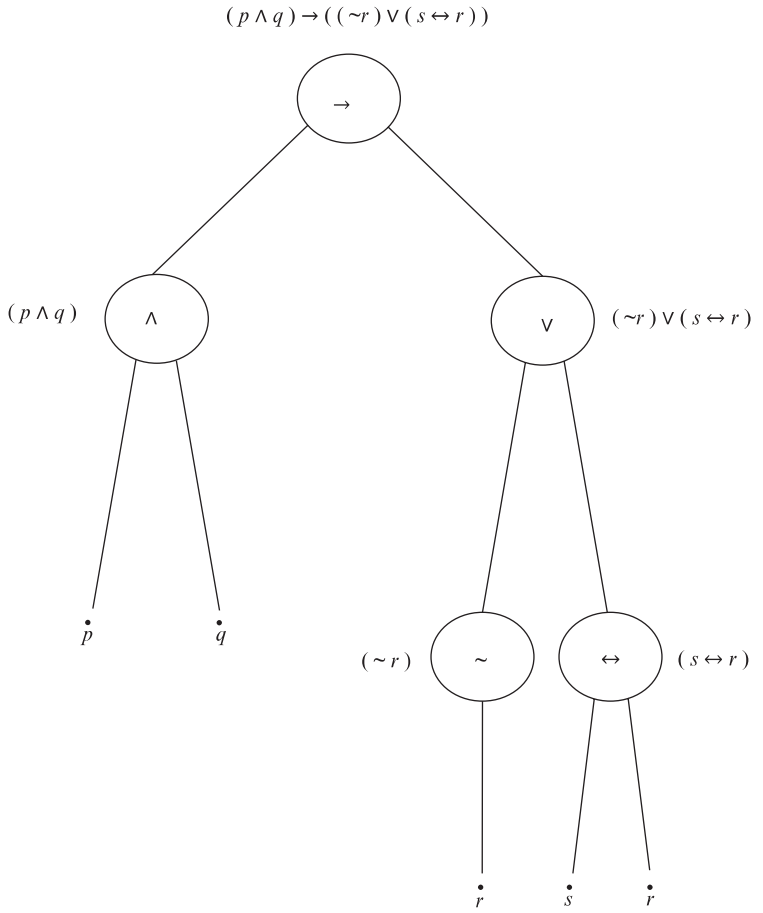
TABLE 2.13 Truth Table for $(p \leftrightarrow q) \rightarrow (\sim p \wedge q)$

p	q	$p \leftrightarrow q$	$\sim p$	$\sim p \wedge q$	$(p \leftrightarrow q) \rightarrow (\sim p \wedge q)$
T	T	T	F	F	F
F	T	F	T	T	T
T	F	F	F	F	T
F	F	T	T	F	F

Example 2.21 Construct the ancestral tree and the truth table for the statement

Solution

$$(p \wedge q) \rightarrow ((\sim r) \vee (s \leftrightarrow r))$$



Construction of the truth table is left as a simple exercise for the reader. ■

Theorem 2.2 (The Law of Excluded Middle) If p is a proposition, then the statement $(p \vee \sim p)$ is always true regardless of the truth value of proposition p.

Proof See Example 2.16. ■

It is worth noting that there are schools of thought in mathematics that take the above theorem with more caution. *Intuitionists*, for example, claim that the Law of Excluded Middle cannot be used without restrictions when dealing with infinitely many objects. "Infinities" do "behave" strangely, or at least they behave in their own specific way, as we have seen in Chapter 1. (Recall, for instance, $\aleph_0 + \aleph_0 = \aleph_0$.)

Sometimes, two or more statements that look different may in fact be equivalent. Usually, one wants to choose the one that is formally easier to deal with. For example, the statements

$$P\colon\ A \setminus B = \varnothing \quad \text{and} \quad Q\colon A \subseteq B$$

are equivalent. Indeed,

$$\begin{aligned}
A \setminus B = \varnothing \quad &\text{is equivalent to} \quad \sim \exists x\ (x \in A \ \wedge\ x \notin B)\\
&\text{which is equivalent to} \quad \forall x\ \sim (x \in A\ \wedge\ x \notin B)\\
&\text{which is equivalent to} \quad \forall x\ (x \notin A\ \vee x \in B)\\
&\text{which is equivalent to} \quad \forall x\ ((x \in A) \rightarrow (x \in B))\\
&\text{which is equivalent to} \quad A \subseteq B
\end{aligned}$$

So we have

Definition 2.18 (Logical Equivalence) We say that two statements P and Q are **logically equivalent** if and only if they have identical truth values for each possible substitution of statements for their statement variables, and we write $P \equiv Q$.

Definition 2.18′ (Logical Equivalence) Two statements P and Q are logically equivalent if and only if $P \leftrightarrow Q$ is a tautology.

This definition can be restated as

Theorem 2.3 Let P and Q be two statements (compound or atomic). We say that P is equivalent to Q, and we write $P \Leftrightarrow Q$ (or $P \equiv Q$) if and only if the biconditional $(P \leftrightarrow Q)$ is a tautology.

Proof $(P \leftrightarrow Q)$ is a tautology if and only if every truth assignment makes $(P \leftrightarrow Q)$ true, which is equivalent to saying that every truth assignment gives P and Q the same truth value, that is, P and Q are logically equivalent.

A rather trivial example of two equivalent statements is (Table 2.14)

$$p \equiv \sim (\sim p)$$

Indeed, ■

TABLE 2.14

p	$\sim p$	$\sim (\sim p)$
T	F	T
F	T	F

The following are examples of two important equivalencies:

Theorem 2.4 (DeMorgan's Laws)

1. $\sim (p \wedge q) \equiv \sim p \vee \sim q$
2. $\sim (p \vee q) \equiv \sim p \wedge \sim q$

Proof Proof for (1): Table 2.15.

Proof for (2) is analogous. ■

TABLE 2.15 Table for $\sim (p \wedge q) \equiv \sim p \vee \sim q$

p	q	$\sim p$	$\sim q$	$p \wedge q$	$\sim (p \wedge q)$	$\sim p \vee \sim q$
T	T	F	F	T	F	F
F	T	T	F	F	T	T
T	F	F	T	F	T	T
F	F	T	T	F	T	T

It might be important to stress again that

(i) $\sim (p \wedge q) \not\equiv \sim p \wedge \sim q$

(ii) $\sim (p \vee q) \not\equiv \sim p \vee \sim q$

With everything said so far it shouldn't be difficult to prove the following:

Theorem 2.5 Let p, q, and r be any statements, and let τ be a tautology and c a contradiction. Then, the following logical equivalences hold:

1. (i) $p \wedge q \equiv q \wedge p$
   (ii) $p \vee q \equiv q \vee p$
2. (i) $p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$
   (ii) $p \vee (q \vee r) \equiv (p \vee q) \vee r$
3. (i) $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$
   (ii) $p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$
4. (i) $p \wedge \tau \equiv p$
   (ii) $p \vee c \equiv p$
5. (i) $p \vee \sim p \equiv \tau$
   (ii) $p \wedge \sim p \equiv c$
6. (i) $p \vee \tau \equiv \tau$
   (ii) $p \wedge c \equiv c$
7. (i) $p \wedge p \equiv p$
   (ii) $p \vee p \equiv p$
8. (i) $\sim \tau \equiv c$
   (ii) $\sim c \equiv \tau$
9. (i) $p \wedge (p \vee q) \equiv p$
   (ii) $p \vee (p \wedge q) \equiv p$

Example 2.22 Write the negation of the following statement S: *The train is late or my watch is broken.*

Solution Let p be: *The train is late,* and q: *My watch is broken*. So,

$$S: p \vee q$$

Therefore, $\sim S : \sim (p \vee q) \equiv \sim p \wedge \sim q$: *The train is not late and my watch is not broken*.

■

Example/Exercise 2.23 State the negation of the following statement S: *4 is an even number and 4 is not an even number.*

Example 2.24 Prove the following logical equivalence:

$$\sim (\sim p \ \wedge \ q) \ \wedge \ (p \ \vee \ q) \equiv p$$

Solution

$$\begin{aligned} \sim (\sim p \ \wedge \ q) \ \wedge \ (p \ \vee \ q) &\equiv (p \vee \sim q) \ \wedge \ (p \ \vee \ q) \\ &\equiv p \vee (\sim q \ \wedge \ q) \\ &\equiv p \ \vee c \\ &\equiv p \end{aligned}$$

■

Example 2.25 Prove the following logical equivalence:

$$\sim (p \ \vee (\sim p \ \wedge \ q)) \equiv \sim p \wedge \ \sim q$$

Solution

$$\begin{aligned} \sim (p \ \vee (\sim p \ \wedge \ q) &\equiv \sim p \wedge \ \sim (\sim p \ \wedge \ q) \\ &\equiv \sim p \ \wedge \ (p \vee \sim \ q) \\ &\equiv (\sim p \ \wedge \ p) \ \vee (\sim p \wedge \ \sim \ q) \\ &\equiv c \ \vee (\sim p \wedge \ \sim \ q) \\ &\equiv \sim \ p \wedge \ \sim q \end{aligned}$$

■

Theorem 2.6 Let p and q be propositions, then

$$p \rightarrow q \equiv \sim p \ \vee \ q$$

Proof See Table 2.16. ■

TABLE 2.16 Table for $p \rightarrow q \equiv \sim p \ \vee \ q$

p	$\sim p$	q	$p \rightarrow q$	$\sim p \ \vee \ q$
T	F	T	T	T
F	T	T	T	T
T	F	F	F	F
F	T	F	T	T

Example 2.26 Write the negation of the following statement S: *If you live in Florida, then you live in Miami.*

Solution Let p be: "*You live in Florida*," and q: "*You live in Miami*." Our statement is formalized as follows:

$$S: p \rightarrow q$$

Therefore,

$$\begin{aligned} \sim S : \sim (p \rightarrow q) &\equiv \sim (\sim p \vee q) \\ &\equiv \sim (\sim p) \wedge \sim q \\ &\equiv p \wedge q : \textit{You live in Florida and you don't live in Miami} \end{aligned}$$

■

Example/Exercise 2.27 Prove that

$$(p \rightarrow q) = \sim (p \wedge (\sim q))$$

Theorem 2.7 Let p and q be propositions, then

$$p \leftrightarrow q \equiv (p \rightarrow q) \wedge (q \rightarrow q)$$

Proof See Table 2.17

TABLE 2.17 Truth Table for $(p \rightarrow q) \wedge (q \rightarrow p)$

p	q	$p \rightarrow q$	$q \rightarrow p$	$p \leftrightarrow q$	$(p \rightarrow q) \wedge (q \rightarrow p)$
T	T	T	T	T	T
F	T	T	F	F	F
T	F	F	T	F	F
F	F	T	T	T	T

■

Example 2.28 Let p and q be propositions. Show that the following statement is true:

$$\sim (p \leftrightarrow q) \equiv (p \wedge \sim q) \vee (q \wedge \sim p)$$

Solution

$$\begin{aligned}\sim (p \leftrightarrow q) &\equiv \sim [(p \rightarrow q) \wedge (q \rightarrow p)] \\ &\equiv \sim (p \rightarrow q) \vee \sim (q \rightarrow p) \\ &\equiv \sim (\sim p \vee q) \vee \sim (\sim q \vee p) \\ &\equiv [\sim (\sim p) \wedge \sim q] \vee [\sim (\sim q) \wedge \sim p] \\ &\equiv (p \wedge \sim q) \vee (q \wedge \sim p)\end{aligned}$$

■

Example/Exercise 2.29 Let p, q, and r be propositions. Show that the following statements are tautologies:

(i) $\sim (p \rightarrow q) \leftrightarrow (p \wedge \sim q)$
(ii) $\sim (p \wedge q) \leftrightarrow (\sim p \vee \sim q)$

Example/Exercise 2.30 Let p and q be propositions. Show that

$$p \leftrightarrow [\sim p \rightarrow (q \wedge \sim q)]$$

is a contradiction.

Theorem 2.8 (Law of simplification) Let p and q be statements. Then,

$$(p \wedge q) \rightarrow q$$

is a tautology.

Proof See Table 2.18.

TABLE 2.18 Table for $(p \wedge q) \rightarrow q$

p	p	$p \wedge q$	$(p \wedge q) \rightarrow q$
T	T	T	T
F	T	F	T
T	F	F	T
F	F	F	T

So, $(p \wedge q) \rightarrow q$ is indeed a tautology. ■

Theorem 2.9 Let p, q, and r be any two propositions. Then the statement

$$[(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r)$$

is a tautology.

Proof The tautology is evident from Table 2.19. ■

TABLE 2.19 Table for $[(p \to q) \wedge (q \to r)] \to (p \to r)$

p	q	r	$p \to q$	$q \to r$	$p \to r$	$(p \to q) \wedge (q \to r)$	$[(p \to q) \wedge (q \to r)] \to (p \to r)$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	T
T	F	T	F	T	T	F	T
T	F	F	F	T	F	F	T
F	T	T	T	T	T	T	T
F	T	F	T	F	T	F	T
F	F	T	T	T	T	T	T
F	F	F	T	T	T	T	T

Example/Exercise 2.31 Let Q be a proposition. Convince yourself that in order to prove that Q is a true statement it suffices to do the following two things:

(i) Find a proposition P such that $P \to Q$ is true statement.
(ii) Show that P is a true statement.

If you had difficulty with the previous example, note that if you take P to be $(p \to q) \wedge (q \to r)$, then $P \to Q$ is true regardless of the truth value of p.

Now you should be able to prove the following:

Theorem 2.10 (*Reductio ad absurdum*) With p, q and r being propositions, the following statement is a tautology regardless of the truth values of p, q, and r.

$$[\sim (p \to q) \to (r \wedge \sim r)] \to (p \to q)$$

Example/Exercise 2.32 Let p, q, and r be propositions. Show that the following statement is a tautology:

$$[(p \wedge \sim q) \to (r \wedge \sim r)] \to (p \to q)$$

Example 2.33 Prove:

$$\forall\, a, b \in \mathbf{R} \text{ if } a = b \text{ then } a^2 = b^2$$

Solution Suppose $a = b$, and $a^2 \neq b^2$ Then, $a = b$ implies $a - b = 0$ and $a^2 \neq b^2$ implies $a^2 - b^2 \neq 0$. Thus, we have

$$(a - b)(a + b) \neq 0$$

and therefore $(a - b) \neq 0$. Our supposition $a = b$ and $a^2 \neq b^2$ gave us a contradiction $(a - b) = 0$ and $(a - b) \neq 0$. Hence, $a = b$ implies $a^2 \neq b^2$. ■

Definition 2.19 Let p and q be propositions and $p \rightarrow q$ the corresponding conditional, then

(i) the **converse** is a statement of the form: $q \rightarrow p$,
(ii) the **inverse** is a statement of the form: $\sim p \rightarrow \sim q$, and
(iii) the **contrapositive** is a statement of the form: $\sim p \rightarrow \sim q$.

Theorem 2.11 The conditional and its contrapositive are logically equivalent.

Proof

$$\begin{aligned} \sim q \rightarrow \sim p &\equiv \sim (\sim q) \vee \sim p \\ &\equiv q \vee \sim p \\ &\equiv \sim p \vee q \\ &\equiv p \rightarrow q \end{aligned}$$

■

Example/Exercise 2.34 Prove Theorem 2.9 by writing out the truth table for the respective conditionals.

Example/Exercise 2.35

(i) Show that a conditional statement is not logically equivalent to its converse.
(ii) Show that a conditional statement is not logically equivalent to its inverse.

Example 2.36 Write the inverse, converse, and contrapositive for the following statements:

(i) If n is prime, then n is odd or n is 2.
(ii) If n is an even integer, then n is the sum of two primes.

Solution Let p be the statement: "n is prime," q: "n is odd," and r: "n *is* 2." Then, our statement formally looks as follows:

$$S\colon\ p \to (q \vee r)$$

Thus, inverse: $\sim p \to \sim (q \vee r)$ reads:

If n is not prime, then it is not the case that n is odd or n is 2. We can also say

$$\sim p \to \sim (q \vee r) \equiv \sim p \to (\sim q \wedge \sim r),$$

that is, *if n is not prime, then n is not odd and n is not 2.*

However, the following is also true:

$$\begin{aligned} \sim p \to \sim (q \vee r) &\equiv \sim (\sim p) \vee \sim (q \vee r) \\ &\equiv p \vee (\sim q \wedge \sim r) \end{aligned}$$

n is prime or n is not odd and n is not 2.

Converse: $(q \vee r) \to p$ reads:

If n is odd or n is 2, then n is prime.

However,

$$\begin{aligned} (q \vee r) \to p &\equiv \sim (q \vee r) \vee p \\ &\equiv (\sim q \wedge \sim r) \vee p \end{aligned}$$

and therefore we say: *n is not odd and n is not 2, or n is prime.*

Contrapositive: $\sim (q \vee r) \to \sim p$: *If it is not the case that n is odd or n is 2, then n is not prime.* Or:

$$\sim (q \vee r) \to \sim p \equiv (\sim q \wedge \sim r) \to \sim p$$

If n is not odd and n is not 2, then n is not prime. ■

Example/Exercise 2.37 Use *reductio ad absurdum* to prove the contrapositive of the statement: For all $a, b \in \mathbf{R}$ if $a = b$, then $a^2 = b^2$.

Example/Exercise 2.38 Convince yourself that all the formulas in propositional calculus can be expressed by using only "$\sim$" and "$\to$" connectives. Hint: Show that

(i) $p \wedge q \equiv \sim (p \to (\sim q))$
(ii) $p \vee q \equiv (\sim p) \to q$
(iii) $p \leftrightarrow q \equiv \sim ((p \to q) \to (\sim (q \to p))$

Why is this thus? What is the reason for this thusness?

Artemus Ward[21]

2.3 ARGUMENTS I

Definition 2.20 An **argument** is a sequence of statements in which we distinguish **premises** as the statements preceding the final statement called the **conclusion.** In other words, by an argument, we mean a set of propositions, of which one is claimed to follow from the others.

Here is one of the most quoted examples:

Example 2.39

If Socrates is a human being, then Socrates is mortal;
Socrates is a human being;
$\therefore$ *Socrates is mortal.*

Using the formalism from the previous section, that is, putting p: "*Socrates is a human being*" and q: "*Socrates is mortal*," our argument in an abstract form looks as follows:

$$\begin{array}{l} p \rightarrow q \\ p \\ \therefore q \end{array}$$ ■

As you may intuit, the formalism above suggests that the argument is valid regardless of the statements substituted in place of variables p and q; as long as the premises are true, then the conclusion is true. Let's state this precisely:

Definition 2.21 We say that an ***argument form*** **is valid** if, no matter what particular statements are substituted for the variables in its premises, making all premises true, then the conclusion is also true.

Definition 2.21′ The ***argument form*** **is valid** iff there exists no case where all premises are true and the conclusion is false.

Definition 2.22 We say that an ***argument*** **is valid** if its form is valid.

Now one very important note is in order: The v*alidity of an argument is in general independent of the truth or falsehood of the premises. It is perfectly possible for a valid argument to have a false conclusion and for an invalid argument to have a true conclusion.*

[21] Artemus Ward (Charles Farrar Browne) (1834–1867), American humorist.

Definition 2.23 The argument is said to be **sound** iff it is valid and all premises are true.

Example 2.40 Here is an example of a valid argument with a false conclusion:

Every man is immortal.
Socrates is a man.
$\therefore$ *Socrates is immortal.* ■

Example 2.41 An example of an invalid argument with a true conclusion:

Every dog is mortal.
Every animal is mortal.
$\therefore$ *Every dog is an animal.* ■

Example 2.42 Here is another example we inherited from the ancients. (It's due to Chrysippus[22]):

Either the first or the second or the third.
Not the first.
Not the second.
Therefore, the third.

As you can see, we could have equally well used the variables p, q, and r and write

$$\begin{array}{l} p \vee q \vee r \\ \sim p \\ \sim q \\ \therefore r \end{array}$$ ■

Example/Exercise 2.43 Analyze and determine the validity of the following argument:

If we have might, then we are right.
We are right.
$\therefore$ *We have might.*

It may appear to you at this point that recognizing a valid argument does not require much thought. Especially, after the formalism is established, one is tempted to conclude that much care is in fact unnecessary – everything seems self-evident. Well, sometimes even "simple" things need more attention.

[22] See DeLong, H., *A profile of Mathematical Logic,* Dover, 1970. (Incidentally, this is a beautiful book, and anyone interested in mathematics should read it.)

Let's pause for a while and digress a little. Consider the following argument:

Bob loves Alice.
Alice is one word.
$\therefore$ *Bob loves one word.*

This sounds like nonsense, right? However, pure formalism of constructing the argument looks good.

Let's look at another example:

Bob speaks French.
French is English for Français.
$\therefore$ *Bob speaks English for Français.*

The problems with the previous two examples are in the second premise. *Alice* is not (just) a word but rather a name of a certain girl who is loved by *Bob*. So, one should say, "*Alice*" is one word and it is a name of a girl Alice, loved by Bob. However, "*Alice*" is the name of "*Alice*," and so on. As you can see, one has to be careful with one's language. It is one thing if we wish to talk about a girl named Alice, and the other if we talk about Alice's name "Alice." (And, by the way, "Alice" is the name of "Alice," etc.) We use English (among other things) to describe/study mathematics – mathematics is the object of our study, but we can equally well use English (and we do) to study English (English grammar, for instance). If we use English to study, say, French, it is easy to recognize English as a metalanguage and French the object (language) of our study. But if we use English to study English, or mathematics or logic to study mathematics, it is hard to draw a boundary between language and metalanguage – between mathematics and metamathematics. We won't be discussing metalanguages any more. We have mentioned those, incidentally, very important issues very briefly, just to provoke curiosity and possibly incite some readers to dig deeper. There is extensive literature on metamathematics, and nothing would excite this author more than to learn that some readers were prompted by this text to embark onto the beautiful field of metamathematics. It might be appropriate to close this digression by the following, very famous, exchange:

"Do you mean that you think you can find out the answer to it?" said the March Hare.
"Exactly so," said Alice.
"Then you should say what you mean," the March Hare went on.
"I do," Alice hastily replied; "at least – at least I mean what I say – that's the same thing, you know."
"Not the same thing a bit!" said the Hatter. "Why you might just as well say that 'I see what I eat' is the same thing as 'I eat what I see'!"[23]

[23] From *Alice in Wonderland,* by L. Carroll (really C.L. Dodgson (1832–1898) writing under pseudonym).

A Valid Argument Forms

Here is a simple algorithm for checking whether a given argument form is valid:

1. List/identify the premises and conclusion of the argument.
2. Construct the truth table showing the truth values of all premises and the conclusion.
3. Identify **critical rows** – rows in which all premises are true.
4. The argument form is valid if and only if all critical rows have a true conclusion.

Example 2.44 Determine whether the following argument is valid or not:

$$p \vee q$$
$$\sim q$$
$$\therefore p$$

Solution The corresponding truth table is Table 2.20.

TABLE 2.20

p	q	$p \vee q$	$\sim q$	p
T	T	T	F	T
F	T	T	F	T
T	F	T	T	T
F	F	F	T	F

Inspecting the table, we notice that the only row with all true premises has the conclusion that is also true, thus the argument is *valid.* ■

Example 2.45 Now, let's show that the following argument is invalid.

$$p \rightarrow q$$
$$\sim p$$
$$\therefore \sim q$$

The corresponding truth table is Table 2.21.

TABLE 2.21

p	q	$p \rightarrow q$	$\sim p$	$\sim q$
T	T	T	F	F
F	T	T	T	F
T	F	F	F	T
F	F	T	T	T

Observe that lines 2 and 4 have true premises but line 2 has a false conclusion, thus the argument is *invalid.* ■

Example 2.46 (Modus Ponens)

If Socrates is human then Socrates is mortal.
Socrates is human.
Therefore, Socrates is mortal.

Formally:

$$p \rightarrow q$$
$$p$$
$$\therefore q$$

(Note that we could have written the same argument as follows: $((p \rightarrow q) \wedge p) \rightarrow q$)

Now, inspecting Table 2.22, we see that all true premises have true conclusion that makes Modus Ponens a valid argument. ■

TABLE 2.22 Modus Ponens

p	q	$p \rightarrow q$	p	q
T	T	T	T	T
F	T	T	F	T
T	F	F	T	F
F	F	T	F	F

Example 2.47 (Modus Tollens)

If Zeus is human then Zeus is mortal.
eus is not mortal.
Therefore, Zeus is not human.

Formally, see Table 2.23

$$p \rightarrow q$$
$$\sim q$$
$$\therefore \sim p$$

$$(\text{Or } ((p \rightarrow q) \wedge \sim q) \rightarrow \sim p)$$ ■

TABLE 2.23 Modus Tollens

p	q	$p \to q$	$\sim q$	$\sim p$
T	T	T	F	F
F	T	T	F	F
T	F	F	T	F
F	F	T	T	T

Example/Exercise 2.48 Convince yourself that the following arguments are valid:

(i) p
$\therefore p \vee q$

(ii) q
$\therefore p \vee q$

(iii) $p \wedge q$
$\therefore p$

(iv) $p \wedge q$
$\therefore q$

(v) p
q
$\therefore p \wedge q$

(vi) $p \vee q$
$\sim p$
$\therefore q$

(vii) $p \vee q$
$\sim q$
$\therefore p$

Example/Exercise 2.49 Convince yourself that the following arguments are valid:

(i) $p \to q$
$q \to r$
$\therefore p \to r$

(ii) $p \vee q$
$p \rightarrow r$
$q \rightarrow r$
$\therefore r$

(iii) $\sim p \rightarrow c$
$\therefore p$

Example 2.50 Convince yourself that the following argument is invalid:

$$p \rightarrow q \vee \sim r$$
$$q \rightarrow p \wedge r$$
$$\therefore p \rightarrow r$$

Solution See Table 2.24. ■

TABLE 2.24

p	q	r	$\sim r$	$q \vee \sim r$	$p \wedge r$	$p \rightarrow q \vee \sim r$	$q \rightarrow p \wedge r$	$p \rightarrow r$
T	T	T	F	T	T	T	T	T
T	T	F	T	T	F	T	F	F
T	F	T	F	F	T	F	T	T
T	F	F	T	T	F	T	T	F
F	T	T	F	T	F	T	F	T
F	T	F	T	T	F	T	F	T
F	F	T	F	F	F	T	T	T
F	F	F	T	T	F	T	T	T

Example/Exercise 2.51 Convince yourself that the following argument is valid:

$$p \vee q$$
$$p \rightarrow r$$
$$q \rightarrow r$$
$$\therefore r$$

Let's summarize: Rules of Inference for Propositional Logic shown in Table 2.25.

TABLE 2.25

Rule of inference	Tautology	Name
p $\therefore p \vee q$	$p \rightarrow (p \vee q)$	*Disjunctive addition*
q $\therefore p \vee q$	$q \rightarrow (p \vee q)$	
$p \wedge q$ $\therefore p$	$(p \wedge q) \rightarrow p$	*Conjunctive simplification*
$p \wedge q$ $\therefore q$	$(p \wedge q) \rightarrow q$	
p q $\therefore p \wedge q$	$((p) \wedge (q)) \rightarrow (p \wedge q)$	*Conjunction*
$p \rightarrow q$ p $\therefore q$	$((p \rightarrow q) \wedge p) \rightarrow q$	*Modus ponens*
$p \rightarrow q$ $\sim q$ $\therefore \sim p$	$((p \rightarrow q) \wedge (\sim q)) \rightarrow \sim p$	*Modus tollens*
$p \rightarrow q$ $q \rightarrow r$ $\therefore p \rightarrow r$	$((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (p \rightarrow r)$	*Hypothetical syllogism*
$p \vee q$ $\sim p$ $\therefore q$	$((p \vee q) \wedge (\sim p)) \rightarrow q$	*Disjunctive syllogism*
$p \vee q$ $\sim q$ $\therefore p$	$((p \vee q) \wedge (\sim q)) \rightarrow p$	

PREDICATE CALCULUS

Consider the following sentence:

All politicians are crooks. (*)

This is not a statement since (regardless of my personal opinion)[24] we definitely don't know all politicians; thus, we cannot decide whether the sentence is true or false. The quantifier "all" forces us to reconsider how to attribute "true" or "false" to the sentence. The sentence (*) could be considered as a statement of generality, and let's suppose that it is either true or false, but not both. What would make the statement (*), if true, false and, if false, true? Well, how about: There exists **a** politician who is not a crook. This does sound like a good negation of the statement (*), doesn't it?

Or, consider another (less politically charged) sentence:

All swans are white. (**)

The sentence (**) is (also demonstrably) false. Even after seeing thousands of white swans one cannot conclude that **all** swans are white. We can still formally negate the statement (**) by saying: There is **a** swan that is not white. One more example to consider:

Example 2.52 (Goldbach[25] Conjecture) *Every even integer (every even integer greater than 2) is the sum of two primes.*

By inspection, you may be tempted to accept this statement as a theorem but the Goldbach conjecture has not yet been proved to be true. Again, that pesky quantifier "*every*" is the one that makes the whole difference in deciding whether a proposition is true or not. ■

Now, let's consider something slightly more intricate (an argument):

All human beings are mortal.
Socrates is a human being (***)
∴Socrates is mortal.

This argument seems correct (cf. Example 2.43), but note again the quantifier "*all*," which prompts us to evaluate the argument somewhat differently from the arguments in propositional calculus.

[24] Artemus Ward again: I am not a politician, and my other habits are good, also.
[25] Christian Goldbach (1690–1764), German mathematician.

It is evident that quantifiers raise subtle points, and one would like to know how to avoid the mistakes that often creep in unnoticed. You have certainly detected that sentences (*) and (**), as well as the first premise of (***), are formally of the same form: "*All x are P*." Why, then, do we still feel some uneasiness when calling (*) and (**) statements, but not with the first premise of (***)? To proceed, let's first make sure this time that we are completely clear with terms the existential and universal quantifiers we have used many times before.

Existential Quantifier "∃"

The existential quantifier "*there exist*s," symbolically "∃," asserts that there is some particular object/mathematical entity x satisfying some specific property(ies)/condition(s) P, and we write: $(\exists x)P(x)$.

Example 2.53 We write the sentence "*There exists a real number greater than 5*" symbolically, using the existential quantifier, as follows:

$$\exists x \in \mathbf{R},\ \ni x > 5$$ ■

Note that "*there exists … such that*" also means "*for at least one … such that*" or "*for some … such that*." In other words, by

$$\exists x \in \mathbf{R}\,,\ \ni x > 5$$

we mean that there exist a set

$$A = \{x \in \mathbf{R} | x > 5\} \neq \varnothing$$

Universal Quantifier "∀"

The universal quantifier addresses a whole set of objects/entities possessing a particular property. In other words, if $P(x)$ asserts that x has a property P, then $(\forall x)P(x)$ means that the property P holds for all x.

Example 2.54 Consider the statement: *For every real number x greater than zero, x^2 is greater than zero.* Symbolically,

$$(\forall x \in \mathbf{R} \setminus \{0\}), \quad (x^2 > 0)$$

by which we mean

$$(x \in \mathbf{R} \setminus \{0\}) \rightarrow (x^2 > 0)$$

So the set

$$\{x \in \mathbf{R} \setminus \{0\} | x^2 > 0\} = \mathbf{R} \setminus \{0\}$$ ■

Example 2.55 To say that X is a subset of Y, we write

$$\forall z \in X(\ (z \in X) \rightarrow (z \in Y))$$

■

Sometimes, we need to use both quantifiers in order to express our thoughts precisely, as the following example illustrates.

Example 2.56 Recall the axiom of the existence of empty set (Chapter 1): *The empty set exists*. We wrote the axiom formally as follows:

$$\exists \varnothing, \forall y(\sim (y \in \varnothing))$$

■

Example 2.57 Suppose we write

$$\forall X, \exists Y(\sim (X \in Y))$$

Equivalently, we could have said

$$\forall X, \exists Y(X \notin Y)$$

■

Example/Exercise 2.58 Write symbolically: *For all natural numbers x, x is less than or equal to* x^2.

Example/Exercise 2.59 What does the following statement say?

$$\forall X, \forall Y((X \in Y \ \wedge \ Y \in Z) \rightarrow X \in Z)$$

Example 2.60 Express the following in formal language:

(i) The set X contains at least one element.
(ii) There is a set with at least one element.
(iii) The set X contains exactly one element.

Solution

(i) $\exists y\ (y \in X)$. Note the difference with
(ii) $\exists y, \exists X(y \in X)$
(iii) is slightly more difficult. We need to express the fact that, assuming $y_1, y_2 \in X$, then $y_1 = y_2$. So we say:

$$\exists y_1(\ y_1 \in X \ \wedge \ \forall y_2 \in X(y_2 \in X \rightarrow y_2 = y_1))$$

■

Example 2.61 What about the truth value of the sentence: *There is a number whose square is negative.*

Note that I said "*sentence,*" not "*statement.*" You immediately sensed, didn't you, that, unless I specify what number(s) I am talking about, no decision could be made about the truth value of the sentence, and, thus, it is not a statement. If we say: "*There is a real number whose square is negative,*" that is

$$P(x)\colon\ \exists x \in \mathbf{R}(x^2 < 0)$$

we immediately recognize a statement which is obviously false. On the other hand, if we say: "*There is a complex number whose square is negative*," that is,

$$Q(x)\colon\ \exists x \in \mathbf{C}(x^2 < 0)$$

the statement is obviously true.[26] This doesn't come as a surprise to you. Recalling Definition 2.5 you recognize the importance of the interpretation of x in propositional functions $P(x)$ and $Q(x)$. ■

Example 2.62 Similarly, we can easily recognize that the statement

$$P(n,m)\colon\ \forall n \in \mathbf{N}, \exists m \in \mathbf{N}\ \ (n \leq m)$$

is true. ■

Example 2.63 Consider this:

Any friend of A is a friend of B.
C is not B's friend.
Therefore, C is not A's friend.

Symbolically, we write this as

$$\begin{aligned} &\forall x(P(x,A) \rightarrow P(x,B)) \\ &\sim P(C,B) \\ \therefore\ &\sim P(C,A) \end{aligned}$$

where $P(x,y)$ stands for $x\ \ is\ \ a\ \ friend\ \ of\ \ y$. ■

Of course, nothing prevents us from constructing sentences involving more than two variables, such as

$$\text{"}x + y = z\text{"},\ \text{"}x + y + \ z \leq 1\text{"},\ \text{"}\ x^2 + y^2 = r^2\text{"},\ \text{and so on}$$

[26] If you are unfamiliar with complex numbers, you can ignore this example.

(Again, we don't consider them as statements/propositions according to our definition of the statement (see Definition 2.4), since we cannot determine "truth" or "falsehood" unless the values of the variables are specified.) However, regardless of the number of variables, we need to distinguish between the "subject" and the "predicate" of our sentences. Consider the following expression: "$x > 5$" (*x is greater than 5*). It has two parts: the first part, the variable x – the subject of the "statement," and the second part "*is greater than*" – the predicate. Analogously, we can construct sentences with several unquantified variables $x, y, \ldots$, which we call *free variables.* We denote such expressions by $P(x, y, \ldots)$, where P stands for the predicate part and $x, y, \ldots$ are the variables that P refers to. Once the values have been assigned to the variables $x, y, \ldots, P(x, y, \ldots)$ attains the truth value, and $P(x, y, \ldots)$ becomes a proposition. Expanding on these concepts, we introduce the following:

Definition 2.24 A **predicate** is a sentence that contains a finite number of variables $x, y, z, \ldots$ and *becomes a proposition* when specific values are substituted for the variables.

Definition 2.25 The **domain** $\mathcal{D}$, or the **universe of discourse** of the predicate variable $x, y, z, \ldots$ is a set of all values that may be substituted in place of the variables, that is $x, y, z, \ldots \in \mathcal{D}$.

Definition 2.26 Given variables $x, y \in \mathcal{D}$, and if $P(x, y)$ is a corresponding predicate, we say that the **truth set** $\mathcal{T}$ of $P(x, y)$ is the set of all elements of $\mathcal{D}$ that make $P(x, y)$ true, that is

$$\mathcal{T} = \{x, y \in \mathcal{D} | P(x, y) \text{ is true}\}$$

Analogously to Definition 2.5 we state

Definition 2.27 If $P(x, y)$ is a propositional function, then an *interpretation* of x, y is the meaning assigned to x and y for which $P(x, y)$ is a proposition.

Example 2.64 Let $P(x, y)$ be the sentence: "x and y *are even integers.*" We can interpret/think of the predicate $P(x, y)$ in two different but equivalent ways:

(i) *For all* x and $y, P(x, y)$
(ii) *There exist* x and y *such that* $P(x, y)$.

Of course, the truth values of (i) and (ii) depend on the domain of the variables x and y.

In both cases, if $\mathcal{D} = \mathbf{N}_e = \{\text{Set of all even natural numbers}\}$, the statement is obviously true. In the case that $\mathcal{D} = \mathbf{N}$, the statement (i) is evidently not true and the statement (ii) is true.

The quantifiers "for all" – $\forall$, and "there exist" – $\exists$, understandably play an important role in deciding on the truth and falsehood of the predicate statements. ■

Example/Exercise 2.65 Notice the difference between the following expressions:

$$(\forall\, x \in \mathbf{N}, x \leq x^2) \text{ and } (\forall x, x \leq x^2)$$

Which one represents a statement?

Definition 2.28 Let $P(x)$ be a predicate with $x \in \mathcal{D}$. A **universal statement** is a statement of the form $\forall x \in \mathcal{D}$, $P(x)$. It is defined to be true iff $P(x)$ is true for every $x \in \mathcal{D}$. It is false iff there exists at least one $x \in \mathcal{D}$ for which $P(x)$ is false.

Example 2.66

$$\forall x \in (\mathbf{R}\backslash\{0\})\ (x^2 > 0)$$

Here, $\mathcal{D} = \mathbf{R}\backslash\{0\}$, and $P(x)\colon\ x^2 > 0$ ■

A dose of caution is often needed, even when some statements look pretty obvious. For instance, we often say: "*if* $n \geq n_0$" meaning "*for all* $n \geq n_0$."

Example 2.67 The statement

$$\text{If } \ n \geq n_0 \ \text{ then } \ \left|\frac{1}{n} - 0\right| < \varepsilon$$

expressed more precisely reads

$$\forall n, \ \text{ if } \ n \geq n_0 \ \text{ then } \ \left|\frac{1}{n} - 0\right| < \varepsilon$$

or

$$(\forall n, n \geq n_0) \rightarrow \left(\left|\frac{1}{n} - 0\right| < \varepsilon\right)$$

■

Definition 2.29 Let $P(x)$ be a predicate with $x \in \mathcal{D}$. An **existential statement** is a statement of the form "$\exists x \in \mathcal{D}$, such that $P(x)$." It is defined to be true iff there exists at least one $x \in \mathcal{D}$ for which $P(x)$ is true. Otherwise, it is false.

Example 2.68 Consider the following predicate $P(x)$: *There exists an odd number that is not prime.* If we assume that $\mathcal{D}$ is a set of all prime numbers we write our predicate formally as

$$\exists x \in \mathcal{D}, P(x)$$

We can establish the fact that the statement is (obviously) true by simply picking **one** example, say, number 9. ■

As you might have anticipated, not all existential statements are that easy. In particular, to prove their truthfulness/falsity could be rather challenging. The next two examples are more interesting.

Example 2.69 Recall Theorem 1.45 where we stated: $\sqrt{2} \notin \mathbf{Q}$. Let's restate the claim that $\sqrt{2}$ is irrational by using the formalism described above:

$$\nexists p, q \in \mathbf{Z} \left(\sqrt{2} = \frac{p}{q}\right) \tag{*}$$

Or, we can say:

$$\sim \exists p, q \in \mathbf{Z} \left(\sqrt{2} = \frac{p}{q}\right) \tag{**}$$

And also

$$\forall p, q \ (p \in \mathbf{Z} \wedge q \in \mathbf{Z}) \rightarrow \left(\sqrt{2} \neq \frac{p}{q}\right) \tag{***}$$

Again, note that the statements (*), (**), and (***) are equivalent. ■

Example 2.70 Consider the following: *If x is any odd prime less than 10, then $x^2 + 4$ is prime.*

Equivalently, we could have said: *For every odd prime less than 10, $x^2 + 4$ is prime.*

Denoting the set of all primes as P, we write formally

$$((\forall x \in \mathrm{P}) \wedge (x = 2k + 1, k \in \mathbf{N}) \wedge (x < 10)) \rightarrow ((x^2 + 4) \in \mathrm{P}) \quad ■$$

Example 2.71 Write the following statement formally: *Everybody loves somebody*.

Solution Let $\mathcal{D}$, domain of our discourse, stand for a set of all humans. Then our statement reads:

$$(\forall x \in \mathcal{D})(\exists y \in \mathcal{D})(P(x, y))$$

where $P(x, y)$ stands for: x loves y.

■

Example/Exercise 2.72 Write the following statement (attributed to William Cowper[27]) formally as a logical expression:

No man can be a patriot on an empty stomach.

Of course, sometimes a statement given formally needs to be "translated" to everyday English. Try this:

Example/Exercise 2.73 Give at least two more informal versions of the statement in Example 2.54. After you have completed Example/Exercise 2.59, give two more informal expressions of it.

Example 2.74 Give two formal versions of the statement: *For every natural number, there is a real number greater than that number.*

Solution

$$(\forall n \in \mathbf{N})(\exists r \in \mathbf{R})(r > n)$$

or, equivalently, we can write

$$(n \in \mathbf{N}) \rightarrow ((\exists r \in \mathbf{R})(r > n))$$

■

Example 2.75 If you have been exposed to any calculus, you probably remember that the definitions of limit

$$\lim_{x \rightarrow x_0} f(x) = L$$

goes like this: *For every real* $\varepsilon > 0$, *there exists a real X such that*

$$\forall x \in \mathbf{R},\ |f(x) - L| < \varepsilon,\ \text{whenever } 0 < |x - x_0| < \delta$$

The formal expression of the above statement is

$$(\forall \varepsilon > 0, \varepsilon \in \mathbf{R}, \exists \delta > 0, \delta \in \mathbf{R}), (\forall x \in \mathbf{R}, 0 < |x - x_0| < \delta) \rightarrow\ |f(x) - L| < \varepsilon)$$

■

As the statements involving quantifiers in general occur so often, the question, naturally arising, is: How do we negate a statement with existential or universal

[27] William Cowper (1731–1800), English poet.

quantifier? In particular, what is a proper negation of a statement involving quantifiers occurring in mathematical proofs? The issue can be rather subtle, so we need to devote some time to it.

Consider again Aristotle's premise: "*All humans are mortal.*" Suppose you negate this sentence by saying "*All humans are immortal,*" or by saying "*All humans are not mortal.*" Is this right? You guessed it, it's not. Although, for a moment, one might be tempted to state the negation that way (and indeed, in everyday parlance one would often say something like that) but your gut feeling tells you that it is incorrect – the mathematician in you is telling you those negations have to be phrased differently. Let's formalize our sentence. Let $\mathcal{D}$ be the domain of our discourse, that is, $\mathcal{D}$ = set of all humans, and let P stand for the predicate "*to be mortal*." Then, Aristotle's sentence formally looks as follows:

$$\forall\, x \in \mathcal{D}, P(x) \tag{*}$$

Because of the quantifier "$\forall$" it is sufficient to find just one $x \in \mathcal{D}$, just one exception that "violates" the claim, and we have "disproved" our statement (*). (cf. Definitions 2.28 and 2.29). Thus, to negate (*) we say

$$\exists\, x \in \mathcal{D}, \sim P(x).$$

What we are saying here is

$$\sim (\forall\, x \in \mathcal{D},\ \ P(x)) \equiv (\exists\, x \in \mathcal{D}, \sim\ P(x))$$

that is, the negation of "*All humans are mortal*" is "*There is* ***a*** *human who is not mortal*." That's what logic requires from us to say. We don't care about the fact that we have not met any immortal human so far.

Formally, the "rules" for the negation of statements involving existential and universal quantifiers are

Theorem 2.12 If $P(x)$ is a predicate with variable x belonging to a certain domain $\mathcal{D}$, then

(i) $\sim (\forall x \in \mathcal{D}, P(x)) \equiv \exists x \in \mathcal{D}, \sim P(x)$
(ii) $\sim\ (\exists x \in \mathcal{D}, P(x)) \equiv \forall x \in \mathcal{D}, \sim P(x)$

Proof (i) Let $\mathcal{D}$ be a domain such that $\sim (\forall x, P(x))$ is true. But that is the case iff $(\forall x, P(x))$ is false, that is, iff the truth set of $\sim P(x)$ is not empty, which furthermore is true iff $(\exists x \in \mathcal{D})$ such that $\sim P(x)$ is true.

Similarly, we prove (ii). ■

Following two examples are more explicit.

Example 2.76 Let $\mathcal{D} = \{x_1, x_2, \ldots, x_n\}$ be a finite domain and let $P(x, y)$ be a predicate on $\mathcal{D}$. Show that

$$\sim (\forall x \in \mathcal{D},\ P(x)) \equiv \exists x \in \mathcal{D}, \sim P(x)$$

Solution

$$\forall x \in \mathcal{D},\ P(x) \equiv P(x_1) \wedge P(x_1) \wedge \cdots \wedge P(x_n)$$

Therefore,

$$\begin{aligned} \sim (\forall x \in \mathcal{D}, P(x)) &\equiv \sim (P(x_1) \wedge P(x_2) \wedge \cdots \wedge P(x_n)) \\ &\equiv ((\sim P(x_1)) \vee (\sim P(x_2)) \vee \cdots \vee (\sim P(x_n))) \\ &\equiv \exists x \in \mathcal{D}, \ni (\sim P(x)) \end{aligned}$$

■

Similarly, you can do

Example/Exercise 2.77 Let $\mathcal{D} = \{x_1, x_2, \ldots, x_n\}$ be a finite domain and let $P(x)$ be a predicate on $\mathcal{D}$. Show that

$$\sim (\exists x \in \mathcal{D}, \ni P(x)) \equiv \forall x \in \mathcal{D},\ \sim P(x)$$

Sometimes, the "size" or the "nature" of the domain $\mathcal{D}$ simplifies the decision on truth and falsity of the universal statement. Here is an illustration.

Example 2.78 Given a set $\mathcal{D} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ then,

$$\forall\, x \in \mathcal{D}, P(x) \colon\ \frac{2x^2 + 5 + (-1)^x}{2} \in \mathcal{D}$$

is false, since by inspection we immediately observe that $P(5) \notin \mathcal{D}$, that is, $\exists x \in \mathcal{D}$ such that $\sim P(x)$. ■

Finally, we have to say a few words about the conditional statements.

Definition 2.30 Let $\mathcal{D}$ be the domain of our discourse and let $x \in \mathcal{D}$. We call the following statement the **universal conditional statement**:

$$\forall x \in \mathcal{D}, (P(x) \to Q(x))$$

Example 2.79 Rewrite the following statement informally:

$$\forall x \in \mathbf{R}, (x > 2) \to (x^2 > 4)$$

Solution One way would be: *If a real number is greater than 2 then its square is greater than 4.*

■

Example/Exercise 2.80 Write formally "*There is a rational number between any two rational numbers.*"

Example 2.81 Prove the statement:

S: If A is any set then $A \subseteq A$ ■

Proof Let $x \in A$, then our statement is

$$(\forall x, x \in A) \rightarrow (x \in A)$$

which is obviously a tautology of the form $P(x) \rightarrow P(x)$, thus the statement is true. ■

Example/Exercise 2.82 Write the negation of the theorem: *No set is an element of itself.*

Example 2.83 We don't need to understand the meaning of the statement in Example 2.75, but we can still write the correct negation of it:

$$\begin{aligned}
&\sim ((\forall \varepsilon > 0, \varepsilon \in \mathbf{R}, \exists \delta > 0, \delta \in \mathbf{R}), (\forall x \in \mathbf{R}, 0 < |x - x_0| < \delta) \\
&\quad \rightarrow (|f(x) - L| < \varepsilon)) \\
&= (\exists \varepsilon > 0, \varepsilon \in \mathbf{R}), \sim (\exists \delta > 0, \delta \in \mathbf{R}), ((\forall x \in \mathbf{R}, (0 < |x - x_0| < \delta) \\
&\quad \rightarrow (|f(x) - L| < \varepsilon)) \\
&= ((\exists \varepsilon > 0, \varepsilon \in \mathbf{R}), (\forall \delta > 0, \delta \in \mathbf{R}), \sim (\forall x \in \mathbf{R}, (0 < |x - x_0| < \delta) \\
&\quad \rightarrow (|f(x) - L| < \varepsilon))) \\
&= ((\exists \varepsilon > 0, \varepsilon \in \mathbf{R}), (\forall \delta > 0, \delta \in \mathbf{R}), (\exists x \in \mathbf{R}, \sim [(0 < |x - x_0| < \delta) \\
&\quad \rightarrow (|f(x) - L| < \varepsilon)])) \\
&= ((\exists \varepsilon > 0, \varepsilon \in \mathbf{R}), (\forall \delta > 0, \delta \in \mathbf{R}), (\exists x \in \mathbf{R}, \sim [\sim (0 < |x - x_0| < \delta) \\
&\quad \vee (|f(x) - L| < \varepsilon)])) \\
&= ((\exists \varepsilon > 0, \varepsilon \in \mathbf{R}), (\forall \delta > 0, \delta \in \mathbf{R}), (\exists x \in \mathbf{R}, (0 < |x - x_0| < \delta) \\
&\quad \wedge |f(x) - L| \geq \varepsilon))
\end{aligned}$$

■

Example/Exercise 2.84 Write the negation of the following statement:

$$\forall x, y \in \mathbf{R}, \forall \varepsilon > 0, \quad \text{if} \quad |x - y| < \varepsilon, \quad \text{then} \quad x = y$$

Definition 2.31 The **inverse** of the universal conditional statement is

$$\forall x \in \mathcal{D}(\sim P(x) \rightarrow \sim Q(x))$$

Definition 2.32 The converse of the universal conditional statement is

$$\forall x \in \mathcal{D}, (Q(x) \rightarrow P(x))$$

Definition 2.33 The contrapositive of the universal conditional statement is

$$\forall x \in \mathcal{D}, (\sim Q(x) \rightarrow \sim P(x))$$

Example 2.85 Suppose a statement $P(x)$ reads as follows:

$$P(x)\colon \forall x \in \mathbf{R} \ ((x > 3) \rightarrow (x^2 > 9))$$

write the inverse, converse, and contrapositive of the statement S.

Solution

Inverse:	$\forall x \in \mathbf{R} \ ((x \leq 3) \rightarrow (x^2 \leq 9))$
Converse:	$\forall x \in \mathbf{R} \ ((x^2 > 9) \rightarrow (x > 3))$
Contrapositive:	$\forall x \in \mathbf{R} \ ((x^2 \leq 9) \rightarrow (x \leq 3))$

■

In case the predicate P involves two variables $x, y \in \mathcal{D}$, we proceed analogously.

Example 2.86 Let the statement S be

$$\exists x, \forall y, P(x, y).$$

Write the negation of S.

Solution

$$\begin{aligned} \sim S &= \sim (\exists x, \forall y, P(x, y)) \\ &\equiv (\forall x, \exists y, \sim P(x, y)) \end{aligned}$$

■

Example 2.87 Write the negation of the statement

S: *Everybody loves somebody.*

Solution $\sim S$ reads: *It is not the case that everybody loves somebody.* However, let's first formalize S and then look for the formal expression of $\sim S$.

Let $P(x, y)$ stand for "x loves y" and let $x, y \in \mathcal{D} =$ All humans.
Our sentence formally written is

$$S\colon\ \forall x \in \mathcal{D}, \exists y \in \mathcal{D}, P(x, y)$$

Thus,

$$\sim S : \sim ((\forall x\ \in \mathcal{D}, \exists y \in\ \mathcal{D}), P(x, y)) \equiv \exists x \in \mathcal{D}, \forall y \in \mathcal{D}, \sim P(x, y)),\ \text{i.e.}$$

There is somebody that doesn't love anybody. (A very pathetic person indeed.)
From the above example, we see that the negation of

$$\forall x, \exists y,\ \ \text{such that } P(x, y)$$

is indeed logically equivalent to

$$\exists x, \forall y,\ \ \text{such that}\ \ \sim P(x, y)$$ ■

The examples above seem pretty much self-evident; however, some extra caution is needed when dealing with multiply quantified statements. Consider the following true statement *S*: *There is an integer that is even and there is an integer that is odd.* Let's first formalize the statement.

$$S\colon\ \forall x \in \mathbf{Z}, (P(x) \vee Q(x)) \qquad (*)$$

where $P(x)$ stands for "*an integer is even,*" and $Q(x)$ stands for "*an integer is odd.*" There are some important subtleties one needs to recognize:
(*) does not imply/is not the same as

$$((\forall x \in \mathbf{Z}, P(x)) \vee\ (\forall x \in \mathbf{Z}, Q(x))) \qquad (**)$$

since (**) reads: *All integers are even or all integers are odd* , which is obviously false.
Similarly, suppose we say

$$((\exists x \in \mathbf{Z},\ P(x))\ \wedge\ (\exists x \in \mathbf{Z}, Q(x)))$$

that is, *there is an integer that is even and there is an integer that is odd*. Evidently, this is true. However, saying

$$\exists x \in \mathbf{Z},\ (P(x)\ \wedge\ Q(x))$$

that is, *there is an integer that is even and odd*, is of course false.

Definition 2.34

(i) By saying that $(\forall x, P(x))$ is a **sufficient condition** for $Q(x)$ we mean that

$$\forall x, (P(x) \rightarrow Q(x))$$

(ii) By saying that $(\forall x, P(x))$ is **a necessary condition** for $Q(x)$ we mean that

$$\forall x,\ (\sim\ P(x) \rightarrow \sim Q(x))$$

(iii) By saying that $(\forall x, P(x)$ ***only if*** $Q(x))$ we mean that

$$\forall x, (\sim Q(x) \rightarrow \sim P(x))$$

Example/Exercise 2.88 Find the alternative expressions for (ii) and (iii) in Definition 2.34.

Example 2.89 Recall the Axiom of Extensionality from Set theory: *For all sets X and Y, if X and Y have the same elements, then* $X = Y$. Rewrite the axiom formally.

Solution

$$\forall X, \forall Y(\forall z(z \in X \leftrightarrow z \in Y) \rightarrow (X = Y))$$ ■

2.4 ARGUMENTS II

By now you are probably in the position to anticipate that the construction of arguments with universal and existential quantifiers can also be formalized. The next two examples, both due to L. Carroll, are classic.

Example 2.90

All lions are fierce.
Some lions don't drink coffee.
$\therefore$ *Some fierce creatures don't drink coffee.*

Let's put $P(x)$: "*x is a lion*," $Q(x)$: "*x is fierce*," and $R(x)$: "*x drinks coffee*"
Then, the formalized argument reads as follows:

$$\forall\ x\ (P(x) \rightarrow Q(x))$$
$$\exists\ x\ (P(x)\ \wedge\ \sim R(x))$$
$$\therefore \exists\ x\ (Q(x) \wedge\ \sim R(x))$$ ■

Example 2.91

All humming birds are richly colored.
No large birds live on honey.
Birds that do not live on honey are dull in color.
Humming birds are small.

Let's put $P(x)$: "x *is a humming bird*"; $Q(x)$: "x *is large*"; $R(x)$: "x *lives on honey*"; $S(x)$: "x *is richly colored.*" Then, the formalized argument reads as follows:

$$\begin{aligned} &\forall x\ (P(x) \to S(x)) \\ \sim&\exists x\ (Q(x)\ \wedge\ R(x)) \\ &\forall x\ (\sim R(x) \to \sim S(x)) \\ \therefore&\forall x\ (P(x) \to \sim Q(x)) \end{aligned}$$ ■

Predictably, these examples, similarly to the rules of inference for propositions, suggest the existence of general rules for handling the arguments with quantified statements, that is,

The Rules of Inference for Quantified Statements

1. *Universal Instantiation*:

$$\begin{aligned} &\forall x \in \mathcal{D}, P(x) \\ &c \in \mathcal{D} \\ \therefore&P(c) \end{aligned}$$

What we are saying here is that the particular $P(c)$ is true as a consequence of a given universal truth of a premise $\forall x \in \mathcal{D}, P(x)$. A classic example you surely remember is

All human beings are mortal.
Socrates is a human being.
$\therefore$ *Socrates is mortal.*

2. *Universal Modus Ponens:*

$$\begin{aligned} &\forall x \in \mathcal{D}\ (P(x) \to Q(x)) \\ &P(c) \\ \therefore&Q(c) \end{aligned}$$

Here we are saying the following: If all $x \in \mathcal{D}$ make $P(x)$ true, then all x make $Q(x)$ also true. Thus, the particular $c \in \mathcal{D}$ makes $P(x)$ true, and, therefore, the same c makes $Q(x)$ also true. Here is an example:

If a number is even then its square is even.
4 is an even number.
$\therefore$ *16 is even.*

3. *Universal Modus Tollens*

$$\begin{aligned} &\forall x \in \mathcal{D}\ \ (P(x) \to Q(x)) \\ &\sim Q(c) \\ &\therefore \sim P(c) \end{aligned}$$

We are saying the following: If all $x \in \mathcal{D}$ make $P(x)$ true, then all x make $Q(x)$ also true. c, however, does not make $Q(x)$ true, and therefore c does not make $P(x)$ true.

Another familiar example:

All humans are mortal.
Zeus is not mortal.
∴ Zeus is not human.

Sometimes, Modus Tollens might be slightly disguised by the additional "no" included in the premises. Consider the following:

Example 2.92 Suppose we say:

No polynomial function has horizontal asymptotes.
Function f has a horizontal asymptote.
∴ function f is not a polynomial.

Before writing the above argument formally, it may be advantageous to restate it as follows:

Let $P(x)$ be: *x is a polynomial function*, and
$Q(x)$: *x* ***does not*** *have a horizontal asymptote*.
Our argument, formally written, is

$$\begin{aligned} &\forall x \in \mathcal{D}, (P(x) \to Q(x)) \\ &\sim Q(f) \\ &\therefore \sim P(f) \end{aligned}$$

Thus, this argument is valid by universal Modus Tollens. ■

Carefully comparing modus ponens and modus tollens in the following two examples will illustrate some common and frequent mistakes.

Example 2.93 (Converse Error) Consider the following argument:

$$\begin{aligned} &\forall x \in \mathcal{D}(P(x) \to Q(x)) \\ &Q(c) \\ \therefore\ &P(c) \end{aligned}$$

Now, recalling that the conditional and its inverse are not logically equivalent, we conclude that the argument is not a valid one. ■

Similarly, we have

Example 2.94 (Inverse Error) The argument

$$\begin{aligned} &\forall x \in \mathcal{D}\ \ (P(x) \to Q(x)) \\ &\sim P(c) \\ \therefore\ &\sim Q(c) \end{aligned}$$

is also invalid since the conditional and its inverse are not logically equivalent. ■

So, to sum up, we have:

Rules of Inference for Quantified Statements

Rule of Inference	Name
$\forall x \in \mathcal{D}, P(x)$ $\therefore P(c), c \in \mathcal{D}$	*Universal Instantiation*
$P(c)$, for an arbitrary $c \in \mathcal{D}$ $\forall x \in \mathcal{D}, P(x)$	*Universal Generalization*
$\exists x, P(x)$ $\therefore P(c), c \in \mathcal{D}$	*Existential Instantiation*
$P(c)$ for some element $c \in \mathcal{D}$ $\therefore \exists x, P(x)$	*Existential Generalization*
$\forall x \in \mathcal{D}, P(x) \to Q(x)$ $P(c), c \in \mathcal{D}$ $\therefore Q(c)$	*Universal Modus Ponens*
$\forall x \in \mathcal{D}, P(x) \to Q(x)$ $\sim Q(c), c \in \mathcal{D}$ $\therefore \sim P(c)$	*Universal Modus Tollens*

2.5 A SHORT REVISIT TO SET THEORY

As an additional exercise in logic (and set theory), and also pursuing our goal of mathematical precision and rigor, let's revisit some concepts from *Set Theory* we discussed in Chapter 1. We can now restate them more formally and with some new conceptual insights. You may remember that at the beginning of our discussion of sets, we mentioned en passant the school of thought in set theory, which holds that everything (mathematical) is a set. By this, we mean that in the *universe* we are about to study x represents a set, and all the elements of x are sets whose elements are again sets, and so on. This is the idea we embrace in the so-called *Zermelo–Frankel Set Theory*[28] or, simply, *Axiomatic Set Theory*. Whether at this point you accept this idea or not, trust me, it is a very efficacious one. However, I suppose you would accept the following Laws of Thought:

(i) *The Law of Identity:* (Any)thing that is is.
(ii) *The Law of Contradiction:* (No)thing can both be and not be.
(iii) *The Law of Excluded Middle*: (Every)thing must either be or not be.

Having those in mind, we will say that an object exists if and only if it is a set. In other words, we think of sets as logical objects that are parts of the formal structure we are studying. Consequently, to say that an object does not exist is the same as saying that it is not a set. Not surprisingly, then, in this section we will use lowercase letters for both the sets and their elements, that is, to write a proposition $(x \in y)$ would not be inappropriate. Therefore, if $(x \in y)$ is a true proposition, we take it that x and y are sets. And, of course, for saying that $(x \notin y)$ is the same as saying $\sim (x \in y)$. So, we start with the following basic assumptions:

1. A set exists if the proposition that asserts its existence is logically true. (If a formal object is not a set, it does not exist.)
2. If the assumption that an object exists leads to a contradiction, we conclude that that object does not exist, that is, it is not a set.

So, in terms of our predicate logic, we would state *set existence* (maybe redundantly) as

$$\exists x(x = x)$$

This is known as *Axiom 0* of set theory.

Now, recall that only primitive concepts of set theory are a "set" and "element" (a "member") of a set. Thus, strictly speaking, since in the above predicate

[28]Ernst Friedrich Ferdinand Zermelo (1871–1953), German mathematician. Abraham Halevi Fraenkel (1891–1965), Israeli/German mathematician.

P: $x = x$, "=" has not been defined, to be consistent we should have said: *the proposition*

$$P\colon\ \exists y \forall x (x \notin y)$$

is true. On the other hand, saying:

$$\forall x (x \notin y)$$

we have defined and an empty set.

Next, we say that if the proposition

$$\forall z (z \in x \rightarrow z \in y)$$

is true, we have got ourselves a concept of subset, that is,

$$x \subseteq y \leftrightarrow \forall z (z \in x \rightarrow z \in y)$$

Axiom of Equality (Axiom of Extensionality) follows: If proposition

$$(x \subseteq y) \ \wedge \ (y \subseteq x)$$

is true, that is,

$$\forall z (z \in x \leftrightarrow z \in y) \rightarrow (x = y)$$

is true, then we have defined what $x = y$ means. Consequently,

$$y = \varnothing \ \leftrightarrow \forall x (x \notin y)$$

As a simple exercise, we prove something that we already know, that is, let's prove

Theorem 2.13 The empty set is unique.

Proof Let $\varnothing_1$ and $\varnothing_2$ be two empty sets. Then, the propositions $x \in \varnothing_1$ and $x \in \varnothing_2$ are both false for all x. Consequently,

$$x \in \ \varnothing_1 \leftrightarrow x \in \ \varnothing_2$$

is true for all x.

From the axiom of equality, we have

$$(\ \forall x \ (x \in \ \varnothing_1 \leftrightarrow x \in \ \varnothing_2)) \rightarrow (\varnothing_1 = \varnothing_2)$$

Thus the proof. ■

2.6 BOOLEAN ALGEBRA

We finish this chapter by briefly discussing a specific structure that has applications in many sciences, including physics and computer science.

Definition 2.35 We say that $\mathcal{B}$ is a *Boolean algebra* if $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$, where B is a set containing at least two elements 0 and 1, two binary operations $+$ and $\cdot$, and one unary operation $^-$ on B, such that for all $a, b, c \in B$ the following axioms hold:

(i) $a + b = b + a$
(ii) $a + (b + c) = (a + b) + c$
(iii) $a \cdot b = b \cdot a$
(iv) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
(v) $a + (a \cdot b) = a$
(vi) $a \cdot (a + b) = a$
(vii) $a + (b \cdot c) = (a + b) \cdot (a + c)$
(viii) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
(ix) $a + 0 = a$
(x) $a \cdot 1 = a$

For every $a \in B$, there exists a *complement* of a, $\overline{a} \in B$, such that

(xi) $a + \overline{a} = 1$
(xii) $a \cdot \overline{a} = 0$

I hope all of this does not look totally unfamiliar. With adequate substitution of operations $+, \cdot$ and $^-$ with analogues in set theory $\cup, \cap$ and c, or $\vee, \wedge$ and $\sim$ in propositional logic, the axioms (i)–(x) are something that we have already encountered. The proofs of the theorems that follow are simple exercises in Boolean algebra.

Example 2.95 Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Prove that for every $a, b \in B$, $a \cdot (a + b) = a$, and $a + (a \cdot b) = a$ (Definition 2.35 (vi) and (vii)). ■

Proof

$$\begin{aligned} a \cdot (a + b) &= (a + 0) \cdot (a + b) \\ &= a + 0 \cdot b \\ &= a + 0 \\ &= a \end{aligned}$$

$$\begin{aligned} a + (a \cdot b) &= (a \cdot 1) + (a \cdot b) \\ &= a \cdot (1 + b) \\ &= a \cdot 1 \\ &= a \end{aligned}$$ ■

Theorem 2.14 Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Then, for all $a \in B$, there is a *unique* $\overline{a}$ such that

$$a\overline{a} = 0 \text{ and } a + \overline{a} = 1 \qquad (*)$$

Proof Suppose there are $\overline{a}_1$, $\overline{a}_2 \in B$ satisfying $(*)$. Then,

$$\begin{aligned} \overline{a}_1 &= \overline{a}_1 \cdot 1 \\ &= \overline{a}_1 \cdot (a + \overline{a}_2) \\ &= (\overline{a}_1 \cdot a) + (\overline{a}_1 \cdot \overline{a}_2) \\ &= 0 + (\overline{a}_1 \cdot \overline{a}_2) \\ &= (a \cdot \overline{a}_2) + (\overline{a}_1 \cdot \overline{a}_2) \\ &= (a + \overline{a}_1) \cdot \overline{a}_2 \\ &= 1 \cdot \overline{a}_2 \\ &= \overline{a}_2 \end{aligned}$$ ■

Theorem 2.15 *Let* $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Then, for every element $a \in B$

$$\overline{(\overline{a})} = a$$

Proof Let a be a complement of $\overline{a}$. By the previous theorem, a complement of $\overline{a}$ is unique thus

$$a = \overline{(\overline{a})}$$ ■

Theorem 2.16 Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Then, for every element $a \in B$

(i) $a \cdot a = a$ and
(ii) $a + a = a$

Proof Let a be any element of B. Then,

$$\begin{aligned}
\text{(i)}\ a &= a \cdot 1 \\
&= a \cdot (a + \overline{a}) \\
&= a \cdot a + a \cdot \overline{a} \\
&= a \cdot a + 0 \\
&= a \cdot a \\
\text{(ii)}\ a &= a + 0 \\
&= a + (a \cdot \overline{a}) \\
&= (a + a) \cdot (a + \overline{a}) \\
&= (a + a) \cdot 1 \\
&= a + a
\end{aligned}$$

■

Theorem 2.17 Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Then, for every $a, b \in B$

(i) $a + \overline{a} \cdot b = a + b$ and
(ii) $a \cdot (\overline{a} + b) = a \cdot b$

Proof

$$\begin{aligned}
\text{(i)}\ a + \overline{a} \cdot b &= (a + \overline{a}) \cdot (a + b) \\
&= 1 \cdot (a + b) \\
&= a + b \\
\text{(ii)}\ a \cdot (\overline{a} + b) &= (a \cdot \overline{a}) + (a \cdot b) \\
&= 0 + (a \cdot b) \\
&= a \cdot b
\end{aligned}$$

■

Example 2.96 Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra, and $a, b \in B$. Show that if $a \cdot b = 1$, then $a = b = 1$.

Solution Let a and b be any two elements of B such that $a \cdot b = 1$. Then,

$$\begin{aligned}
1 &= a \cdot b \\
&= (a \cdot a) \cdot b \\
&= a \cdot (a \cdot b) \\
&= a \cdot 1 \\
&= a
\end{aligned}$$

Also,

$$\begin{aligned} 1 &= a \cdot b \\ &= a \cdot (b \cdot b) \\ &= (a \cdot b) \cdot b \\ &= 1 \cdot b \\ &= b \end{aligned}$$

Thus, $a = b = 1$ as claimed. ■

Theorem 2.18 (DeMorgan's Laws) Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Then, for every $a, b \in B$

(i) $\overline{(a + b)} = \overline{a} \cdot \overline{b}$ and
(ii) $\overline{(a \cdot b)} = \overline{a} \cdot \overline{b}$

Proof The proof of DeMorgan's laws in Boolean algebra is a little trickier than the relatively simpler proofs we have encountered in set theory and logic.

(i) We reason the following way: if it is true that $\overline{(a + b)} = \overline{a} \cdot \overline{b}$, then $(a + b)$ and $(\overline{a} \cdot \overline{b})$ should be complementary, that is,

$$(a + b) + (\overline{a} \cdot \overline{b}) = 1 \quad \text{and} \quad (a + b) \cdot (\overline{a} \cdot \overline{b}) = 0$$

Then,

$$\begin{aligned} (a + b) + (\overline{a} \cdot \overline{b}) &= (a + b + \overline{a}) \cdot (a + b + \overline{b}) \\ &= (1 + b) \cdot (a + 1) \\ &= 1 \end{aligned}$$

Also,

$$\begin{aligned} (a + b) \cdot (\overline{a} \cdot \overline{b}) &= (a \cdot \overline{a} \cdot \overline{b}) + \overline{(b \cdot \overline{a} \cdot \overline{b})} \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

Similarly, for

(ii) If it is true that $\overline{(a \cdot b)} = \overline{a} + \overline{b}$, then $(a \cdot b)$ and $\overline{a} + \overline{b}$ should be complementary, that is,

$$(a \cdot b) + (\overline{a} + \overline{b}) = 1 \text{ and } (a \cdot b) \cdot (\overline{a} + \overline{b}) = 0$$

So, we have

$$\begin{aligned}(a \cdot b) + (\overline{a} + \overline{b}) &= (a + \overline{a} + \overline{b}) \cdot (b + \overline{a} + \overline{b}) \\ &= (1 + \overline{b}) \cdot (1 + \overline{a}) \\ &= 1 \cdot 1 \\ &= 1\end{aligned}$$

and

$$\begin{aligned}(a \cdot b) \cdot (\overline{a} + \overline{b}) &= (a \cdot b \cdot \overline{a}) + (a \cdot b \cdot \overline{b}) \\ &= 0 + 0 \\ &= 0\end{aligned}$$

Thus the proof. ■

2.7 SUPPLEMENTARY PROBLEMS

1. Which of the following sentences are statements?

(i) $2 + 3 = 3 + 2$

(ii) He is a math student.

(iii) $x + 3 = 12$

(iv) $2 + 3 \neq 3 + 2$

2. Construct the truth tables for the following statements:

(i) $p \vee (q \rightarrow r)$

(ii) $(p \rightarrow q) \rightarrow (r \wedge q)$

(iii) $(p \wedge (q \rightarrow r)) \rightarrow ((r \rightarrow p) \wedge q)$

3. Construct truth tables for the following statements:

(i) $(\sim p \leftrightarrow \sim q) \leftrightarrow (p \leftrightarrow q)$

(ii) $(p \rightarrow q) \wedge (\sim p \rightarrow r)$

(iii) $(p \leftrightarrow q) \leftrightarrow (r \leftrightarrow s)$

4. Determine whether each of the following statements is true or false:
(i) *If* $2+2=4$ *then* $2+3=7$
(ii) *If pigs can fly then* $2\times 2=22$
(iii) (*If* $2+2=4$ *then* $2+3=7$) $\rightarrow$ (*If pigs can fly then* $2\times 2=22$)

5. State the negation of the following statements:
(i) *Big girls don't cry.*
(ii) *Something bothers me.*

6. Show that

$$\sim(p\wedge\sim q) \text{ and } (\sim p\vee q) \text{ are logically equivalent to } (p\rightarrow q)$$

7. Given a statement

$$S\colon\ p\vee(p\rightarrow(q\wedge\sim p)),\quad \text{find}\quad \sim S.$$

8. Write the inverse, converse, and contrapositive of the statement

S: *If I study hard, I will succeed in mathematics.*

9. Determine the truth value of the statement

$$S\colon\ \textit{If pigs can fly then } 2+2=22$$

10. Show that for two statements p and q the following is true:
(i) $\sim(p\rightarrow q)\equiv p\wedge\sim q$
(ii) $\sim(p\leftrightarrow q)\equiv(p\wedge\sim q)\vee(q\wedge\sim p)$

11. Suppose the truth value of $(p\rightarrow q)$ is F, what is the truth value of
(i) $\sim p\vee(p\rightarrow q)$
(ii) $\sim(p\wedge q)\leftrightarrow(\sim p\rightarrow\sim q)$

12. Write the negations of the following statements:
(i) $p\rightarrow(q\vee r)$
(ii) $p\rightarrow(q\rightarrow r)$
(iii) $(p\rightarrow q)\rightarrow r$
(iv) $\sim(p\rightarrow r)\rightarrow[q\rightarrow(p\wedge r)]$

13. Suppose the truth value of $p\leftrightarrow q$ is T, what are the truth values of
(i) $p\wedge q$
(ii) $p\vee q$
(iii) $(p\wedge r)\leftrightarrow(q\wedge r)$
Repeat the problem if $p\leftrightarrow q$ is F.

14. Prove all the statements of Theorem 2.3.

15. Determine whether each of the following statements is logically equivalent to the statement

S: *If you study hard, then you will succeed in mathematics.*

(i) *If you don't study hard, you won't succeed in mathematics.*
(ii) *If you don't succeed in mathematics, you didn't study hard.*
(iii) *Your succeeding in mathematics implies you studied hard.*
(iv) *Your studying hard is necessary for you to succeed in mathematics.*

16. Prove the following theorem:
If p and $(p \to q)$ are tautologies then so is q.

17. If $p \leftrightarrow q$ is F, what can be said about the truth values of the following?

(i) $p \wedge q$
(ii) $p \vee q$
(iii) $p \to q$
(iv) $(p \wedge r) \leftrightarrow (q \wedge r)$

18. Which of the following statements are logically implied by $(p \wedge q)$?

(i) p
(ii) q
(iii) $(p \vee q)$
(iv) $(p \to q)$
(v) $(p \leftrightarrow q)$
(vi) $(\sim p \to \sim q)$

19. Determine whether the following are tautologies:

(i) $(p \to q) \vee (q \to p)$
(ii) $(p \leftrightarrow q) \leftrightarrow (p \leftrightarrow (q \leftrightarrow p))$
(iii) $(p \to (q \vee r)) \vee (p \to q)$

20. Determine whether each of the following is a tautology, contradiction or neither:

(i) $p \wedge (\sim (p \vee q))$
(ii) $p \leftrightarrow (p \vee p)$
(iii) $(p \leftrightarrow (p \leftrightarrow q)) \to q$
(iv) $(p \to q) \to ((q \to \sim r) \to (\sim r \to p))$

21. Show that the following statements are tautologies:

(i) $(p \vee (q \wedge r)) \leftrightarrow ((p \vee q) \wedge (p \vee r))$

(ii) $((p \wedge \sim q) \rightarrow (r \wedge \sim r)) \rightarrow (p \rightarrow q)$

22. Prove:

If A is a finite set and B is a proper subset of A, then B is not equivalent to A.

23. Convince yourself that the following propositions are other variants of the method of *reductio ad absurdum*

(i) $((p \wedge \sim q) \rightarrow \sim p) \rightarrow (p \rightarrow q)$

(ii) $((p \wedge \sim q) \rightarrow q) \rightarrow (p \rightarrow q)$

24. Suppose $(p \rightarrow q)$ is F, and $\sim (p \wedge q)$ is T. What are the truth values of the following statements?

(i) $\sim p \vee (p \rightarrow q)$

(ii) $\sim (p \wedge q) \leftrightarrow (\sim p \rightarrow \sim q)$

25. If p and q are true and r is false, what are the truth values of the following statements?

(i) $(q \wedge \sim p) \leftrightarrow (p \leftrightarrow r)$

(ii) $(q \rightarrow p) \rightarrow ((p \rightarrow \sim r) \rightarrow (\sim r \rightarrow q))$

26. Here is another famous paradox you may want to consider: Socrates says,

"*What Plato says is false*."

Plato says,

"*What Socrates says is true*."

So, what do you think: Is what Socrates says true?

27. Consider the statement

S: *There exists a man without a country.*

What is the negation of this statement?

28. Write the negation of the statement: *No set is element of itself.* Can you prove it, or its negation? (Hint: see Chapter 1.)

29. What about the truthfulness or falsity of the following statements?

(i) *Every statement is true or false.*

(ii) *No statement is both true and false.*

30. Prove (ii) of Theorem 2.12.

31. Rewrite formally, and carefully (!) examine, the statements

S_1: *All integers are even.*

S_2: *All integers are not even.*

Determine the truth value of each.

32. Determine whether the following propositions are true or false:

(i) $\exists n \in \mathbf{N}, \ni \forall m \in \mathbf{N}, n \leq m$

(ii) $\exists p \in \mathbf{Q}, \ni \forall q \in \mathbf{Q}, p \leq q$

33. Determine whether *the negations* of the following statements are true or not:

(i) $\exists x \in \mathbf{R}, \ni \forall y \in \mathbf{R},\ ((x < y) \vee (y < x))$

(ii) $\exists x \in \mathbf{R}, \ni \forall y \in \mathbf{Z},\ (\,(x < y) \vee (y < x))$

34. Determine whether the following predicates are true:

(i) $(\forall x \in \mathbf{R}, x = \sqrt{x^2})$

(ii) $(\,\forall x \in \mathbf{R}, x > 0\,)$

35. Determine the validity of the following arguments:

(i) *If a number is even, then twice that number is even*
The number 2n is even.
Therefore, n is even.

(ii) *If a product of two numbers is* 0, *then at least one of the numbers is* 0.
Neither $(x - 1)$ *nor* $(x + 1)$ *equals* 0.
Therefore, $(x - 1)\ (x + 1)$ *is not* 0.

36. Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra. Find

(i) $1 \cdot \overline{0}$

(ii) $1 + \overline{1}$

(iii) $\overline{0} \cdot 0$

(iv) $\overline{(1 + 0)}$

37. Let $\mathcal{B} = (B, +, \cdot, ^-, 0, 1)$ be a Boolean algebra, and let $a, b \in B$. Find the complements of

(i) $a + b$

(ii) $\overline{a} \cdot \overline{b}$

3

PROOFS

> Plato says that the Good, the Beautiful and the True are all and the same thing. I find this idea both beautiful and good, but is it true?
>
> Raymond M. Smullyan[1]

> Is there any knowledge in the world which is so certain that no reasonable man could doubt it? This question, which at first sight might not seem difficult, is really one of the most difficult that can be asked.
>
> Bertrand Russell[2]

> I state my case, even though I know it is only part of the truth, and I would state it just the same if I knew it was false, because certain errors are stations on the road to the truth. I am doing all that is possible to a definite job at hand.
>
> Robert Musil[3]

3.1 INTRODUCTION

Ever since the time of the Greeks, mathematics has involved proof; and it is even doubted by some whether proof, in the precise and rigorous sense which

[1]Smullyan, R. M., *A Spiritual Journey,* Praxis International, Inc., West Chester, Pennsylvania, 2009.
[2]Russell, B., *The Problems of Philosophy,* Oxford University Press, 1912.
[3]Robert Musil (1880–1942), Austrian writer.

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

the Greeks gave to this word, is to be found outside mathematics. We may fairly say that this sense has not changed, because what constitute a proof for Euclid is still a proof for us; and in times when the concept has been in danger of oblivion, and consequently mathematics itself has been threatened, it is to the Greeks that man have turned again for models of proof.[4]

We need to step back, pause for a while, and reflect upon some preoccupying conceptual questions that have been with us all along, whether explicitly stated or not. The very title of this chapter might be looked upon with suspicion by a meticulous reader: Why devote a whole separate chapter to something that we have been doing all along – proving "things?" We need to step back indeed.

Mathematics is one of the few disciplines whose object of research, let alone its importance, is difficult to explain to the uninitiated. I can think of no other field, except for physics and philosophy, which has similarly difficult problem. So, what is mathematics indeed? Or, to put it differently: what do mathematicians do? Since mathematics is certainly a technical, but also philosophical and in some sense artistic discipline, some formalists may simply tell you that "mathematics is what mathematicians do." I don't like that, since it leads to a silly tautological conclusion: mathematicians do what mathematicians do. The other answer that "mathematicians prove theorems" is absolutely true, but it still doesn't tell you much, unless you are a mathematician of course. As you might have anticipated it is even more difficult to explain what mathematics is (all about). My PhD adviser, who is one of the leading knot theorists in the world, was once approached at a party by a stranger with a question: "So, what do you do for a living?" You can only imagine the reaction his answer "I study knots" provoked. Why am I telling you all of this? Well, although we have been proving things all along, the purpose of those (proofs) was to establish the truthfulness of the claims we make regarding sets, functions, you name it. But the ultimate question is: "What is truth and how do we recognize it?" Sometimes we hear: "Truth is what is factual" and then also: "factual is what is true." That doesn't help much, does it? Slightly more serious sounding answer we get is: "A statement is true if it accurately represents or picture reality or the world the way it is."[5] What on earth does that mean? What is "reality" we may ask. Reality of a chair I am sitting on while writing these words on my laptop, or reality of a number 2, for instance, or reality of your thoughts and feelings while reading these lines. No wonder that question(s) of reality and/or truth has been with us for centuries. Aristotle famously said:

> "*To say that that which is is not or that which is not is, is a falsehood; and to say that which is is and that which is not is not, is true.*"[6]

I am not going to even try to answer the question of what truth is. I am simply going to assume that *it* exists and, without dwelling into profound philosophical

[4] Bourbaki, N., *Elements of Mathematics – Theory of Sets,* Springer, 2004.
[5] This answer is known to belong to Correspondence Theory of Truth.
[6] Aristotle, *Metaphysics*, Oxford University Press, 1993.

questions, I am also going to assume that the nature of reason is such that we can occasionally, at least in mathematics and sciences, recognize it. I am not saying that our mental faculties are such that we can know *all* the truths. What I am saying is that unless we assume that truth exists, that we are smart enough, and unless we try hard enough we will never know.

Well then, since by definition we demand a higher level of certainty in mathematics, we are constantly searching for new ways to "discover" and prove our discoveries. Eventually, we have to express those discoveries and proofs thereof in our natural language. Thus, one might be inclined to assume that our natural language is primary and that logic is derived from it. However, that cannot be. There are many languages, but the logic for a Chinese native speaker, say, and an English one has to be the same. So, if logic and mathematics are absolute, in order to prove that something is true we turn to mathematics and the *proofs* therein, that is, we scrutinize the proofs themselves. How do we know that our proof is indeed THE proof? Well, admitting some circularity in the claim, we will say: above all, the proof has to be "*logically necessary*." We won't go very deep into metamathematics, if at all, but the idea might be a little clearer to you now – we are concerned with "studying proofs." (This is actually exactly what Turing[7] was interested in. His study of computation theory which led to computers, so ubiquitous nowadays, was in fact motivated not by the desire to build a computer but by the issues of the nature of mathematical proof.) And even that is not completely a valid statement. As a matter of fact, let's admit it right away that the purpose of this essay, and the whole book for that matter, is to hopefully provoke your curiosity and to put you on the right track of thinking mathematics. Maybe another way of approaching the whole endeavor, as an artistic enterprise, can give you a better feel for the essence of "proofs." Namely, although there are some specific and rigorous rules one needs to follow when doing mathematics, there is still not a universal recipe that will give you a clear and secure way to reach your goal. *A clear and universal recipe* (*if it exists*) *to reach our goals* – that's what we are looking for. For instance, we may ask:

1. *Is there a finite list of logical axioms and rules of inference that summarizes all of the logic used in mathematical reasoning?* (The answer happens to be *yes.*)
   Or, we can ask:
2. *Is arithmetic consistent and complete?* (Surprisingly the answer is *no.*)
   But, how do we know that? How do we prove these claims?

So, here comes *Proof Theory,* as part of mathematical logic that studies mathematical proofs. For it we need precision and rigorous mathematical framework so that all the concepts we are using are clear, that is, precisely

[7] Alan Mathison Turing (1912–1954), English mathematician, logician, cryptanalyst, and computer scientist.

defined. (We have already used many of those concepts before, tacitly assuming their validity and precision.)

Definition 3.1 Let S be a set of symbols. An **expression** or **word** in S is a finite sequence of symbols of S.

Example 3.1 Let $S = \{a, b, c, \ldots, w, y, z, \wedge, \in, \rightarrow, \forall, \exists,\}$ be a set of symbols we call the **alphabet**. Then, $aabb \rightarrow bcd$ and $axb \wedge cc$ would be two words/expressions in S.

Definition 3.2 An expression in the alphabet S which has meaningful interpretation is a **formula** φ.

Example 3.2 Let $S = \{a, b, c, \ldots, w, y, z\}$ be a set of 26 letters of English alphabet, then "*house*" and "*brohuse*" are both expressions but only "*house*" is a formula. ■

Definition 3.3 A set of symbols S with selection of formulas is called a **formal language** $\mathcal{L}$.

Definition 3.4 A collection of carefully selected formulas that are considered *self-evident,* or *obviously true,* we call **axioms**.

Definition 3.5 We say that $\mathcal{F} = (\mathcal{L}, \mathcal{A}, R)$ is a **formal system** where $\mathcal{L}$ is a formal language, $\mathcal{A}$ is a set of axioms, and R are the rules of inference.

Example 3.3 Let's consider a formal system $\mathcal{F}$ with language $\mathcal{L}$ consisting of the alphabet $S = \{+, =, *\}$, and with a formula defined as any expression of the form $a + b = c$, where $a, b,$ and c stand only for the symbol $*$. Furthermore, let the only axiom be: $* + * = **$. Then, for instance,

$* + ** = ***$ is a formula and $* ++$ is not.

In addition, let's say the following two rules of inference are given:

1. $a * + b = c *$
2. $b + a = c$

Let's now prove that the formula φ: $** + *** = ****$ is a theorem.

By axiom $* + * = **$

And then by rules of inference, we have:

$$\begin{aligned} ** + * &= *** && \text{by (1)} \\ *** + * &= **** && \text{by (1)} \\ * + *** &= **** && \text{by (2)} \\ ** + *** &= ***** && \text{by (1)} \end{aligned}$$

Let's consider something more intriguing.

Goldbach's Conjecture[8]: Every even number greater than 2 is the sum of two primes.

What we are saying here is: *Every even natural number n has the property P,* or

$$\forall n = 2k \in \mathbf{N},\ P(n)$$

where by $P(n)$ we mean "*even number greater than 2 is the sum of two primes*."

It is not difficult to pick a couple of even numbers greater than 2 and verify that the above statement holds. Say, we take $4 = 2 + 2$, or $6 = 3 + 3$, or $12 = 5 + 7$, and so on. Does this prove the conjecture? What about if we check a thousand more examples and observe (which we would) that each one of them is indeed expressible as a sum of two primes? In other words, after thousand tries we were not able to find any $n \in \mathbf{N}$, such that $\sim P(n)$. Could we consider this as a satisfactory proof? Let's look at another conjecture not involving primes:

Collatz's Conjecture[9]: *Starting with any natural number n compute* $n/2$ *if n is even, or* $3n + 1$*, if n is odd. Continue the process applying the same rule to the newly obtained number. Eventually, we end up with 1.*

For this conjecture, we may need an example. Let's start with 7. We get:

$$7, 22, 11, 34, 17, 52, 26, 13, 50, 20, 10, 5, 16, 8, 4, 2, 1$$

Again we ask the same question: Does this constitute the proof? In both cases, the answer is categorically NO! It does not. (By the way, both conjectures are still unsolved mathematical problems.) The point here is that simply giving a finite number of examples is not, cannot be, considered as the proof. Thus, in order to prove a mathematical statement, to make it a theorem, we need to do more. We

[8]Christian Goldbach (1690–1764), German mathematician.

[9]Lothar Collatz (1910–1990), German mathematician. Incidentally, the conjecture is also known as the $3n + 1$ conjecture, the Ulam conjecture (after Stanislaw Ulam), Kakutani's problem (after Shizuo Kakutani) and by some other names.

need to prove it "in general." How do we do that? One usually starts by using the tools existing within the framework of a certain axiomatic system, including the definitions and other theorems already proven within the same system. Often new hypotheses, conjectures, and techniques are needed. One thing, however, is certain – to be accepted as the "proof" a proof has to be finite. That is, although mathematical conjectures generally pertain to infinitely many cases, the proof of them must be accomplished in a finite number of steps and using a finite number of axioms. Another thing worth mentioning is this: Very often, we come to the point where a criterion that is difficult to explain, the "esthetic" criterion (for the lack of better words) enters the picture, in fact, it becomes indispensable. It is definitely true – it is a historical fact – that the search for elegance, beauty, and harmony has been a guiding principle of all sciences, mathematics in particular. Theories or proofs that are lacking this esthetic component always turn out to be wrong. When the right proof is found we recognize it as the most beautiful, the most elegant, and we all eventually say: how could it have been otherwise?[10] Starting from the premise that the human reason could be subject to a rigorous scientific analysis itself, it becomes obvious that the primary challenge we are faced with is to obtain proper understanding of reasoning. As someone fancifully stated, we are *thinking about thinking.* So, again, how do we do that? Can we assume that mathematics per se is the quintessence of precision, rigor, and truth, thus to reason soundly means to reason mathematically, and vice versa, to understand mathematics we need to reason soundly? If all of this sound too "l'art pour l'art-tistic" to you, so be it. Still, the importance of "it all" is undeniable. Certainly nobody doubts the importance of applications of mathematics. The precision and value of (mathematical) proofs present itself as a *conditio sine qua non* of any claim aspiring to be verifiably accurate. Of course, you say, that is something to be expected – mathematical results are certain, mathematics is objective. However, a mathematician's ambitions, like a philosopher's, go beyond that – s/he is interested in *truth(s).* Again, I won't even attempt to answer what the truth is. I leave it to the reader to ponder. I would just timidly offer *that it is not the case that only provable is true. But it is the case that provable is true.* Now, going back to "*mathematicians prove theorems*", by a *theorem* we mean a proposition/statement that is known (proven) to be true. And *proof* is a justification of the truth of a statement making it a theorem. Aha, say you, here we go again – what is "*true*?" I better leave this discussion abruptly and show you some rather revealing examples. (Actually, all the examples that follow, the whole chapter for that matter, may be considered as additional exercise in proving statements.) First, two "simple" ones; the first one, being not exactly a "proof" is here just as a warm up for the things that follow.

[10]"Behind it all is surely an idea so simple, so beautiful, that when we grasp it – in a decade, a century, or a millennium – we will all say to each other, how could it have been otherwise." Wheeler, J. A. New Techniques and Ideas in Quantum Measurement Theory, pages xv-xvii, *Annals of the New York Academy of Sciences*, 480, 1986.

Example 3.4 Prove that

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}$$

Proof Consider the following "sum"

$$\begin{aligned} \Sigma &= 1 + x + x^2 + x^3 + \cdots \qquad (*) \\ &= 1 + x(1 + x + x^3 + \cdots) \\ &= 1 + x \cdot \Sigma \end{aligned}$$

Thus,

$$\Sigma(1 - x) = 1$$

which, assuming $x \neq 1$ gives us (we've proved(?))

$$\Sigma = \frac{1}{1-x} \qquad (**)$$

So far so good. Now, let's examine our sum. First, you would agree that (*) and (**) represent the same sum. Well, if we put $x = -1$ in (*) the sum is

$$\Sigma = 1 - 1 + 1 - 1 + 1 - \cdots \qquad (***)$$

Therefore, most people would agree that $\Sigma = 0$, right?

However, consider this

$$\begin{aligned} 0 &= 0 + 0 + \cdots + 0 + \cdots \\ &= (1 - 1) + (1 - 1) + \cdots + (1 - 1) + \cdots \qquad (****) \\ &= 1 - (1 - 1) - (1 - 1) - \cdots - (1 - 1) - \cdots \\ &= 1 - 0 - 0 - \cdots - 0 \cdots = 1 \end{aligned}$$

Now, you would agree that (***) is equal to (****), wouldn't you? But certainly you cannot accept that $0 = 1$!

Now, substitute again $x = -1$ into (**) and it turns out that $\Sigma = 1/2$. So, what did we prove? Or, take $x = -1/2$. Equation (*) gives you

$$\Sigma = 1 - \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \cdots$$

while (**) yields $\Sigma = 2/3$. So, does the sum Σ equals 2/3? You can now experiment on your own and try, say, $x = 1/2$. Which of the results would you take as "true?" ■

How about if our sum has a finite number of terms? Consider

Example 3.5 Take the sum as in the previous example, but this time with a finite number of terms, that is

$$S_n = 1 + x + x^2 + \cdots + x^n, \quad n \in \mathbf{N} \tag{*}$$

and find the formula for S_n.

Solution Let's multiply both sides of (*) by x:

$$xS_n = x + x^2 + x^3 + \cdots + x^n + x^{n+1}$$

and consider the difference

$$\begin{aligned} xS_n - S_n &= (x-1)S_n \\ &= (x + x^2 + x^3 + \cdots + x^n + x^{n+1}) - (1 + x + x^2 + \cdots + x^n) \\ &= x^{n+1} - 1 \end{aligned}$$

Thus,

$$S_n = \frac{x^{n+1} - 1}{x - 1}, \quad \forall x \neq 1$$ ■

This looks pretty reasonable. What difference infinity can make!
Let's try one more

Example 3.6 Take the final result from the previous problem

$$S_n = \frac{x^{n+1} - 1}{x - 1}$$

and rewrite it as

$$S_n = \frac{x^{n+1}}{x - 1} - \frac{1}{x - 1}$$

Reshuffling things around a little we get

$$\frac{1}{1-x} = \underbrace{1 + x + x^2 + \cdots + x^n}_{S_n} + \frac{x^{n+1}}{1-x} \tag{*}$$

Thus, taking the results from Examples 3.4–3.6 all together

$$\Sigma = S_n + \frac{x^{n+1}}{1-x} \tag{**}$$

Would you say that the difference between Σ and S_n is $\frac{x^{n+1}}{(1-x)}$? If all this looks too confusing take, say, $x = 2$. The left-hand side of (*) is then well defined and equals -1, whereas the right-hand side equals

$$1 + 2 + 2^2 + \cdots + 2^n + \frac{2^{n+1}}{1-2}$$

So we have

$$2^{n+1} - 1 = 1 + 2 + 2^2 + 2^3 + \cdots + 2^n$$

And so on, and so forth. You can experiment some more on your own, and maybe at some point you can make sense of what's going on here. ■

Here is another neat example for which no special knowledge of number theory is required, only a little trick is sufficient to reveal its rather nontrivial nature.

Example 3.7 Show that $0.\overline{9} = 0.99999\ \ldots\ = 1$.

Solution Let $x = 0.\overline{9}$. Then,

$$10x = 9.\overline{9}$$

and thus,

$$\begin{aligned} 10x - x &= 9 \\ 9x &= 9 \\ x &= 1 \end{aligned}$$

So, $0.\overline{9} = 1$. ■

Now it is a little bit more clear what is meant by the "*proof*." Nevertheless, let's state it formally.

Definition 3.6 A **proof** is a logical argument that establishes (in finite number of steps) the truth of a statement beyond any doubt.

More precisely,

Definition 3.7 A **theorem** is a mathematical statement for which the truth can be (has been) established in a finite number of steps using logical reasoning.

The next example is a classical theorem that some might consider prematurely brought up at this introduction. However, its proof contains some subtleties that you, as some mathematicians do, may find unsettling; so, why not incite your curiosity even more at the beginning by considering something rather nontrivial.

Theorem 3.1 There are irrational numbers x, y, such that x^y is rational.

Proof Consider $(\sqrt{2})^{\sqrt{2}}$. We know that $\sqrt{2}$ is irrational (we have proved it in Chapter 1), so if $(\sqrt{2})^{\sqrt{2}}$ is rational there is nothing to prove, we have our result. However, if $(\sqrt{2})^{\sqrt{2}}$ is irrational, then $((\sqrt{2})^{\sqrt{2}})^{\sqrt{2}} = (\sqrt{2})^2 = 2$, which is rational, and we have our proof. So, how do you feel about this proof? Do you find anything troubling about it? Note that in our proof we have relied on the law of excluded middle (LEM): Either $(\sqrt{2})^{\sqrt{2}}$ is rational or not. If it is – fine, we are done. If not, that is, if $(\sqrt{2})^{\sqrt{2}}$ is irrational, then $((\sqrt{2})^{\sqrt{2}})^{\sqrt{2}}$ certainly is. Thus the proof. ■

Although I would consider the above proof legitimate, I sympathize with those who are troubled by the "nature" of $(\sqrt{2})^{\sqrt{2}}$. Namely, they say, since we cannot construct in finite number of steps $(\sqrt{2})^{\sqrt{2}}$, we cannot assume that $(\sqrt{2})^{\sqrt{2}}$ is either rational or irrational, and that assumption is crucial for the proof. Therefore, the above "proof" is not *the* proof.

Now you may want to revisit the proof that $\sqrt{2} \notin \mathbf{Q}$ (or any other proof in the previous chapters) and compare it with the proof above. It might give you a sense, a "flavor," of things we want to study in this section.

And finally, the most intriguing of all, the so-called Banach–Tarski Theorem (Paradox), which we will have more to say about later. For now just a tease: it has been "proved" that if one accepts one of the most important axioms of mathematics, *The Axiom of Choice* (which most mathematicians do; see Chapter 1), then the following can be done: Take **any** two spheres S_1 and S_2 in our three-dimensional Euclidean space (any two means: one, say, S_1 can be very small, the size of a pea, and the other, S_2, as big as Jupiter). You can decompose S_1 into finitely many parts and, with proper manipulations of the parts, recompose it back so that it can fill out S_2 – the Jupiter. Speaking of two peas in a pot. Crazy?! Sure. Beautiful?! Absolutely!

Leaving for the time being the issue of empirical verification, you may be willing to accept the claim that mathematics is a (special) kind of science. Contemplating the questions of proofs we are evidently faced with that "artistic" nature of mathematics. The idea might not look so farfetched any more. Think about it! Given the *abstract* concepts A, B, C, … you establish (you discover) that D must follow. The hidden, unexpected, relationship between A, B, C, and D reveals itself. How did that come about? Admittedly, some "common sense" faculties you are in possession of are critical (leaving aside for the time been what that means exactly). Moreover, certain (mathematical) tools and techniques you have acquired through your formal training in school, or otherwise, are indispensable. And, as you might have anticipated, there is more, much more. Indeed, proof theory is a beautiful and rather complex mathematical discipline. We will, however, discuss only the examples from elementary number theory.

> Mathematical proofs, like diamonds, are hard as well as clear, and will be touched with nothing but strict reasoning.
>
> John Locke

3.2 DIRECT PROOF

In this section, we discuss one type of universal statements that occur very frequently in mathematics, logic, and, you may say, in every reasonable discourse. It is a method of generalizing from generic to particular, that is, the arguments follow directly from the hypothesis to the conclusion. We take an arbitrary element x from a domain $\mathcal{D}$ and try to show that it satisfies a certain expected property of the form "If $P(x)$ then $Q(x)$." Recall that the conditional $P(x) \rightarrow Q(x)$ is false only in case when $P(x)$ is true and $Q(x)$ is false. Hence, we argue, assuming $P(x)$ (the antecedent) is true for every x from $\mathcal{D}$, one needs to prove that $Q(x)$ (the consequent) is also true, which would furthermore yield that $P(x) \rightarrow Q(x)$ is true.

In other words, by direct proof we show that the universal statement holds regardless of the size of the domain over which the statement is quantified, that is, we are proving the proposition

$$\forall x \in \mathcal{D}, (P(x) \rightarrow Q(x)) \qquad (*)$$

So, we say: Suppose $P(x)$ is true for a particular but arbitrarily chosen $x \in \mathcal{D}$. We need to show that $Q(x)$ is true.

Example 3.8 As a simple conceptual illustration of a distinction between *direct proofs* and *indirect proofs* (of the next section) let's prove a statement that we recall from Chapter 1, namely: *There exist a unique multiplicative identity in the set* $\mathbf{R}$.

Proof **(Direct)** First, recall one of the properties of $\mathbf{R}$:

$\forall x \in \mathbf{R}$, $\exists\, e \in \mathbf{R}$, such that $e \cdot x = x \cdot e = x$

Suppose there exist another element of $\mathbf{R}$, say, e' such that $e' \cdot x = x \cdot e' = x$, $\forall x \in \mathbf{R}$. Then, since both claims involve universal statement ($\forall x \in \mathbf{R}$) we have

$$e \cdot e' = e' \cdot e = e'$$

but also

$$e' \cdot e = e \cdot e' = e$$

Thus, any two multiplicative identities in $\mathbf{R}$ must be equal. We say $e = e' (= 1)$. ■

Proof **(Indirect)** Suppose, contrary to the claim, that there is no unique multiplicative identity in **R**. Then, either there is more than one multiplicative identity or there is no multiplicative identity at all. However,

$$1 \cdot x = x \cdot 1 = x, \quad \forall x \in \mathbf{R}$$

thus, one identity exists so there are more than one. Suppose $e \in \mathbf{R}$ is another identity. Then,

$$1 \cdot e = e \cdot 1 = e$$

and also

$$e \cdot 1 = 1 \cdot e = 1$$

Thus, $e = 1$. This contradicts our assumption that there is more than one multiplicative identity.

We will have to say more about indirect proof shortly. ■

Example 3.9 Let A, B, and C, be sets such that $|A| = 1$, $A \cap B \neq \emptyset$, and $A \cap C \neq \emptyset$. Then,

$$B \cap C \neq \emptyset$$

Proof Since $|A| = 1$ there exists $x \in A$, and since $A \cap B \neq \emptyset$, $x \in A \cap B$. Formally,

$$\exists x(x \in A \wedge x \in B)$$

Similarly, since $A \cap C \neq \emptyset$,

$$\exists x(x \in A \wedge x \in C)$$

So, $x \in B$ and $x \in C$ implies $x \in B \cap C$. Thus, $B \cap C \neq \emptyset$, as claimed, and we write

$$(\exists x(x \in A \ \wedge \ x \in B) \wedge \ \exists x(x \in A \ \wedge \ x \in C)) \rightarrow (x \in B \cap C)$$

■

For the next example, we need to recall

Definition 3.8 We say that an integer x is **even** iff $x = 2k$, for some integer k, that is

$$x \text{ is even iff } \exists k \in \mathbf{Z}, \text{ s.t. } x = 2k$$

We say that an integer x is **odd** iff $x = 2k + 1$, for some integer k, that is

$$x \text{ is odd iff } \exists k \in \mathbf{Z}, \text{ s.t. } x = 2k + 1$$

The next example, although rather trivial, illustrates "the way of thought" and can be thought of as a segue to several proofs that follow.

Example 3.10 Prove that if $x \in \mathbf{Z}$ is odd, then $x + 1$ is even.

Proof Obviously the argument is of the form

$$P(x) \rightarrow Q(x)$$

where, of course, $P(x)$ stands for "*x is an odd integer*" and $Q(x)$ for "*$x + 1$ is an even integer*."

Suppose now that $x \in \mathbf{Z}$ is an odd integer. By definition, the integer x is odd if $x = 2k + 1$ for some $k \in \mathbf{Z}$. Using this definition, we are in a position to handle our statements $P(x)$ and $Q(x)$ by an equation, and we write

$$\begin{aligned} x + 1 &= (2k + 1) + 1 \\ &= 2(k + 1) \end{aligned}$$

Now, $k + 1$ is also an integer (the sum of two integers is an integer), thus $2(k + 1)$ is an even integer (by definition of an even integer). We have deduced $Q(x)$. ■

The following examples are similar.

Example 3.11 Prove the following proposition:
For all $x, y \in \mathbf{Z}$, if x and y are odd, then $x + y$ is an even integer.

Proof Let $x, y \in \mathbf{Z}$, and let $x = 2k + 1$, $y = 2l + 1$, with $k, l \in \mathbf{Z}$. Then,

$$x + y = 2k + 1 + 2l + 1 = 2k + 2l + 2 = 2(k + l + 1)$$

Now, since $(k + l + 1)$ is obviously an element of $\mathbf{Z}$, we see that $x + y$ is indeed an even integer. ■

Example 3.12 Prove the following proposition:
For all $x, y \in \mathbf{Z}$, if the sum $x + y$ is even, so is the difference $x - y$.

Proof Let $x, y \in \mathbf{Z}$, such that $x + y = 2k, k \in \mathbf{Z}$. It follows that $x = 2k - y$. Consider now

$$\begin{aligned} x - y &= (2k - y) - y \\ &= 2k - 2y \\ &= 2(k - y) \end{aligned}$$

Since a difference of two integers $k - y$ is an integer, $2(k - y)$ is an even integer, that is, $x - y$ is even, as claimed. ■

Example/Exercise 3.13 Convince yourself that zero is an even number.

Example 3.14 Prove that the sum of two consecutive integers is an odd integer.

Proof Take any $n, (n + 1) \in \mathbf{Z}$. Obviously, if n is even then $n + 1$ is odd, and vice versa. Without loss of generality let's take

$$n = 2k, \quad k \in \mathbf{Z}$$

Then,

$$\begin{aligned} n + (n + 1) &= 2k + 2k + 1 \\ &= 4k + 1 \\ &= 2(2k) + 1 \\ &= 2l + 1, \quad l = 2k \in \mathbf{Z} \end{aligned}$$

So, the sum of two consecutive integers is indeed an odd integer. ■

Now you shouldn't have any difficulty proving the statement in the following two examples.

Example/Exercise 3.15 The square of an even number is an even number.

Example/Exercise 3.16 Show that if $n \in \mathbf{Z}$ is an even integer, then $3n^5$ is an even integer too.

Example 3.17 Show that if $n \in \mathbf{Z}$ is an odd integer, then n^2 is of the form $8m + 1$, where $m \in \mathbf{Z}$.

Solution Suppose n is an odd integer. Then, there exists $k \in \mathbf{Z}$ such that $n = 2k + 1$. So,

$$\begin{aligned} n^2 = (2k + 1)^2 &= 4k^2 + 4k + 1 \\ &= 4k(k + 1) + 1 \end{aligned}$$

k and $k + 1$ are obviously two consecutive integers; thus, one must be even and the other odd. So, if k is even, that is, $k = 2i, (i \in \mathbf{Z})$, we have

$$n^2 = 4k(k + 1) + 1$$

$$= 4(2i)(2i+1)+1$$
$$= 8[i(2i+1)]+1$$
$$= 8m+1$$

where $m = i(2i+1)$.

If k is odd, that is, $k = 2i+1, (i \in \mathbf{Z})$, we have

$$n^2 = 4k(k+1)+1$$
$$= 4(2i+1)(2i+1+1)+1$$
$$= 8[(i+1)(2i+1)]+1$$
$$= 8m+1$$

Where, this time, $m = (i+1)(2i+1)$.

Hence, in both cases, n^2 is of the form $8m+1$. ■

Example 3.18 Show that given a set $S = \{1, 2, 3\}$, and $n \in S$, if

$$\frac{n(n+3)}{2} \text{ is even, then } \frac{(n+2)(n-5)}{2} \text{ is even}$$

Solution First, observe that the only elements of S we need to consider are those that make $(n(n+3))/2$ even. Fortunately, the set S has only three elements; so, rather than assuming that the statement is true for an arbitrary element of S, we can find the right elements by direct inspection: 1 is the only element with desired property. By means of direct proof, we see immediately that the statement is true. ■

Example/Exercise 3.19 Revisit Theorem 1.20.

Example/Exercise 3.20 Revisit Theorem 1.21.

More intriguing is

Example 3.21 Show that if $n \in \mathbf{N}$ is a natural number with two or more digits, then the product of the digits is less than n.

Solution If n is a natural number, we can write it as

$$n = d_k 10^k + d_{k-1} 10^{k-1} + \cdots + d_1 10 + d_0$$

where d_i $(i = 0, 1, 2, \ldots, k)$, is the digit in the 10^ith place of n. Thus, the product of the digits od n is

$$d_k \cdot d_{k-1} \cdot \ldots \cdot d_0$$

Note that the product of k factors

$$d_{k-1} \cdot \ldots \cdot d_0 < 10^k, \quad k \geq 1$$

Therefore,

$$\begin{aligned} d_k(d_{k-1} \cdot \ldots \cdot d_0) &< d_k 10^k \\ &\leq d_k 10^k + d_{k-1}10^{k-1} + \cdots + d_1 10 + d_0 = n \end{aligned}$$

■

Let's now look at somewhat more intricate proofs of the form

$$P(x) \rightarrow (Q_1(x) \vee Q_2(x))$$

or

$$(P_1(x) \vee P_2(x)) \rightarrow Q(x)$$

Similarly, we may ponder the proofs of the form

$$P(x) \rightarrow (Q_1(x) \wedge Q_2(x))$$

or

$$(P_1(x) \wedge P_2(x)) \rightarrow Q(x)$$

Example 3.22 Let $x \in \mathbf{R}$ be any real number. Prove that $-|x| \leq x \leq |x|$.

Proof First, recall the definition of the absolute value of $x \in \mathbf{R}$:

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

We have to consider two cases:

(i) Suppose $x \geq 0$. Then, $|x| = x$. So, $-x \leq x$. Thus, $-x \leq x \leq x$, that is

$$-|x| \leq x \leq |x|$$

(ii) However, if $x < 0$, then $|x| = -x$. So, $x \leq -x$ and we have $x \leq x \leq -x$ or $-(-x) \leq x \leq -x$, that is

$$-|x| \leq x \leq |x|$$

Hence, we have our proof for both cases. ■

Example 3.23 Prove that if $x, y \in \mathbf{R}$ then, $|x - y| \geq |x| - |y|$

Solution Let $x, y \in \mathbf{R}$. Consider

$$\begin{aligned} |x| &= |x + 0| = |x + (-y + y)| \\ &= |(x - y) + y| \leq |x - y| + |y| \end{aligned}$$

Then,

$$|x| - |y| \leq |x - y|$$

that is

$$|x - y| \geq |x| - |y|, \quad \forall x, y \in \mathbf{R}$$

■

Example 3.24 Prove that if $x^2 \leq 1, x \in \mathbf{R}$, then, $x^2 - 7x > -10$.

Solution Let's rewrite $x^2 - 7x > -10$ as

$$x^2 - 7x + 10 > 0$$

that is

$$(x - 5)(x - 2) > 0$$

which is true if both $(x - 5)$ and $(x - 2)$ are positive or negative.

From our assumption that $x^2 \leq 1$, that is, $-1 \leq x \leq 1$, it follows

$$x - 5 < 0 \quad \text{and} \quad x - 2 < 0$$

Therefore,

$$(x - 5)(x - 2) > 0$$

Thus,

$$x^2 - 7x + 10 > 0$$

In other words,

$$x^2 - 7x > -10$$

as claimed. ■

Before proceeding, it may be useful to recall two equivalences from Chapter 2 as rather convenient tools when dealing with proofs of the form

$$P(x) \rightarrow (Q_1(x) \vee Q_2(x))$$

First, let's verify the following two equivalences:

(i) $P \to (Q_1 \vee Q_2) \equiv (P \wedge \sim Q_1) \to Q_2$
Observe that

$$\begin{aligned} P \to (Q_1 \vee Q_2) &\equiv \sim P \vee (Q_1 \vee Q_2) \\ &\equiv \sim P \vee Q_1 \vee Q_2 \end{aligned}$$

However,

$$\begin{aligned} (P \wedge \sim Q_1) \to Q_2 &\equiv \sim (P \wedge \sim Q_1) \vee Q_2 \\ &\equiv \sim P \vee Q_1 \vee Q_2 \end{aligned}$$

So, the equivalence (i) holds.

Next, consider

(ii) $P \to (Q_1 \vee Q_2) \equiv (P \wedge (\sim Q_2)) \to Q_1$
Indeed,

$$P \to (Q_1 \vee Q_2) \equiv \sim P \vee Q_1 \vee Q_2$$

However,

$$\begin{aligned} (P \wedge (\sim Q_2)) \to Q_1 &\equiv \sim (P \wedge (\sim Q_2)) \vee Q_1 \\ &\equiv \sim P \vee Q_2 \vee Q_1 \end{aligned}$$

Thus, the equivalence (ii) is also satisfied.

Example/Exercise 3.25 You may want to revisit the proof of the distributive law of "$\cup$" over "$\cap$" of three sets, in Chapter 1, and see how it fits our formal logical analysis of proofs.

Example 3.26 Prove that for all finite sets X and Y, if $|X| = m, |Y| = n$ and $|X \cap Y| = k$ then,

$$|X \cup Y| = m + n - k$$

Proof Let X and Y be sets such that $|X| = m$ and $|Y| = n$.

We can split the set X into two parts:

$$X \cap Y = \{x \in X | x \in Y\}$$

and

$$X \setminus Y = \{x \in X | x \notin Y\}$$

In other words,

$$X = (X \cap Y) \cup (X \setminus Y)$$

Similarly, we can split Y into

$$Y \cap X = \{y \in Y \mid y \in X\}$$

and

$$Y \setminus X = \{y \in Y | y \notin X\}$$

that is

$$\begin{aligned} Y &= (Y \cap X) \cup (Y \setminus X) \\ &= (X \cap Y) \cup (Y \setminus X) \end{aligned}$$

Since $|X \cap Y| = k$ and $|X| = m$

$$\begin{aligned} |X \setminus Y| &= |X| - |X \cap Y| \\ &= m - k \end{aligned}$$

Moreover,

$$\begin{aligned} |Y \setminus X| &= |Y| - |X \cap Y| \\ &= n - k \end{aligned}$$

Consider now

$$\begin{aligned} X \cup Y &= \{x | x \in X \text{ or } x \in Y\} \\ &= (X \setminus Y) \cup (X \cap Y) \cup (Y \setminus X) \end{aligned}$$

Thus,

$$\begin{aligned} |X \cup Y| &= |X \setminus Y| + |X \cap Y| + |Y \setminus X| \\ &= m - k + k + n - k \\ &= m + n - k \end{aligned}$$

as was to be shown. ■

In the gist of the previous discussion, it may be the right time to justify the so often used phrase "*if and only if*." We start with

Example 3.27 Recall, in Chapter 1, we defined equality of two sets A and B by saying: $A = B$ iff $A \subseteq B$ and $B \subseteq A$. By saying $A \subseteq B$, we are actually saying: if $x \in A$, then $x \in B$, that is, we are stating conditional: $P \to Q$. Similarly, $B \subseteq A$ means: if $x \in B$ then, $x \in A$, that is, we have another conditional: $Q \to P$. Thus,

$(P \to Q) \wedge (Q \to P) \equiv P \leftrightarrow Q$ is exactly our "iff" in the definition of equality of two sets.

Simply put, the (obvious) rule of thumb for proving $P \leftrightarrow Q$ is this:

(i) Show $P \to Q$ by any method.

(ii) Show $Q \to P$ by any method.

Therefore, $P \leftrightarrow Q$.

The proofs of the following two theorems illustrate this.

Theorem 3.2 The integer x is odd iff x^2 is odd.

Proof Easy! Note that the first part of the "if and only if" statement above has already been proved in Example 3.17. You only need to prove the "only if" part. ■

Theorem 3.3 An integer is odd iff it may be written as the sum of two consecutive integers.

Proof Suppose $n \in \mathbf{Z}$ is an odd integer. Then, we can write $n = 2k + 1$ for some $k \in \mathbf{Z}$. Since k and $k + 1$ are two consecutive integers, $n = k + (k + 1)$ is the sum of two consecutive integers indeed (cf. Example 3.14). ■

As prime numbers are ubiquitous in many different fields of mathematics we want to define them more formally.

Definition 3.9 A natural number $p > 1$ is **prime** iff for all positive integers r and s if

$p = r \cdot s$, then $r = 1$ or $s = 1$, that is

$p \in \mathbf{N}$ is prime iff $\forall r, s \in \mathbf{N}, (p = r \cdot s) \to ((r = 1) \vee (s = 1))$

In other words, a natural number $p > 1$ is **prime** iff the only positive divisors of p are 1 and p.

Consequently, an integer is **composite** iff $p = r \cdot s$ for some positive integers r and s, such that $r \neq 1$ and $s \neq 1$, that is, $p \in \mathbf{N}$ is composite iff $\exists\, r, s \in \mathbf{N}$, $r, s > 0$, s.t. $(p = r \cdot s) \wedge (r \neq 1 \wedge s \neq 1)$.

Example/Exercise 3.28 Prove/disprove: For all $x \in \mathbf{Z}$, if $x > 2$, then $x^2 - 4$ is composite.

Definition 3.10 Let $n, d \in \mathbf{Z}$, and let $d \neq 0$. We say that n is **divisible by** d, or that d **divides** n iff $n = d \cdot k$, for some $k \in \mathbf{Z}$, and we write $d|n$. Symbolically,

$$d|n \quad \text{iff} \quad \exists k \in \mathbf{Z}, \ \text{s.t.} \ \ n = d \cdot k$$

Negating the existential statement in the definition above, we get

Definition 3.11 For all $n, d \in \mathbf{Z}, d \neq 0$, we say that d does not divide n, and we write $d \nmid n$ iff n/d is not an integer.

Example 3.29 Let $x, y \in \mathbf{N}$, such that $x|y$. Prove that $x \leq y$.

Proof If $x|y$, then $\exists k \in \mathbf{Z}$, such that $y = k \cdot x$. Since $x, y \in \mathbf{N}$, k has to be a positive integer too. Thus, $1 \leq k$. It follows then

$$x \leq k \cdot x = y$$

So, yes, $x \leq y$. ■

To proceed further, we need to digress a little and remind ourselves of a few additional concepts.

Definition 3.12 A binary relation "$\leq$" on a set X is a **partial ordering** if it is reflexive, transitive, and antisymmetric, that is

$$\forall x, y, z \in X$$

(i) $x \leq x$
(ii) $(x \leq y \wedge y \leq z) \Rightarrow x \leq z$
(iii) $(x \leq y \wedge y \leq x) \Rightarrow x = y$

A **partially ordered set** or **poset** is a set with a partial order on it.

Definition 3.13 A partial ordering "$\leq$" is called a **total ordering** (sometimes **linear**) if, in addition of being a partial ordering, any two elements of X are **comparable** in terms of "$\leq$," that is

$$(\forall x, y \in X)((x < y) \vee (x = y) \vee (y < x))$$

A **totally ordered set** or **linearly ordered set** is a set with total order on it.

Definition 3.14 A binary relation "$\leq$" on X is a **well-ordering** on X, if it is a total ordering on X and if every nonempty subset of X has a least element, that is

$$(\forall Y \subseteq X)((Y \neq \varnothing \Rightarrow (\exists x \in Y) \ni (\forall y \in Y, x \leq y)))$$

A **well-ordered set** is a set with well ordering on it.

Theorem 3.4 If X is well-ordered set then

(i) Every subset of X is well ordered.
(ii) If there exists an order preserving bijection $X \xrightarrow{f} Y$, then Y is well ordered.

This theorem, and the corresponding definitions above, agrees with our intuitive notion of integers and natural numbers. Consequently, we also have

Theorem 3.5 For every integer $n \in \mathbf{Z}$, the set

$$X = \{x \in \mathbf{Z} | x \geq n\}$$

is well ordered.

Now we can restate the definitions of the set of integers as follows:

Definition 3.15 The set $\mathbf{Z}$, which is discrete and totally ordered, without the first and last element and with addition and multiplication defined the usual way (cf. Chapter 1), is said to be the set of integers.

Example 3.30 Consider set $\mathbf{N} \subseteq \mathbf{Z}$. The order relation "$\leq$" is defined by the following equivalence:

$$(\forall n, m \in \mathbf{N}, n \leq m) \Leftrightarrow ((\exists s \in \mathbf{N}) \text{ such that } n + s = m)$$

Predictably, the **well-ordering principle (WOP) for natural numbers (nonnegative integers)** is:

Definition 3.16 We say that a set $X \subseteq \mathbf{N}$, ($X \subseteq \mathbf{Z}^+$) has a least element if $X \neq \varnothing$ and all elements of X are greater than some fixed number, that is, *the set of natural numbers is well ordered.*

Now it is not difficult to prove

Theorem 3.6 If a set $X \subseteq \mathbf{R}$ has a least element, then that element is unique.

Proof Let l_1 and l_2 be the least elements of X. Since l_1 is a least element it has to be less than, or equal to, any element of X, in particular, $l_1 \leq l_2$. But m_2 is also a least element, so $l_1 \geq l_2$. Thus, $l_1 = l_2$. ■

Let's illustrate the efficacy of the fact *that every nonempty set of natural numbers (nonnegative integers) has a least element.*

Example 3.31 With $\mathbf{N} \subseteq \mathbf{Z}$, we recall the following:

(i) $\mathbf{N}$ has a first element, 1

(ii) $\forall n \in \mathbf{N}, \exists (n + 1) \in \mathbf{N}$, a unique successor

Furthermore, let X be the following subset of $\mathbf{Z}$:

$$X = \{x |\ x \geq n,\ n \in \mathbf{Z}\}$$

and let $S \subseteq X$ be such that

(iii) $n \in S$

(iv) $\forall x \in X$ if $x \in S$ then $x + 1 \in S$
Then, $X = S$

If $n = 1$ then $S = \mathbf{N}$.

Theorem 3.7 (The division algorithm) Let $a, b \in \mathbf{N}$ be natural numbers. Then, there exist unique integers $q, r \in \mathbf{Z}$ with $0 \leq r < a$, such that $b = aq + r$.

Proof Consider a set

$$X = \{b - ax | x \in \mathbf{Z},\ b - ax \geq 0\}$$

Note that $X \neq \varnothing$ (since, even if we let $x = 0$, we still have $b \in X$). By Theorem 3.5, X has the smallest element $r \geq 0$. But, since $r \in X$, there exists some $q \in \mathbf{Z}$ such that $r = b - aq$. Thus, $b = aq + r$. We need to show that $r < a$. Suppose not, that is, suppose $r \geq a$, and let $r - a = t$. Then, of course, $t \geq 0$. By our hypothesis $a > 0$, thus $t < r$. So, we have

$$\begin{aligned} t = r - a &= (b - aq) - a \\ &= b - (aq + a) \\ &= b - a(q + 1) \end{aligned}$$

That means that $t \in X$, which cannot be since we have assumed that r is the smallest element of X. Therefore, $r < a$. Finally, we need to show uniqueness, that is, we need to show that q and r are the only integers satisfying

$$b = aq + r \quad \text{with} \quad 0 \leq r < a$$

Suppose not. Suppose that there exist $q', r' \in \mathbf{Z}$, such that

$$b = aq' + r' \quad \text{with} \quad 0 \leq r' < a$$

and, without the loss of generality, let's assume that $r' > r$. Then,

$$aq + r = aq' + r'$$

from which follows that

$$a(q - q') = r' - r$$

Now, since $q - q' \in \mathbf{Z}, a|(r' - r)$. However, since $r' > r$, and $0 < r' - r < a$, we must have $r' - r = 0$, that is, $r' = r$. Furthermore, since $a \neq 0$, we have $a(q - q') = 0$, hence $q = q'$. Thus the proof. ■

Theorem 3.8 For all $x, y, w \in \mathbf{Z}$, if $x|y$ and $y|w$ then $x|w$.

Proof Let $x, y, w \in \mathbf{Z}$, such that $x|y$ and $y|w$, that is

$$y = k \cdot x \text{ and } w = l \cdot y, \quad k, l \in \mathbf{Z}$$

So, we have

$$\begin{aligned} w &= l \cdot y \\ &= l \cdot k \cdot x \\ &= n \cdot x \end{aligned}$$

where $n = l \cdot k$. Since a product of two integers is also an integer, we conclude that $x|w$, as was to be shown. ■

Example 3.32 Show that if, $x, y \in \mathbf{Z}$ and $x|y$, then $x|y^2$.

Solution Let $x, y \in \mathbf{Z}$, and suppose that $x|y$. Then, $y = xd$ for some $d \in \mathbf{Z}$.

Consider $y^2 = (xd)^2 = x^2d^2 = x(xd^2) = xd'$, $d' = xd^2 \in \mathbf{Z}$. So $x|y^2$, as claimed. ■

Example 3.33 Show that if $x, y, w \in \mathbf{Z}$, and $x|y$ and $x|w$, then $x|(y + w)$.

Solution Let $x, y, w \in \mathbf{Z}$. Suppose $x|y$ and $x|w$. That means that $y = xa$ and $w = xb$ for some $a, b \in \mathbf{Z}$. Now, consider

$$\begin{aligned} y + w &= xa + xb \\ &= x(a + b) \\ &= xc, \quad \text{where } c = a + b \in \mathbf{Z} \end{aligned}$$

Thus, $x|(y + w)$, as claimed. ■

Example/Exercise 3.34 Show that if $x, y, w \in \mathbf{Z}$, and $x|y$ and $x|w$, then $x|(ay + bw)$, $\forall a, b \in \mathbf{Z}$

Example 3.35 Let $x, y, w \in \mathbf{Z}$. Show that if $x|y$ and $x|(y + w)$, then $x|w$.

Solution Let $x, y, w \in \mathbf{Z}$ such $x|y$ and $x|(y + w)$. That means that $y = xa$ and $y + w = xb$ for some $a, b \in \mathbf{Z}$.

From $y + w = xb$, it follows that $w = xb - y$, so we have

$$\begin{aligned} w &= xb - y = xb - xa \\ &= x(b - a) = xc \end{aligned}$$

where $c = b - a \in \mathbf{Z}$.

Hence, $x|(y + w)$ as claimed. ■

Definition 3.17 Let $x, y \in \mathbf{Z}$ be two integers, both not equal to zero. Then, the g*reatest common divisor* – $\boldsymbol{gcd(x, y)}$ of x and y is the greatest positive integer that divides both x and y.

Definition 3.18 Two integers $x, y \in \mathbf{Z}$ are said to be **relatively prime** if $gcd(x, y) = 1$.

Theorem 3.9 Let $x, y \in \mathbf{Z}$ such that both are not 0. Then, the g*reatest common divisor* of x and y is the least positive integer that is a linear combination of x and y.

Proof Let

$$S = \{ax + by | a, b \in \mathbf{Z},\ ax + by > 0\}$$

be a set of all positive integers that are linear combination of x and y. Since, at least one of x and y is different from zero,

$$x^2 + y^2 > 0$$

So, we can write

$$x^2 + y^2 = x \cdot x + y \cdot y > 0$$

Thus, $S \neq \varnothing$ and, evidently, $S \subseteq \mathbf{N}$. By WOP we know that S contains a least element, say, l. But then, there should exist $a_0, b_0 \in \mathbf{Z}$, such that

$$l = a_0 x + b_0 y$$

We would like to show that $l = \gcd(x, y)$. By Theorem 3.7, we have

$$x = lq + r, \quad 0 \leq r < l$$

Thus,

$$\begin{aligned} r = x - lq &= x - q(a_0 x + b_0 y) \\ &= (1 - qa_0)x + (-qb_0)y \end{aligned}$$

Clearly, r is a linear combination of x and y. Now, if $r > 0$ then, necessarily, $r \in S$. But that contradicts our hypothesis that l is the least element of S. Thus, $r = 0$, implying that $l | x$. Similarly, we show that $l | y$, and therefore l is a common divisor of x and y. It remains to be seen whether it is the gcd. Let $l_0 \in \mathbf{N}$ be another divisor of x and y. By Theorem 3.8 and Examples 3.33 and 3.34, l_0 divides every linear combination of x and y. In particular,

$$l_0 | l = a_0 x + b_0 y$$

Since both l_0 and l are positive, it follows that $l_0 \leq l$ and therefore $l = \gcd(x, y)$. ■

Theorem 3.10 Let $x, y \in \mathbf{Z}$, such that both are not equal to zero. Then, $l = \gcd(x, y)$ iff $l \in \mathbf{N}$ such that the following two conditions are satisfied:

(i) l is the common divisor of x and y.
(ii) If l_0 is any other common divisor of x and y then $l_0 | l$.

Proof Suppose that $l = gcd(x, y)$. Then, (i) is obviously satisfied. We need to prove (ii). Let l_0 be any other common divisor of x and y. Since $l = gcd(x, y)$, we can express it (cf. Theorem 3.8) as

$$l = a_0 x + b_0 y, \quad a_0, b_0 \in \mathbf{Z}$$

However, since $l_0 | x$ and $l_0 | y$, by Theorem 3.7 and Examples 3.33 and 3.34, $l_0 | a_0 x + b_0 y = l$. So we also proved (ii).

Conversely, suppose l satisfies (i) and (ii). We only need to prove that $l = gcd(x, y)$. Again, let $l_0 \in \mathbf{N}$ be any common divisor of x and y. Since l satisfies (ii) it follows that $l_0 | l$. However, both l and l_0 are positive, and so $l_0 \leq l$, which implies that $l = gcd(x, y)$. ■

Theorem 3.11 Let x be any positive integer, then $gcd(x, 0) = x$.

Proof Trivial. ■

Theorem 3.12 Let $x, y, w \in \mathbf{Z}$, such that $x \neq 0$. If $x | yw$ and $gcd(x, y) = 1$, then $x | w$.

Proof Suppose $x | yw$, that is, suppose $yw = k \cdot x$ for some $k \in \mathbf{Z}$. We assume that x and y are relatively prime (otherwise, there will be nothing to prove) so we know that there exist $a, b \in \mathbf{Z}$ such that $a \cdot x + b \cdot y = 1$. Consider

$$\begin{aligned} w &= w \cdot 1 = w(a \cdot x + b \cdot y) \\ &= x(aw) + (wy)b \\ &= x(aw) + (kx)b \\ &= x(aw + kb) \end{aligned}$$

But, since $(aw + kb) \in \mathbf{Z}$ it follows that $x | w$, what was to be shown. ■

The following exercise addresses the corollary of the previous theorem.

Example/Exercise 3.36 Prove that for $x, y \in \mathbf{Z}$ and p prime, if $p | xy$, then $p | x$ or $p | y$.

Theorem 3.13 Any natural number $n > 1$ is divisible by a prime number.

Proof Let $n \in \mathbf{N}$ be any number greater than 1. If n is prime, then it is divisible by itself. So, suppose that n is composite, then

$$n = k_0 \cdot l_0$$

where $k_0, l_0 \in \mathbf{N}$, such that $1 < k_0 < n$, and $1 < l_0 < n$.

It follows by definition of divisibility that $k_0|n$. If k_0 is prime, we are done. If not, then as any other composite, k_0 can be expressed in the form

$$k_0 = k_1 \cdot l_1$$

where $k_1, l_1 \in \mathbf{N}$, such that $1 < k_1 < k_0$ and $1 < l_1 < k_0$. We conclude that $k_1|k_0$. Since $k_0|n$, it follows from Theorem 3.7 that $k_1|n$. Again, if k_1 is prime, we are done. If not, we proceed similarly and express k_1 as

$$k_1 = k_2 \cdot l_2$$

We continue the process factoring the successive factors of x until we hit the prime factor, say, k_p. The process is obviously finite, since by definition $1 < k_i < k_{i+1}$, that is

$$1 < k_p < k_{p-1} < \cdots < k_0 < x$$

Hence, k_p is prime and divides x. ■

Theorem 3.14 Let $x, y \in \mathbf{Z}$ be integers, both not equal to zero. Then, $gcd(x, y) = 1$ iff there exist $a, b \in \mathbf{Z}$ such that $1 = ax + by$.

Proof If $gcd(x, y) = 1$, we are done. We have already seen in Theorem 3.12 that there exist $a, b \in \mathbf{Z}$, such that $1 = ax + by$. Let's assume that $x, y \in \mathbf{Z}$ are integers, both not equal to zero, for which there exist $a, b \in \mathbf{Z}$ such that $1 = ax + by$.

Theorem 3.12 revealed that $gcd(x, y)$ is the smallest positive integer which is a linear combination of x and y. By our assumption, 1 is a linear combination of x and y; therefore, $gcd(x, y) = 1$. ■

Theorem 3.15 Let x, y, w be three integers such that $x \neq 0$ and $y \neq 0$ are relatively prime. Then, if $x|w$ and $y|w$, then $xy|w$.

Proof If $x|w$ and $y|w$ then $\exists\, k, l \in \mathbf{Z}$, such that $w = k \cdot x$ and $w = l \cdot y$. However, since x and y are relatively prime, $\exists\, s, t \in \mathbf{Z}$ such that

$$1 = s \cdot x + t \cdot y$$

Multiplying by w we obtain

$$\begin{aligned} w &= w \cdot 1 \\ &= w(sx + ty) \\ &= w(sx) + w(ty) \\ &= (ly)(sx) + (kx)(ty) \\ &= xy(ls + kt) \end{aligned}$$

Since $(ls + kt) \in \mathbf{Z}$, it follows that $xy|w$. ■

Theorem 3.16 Let $x \in \mathbf{Z}$ be any integer and p any prime then, if $p|x$, then $p \nmid (x + 1)$.

Proof Suppose the opposite, that is, suppose that there exist an integer x and a prime$p$, such that $p|x$ and $p|(x + 1)$. By definition of divisibility, it follows that

$\exists\, a, b \in \mathbf{Z}$, such that $x = a \cdot p$, and $x + 1 = b \cdot p$

Obviously, we can write

$$\begin{aligned} 1 &= (x + 1) - x \\ &= b \cdot p - a \cdot p \\ &= (b - a)p \end{aligned}$$

$b - a$ is an integer, thus $p|1$. But $p > 1$ is prime, and the only integer divisors of 1 are 1 and (-1). We have a contradiction! Our assumption was wrong, thus the claim of the theorem is true. ■

Theorem 3.17 If $n \in \mathbf{Z}^+$, then $\sqrt{n}$ is a rational number iff $\sqrt{n}$ is an integer.

Proof If $\sqrt{n}$ is an integer there is nothing to prove – $\sqrt{n}$ is a rational number. We only need to prove the converse. So, suppose that there exists $n \in \mathbf{Z}^+$, such that $\sqrt{n}$ is a rational number but not an integer. Then $\sqrt{n} = p/q$, where $p, q \in \mathbf{Z}^+$. Without loss of generality, we may assume that p and q are relatively prime, that is, $gcd(p, q) = 1$. Since p/q is not an integer, $q \geq 2$, and furthermore, $n = p^2/q^2$, hence $p^2 = nq^2$. By Theorem 3.13, q has a prime factor, say, a and so $a|nq^2$. But that implies that $a|p^2$ and therefore $a|p$. So, we obtained that $a|q$ and $a|p$, which contradict our assumption that $gcd(p, q) = 1$. Thus the proof. ■

Theorem 3.18 If $n \in \mathbf{N}$ is composite, then n has a prime factor p such that $p \leq \sqrt{n}$.

Proof n, being composite, means that $n = xy$ such that $1 < x < n$ and $1 < y < n$. Let's say that $x \leq y$. Then, $x^2 \leq xy = n$, from which it follows that $x \leq \sqrt{n}$.

Since $x > 1$ there should be a prime factor of x, say, p. But, since p is a factor of x, and at the same time x is a factor of n, it follows that p is a factor of n as well. Hence, $p \leq x \leq \sqrt{n}$. ■

Theorem 3.19 For every $n \in \mathbf{N}$ there exists a sequence of n consecutive nonprime natural numbers.

Proof For this proof a bit of ingenuity is needed in order to choose "the right number" to start with. So, for some $n \in \mathbf{N}$ let's try $x = (n+1)! + 2$. We will show that none of the numbers

$$x, x+1, x+2, x+3, \ldots, x+(n-1)$$

is prime.

First, note that

$$\begin{aligned} x &= 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n+1) + 2 \\ &= 2 \cdot (1 \cdot 3 \cdot 4 \cdot \cdots \cdot (n+1) + 1) \end{aligned}$$

obviously is not prime. Similarly,

$$\begin{aligned} x+1 &= 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n+1) + 3 \\ &= 3 \cdot (1 \cdot 2 \cdot 4 \cdot \cdots \cdot (n+1) + 1) \end{aligned}$$

is not prime either. Continuing in the same manner we see that any number $x + i$, where $0 \leq i \leq n-1$, is not a prime. Indeed,

$$\begin{aligned} x+i &= 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n+1) + 2 + i \\ &= (i+2)(\, 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (i+1) \cdot (i+3) \cdot \cdots \cdot (n+1) + 1) \end{aligned}$$

So we obtained the sequence of n consecutive nonprime natural numbers. ■

Let's illustrate this with an example when $n = 100$.

Example 3.37 Show that there exists a sequence of a 100 consecutive natural numbers containing no primes.

Solution Consider the following 100 consecutive numbers:

$$101! + 2, \;\; 101! + 3, \;\; 101! + 4, \ldots, \;\; 101! + 101$$

Observe that both $101!$ and 2 contain a factor of two, thus $101! + 2$ is not prime. Similarly, $101!$ and 3 contain a factor of 3, and so $101! + 3$ is not prime either. Continuing in this fashion, we conclude that for every

$n \in \{2, 3, \ldots, 101\}$, $101! + n$ is not prime. Hence, we have obtained a sequence of a 100 nonprime natural numbers as claimed. ■

Theorem 3.20 Let $n \in \mathbf{N}$ be a nonprime number. Then, $2^n - 1$ is not prime.

Proof Since n is not prime, there are natural numbers $a < n$ and $b < n$, such that $n = ab$. Let

$$x = 2^b - 1$$

and

$$y = 1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}$$

Then,

$$\begin{aligned} xy &= (2^b - 1)(1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}) \\ &= 2^b(1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}) - (1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}) \\ &= (2^b + 2^{2b} + 2^{3b} + \cdots + 2^{ab}) - (1 + 2^b + 2^{2b} + \cdots + 2^{(a-1)b}) \\ &= 2^{ab} - 1 \\ &= 2^n - 1 \end{aligned}$$

Hence $2^n - 1$ is not prime. ■

Just for the fun of it, let's mention the famous

Marsenne[11]**Conjecture**: There are infinitely many primes of the form $2^p - 1$, where p is a prime number.

Proof I don't know it. ■

3.3 INDIRECT PROOF

Direct proofs seem to be the most "natural" proofs. Unfortunately, that is not the case. Notwithstanding the fact that the method "*if P then Q*" seems to be the most efficient way to establish the truth of a statement, often it does not work or, at least, it is not the most "economical" way to reach one's goal.

As you would recall, one of the characteristic of the two-valued logic is that either a proposition is true or its negation is true (i.e., $P \vee \sim P = \tau$) thus, disproving something by proving a counterexample may, occasionally, be the best way to go. Consider the proposition we have started with

$$\forall x \in \mathcal{D}, \quad (P(x) \rightarrow Q(x)) \qquad (*)$$

[11] Mari Marsenne (1588–1648).

To disprove the proposition (*) we need to prove that its negation is true. Since our proposition contains the universal statement "for all," in order to disprove it, it is sufficient to find at least one $x \in \mathcal{D}$ for which the negation of (*) is true. In other words, we need to prove

$$\begin{aligned}\exists\, x \in \mathcal{D},\ \sim(P(x) \to Q(x)) &= \sim(\sim P(x) \vee Q(x)) \\ &= P(x) \wedge \sim Q(x)\end{aligned}$$

That is, to disprove (*) we are looking for $x \in \mathcal{D}$ for which $P(x)$ is true and $Q(x)$ is false.

Example 3.38 Prove that the following claim is false:

For all $x \in \mathbf{R}$, if $x < x^2$, then $x^2 < x^3$.

Proof Clearly, in this case the domain $\mathcal{D}$ is the set of all real numbers $\mathbf{R}$, $P(x)$ stands for "$x < x^2$" and $Q(x)$ for "$x^2 < x^3$." So, our statement is in the form

$$\forall x \in \mathcal{D}, (P(x) \to Q(x))$$

To prove that the claim is false it is sufficient to find a counterexample, that is, to show that $\exists\, x \in \mathbf{R}$ such that $x < x^2$ and $x^2 \geq x^3$. Well, take any negative number, say, $x = -2$. Then $(-2) < (-2)^2$ and $(-2)^2 > (-2)^3$, and we are done. Observe that our proof is indeed in the desired form

$$\exists x \in \mathcal{D}, (\sim(P(x) \to Q(x)) = \ P(x) \wedge \sim Q(x)\ ) \qquad \blacksquare$$

Example/Exercise 3.39 Determine whether the following statement is true:

For all $x, y \in \mathbf{R}$, if $x^2 = y^2$, then $x = y$.

In other words, consider the statement

$\forall x \in \mathcal{D}, (P(x) \to Q(x))$, where $\mathcal{D} = \mathbf{R}$, $P(x)\colon\ x^2 = y^2$, and $Q(x)\colon\ x = y$,

and determine whether there exists an $x \in \mathbf{R}$ such that $\sim(P(x) \to Q(x))$.

Example 3.40 Prove/disprove: All primes are odd.

Solution Following the previous examples, we conclude as follows:

2 is a prime number and it is also even; therefore, the claim is disproved. $\blacksquare$

Example/Exercise 3.41 Prove/disprove: For all $x \in \mathbf{Z}$, $x^2 - x + 7$ is a prime number.

Another way to approach the proof indirectly is a technique called **proof by contradiction (*reductio ad absurdum*)**. We reason as follows:

Suppose that the statement P to be proved is false. Therefore, its negation $\sim P$, has to be true. If this assumption leads to a contradiction, then the original statement has to be true.

Here is a classic example.

Theorem 3.21 There is no greatest integer.

Proof Suppose the statement is false, that is, suppose there is the greatest integer N. What about $M = N + 1$? M has to be an integer since it is a sum of two integers. But M is evidently greater than N, which contradicts our supposition. Thus, there is no greatest integer. ■

Another theorem where proof by contradiction works nicely is

Theorem 3.22 The set $\mathbf{N}$ of natural numbers is infinite.

Proof Clearly $\mathbf{N} \neq \varnothing$. Suppose that $\mathbf{N}$ is finite. Then, for some $k \in \mathbf{N}$, there exist a bijection $f : \mathbf{N}_k \to \mathbf{N}$. Let $n = f(1) + f(2) + \cdots + f(k) + 1$. Evidently, $n > f(i)$ for any $i \in \mathbf{N}_k$, thus $n \notin \mathcal{R}(f)$. Therefore, f is not onto, which contradicts our assumption that f is a bijection. Hence, $\mathbf{N}$ is infinite. ■

Example 3.42 Prove the following statement:

For all $x \in \mathbf{Z}$, if $3x + 2$ is odd, then x is odd.

Proof Suppose that x is even, that is, suppose that $x = 2k$, where $k \in \mathbf{Z}$. Then,

$$3x + 2 = 3(2k) + 2 = 6k + 2 = 2(3x + 1)$$

Thus, $3x + 2$ is even, which contradicts our assumption. We conclude that the original claim is true. ■

Example 3.43 Prove that for all $m, n, p \in \mathbf{Z}$, such that $m > n$,

$$(mp \leq np) \to (p \leq 0)$$

Proof Our statement is obviously in the form $P \to Q$. Since we don't know whether $p \lesseqgtr 0$, it is not evident how to make use of multiplicative law of inequalities. So, we will try to prove the statement by using a method of proof by contradiction. Recall that if the conditional is false, then the antecedent must be true while the consequent is false. We will have our proof if we succeed in proving that the case when P is true and Q is false is a contradiction, thus $P \to Q$ must be true.

So, suppose $\forall m, n, p \in \mathbf{Z}$, with $m > n$, $mp \leq np$ but $p > 0$. Then, $m > n$ implies that $mp > np$, contradicting the statement $mp \leq np$. Therefore, our assumption that $p > 0$ must be wrong (i.e., $p \leq 0$). Thus, the claim

$$\forall m, n, p \in \mathbf{Z}, \text{ such that } m > n, (mp \leq np) \rightarrow (p \leq 0)$$

is true. ■

Example 3.44 There is no odd integer that can be expressed both as $4k - 1$ and $4l + 1$ for some $k \neq l \in \mathbf{Z}$.

Proof Suppose the contrary, that is, suppose $n \in \mathbf{Z}$ is odd and it can be expressed as $n = 4k - 1$ and also as $n = 4l + 1, k \neq l \in \mathbf{Z}$. Then, certainly, we could write

$$4k - 1 = 4l + 1$$

which gives us

$$2(k - l) = 1$$

But this is obviously a contradiction since the left side is even and the right side is odd. ■

Example/Exercise 3.45 Prove that there is no integer that is both even and odd.

Example 3.46 Prove that a sum of rational and irrational numbers is irrational.

Proof Let $r \in \mathbf{Q}$ be any rational number and $i \in \mathbf{R} \setminus \mathbf{Q}$ any irrational number. Suppose, furthermore, that $r + i$ is rational. Since the difference (or a sum) of any two rational numbers is rational it follows that $(r + i) - r = i$ should be rational. But, since i is irrational by definition, we have a contradiction. Therefore, $r + i$ must be irrational. ■

Example 3.47 Prove that for all integers n, if n^2 is even then n is even.

Proof Suppose there exists $n \in \mathbf{Z}$ such that n^2 is even and n is odd. Then, by definition of odd,

$$n = 2k + 1, \quad k \in \mathbf{Z}$$

and

$$\begin{aligned} n^2 &= (2k + 1)^2 = 4k^2 + 4k + 1 \\ &= 2(2k^2 + 2k) + 1 \end{aligned}$$

$2k^2 + 2k$ is obviously an integer, and therefore n^2 is an odd integer. We have that n^2 is both even and odd. But that is a contradiction, thus our statement is true. ■

Example 3.48 Show that $3|(2^{2n}-1)$ for every $n \in \mathbf{Z}^+$.

Solution Suppose not. Suppose that there exist some nonnegative integers n, such that $3 \nmid (2^{2n}-1)$. By Theorem 3.5, there is a smallest among them, say, m. So, $3 \nmid (2^{2m}-1)$, while $3|(2^{2n}-1)$ for every $n \in \mathbf{Z}^+$ for which $0 \leq n < m$. We recognize immediately that $m \geq 1$ (since for $n = 0$ $3|(2^{2n}-1)$). So, let $m = k+1$, where $0 \leq k < m$. Thus, $3|(2^{2k}-1)$, and therefore $(2^{2k}-1) = 3x$ for some $x \in \mathbf{Z}$. Consequently, $2^{2k} = 3x+1$. Now we have

$$\begin{aligned} 2^{2m}-1 &= 2^{2(k+1)}-1 = 2^2 \cdot 2^k - 1 \\ &= 4(3x+1)-1 \\ &= 12x+3 \\ &= 3(4x+1) \end{aligned}$$

implying that $3|(2^{2m}-1)$, which is a contradiction. Hence, we conclude that $3|(2^{2n}-1)$ for every $n \in \mathbf{Z}^+$. ■

Theorem 3.23 (Euclid) There are infinitely many primes.

***Proof* 1** Let $P = \{p_1, p_2, \ldots, p_n\}$ be a finite set of **all** primes. Consider the integer

$$N = p_1 \cdot p_2 \cdot \ldots \cdot p_{i-1} \cdot p_i \cdot p_{i+1} \cdot \ldots \cdot p_n + 1$$

It is either a prime or not. Suppose it is a prime. Then, it is greater than any prime in P, and thus P does not contain all prime numbers, contrary to the assumption that all primes are in the set $P = \{p_1, p_2, \ldots, p_n\}$. Hence, P must be infinite.

However, if N is a composite it has to be divisible by at least one prime, and that prime has to be an element of P, say, p_i. But that cannot be since

$$N = qp_i + 1, \quad \forall p_i \in P$$

Thus, there has to be another prime $p \notin P$ that is a factor of N. But that contradicts our assumption that P contains all prime numbers. In other words, we have

$$p \in P \quad \text{and} \quad p \notin P$$

which cannot be, so we conclude again that P is infinite, that is, there is infinitely many prime numbers. ■

The equivalent is

***Proof* 2** Suppose, contrary to the claim, that there are only a finite number of primes:

$$P = \{p_1, p_2, \ldots, p_n\}$$

Consider the number

$$N = p_1 \cdot p_2 \cdot \ \ldots \ \cdot p_n + 1$$

By Theorem 3.13, N is divisible by some prime number p and, since set P contains all primes, $p \in P$. But being one of the primes from P means that $p|(p_1 \cdot p_2 \cdot p_3 \cdot \ \ldots \ \cdot p_n)$. In that case, Theorem 3.16 requires that $p \nmid (p_1 \cdot p_2 \cdot p_3 \cdot \ \ldots \ \cdot p_n + 1)$. So we have that $p \nmid N$ and $|N$, which cannot be. Thus, our assumption that there are finitely many primes led us to a contradiction, and therefore the theorem is true. ■

Another form of indirect proof, **the proof by contraposition**, is based on the equivalence of implication and its contrapositive. Thus, we may try to prove the statement by showing that its contrapositive is true. That is, suppose the statement to be proved is in the form

$$\forall x \in \mathcal{D},\ (P(x) \rightarrow Q(x))$$

Sometimes, it may be easier to prove the equivalent statement

$$\forall x \in \mathcal{D},\ (\sim Q(x) \rightarrow \sim P(x))$$

So, we proceed proving the contrapositive by a direct proof, that is, with the assumption that $Q(x)$ is false, we prove that $P(x)$ is also false for every $x \in \mathcal{D}$.

Example 3.49 Let $n \in \mathbf{N}$. Prove that if $2 \nmid n$, then $4 \nmid n$ either.

Solution The contrapositive of the above statement is:

$$\text{If } \ 4|n \quad \text{then} \quad 2|n$$

Well, if $4|n$, then there exists a $k \in \mathbf{N}$ such that $n = 4k = 2(2k)$, implying that $2|n$. Since conditional and its contrapositive are equivalent, we have the proof of our statement. ■

Example 3.50 Let's prove the statement from Example 3.43, this time using the method of proof by contrapositive.

Observe that the contrapositive of

$$(mp \leq np) \ \rightarrow \ (p \leq 0),\ m, n, p \in \mathbf{Z},\ \ m > n$$

is the statement

$$(p > 0) \rightarrow (mp > np)$$

But this is evidently true, because we are given that $m > n$. Thus, the statement is true indeed.

Example 3.51 Let $x, y \in \mathbf{R}$ be two real numbers such that $x < 2y$. Show that

$$\text{If } 7xy \leq 3x^2 + 2y^2, \text{ then } 3x \leq y$$

Solution Suppose $x, y \in \mathbf{R}$ are such that $x < 2y$ and $3x > y$. Then, $2y - x > 0$ and $3x - y > 0$. It follows that

$$(2y - x)(3x - y) = 7xy - 3x^2 - 2y^2 > 0$$

and therefore,

$$7xy > 3x^2 + 2y^2$$

Thus, by contraposition, if

$$7xy \leq 3x^2 + 2y^2, \text{ then } 3x \leq y$$ ■

Before the following theorem you may want to revisit Theorem 3.17.

Theorem 3.24 If p is prime, then $\sqrt{p}$ is irrational.

Proof Suppose not, that is, suppose there exists a prime p such that $\sqrt{p} \in \mathbf{Q}$, then, by Theorem 3.17, $\sqrt{p} = n \in \mathbf{Z}, n \geq 2$. It follows that $p = n^2$, which is a contradiction since n^2 is evidently a composite. ■

To go one step further in our "proof techniques," we utilize the unique structure of natural numbers to devise a special rigorous proof method called ***mathematical induction***. We will study this in the next section.

3.4 MATHEMATICAL INDUCTION

As you may already know, our word "induction" is derived from the Latin "*inductio*" which, in turn, was a translation of the Greek word "*epagoge*," a term that Aristotle used to indicate "*progress from particular to universal*." Can we use it in mathematics? Without getting into philosophical discussion of the validity, or how much one should be relying on the path "*particular* → *universal*" in the sciences, the fact is that this method of proving statements in mathematics very often turned out to be a very powerful one.

Let's recall some things we have encountered before: We say that a nonempty set $A \subseteq X$ has a unique least element (the smallest element) $m \in X$ if $a \geq m$ for every $a \in A$. Consequently, some nonempty subsets of real numbers have a least element, and some do not. For instance, the set $\mathbf{N}$ has a least element, namely, 1, while the set $\mathbf{Z}$, obviously, does not.

We also said (Definition 3.16) that the set **N** of natural numbers is well ordered.

Similarly, the closed interval [2, 7] has the smallest element, and the open interval (2, 7) does not.

Finally, observe the following:

Let S be a subset of **N** with the following properties:

(i) $1 \in S$

(ii) If $n \in S$ then $n + 1 \in S$

Then $S = \mathbf{N}$.

The consequence of WOP is the foundation of a powerful proof technique, namely

Theorem 3.25 (Principle of mathematical induction (PMI) I) Let $P(n)$ be a predicate expressing a property defined for a positive integer $n \in \mathbf{Z}^+$, and let $a \in \mathbf{Z}^+$ be another fixed number, if

(i) $P(a)$ is true.

(ii) $\forall k \geq a$, if $P(k)$ is true, then $P(k + 1)$ is also true.

Then, $P(n)$ is true $\forall n \geq a$.

Proof Assume to the contrary, that is, assume that conditions (i), and (ii) are satisfied but there exist some subset of $\mathbf{Z}^+$ for which $P(n)$ is false, that is

$$S = \{n \in \mathbf{Z}^+ | P(n) \ \textit{is false}\} \subseteq \mathbf{Z}^+$$

Since $S \neq \emptyset$, it follows by WOP that S contains a least element s. By our hypothesis, there exists $a \in \mathbf{Z}^+$, such that $P(a)$ is true, so $a \notin S$. Thus, $s \geq a$, and also $s - a \in \mathbf{Z}^+$. Therefore, $s - a \notin S$, and so $P(s - a)$ is a true statement. But, by (ii), $P(s)$ is a true statement too, implying that $s \notin S$. This, however, contradicts our assumption that $s \in S$. Hence the proof of the theorem. ■

Although we will be mostly using PMI in the study of proofs involving nonnegative integers, it is worth mentioning that mathematical induction is a powerful technique that extends far beyond the study of natural numbers. Let's recall the following:

Theorem 3.26 Suppose A is a partially ordered set. Then any **finite**, nonempty set $B \subseteq A$ has a minimal element.

Proof We show by PMI that $\forall n \in \mathbf{N}$, $n > 1$, every subset of A has a minimal element.

$P(1)$: $B \subseteq A$ and $|B| = 1$. So $\exists\, b \in B$. In other words, $\nexists x \in B(x \neq b)$, that is, $\nexists x \in B(xRb \wedge x \neq b)$. Thus, b is a minimal element

$P(k > 1)$: Suppose, now, that every subset $C \subseteq A$ with k elements has a minimal element and consider

$P(k + 1)$: Let $C_1 \subseteq A$ be an arbitrary subset such that $|C_1| = k + 1$, and let c_1 be any element of C_1

Then, $C_1' = C_1 \setminus \{c_1\} \subseteq A$ has k elements. By inductive hypothesis, C_1' has a minimal element, say, c_0'. So, either $c_1 \leq c_0'$ or $c_1 \not\leq c_0'$.

We claim that if $c_1 \not\leq c_0'$, then c_1 is a minimal element of C_1. Suppose it isn't. Then, we can choose some $x \in C_1$ such that $x \leq c_1$ and $x \neq c_1$. But, since $x \neq c_1$, $x \in C_1'$. Because $x \leq c_1$ and $c_1 \leq c_0'$, it follows that $x \leq c_0'$. However, c_0' is a minimal element of C_1' and we must have $x = c_0'$. We obtained $c_1 \leq c_0'$ and $c_0' \leq c_1$, which would imply (remember $\leq$ is a partial order) that $c_1 = c_0'$. But this is impossible since $c_0' \in C_1' = C_1 \setminus \{c_1\}$. Thus, c_1 must be a minimal element of C_1.

If, however, $c_1 \not\leq c_0'$, then, c_0 is a minimal element of C_1. Suppose not. Then, we can choose some $x \in C_1$ such that xRc_0 and $x \neq c_0'$. However, since c_0' is a minimal element of C_1', x cannot be an element of C_1', so the only other possibility is $x = c_1$. Thus, we have $x \not\leq c_0'$ and $c_1 \not\leq c_0'$. But that contradicts our assumption that $c_1 \not\leq c_0'$. Hence, c_0' is a minimal element of C_1. We proved our claim for $k + 1$ and therefore, by PMI, it follows that $\forall n \in \mathbf{N}$, $n > 1$, every subset of A has a minimal element. ■

Theorem 3.27 The Principle of Mathematical Induction (PMI) is equivalent to WOP, that is, if N is a *nonempty* subset of $\boldsymbol{N}$, then N contains a least element.

Proof Suppose N has no least element. Furthermore, suppose that M is a set of all natural numbers which are less than every element of N. Then $1 \in M$. Otherwise, $1 \in N$, and thus 1 would be a least element of N. Take some $k \in M$. Then, k is less than every element of N. Therefore, $k + 1 \in M$. Otherwise, $k + 1$ would be a least element of N. By PMI, it follows that M contains every natural number which makes N empty. But that contradicts our hypothesis that N is nonempty and so the assumption that N has no least element cannot be true. Thus, the theorem is true. ■

The rest is going to be much easier.

Example 3.52 Show that for all $n \in \mathbf{N}$,

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

Solution First, we recognize that our statement is

$$P(n): \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

Our next step is to verify that $P(1)$ is true. We explicitly check:

$$P(1): \frac{1(1+1)}{1} = 1$$

Finally, assuming that $P(k)$ is also true for some integer $k \geq 1$, we need to prove that $P(k+1)$ is also true. Well, let's see:

$$\begin{aligned} P(k+1): \quad \sum_{i=1}^{k+1} i &= 1 + 2 + 3 + \cdots + k + (k+1) \\ &= \frac{(k+1)[(k+1)+1]}{2} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned} \tag{*}$$

Note that the left-hand side of our equation actually equals

$$\begin{aligned} \frac{k(k+1)}{2} + (k+1) &= \frac{k(k+1) + 2(k+1)}{2} \\ &= \frac{(k+1)(k+2)}{2} \end{aligned}$$

And that is exactly the (*) – the right-hand side of $P(k+1)$, as we expected.

Thus, $P(k+1)$ is true whenever $P(k)$ is true, so $P(n)$ is true for all $n \in \mathbf{N}$. ■

Example 3.53 Prove that if

$$S_n = \sum_{i=1}^{n} i, \quad n \in \mathbf{N}$$

Then

$$S_n + S_{n-1} = n^2$$

Proof From the previous example, it follows that

$$\begin{aligned} S_n + S_{n-1} &= \frac{n(n+1)}{2} + \frac{(n-1)n}{2} \\ &= \frac{n}{2}((n+1) + (n-1)) = n^2 \end{aligned}$$

■

Example 3.54 Show that $\sum_{i=1}^{2n} i = 2n^2 + n$

Solution Observe that

$$\sum_{i=1}^{2n} i = \sum_{i=1}^{N} i$$

where, of course, $N = 2n$, then follow the steps as in Example 3.53. ■

Example 3.55 Show that

$$\sum_{i=1}^{n} (2i - 1) = n^2, \quad \forall n \in \mathbf{N}$$

Solution

$P(n)$: $\sum_{i=1}^{n} (2i - 1) = n^2, \quad \forall n \in \mathbf{N}$

$P(1)$: $2 \cdot 1 - 1 = 1$ is obviously true. Assuming that

$P(k)$: $\sum_{i=1}^{k} (2i - 1) = k^2, \quad k \in \mathbf{N}$we check

$$P(k+1): \sum_{i=1}^{k+1} (2i - 1) = \sum_{i=1}^{k} (2i - 1) + [2(k + 1) - 1]$$
$$= k^2 + 2k + 1$$
$$= (k + 1)^2$$

So, $P(k + 1)$ is also true whenever $P(k)$ is true, thus $P(n)$ is true for all $n \in \mathbf{N}$. ■

Example 3.56 Show that

$$\sum_{i=1}^{n} i^2 = \frac{n(n + 1)(2n + 1)}{6}, \quad \forall n \in \mathbf{N}$$

Solution

$P(n)$: $\sum_{i=1}^{n} i^2 = \frac{n(n + 1)(2n + 1)}{6}, \quad \forall n \in \mathbf{N}$

$P(1)$: $1^2 = \frac{1 \cdot 2 \cdot 3}{6}$ is evidently true. Assume that

$P(k)$: $1^2 + 2^2 + \cdots + k^2 = \frac{k(k + 1)(2k + 1)}{6}$, $k \in \mathbf{N}$, is also trueLet's check

$$P(k+1)\colon\ 1^2+2^2+\cdots+k^2+(k+1)^2=\frac{k(k+1)(2k+1)}{6}+(k+1)^2$$
$$=\frac{k(k+1)(2k+1)+6(k+1)^2}{6}$$
$$=\frac{(k+1)[k(2k+1)+6(k+1)]}{6}$$
$$=\frac{(k+1)(2k^2+7k+6)}{6}$$
$$=\frac{(k+1)(k+2)(2k+3)}{6}$$

which is exactly the desired result. Since $P(k+1)$ is true, whenever $P(k)$ is true, we conclude that $P(n)$ is true for all $n\in\mathbf{N}$. ■

Example 3.57 Show that

$$\sum_{i=1}^{n}\frac{1}{(i+1)(i+2)}=\frac{n}{2n+4},\quad \forall n\in\mathbf{N}$$

Solution As before, we have

$P(1)$: $\dfrac{1}{2\cdot 3}=\dfrac{1}{2\cdot 1+4}$ is true. Assume that

$P(k)$: $\dfrac{1}{2\cdot 3}+\dfrac{1}{3\cdot 4}+\cdots+\dfrac{1}{(k+1)(k+2)}=\dfrac{k}{2k+4},\ k\in\mathbf{N},$ is also true

Then,

$$P(k+1)\colon\ \frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{(k+1)(k+2)}+\frac{1}{(k+2)(k+3)}=\frac{k+1}{2(k+1)+4}$$
$$=\frac{k+1}{2k+6}$$

Observe that

$$\left[\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\cdots+\frac{1}{(k+1)\,(k+2)}\right]+\frac{1}{(k+2)(k+3)}$$
$$=\frac{k+1}{2(k+1)+4}=\frac{k}{2k+4}+\frac{1}{(k+2)(k+3)}=\frac{k}{2(k+2)}+\frac{1}{(k+2)(k+3)}$$
$$=\frac{k(k+3)+2}{2(k+2)(k+3)}=\frac{k^2+3k+2}{2(k+2)(k+3)}$$
$$=\frac{(k+1)(k+2)}{2(k+2)(k+3)}=\frac{k+1}{2(k+3)}=\frac{k+1}{2k+6}$$

Since $P(k+1)$ is exactly as desired, we conclude that $P(n)$ is true for all $n\in\mathbf{N}$. ■

Example 3.58 Show that for any $n \in \mathbf{N}$, $n^2 + 3n + 2$ is even.

Solution Observe that the statement

$P(n)$: $n^2 + 3n + 2$ is even, $\quad \forall\, n \in \mathbf{N}$

is obviously true for $n = 1$, that is

$P(1)$: $1 + 3 + 2 = 6$

Assume that

$P(k)$: $k^2 + 3k + 2$ is true for all $k \in \mathbf{N}$, and consider

$$\begin{aligned} P(k+1)\colon (k+1)^2 + 3(k+1) + 1 &= k^2 + 2k + 1 + 3k + 3 + 2 \\ &= (k^2 + 3k + 2) + 2k + 4 \\ &= (k^2 + 3k + 2) + 2(k + 2) \end{aligned}$$

The first term, $(k^2 + 3k + 2)$, is even by the inductive hypothesis while the second term, $2(k + 2)$, being a multiple of 2, is obviously true. So, $P(k + 1)$ as a sum of two even numbers (thus, an even number) is also true. By PMI we conclude that $P(n)$ is true for any $n \in \mathbf{N}$. ■

Example/Exercise 3.59 Prove by induction that for all $x \in \mathbf{R}$

(i) $x^n y^n = (xy)^n$
(ii) $x^{m+n} = x^m x^n$
(iii) $(x^m)^n = x^{mn}$

Example 3.60 Show that $3^{2n} - 1$ is divisible by 8 for all $n \in \mathbf{N}$.

Solution Our proposition is

$P(n)$: $8|(3^{2n} - 1), \quad \forall n \in \mathbf{N}$

Observe that

$P(1)$: $8|(3^{2\cdot 1} - 1)$ is obviously true.

Suppose $P(k)$ is also true, that is, suppose $8|(3^{2k} - 1)$. We need to prove that 8 divides $(3^{2(k+1)} - 1)$.

Well, if $8|(3^{2k}-1)$, then there exists $x \in \mathbf{N}$ such that

$$3^{2k}-1 = 8x \qquad (*)$$

Let's multiply both sides of (*) by 3^2:

$$\begin{aligned} 3^{2k+2}-9 &= 9\cdot 8x \\ 3^{2k+2} &= 9+9\cdot 8x \\ &= 1+8+8\cdot 9\cdot x \\ &= 1+8(1+9x) \end{aligned}$$

Thus,

$$3^{2k+2}-1 = 8(1+9x)$$

Therefore, $8|(3^{2(k+1)}-1)$. Thus, we have proved the statement $P(n)$. ■

Example/Exercise 3.61 Show that $4|(5^n-1)$ for all $n \in \mathbf{N}$.

Example 3.62 Show that $3|10^{n-1}$, $\forall n \in \mathbf{N}$.

Solution

$P(n)$: $3|10^{n-1}$, $\forall n \in \mathbf{N}$
$P(1)$: $3|10^{1-1}$ is obviously true

Let's assume that $P(k)$ is also true, that is, assume that

$P(k)$: $3|10^{k-1}, k \in \mathbf{N}$, is true. We need to prove that $3|10^{k+1-1}-1$

Consider

$$\begin{aligned} 10^{k+1-1} &= 10^k - 1 \\ &= 10^k - 1 + 10^{k-1} - 10^{k-1} \\ &= (10^k - 10^{k-1}) + 10^{k-1} - 1 \\ &= 10^{k-1}(10-1) + 10^{k-1} - 1 \\ &= 9\cdot 10^{k-1} + 10^{k-1} - 1 \\ &= 3(3\cdot 10^{k-1}) + (10^{k-1}-1) \end{aligned}$$

The first term is obviously divisible by 3, and the second is divisible by 3 by inductive hypothesis, thus $3|10^{n-1}$ for all $n \in \mathbf{N}$. ■

You may wonder if it would it be possible to apply similar reasoning in the instance when the initial case in induction proof was not $n = 1$. Well, depending on a particular case and with some caution we can start with any integer, even a negative one or zero, and consider the truthfulness of $P(n)$for every integer $n \geq m$.

Recalling that for every integer m, a set $S = \{n \in \mathbf{Z} | n \geq m\}$ is well ordered, we extend Theorem 3.25 and state

Theorem 3.28 (Principle of mathematical induction II) Let $S = \{a \in \mathbf{Z} | a \geq m, m \in \mathbf{Z}\}$ and let $P(n)$ be a statement for every $n \in S$.

If

(i) $P(m)$ is true.

(ii) if $\forall k \in S$, $P(k)$ is true and $P(k+1)$ is true then, $P(n)$ is true for every integer $n \in S$.

Example 3.63 Let's redo Example 3.48, this time by mathematical induction, that is, let's show that $3|(2^{2n} - 1)$ for every $n \in \mathbf{Z}^+$.

Solution The statement we are about to prove is

$P(n)$: $3|(2^{2n} - 1)$ for every $n \in \mathbf{Z}^+$

Obviously,

$P(0)$: $3|(2^0 - 1)$ is true. Assume that
$P(k)$: $3|(2^{2k} - 1)$ with $k \in \mathbf{Z}^+$ is true. We need to show that
$P(k+1)$: $3|(2^{2(k+1)} - 1)$ is also true. Well, if

$3|(2^{2k} - 1)$, then there exists an integer x such that $2^{2k} - 1 = 3x$, and so $2^{2k} = 3x + 1$. Now, $2^{2(k+1)} - 1 = 2^2 \cdot 2^{2k} - 1 = 4(3x + 1) - 1 = 12x + 3 = 3(4x + 1)$, which is evidently divisible by 3. By PMI we conclude that

$3|(2^{2n} - 1)$ for every $n \in \mathbf{Z}^+$ ■

Example/Exercise 3.64 Show that for all $x \in \mathbf{Z}, x \neq 1, (x - 1)|(x^n - 1)$.

Hint: $x^{n+1} - 1 = x(x^n - 1) + (x - 1)$

Example/Exercise 3.65 Prove that for all $n \in \mathbf{N}$, $3|(2n^3 + 4n + 9)$.

Example 3.66 Show that $n < 2^n$ for all $n \in \mathbf{Z}^+$.

Solution We need to prove the proposition

$$P(n)\colon\ n < 2^n,\ \forall n \in \mathbf{Z}^+$$

First note that

$$P(0)\colon\ 0 < 2^0 = 1$$

is obviously true. We assume that $P(k)$ is also true for some $k > 0$, that is$P(k)$: $k < 2^k$ (*) Let's see whether $P(k+1)$ is also true.

Adding 1 to both sides of (*) we get:

$$\begin{aligned} k + 1 < 2^k + 1 \leq 2^k + 2^k &= 2 \cdot 2^k \\ &= 2^{k+1} \end{aligned}$$

Since our proposition is true for $k + 1$ it is true for all $n \in \mathbf{Z}^+$. ■

Example 3.67 Show that $2^n > n^2$ for every integer $n \geq 5$.

Solution We need to prove the statement

$P(n)\colon\ 2^n > n^2,\ n \geq 5,\ \forall n \in \mathbf{Z}$

Observe that $2^5 > 5^2$, thus $P(5)$ is true. Assuming that $P(k)$ is true we need to show that $P(k+1)$ is also true, that is, that

$$2^{k+1} > (k+1)^2 \ \text{ where } \ k \geq 5$$

Well, let's see.

$$\begin{aligned} 2^{k+1} = 2 \cdot 2^k > 2 \cdot k^2 &= k^2 + k^2 \geq k^2 + 5k \\ &= k^2 + 2k + 3k \\ &\geq k^2 + 2k + 15 \\ &> k^2 + 2k + 1 = (k+1)^2 \end{aligned}$$

So, $2^{k+1} > (k+1)^2$, and therefore $2^n > n^2$ for every integer $n \geq 5$. ■

Example/Exercise 3.68 Revisit the proof of Theorem 1.11.

Example/Exercise 3.69 Show that for all $n \in \mathbf{N}$, where $n \geq 4, n! > 2^n$.

Example 3.70 Show that for all $n \in \mathbf{N}$, where $n \geq 3, 2n + 1 < 2^n$.

Solution

$P(n)\colon\ \forall\, n \in \mathbf{N}, n \geq 3, 2n + 1 < 2^n$
$P(3)\colon\ 2 \cdot 3 + 1 < 2^3 = 8$

is obviously true. Assume it is also true for $n = k$. Let's see whether it is true for $n = k + 1$.

$$2(k + 1) + 1 = 2k + 3 = (2k + 1) + 2$$

Since $2k + 1 < 2^k$, by inductive hypothesis

$$2k + 3 < 2^k + 2^k = 2 \cdot 2^k = 2^{k+1}$$

It follows that $2k + 3 < 2^{k+1}$, what was to be shown. ■

Example/Exercise 3.71 Show that for all $n \in \mathbf{Z}^+, 1 + 2 + 2^2 + 2^3 + \cdots + 2^n = 2^{n+1} - 1$.

Example/Exercise 3.72 Prove that for every $n \in \mathbf{N}$,

$$\sum_{i-1}^{n}(3i - 2) = \frac{n}{2}(3n - 1)$$

For the next few examples, we need to digress a little and introduce another concept.

Definition 3.19 A sequence $\{a_n\}_{n\in\mathbf{N}}$ is a function from $\mathbf{N}$ into $\mathbf{R}$. A sequence $\{a_n\}_{n\in\mathbf{N}}$ is bounded iff the set $\{a_n | n \in \mathbf{N}\}$ is bounded. A finite sequence of length n in the set S, is a function $f : \{1, 2, \ldots, n\} \to S$. An infinite sequence in the set S is a function $f : \mathbf{N} \to S$.

Example 3.73 One can think of a finite sequence as an n-tuple

$$(a_1, a_2, \ldots, a_n) = (f(1), f(2), \ldots, f(n))$$

and equivalently, an infinite sequence

$$\{a_i\}_{i=1}^{\infty} = (a_1, a_2, \ldots), \quad a_i \in S, \ \ i \in \mathbf{N}$$

Thus, for instance

$$1, 2, 3, 4, 5, \ldots$$

might be thought of as a sequence $f : \mathbf{N} \to \mathbf{N}$. ■

Example 3.74 The first three terms of a sequence

$$\{a_i\} = \frac{i}{i + 1}$$

are

$$a_1 = \frac{1}{1 + 1} = \frac{1}{2}, \ a_2 = \frac{2}{2 + 1} = \frac{2}{3}, \ a_3 = \frac{3}{3 + 1} = \frac{3}{4}, \ \ldots$$

■

Example/Exercise 3.75 Write the first five terms of the following sequences:

(i) $\{a_n\}_{n=-3}^{\infty}$, where $a_n = (-2)^n$.
(ii) $\{b_m\}_{m=-1}^{\infty}$, where $b_m = 2m^2 - m + 1$.

Now we continue with our proofs.

Theorem 3.29 (Sum of geometric sequence) For all $r \in \mathbf{R}$, $r \neq 1$, and any $n \in \mathbf{Z}^+$

$$\sum_{i=0}^{n} r^i = \frac{r^{n+1} - 1}{r - 1}$$

Proof Let $r \in \mathbf{R}$ be any real number different from 1, then

$$P(n): \quad \sum_{i=0}^{n} r^i = \frac{r^{n+1} - 1}{r - 1}$$

We immediately see that

$$P(0): \quad \sum_{i=0}^{0} r^i = \frac{r^{0+1} - 1}{r - 1} = 1$$

Suppose $P(k)$ is also true for some $k \geq 0$. We need to show that $P(k+1)$ is also true, that is, we must show that

$$P(k+1): \quad \sum_{i=0}^{k+1} r^i = \frac{r^{(k+1)+1} - 1}{r - 1}$$

Observe that

$$\begin{aligned}
\sum_{i=0}^{k+1} r^i &= \sum_{i=0}^{k} r^i + r^{k+1} \\
&= \frac{r^{k+1} - 1}{r - 1} + r^{k+1} \\
&= \frac{(r^{k+1} - 1) + r^{k+1}(r - 1)}{r - 1} \\
&= \frac{r^{k+2} - 1}{r - 1}
\end{aligned}$$

And this is exactly what we had to show. ■

There are situations where more consideration is needed for step II ("$n = k$") in order to successfully complete step III ("$n = k + 1$") of mathematical induction. Namely, after explicitly verifying the basic step I, we use it to check step II and then both of them for step III. We can use as many of the first "k" steps as we need to verify $(k + 1)$th. In other words, if we have the statement $P(n)$, with $n \in \mathbf{Z}^+$, such that $P(1), P(2), \ldots, P(k)$ are all true, and

$$(P(1) \wedge P(2) \wedge \ \ldots \ \wedge P(k)) \text{ implies that } P(k + 1) \text{ is true}$$

then $P(n)$ is true for all $n \in \mathbf{Z}^+$.

This is formally stated in the following:

Theorem 3.30 (The strong principle of mathematical induction – SPMI) For each $n \in \mathbf{Z}^+$, let $P(n)$ be a statement, such that

(i) $P(m)$ is true, and the implication.
(ii) *If* $P(i)$ *then* $P(k + 1)$ is true, for every $i, m \leq i \leq k$.

Then, $P(n)$ is true for every $n \in \mathbf{Z}^+$.

Example 3.76 Let's prove Theorem 3.13: Any natural number greater than 1 is divisible by a prime number this time using SPMI.

Proof

$P(n)$: *If* $n \in \mathbf{N}$ *and* $n > 1$, *then n is divisible by a prime number*

We need to show that $P(2)$ is true. But 2 is divisible by 2, which is prime, so $P(2)$ is true.

Now we check whether for all $k \geq 2$, if $P(i)$ is true for all numbers i between 2 and k, then $P(k + 1)$ is also true. Take any $k \geq 2$ and suppose that i is divisible by a prime number for all $2 \leq i \leq k$. We need to show that $k + 1$ is also divisible by a prime number. In case that $k + 1$ is prime itself, we are done. If $k + 1$is a composite, we can express it as

$$k + 1 = ab$$

where $1 < a < k + 1$ and $1 < b < k + 1$. But then, in particular, $2 \leq a \leq k$ and thus, by our inductive hypothesis, a is divisible by a prime number p. However, since $k + 1 = ab$ it follows that $k + 1$ is also divisible by a prime p. Hence, $P(k + 1)$ is true and therefore our statement $P(n)$ is true. ■

Theorem 3.31 (The fundamental theorem of arithmetic) Every integer $n \geq 2$ is either prime or it can be uniquely expressed as a product of primes.

Proof Since 2 is prime, the statement is certainly true for $n = 2$. Next, we will assume that every integer i, with $2 \leq i \leq k$, is either prime or can be expressed as a product of primes. We need to prove the same for $k + 1$. Of course, if $k + 1$ is prime, we are done – there is nothing to prove. So, let's assume that $k + 1$ is composite. Then, there exist $a, b \in \mathbf{Z}$, such that $k + 1 = ab$, where $2 \leq a \leq k$ and $2 \leq b \leq k$. By induction hypothesis, both a and b are either prime or can be expressed as a product of primes. In any case, $k + 1 = ab$ is a product of primes. Thus, by SPMI, every integer $n \geq 2$ is either prime or can be expressed as a product of primes. It remains to be proved that this factorization is unique. Suppose it is not, that is, suppose there are two different factorizations:

$$\begin{aligned} n &= p_1 \cdot p_2 \cdot \ \ldots \ \cdot p_s \\ &= q_1 \cdot q_2 \cdot \ \ldots \ \cdot q_t \end{aligned}$$

where all p_i, q_j are primes, and we arrange them as follows:

$$p_1 \leq p_2 \leq \cdots \leq p_s; q_1 \leq q_2 \leq \cdots \leq q_t \qquad (*)$$

Now, since factorizations are different, there must be a smallest $r \in \mathbf{N}, r \geq 2$, such that $p_r \neq q_r$. It follows that $p_i = q_i$ for every i with $1 \leq i \leq r - 1$.

Therefore, we can write

$$p_r \cdot p_{r+1} \cdot \ \ldots \ \cdot p_s = q_r \cdot q_{r+1} \cdot \ \ldots \ \cdot q_t \qquad (**)$$

In case that $s = r$, the left side of equation (**) is simply p_r. If, however, $s > r$, then $p_{r+1} \cdot \ \ldots \ \cdot p_s$ is an integer that is a product of $s - r$ primes. In either case

$$p_r | q_r \cdot q_{r+1} \cdot \ \ldots \ \cdot q_t$$

But, since $q_r, q_{r+1}, \ldots, q_t$ are primes, $p_r = q_j$ for some j with $r \leq j \leq t$. Remember, we arranged our p's and q's in nondecreasing order (*), so $q_r \leq q_j$ and therefore $q_r \leq p_r$. By the similar argument, we can consider the left-hand side of equation (**) and show that $p_r \leq q_r$. Hence, $p_r = q_r$, and we have proved the uniqueness of factorization. ■

Theorem 3.32 If n is composite, then n has a prime factor $p \leq \sqrt{n}$.

Proof Being composite, n is expressible as $n = ab$, where $1 \leq a \leq n; 1 \leq b \leq n$.

Let's assume that $a \leq b$. Then, $a^2 \leq ab = n$, and therefore $\leq \sqrt{n}$. Now, since $a > 1, a$ has a prime factor, say, p. But, since $n = ab, p$ is also a factor of n. Thus, $p \leq a \leq \sqrt{n}$. ■

Definition 3.20 A sequence $\{a_n\}_{n=1}^{\infty}$ is called a recursive sequence if, from some point on each term, a_k may be expressed as a function of some previous terms, that is

$$a_{n+k} = f(a_{n+k-1}, \ldots, a_{n+1}, a_n), \ \forall n \in \mathbf{N}$$

In other words, a recursive sequence $\{a_n\}_{n=1}^{\infty}$ is determined by some *initial values* $a_1, \ldots, a_k$ and a *recursive relation,* expressing each term beyond the initial k terms as a function of the previous k terms.

Example 3.77 The well-known classical example is the Fibonacci[12] sequence defined recursively by

$$f_{n+2} = f_{n+1} + f_n, \ \forall n \in \mathbf{N}$$

with initial values $f_1 = 1$ and $f_2 = 1$.

Consequently, we have

$$\{f_n\}_{n=1}^{\infty} = (1, 1, 2, 3, 5, 8, 13, 21, \ldots) \qquad \blacksquare$$

Example 3.78 Let $\{s_n\}$ be a sequence defined recursively by

$$s_1 = 1, s_2 = 3, \text{ and } s_n = 2s_{n-1} - s_{n-2}, \ \forall n \geq 3, n \in \mathbf{N}$$

Then

$P(n)$: $s_n = 2n - 1, \ \forall n \in \mathbf{N}$

Proof We immediately see that the formula holds for $n = 1$:

$P(1)$: $s_1 = 2 \cdot 1 - 1 = 1$

Assume that for an arbitrary $k \in \mathbf{N}$ and for all $i \in \mathbf{N}$, such that $1 \leq i \leq k$,

$P(i)$: $s_i = 2i - 1$

We need to show that

$P(k+1)$: $s_{k+1} = 2(k+1) - 1$

also holds.

Since

$$\begin{aligned} s_{k+1} &= 2(k+1) - 1 \\ &= 2k + 1 \end{aligned}$$

[12]Leonardo of Pisa (Leonardo Pisano) "filius Bonnaci" (cf. 1180–1250).

for $k = 1$ we have

$$s_2 = 2 \cdot 1 + 1 = 3$$

as expected (by sequence recursive definition). Thus, we can proceed assuming that $k \geq 2$. But in that case, $k + 1 \geq 3$, and we have

$$\begin{aligned} s_{k+1} &= 2s_k - s_{k-1} \\ &= 2(2k - 1) - (2k - 3) \\ &= 2k + 1 \end{aligned}$$

which is the desired result. By SPMI we conclude that $P(n)$ is true. ■

Example 3.79 Let $\{s_n\}$ be a sequence defined recursively as follows:

$$s_0 = 0, \quad s_1 = 4, \quad s_n = 6s_{n-1} - 5s_{n-2}, \quad \forall n \geq 2, \quad n \in \mathbf{N}$$

Then,

$$P(n): \quad s_n = 5^n - 1, \ \forall n \geq 2, n \in \mathbf{N}$$

Proof Let's first convince ourselves that $P(0)$ and $P(1)$ hold:

$$P(0): \quad s_0 = 5^0 - 1 = 0$$

$$P(1): \quad s_1 = 5^1 - 1 = 4$$

Assuming that for all integers $k \geq 1$, $P(i)$ is true for all $0 \leq i \leq k$, we need to show that $P(k + 1)$ is also true. So, consider

$$s_{k+1} = 5^{k+1} - 1$$

and observe that since $k \geq 1$, $k + 1 \geq 2$. Thus,

$$\begin{aligned} s_{k+1} &= 6s_k - 5s_{k-1} \\ &= 6(5^k - 1) - 5(5^{k-1} - 1) \\ &= 6 \cdot 5^k - 6 - 5^k + 5 \\ &= 5^k(6 - 1) - 1 \\ &= 5^k \cdot 5 - 1 \\ &= 5^{k+1} - 1 \end{aligned}$$

which is exactly the result we desired. So, $P(k + 1)$ is true. Hence by PSMI $P(n)$ is true for all natural numbers greater or equal to 2. ■

Example 3.80 Let $\{s_n\}$ be a sequence defined recursively as follows:

$$s_1 = 1,\ s_2 = 4,\ s_n = 2s_{n-1} - s_{n-2} + 2,\ \ \forall n \geq 3,\ \ n \in \mathbf{N}$$

Then,

$P(n)$: $\quad s_n = n^2,\ \forall n \in \mathbf{N}$

Proof Formula holds for $n = 1$:

$P(1)$: $\quad s_1 = 1 = 1^2$

Assuming that for any $k \in \mathbf{N}$ and $1 \leq i \leq k$

$P(i)$: $\quad s_i = i^2$

we show that

$P(k+1)$: $\quad s_{k+1} = (k+1)^2$

Observe that, according to definition, $s_2 = 4$, that is, for $k = 1, s_{k+1} = (k+1)^2$, and we may assume that $k \geq 2$, thus $k + 1 \geq 3$. Then,

$$\begin{aligned} s_{k+1} &= 2s_k - s_{k-1} + 2 \\ &= 2k^2 - (k-1)^2 + 2 \\ &= 2k^2 - k^2 + 2k - 1 + 2 \\ &= k^2 + 2k + 1 = (k+1)^2 \end{aligned}$$

Hence $P(k+1)$ is true, and by SPMI we conclude that $P(n)$ is also true. ■

Example 3.81 Rewriting Fibonacci numbers introduced in Example 3.77 as follows:

$$\begin{aligned} f_1 &= 1 \\ f_2 &= 1 \\ f_n &= f_{n-1} + f_{n-2} \end{aligned}$$

where $n \in \mathbf{N}$ and $n \geq 2$, we can extend the "usual" *Fibonacci sequence* (whose first nine members are $1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$), also to $n = 0, -1, -2, \ldots$. We obtain the following sequence:

$$\begin{aligned} &(\ldots f_{-7}, f_{-6}, f_{-5}, f_{-4}, \ldots, f_{-1}, f_0, f_1, f_2, f_3, \ \ldots) \\ &\quad = (\ldots\ 13, -8, 5, -3, 2, -1, 1, 0, 1, 1, 2, \ \ldots) \end{aligned}$$

Show that Fibonacci numbers are given by the following expression (Binet[13] formula):

$$f_n = \frac{\alpha^n - \beta^n}{\sqrt{5}}$$

where $\alpha = (1 + \sqrt{5})\ /2$ and $\beta = (1 - \sqrt{5})/2$.

Solution Before the actual proof, let me digress a little and show you something you might have heard about in one way or another before, namely, the relationship of Fibonacci numbers, the "*golden ratio*" and the following quadratic equation:

$$t^2 - t - 1 = 0 \qquad (*)$$

The Ancient Greeks considered the following ratio as an exceptionally interesting (and the most pleasing one):

Take a segment L and divide it into two parts, l and x, that is, let $L = l + x$, such that the following ratio holds:

$$\frac{l+x}{l} = \frac{l}{x}$$

or

$$1 + \frac{x}{l} = \frac{l}{x}$$

With

$$\frac{l}{x} = t$$

we get

$$1 + \frac{1}{t} = t$$

In other words,

$$t + 1 = t^2$$

And finally,

$$t^2 - t - 1 = 0$$

which is equation (*).

The solutions of this equation (you should check this) are $\alpha = (1 + \sqrt{5}\ )/2$ and $\beta = (1 - \sqrt{5}\ )/2$.

Therefore,

$$\alpha^2 = \alpha + 1 \qquad (3.1)$$

$$\beta^2 = \beta + 1 \qquad (3.1')$$

[13] J.P.M. Binet (1786–1856).

After multiplying both sides of (3.1) by α, we get

$$\begin{aligned} \alpha^3 &= \alpha^2 + \alpha \\ &= \alpha + 1 + \alpha \\ &= 2\alpha + 1 \end{aligned} \tag{3.2}$$

Let's do it again, that is, let's multiply (3.2) by α

$$\begin{aligned} \alpha^4 &= 2\alpha^2 + \alpha(3) \\ &= 2(\alpha + 1) + \alpha \\ &= 2\alpha + 2 + \alpha \\ &= 3\alpha + 2 \end{aligned} \tag{3.3}$$

Multiplying one more time will enable you to see the pattern (if not, continue multiplying the same way a couple of times more until you get it)

$$\begin{aligned} \alpha^5 &= 3\alpha^2 + 2\alpha(4) \\ &= 3(\alpha + 1) + 2\alpha \\ &= 3\alpha + 3 + 2\alpha \\ &= 5\alpha + 3 \end{aligned} \tag{3.4}$$

Observe that the leading coefficients of (3.1)–(3.4) are numbers $1, 2, 3, 5$, that is, the Fibonacci numbers.

Now, we do our proof.

$P(n)$: $\quad f_n = \dfrac{\alpha^n - \beta^n}{\sqrt{5}}$

where $\alpha = (1 + \sqrt{5})\ /2$ and $\beta = (1 - \sqrt{5})/2$.

$$\begin{aligned} P(1)\text{:} \quad f_1 &= \frac{\alpha - \beta}{\sqrt{5}} \\ &= \frac{(1 + \sqrt{5}) - (1 - \sqrt{5}\,)}{2\sqrt{5}} \\ &= 1 \end{aligned}$$

Next, suppose that the formula holds for all natural numbers $k \geq 2$. Then,

$$P(k+1): \quad f_{k+1} = f_{k-1} + f_k$$
$$= \frac{\alpha^{k-1} - \beta^{k-1} + \alpha^k - \beta^k}{\sqrt{5}}$$
$$= \frac{\alpha^{k-1}(1+\alpha) - \beta^{k-1}(1+\beta)}{\sqrt{5}}$$

using (3.1)

$$= \frac{\alpha^{k-1} \cdot \alpha^2 - \beta^{k-1} \cdot \beta^2}{\sqrt{5}}$$
$$= \frac{\alpha^{k+1} - \beta^{k+1}}{\sqrt{5}}$$

Thus, the formula holds for all $n \in \mathbf{N}$. ■

Example/Exercise 3.82 Define a sequence recursively as follows:

$$a_0 = 0,\ a_1 = 1,\ a_2 = 1, \ldots,\ a_n = \frac{1}{2}a_{n-3} + \frac{3}{2}a_{n-2} + \frac{1}{2}a_{n-1},\ \forall n \geq 3,\ n \in \mathbf{N}$$

Show that $a_n = f_n$.

Theorem 3.33 Let α be a solution of equation (*) from Example 3.81. Then,

$$\alpha^n = f_{n-1} + \alpha f_n,\ \forall n \in \mathbf{Z}$$

Proof For $n = -1, n = 0,$ and $n = 1$ the statement is demonstrably true (cf. Example 3.81). Indeed,

$$\alpha^{-1} = f_{-2} + \alpha f_{-1}$$
$$= -1 + \alpha$$
$$\alpha^0 = f_{-1} + \alpha f_0$$
$$= 1 + \alpha \cdot 0$$
$$\alpha^1 = f_0 + \alpha f_1$$

Let's assume that the statement is true for all $n = 1, 2, \ldots, k$. We need to show that the statement holds for $n = k + 1$. Let's see.

$$\begin{aligned}
\alpha^{k+1} &= \alpha\alpha^k \\
&= \alpha(f_{k-1} + \alpha f_k) \\
&= \alpha f_{k-1} + \alpha^2 f_k \\
&= \alpha f_{k-1} + (\alpha + 1) f_k \\
&= f_k + \alpha(f_{k-1} + f_k) \\
&= f_k + \alpha f_{k+1}
\end{aligned}$$

which is the desired result. Remember, the original statement pertains to all integers. So, assuming that the statement holds for $n = -1, -2, \ldots, -k$, we need to show that it also holds for $n = -(k + 1)$. In other words, we would like to show that

$$\alpha^{-(k+1)} = f_{-(k+1)-1} + \alpha f_{-(k+1)}$$

We proceed as before:

$$\begin{aligned}
\alpha^{-(k+1)} &= \alpha^{-1}\alpha^{-k} \\
&= \alpha^{-1}(f_{-(k+1)} + \alpha f_{-k}) \\
&= \alpha^{-1} f_{-(k+1)} + f_{-k} \\
&= (\alpha - 1) f_{-(k+1)} + f_{-k} \\
&= f_{-k} - f_{-(k+1)} + \alpha f_{-(k+1)} \\
&= f_{-(k+2)} + \alpha f_{-(k+1)}
\end{aligned}$$

Hence, by PMI, we conclude that $\alpha^n = f_{n-1} + \alpha f_n, \ \forall n \in \mathbf{Z}$. ■

Example/Exercise 3.83 If $\beta = (1 - \sqrt{5})/2$ is the other solution of equation (*) from Example 3.81 show that

$$\beta^n = f_{n-1} + \beta f_n, \ \forall n \in \mathbf{Z}$$

Theorem 3.34 Let α be a solution of equation (*) from Example 3.81. Then, $f_n \leq \alpha^{n-1}$ for all natural numbers $n \geq 1$.

Proof

$P(n)$: $f_n \leq \alpha^{n-1}$ for all natural numbers $n \geq 1$

Since $f_1 = f_2 = 1$, the claim is evidently true for $n = 1$ and $n = 2$.
Assume that

$P(k)$: $f_k \leq \alpha^{k-1}$ for some $k > 2, k \in \mathbf{N}$. Let's see whether $P(k+1)$ is also true.
Consider

$$\begin{aligned} f_{k+1} &= f_k + f_{k-1} \\ &\leq \alpha^{k-1} + \alpha^{k-2} = \alpha^{k-2}(\alpha + 1) \\ &= \alpha^{k-2}\alpha^2 \\ &= \alpha^k \end{aligned}$$

That is, $P(k+1)$ is true if $P(k)$ is true – exactly the desired result.
By PMI, we conclude that $P(n)$ is true, that is, $f_n \leq \alpha^{n-1}$ for every $n \geq 1$. ■

Theorem 3.35 Let f_n be the nth Fibonacci number. Then $f_n < 2^n, \ \forall n \in \mathbf{N}$

Proof

$P(n)$: $f_n < 2^n, \ \forall n \in \mathbf{N}$

$P(1)$: $f_1 = 1 < 2$ is obviously true. Let's assume that

$P(2), \ldots, P(k)$, are also true for all $k \in \mathbf{N}$, and consider

$P(k+1)$: $$\begin{aligned} f_{k+1} &= f_k + f_{k-1} \\ &< 2^k + 2^{k-1} < 2^k + 2^k = 2 \cdot 2^k = 2^{k+1} \end{aligned}$$

Thus, $P(k+1)$ is true if $P(k)$ is true, and therefore $P(n)$ is true for any $n \in \mathbf{N}$. ■

Theorem 3.36 Let f_n be the nth Fibonacci number. Then,

$$\sum_{i=1}^{n} f_i = f_{n+2} - 1, \quad n \in \mathbf{N}$$

Proof

$P(n)$: $\sum_{i=1}^{n} f_i = f_{n+2} - 1, n \in \mathbf{N}$

$P(1)$: $f_3 - 1 = 2 - 1 = 1$ is obviously true. Assume that

$P(k)$: $\sum_{i=1}^{k} f_i = f_{k+2} - 1$ is also true

Then,

$$P(k+1): \quad \sum_{i=1}^{k+1} f_k = f_{k+1+2} - 1 = f_{k+3} - 1$$

$$= \sum_{i=1}^{k} f_i + f_{k+1} = f_{k+2} - 1 + f_{k+1}$$

$$= f_{k+1} + f_{k+2} - 1$$

$$= f_{k+3} - 1$$

which is exactly what we had hoped for. So, $P(k+1)$ is true whenever $P(k)$ is true, thus$P(n)$ is true $\forall n \in \mathbf{N}$. ■

Theorem 3.37 For any $n \in \mathbf{N}$,

$$\sum_{i=1}^{n} f_i^2 = f_n f_{n+1}$$

Proof

$$P(n): \quad \sum_{i=1}^{n} f_i^2 = f_n f_{n+1}$$

It is easy to verify that this is true for $n = 1$. Indeed, since $f_1 = f_2 = 1$,

$$f_1^2 = f_1 f_2$$

Assume that

$$P(k): \quad \sum_{i=1}^{k} f_i^2 = f_k f_{k+1} \text{ is true, and consider}$$

$$P(k+1): \quad \sum_{i=1}^{k+1} f_i^2 = \sum_{i=1}^{k} f_k^2 + f_{k+1}^2$$

$$= f_k f_{k+1} + f_{k+1}^2$$

$$= f_{k+1}(f_k + f_{k+1})$$

$$= f_{k+1} f_{k+2}$$

Thus, $P(k+1)$ is true when $P(k)$ is true, and therefore $P(n)$ is true $\forall n \in \mathbf{N}$. ■

Example/Exercise 3.84 Prove that $2|f_n$ iff $3|n$.

Theorem 3.38 Consecutive Fibonacci numbers are relatively prime.

Proof Suppose not. Suppose there is $d > 1$, such that $d|f_n$ and $d|f_{n+1}$. It follows that $d|(f_{n+1} - f_n = f_{n-1})$. But, since $|f_n$ and $d|f_{n-1}$, that also implies that $d|(f_n - f_{n-1} = f_{n-2})$. Continuing this process we reach the point that $d|f_1$, which contradicts our assumption that $d > 1$. Thus the proof. ■

Example/Exercise 3.85 Show that for any $n \in \mathbf{N}$, f_n and f_{n+2} are relatively prime.

3.5 SUPPLEMENTARY PROBLEMS

1. Let $n,\ m \in \mathbf{Z}$ be integers. Show that if n is even and m is odd, then nm is even.

2. Let $x, y \in \mathbf{R}$ be real numbers. Show that
(i) $|xy| = |x||y|$
(ii) $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$

3. Let $m, n, p,\ q \in \mathbf{Z}$ be integers. Prove that
(i) if $m|n$ and $p|q$ then $mp|nq$,
(ii) if $m, n \geq 0$, and $m|n$ and $n|m$ then $m = n$.

4. Let $p \in \mathbf{Z}$, and $p \geq 2$ be such that for every pair $a, b \in \mathbf{Z}$ if $p|ab$ then $p|a$ or $p|b$ which implies that p is prime.

5. Let $n \in \mathbf{N}$ be a natural number. Prove that $n^2 + n + 3$ is odd.

6. Prove by contradiction from trichotomy law that for any $x, y \in \mathbf{R}$,

$$((x \leq y) \wedge (y \leq x)) \rightarrow (x = y)$$

7. Prove that for all $x, y \in \mathbf{R}$,
(i) $|x| < |y|$ iff $x^2 < y^2$
(ii) $|x| = |y|$ iff $x^2 = y^2$

8. Prove that for all $x, y, z \in \mathbf{R}$

$$|x - y| \leq |x - z| + |x - y|$$

9. Prove that for all $x, y, z \in \mathbf{R}$

$$||x| - |y|| \leq |x| - |y|$$

10. Suppose $x, y \in \mathbf{R}$ and $x, y \geq 0$. Prove that

$$\frac{x+y}{2} \geq \sqrt{xy}$$

11. Prove that $\sqrt[3]{2}$ is irrational.

12. Suppose $n \in \mathbf{Z}$ is an integer. Prove that the following statements are equivalent:
(i) n is odd; (ii) $3n+1$ is even; (iii) $\frac{n+1}{2} \in \mathbf{Z}$.

13. Prove that if $a \in \mathbf{Q}$ and $b \in \mathbf{I}$ then
(i) $a+b \in \mathbf{I}$
(ii) $a-b \in \mathbf{I}$
(iii) $a \cdot b \in \mathbf{I}, \quad a \neq 0$
(iv) $a-b \in \mathbf{I}$
(v) $\frac{a}{b} \in \mathbf{I}, \quad a \neq 0$
(vi) $\frac{b}{a} \in \mathbf{I}, \quad a \neq 0$

14. Prove that if $m, n \in \mathbf{Z}$, then the following three statements are equivalent:
(i) $m^2 - n^2$ is even, (ii) $m-n$ is even, (iii) m^2+n^2 is even.

15. Let $x, y \in \mathbf{R}$ be two real numbers. Prove that

$$(x-y)^5 + (x-y)^3 = 0, \text{ iff } x = y$$

16. Prove that the square of no integer is of the form $2n-1$, where $n \in \mathbf{Z}$.

17. Prove that for every odd integer n there exists $k \in \mathbf{Z}$, such that $n^2 = 4k+1$.

18. Prove that if $n = k^3 + 1 \geq 3$, where $k \in \mathbf{Z}$, then n is not prime.

19. Prove that $\forall n \in \mathbf{N}$,

$$\sum_{i=1}^{n} (2i-1) = n^2$$

20. Prove that $\forall n \in \mathbf{N}$

$$\prod_{i=1}^{n} (4i-2) = \frac{(2n)!}{n!}$$

21. Prove that

$$1 + 4 + 7 + \cdots + (3n-2) = \frac{n(3n-1)}{2}, \ \forall n \in \mathbf{N}$$

22. Show that $9|(4^{3n} - 1)$ for every $n \in \mathbf{Z}^+$.

23. Prove that

$$x + x^3 + x^5 + \cdots + x^{2n-1} = \frac{x^{2n+1} - x}{x^2 - 1}, \quad x \in \mathbf{R}, x \neq \pm 1, n \in \mathbf{N}$$

24. Prove that

$$1 \cdot 3 + 2 \cdot 4 + 3 \cdot 5 + \cdots + n(n+2) = \frac{n(n+1)(2n+7)}{6}, \ \forall n \in \mathbf{N}$$

25. Prove that

$$\frac{1}{3 \cdot 4} + \frac{1}{4 \cdot 5} + \frac{1}{5 \cdot 6} + \cdots + \frac{1}{(n+2)(n+3)} = \frac{n}{3n+9}, \ \forall n \in \mathbf{N}$$

26. Prove that for every $n \in \mathbf{N}$,

$$\sum_{i=1}^{n} 2^i = 2^{n+1} - 2$$

27. Prove that for every $n \in \mathbf{N}$,

$$\sum_{i=1}^{n} \frac{1}{i^2} \leq 2 - \frac{1}{n}$$

28. Prove that for ever $n \in \mathbf{N}$,

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \cdots + \frac{1}{\sqrt{n}} \geq n$$

29. Prove that for every $n \in \mathbf{N}$,

$$\prod_{i=1}^{n} (2i - 1) = \frac{(2n)!}{n! 2^n}$$

30. Prove that for every natural number $n \geq 2$

$$\prod_{i=2}^{n} \frac{i^2 - 1}{i^2} = \frac{n+1}{2n}$$

31. Suppose the quadratic equations $x^2 + bx + c = 0$ is such that $b, c \in \mathbf{Q}$. Prove that, if one solution of this equation is rational, then the other solution is rational too.

32. Prove that for every $x \in \mathbf{R}$ if $x^2 \geq x$, then either $x \leq 0$ or $x \geq 1$.

33. Prove Bernoulli's inequality

$$(1+x)^n \geq 1+nx, \ \forall x \in \mathbf{Z}^+$$

34. Let $a, b, c, d \in \mathbf{Z}$, with $a, c \neq 0$. Prove that if $a|b$ and $c|d$ then $ac|(ad+bc)$.

35. Is it true that for all $a, b \in \mathbf{Z}$ if $|b^2$ then $a|b$?

36. Show that $3|(n^3-n)$ for every $n \in \mathbf{Z}$.

37. Show that $3|(4^n-1)$ for every $n \in \mathbf{N}$.

38. Prove that $8|(5^{2n}+7)$for every $n \in \mathbf{Z}^+$.

39. Prove that $5|(3^{3n+1}+2^{2n+1})$ for every $n \in \mathbf{Z}^+$

40. Prove that $9|(4^{3n}-1)$ for every $n \in \mathbf{Z}^+$.

41. Prove that for every $n \in \mathbf{N}, 7|(11^n-4^n)$

42. Prove that $(x-y)|(x^3-y^3)$.

43. Prove that for every $n \in \mathbf{N}, 7|(11^n-4^n)$.

44. Prove that for every $n \in \mathbf{Z}, 6|n$ iff $2|2$ and $3|n$.

45. Prove that

(i) $\forall x, y\mathbf{N}$, and $x > y$, $(x-y)|(x^n-y^n)$

(ii) $\forall x \in \mathbf{Z}$, $x \neq 1$, $(x-1)|(x^n-1)$

46. Prove that if p and q are prime with $p \geq q \geq 5$, then $24|(p^2-q^2)$.

47. Prove that $n! > 2^n$, for every integer $n \geq 4$.

48. Prove that $3^n \geq 1+2^n$, for every $n \in \mathbf{N}$.

49. Prove that $\sqrt{6} \notin \mathbf{Q}$.

50. Let p be a prime and let $n \in \mathbf{Z}$. Prove that $p^{1/n} \notin \mathbf{Q}$.

51. Let f_i be Fibonacci numbers. Find the formula for

(i) $f_{n+3} - f_{n+1}$

(ii) $f_n^2 - f_{n-1}^2$

(iii) $f_n^2 + f_{n-1}^2$

52. Let $f_{3n}, f_{3n+1}, f_{3n+2}$ be Fibonacci numbers. Show that if f_{3n} is even, then both f_{3n+1} and f_{3n+2} are odd for all $n \in \mathbf{N}$.

53. Show that if $m \in \mathbf{N}$ is any natural number, then

$$f_m f_n + f_{m+1} f_{n+1} = f_{m+n+1}$$

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

54. Let f_n be the nth Fibonacci number. Prove that

$$\sum_{i=1}^{n} f_{2i+1} = f_{2n}, \ \forall n \in \mathbf{N}$$

4

FUNCTIONS

> "Take some more tea" the March Hare said to Alice, earnestly.
>
> "I've had nothing yet," Alice replied in an offended tone, "so I can't take more."
>
> "You mean you can't take less," said the Hatter: "it's very easy to take more than nothing." [1]

4.1 INTRODUCTION

We have mentioned in Chapter 1 that one of the most important concepts in all of mathematics is the concept of a function. It is also the most ubiquitous. Think about it. If you accept the premise that mathematics is, among other things, a system of thought, then it is easy to see how everything you do, and everything you think about, involves some relationship between different sets of "objects." While you are reading these lines, your mind associates letters of the Latin alphabet appearing in "clusters" in front of you – the English words familiar to you, which, further on, represent the concepts your mind is more or less familiar with; then you/your mind arranges those concepts, according to specific rules, to more

[1] Lewis Carroll.

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

complex sets forming your thoughts, opinions, and so on. And then, you can go back and reflect on your thoughts and your thinking, creating another layer of "associations" between "thinking" and thinking about that "thinking." All those "associations," those "relations," between one set of objects and the other set are the *functions* we want to study. Of course, you may object and ask why this introduction has to be so complicated. Why not introduce the concept of a function simply as a relation between two "objects," say, the size of a pizza and its price? That would also be a perfectly good example of a function. But we will need more, much more. So, let's start, as always, slowly and from the beginning.

> … Leibniz speaks of a "relation between L and M, without consideration as to which member is preceding or succeeding, which is subject or object" … It must be said, therefore, that the relation is something outside of subjects; but since it is neither subject nor accidents it must be something purely ideal, which is nevertheless well worthy of examination.[2]

4.2 RELATIONS

In all the branches of mathematics, and at all levels, one encounters different varieties of relations, usually without thinking about their general features, as well as the properties they have in common and, in particular, their "set-theoretical" nature. Since those features appear frequently, it is useful to study them on their own. On the other hand, functions that often look quite different from relations turn out to be nothing more than a particular kind of a relation too.

Definition 4.1 Let X and Y be any two sets. We say that a **(binary) relation** R between sets X and Y (or a relation from set X to set Y) is a subset of $X \times Y$.

Let $x \in X$ and $y \in Y$. We say that the *ordered pair* $(x, y) \in X \times Y$, or simply $(x, y) \in R$, indicate that x is somehow related to y, and we write

$$xRy \text{ iff } (x, y) \in R \subseteq X \times Y$$

By *ordered pair* (x, y), as you would recall, we mean a pair in which we distinguish the "first" element, x, and the "second," y, so that

$$(x, y) = (w, z) \text{ iff } x = w \text{ and } y = z$$

Often we say that the *domain* $\mathcal{D}(R)$ of a relation R from X to Y is a set of all first elements $x \in X$ of the ordered pairs, that is

$$\mathcal{D}(R) = \{x \in X | \exists y \in Y \text{ s.t. } xRy\}$$

[2] Weyl, H., *Philosophy of Mathematics and Natural Science*, Princeton University Press, 1949, 2009.

and likewise, the *range* $\mathcal{R}(R)$ of R is a set of all second elements $y \in Y$ of the ordered pairs, that is

$$\mathcal{R}(R) = \{y \in Y | \exists x \in X \text{ s.t. } xRy\}$$

Equivalently we have,

Definition 4.2 A relation on sets X and Y (or a relation between sets X and Y) is a subset R of $X \times Y$, and we say that $x \in X$ and $y \in Y$ are related if $(x, y) \in R$.

If $X = Y$, then we speak of a relation $R \subseteq X \times X$ on X.

Of course, if x and y are not related, we write $x \not R y$.

Example 4.1 Let "$\leq$" be the usual ordering on $\mathbf{Q}$. Then "$\leq$" is a relation on $\mathbf{Q}$. For instance, we say that $(1/2)R2$ since $1/2 \leq 2$. ■

Example 4.2

(i) *Less than* ($<$), *equals* ($=$), and *divides* ($|$) can be defined as relations on $\mathbf{N}$ or $\mathbf{Z}$.

(ii) For instance, we can define a relation R on $\mathbf{Z}$ by setting

$$R = \{(x, y) | x, y \in \mathbf{Z}, \text{ and } x < y\}$$

(iii) In Chapter 1 we encountered relations such as: subset "$\subseteq$," intersection "$\bigcap$", each of them creating a new set according to particular rules. For instance, "$\subseteq$" can be a relation on the power set $\mathcal{P}(A)$ of a set A.

(iv) Suppose you are considering a set of all the lines in the Euclidean plane,[3] then a geometric relations *parallel* "$\|$," or *perpendicular* "$\perp$" is another example of two relations on the set of all lines in the plane.

■

Example/Exercise 4.3 Convince yourself that "$\subseteq$" is a relation on any collection of sets.

Example 4.4 Consider $R \subseteq \mathbf{N} \times \mathbf{N}$, such that $R = \{x, y \in \mathbf{N} | x \leq y\}$.

In other words, we say that xRy if $\forall x, y \in \mathbf{N}, x \leq y$. Considering x and y as coordinates of the points on the Euclidean XY plane, we obtain a collection of points above the line $y = x$. For instance, $1R3$, $3R7$, and so on, but $5 \not R 3$, $8 \not R 6$, and so on (see Figure 4.1).

[3] If you are unfamiliar with the term "Euclidean plane" think of it, for the time being, as just an "ordinary" plane without boundaries.

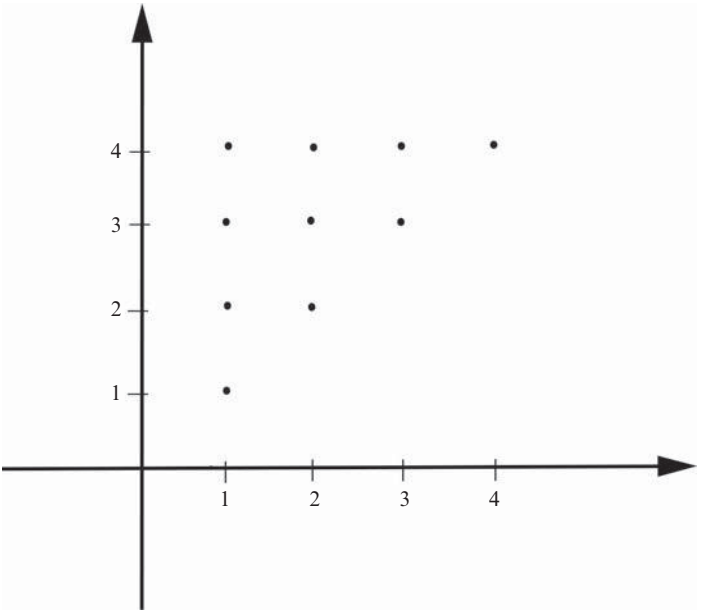


Figure 4.1 xRy iff $x \leq y$

Example 4.5 Define a binary relation R from **R** to **R** as follows:

$$(x, y) \in R \quad \text{iff} \quad x^2 + y^2 = 1$$

It is easy to see that $(0, 1) \in R$ and $(0, 0) \notin R$, hence $0R1$ and $0\not R0$. ■

Example 4.6 Some relations can be easily "visualized," as illustrated below:

(i) If a binary relation on **R** is defined as follows:

$$(x, y) \in R \quad \text{iff} \quad x^2 + y^2 < 25$$

then we can sketch it as in Figure 4.2(i).

(ii) If a binary relation on **R** is defined as follows:

$$(x, y) \in R \quad \text{iff} \quad x^2 - 4y^2 \geq 9$$

then we can sketch it as in Figure 4.2(ii).

(iii) If a binary relation on **R** is defined as follows

$$(x, y) \in R \quad \text{iff} \quad y > x^3$$

then we can sketch it as in Figure 4.2(iii). ■

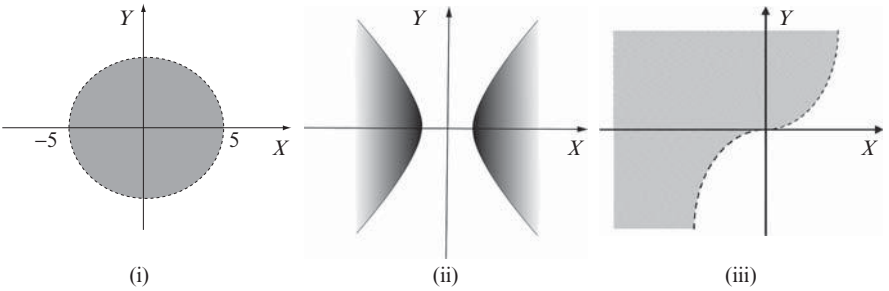


Figure 4.2

Here is another noteworthy example.

Example 4.7[4] Let A and B be any two sets. A set $R = \varnothing \subseteq A \times B$ is the relation of *chaos!* Nothing is related to anything. A set $R = A \times B$ is the relation of *boredom!* Everything is related to everything else and (on this abstract level) there is nothing more to be said. ■

Definition 4.3

A relation on a set X is **reflexive** iff $\forall x \in X, xRx$.

A relation on a set X is **symmetric** iff $\forall\, x, y \in X$, if xRy then yRx.

A relation on a set X is **transitive** iff $\forall\, x, y, z \in X$, if xRy and yRz, then xRz.

A relation on a set X is **antisymmetric** iff $\forall\, x, y \in X$, $x\,R\,y$ and yRx then $x = y$.

Example 4.8 Consider the following relations on the set $X = \{1, 2, 3, 4\}$.

(i) $R = \{(1, 1), (1, 2), (2, 3), (1, 3), (4, 4)\}$

(ii) $S = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$

Determine which of the relations are reflexive, and/or symmetric, and/or antisymmetric, and/or transitive.

Solution R is not reflexive since $(2, 2), (3, 3) \notin R$. R is not symmetric since $(1, 2), (1, 3), (2, 3) \in R$ but $(2, 1), (3, 1), (3, 2) \notin R$. R is transitive and antisymmetric.

[4]Sentilles, D., A Bridge to Advanced Mathematics, The Williams &Wilkins Company, Baltimore, 1975

S is symmetric but not antisymmetric since, for instance, $(1,2),(2,1) \in S$ but $1 \neq 2$. S is reflexive and transitive. ■

Definition 4.4 We say that a relation R satisfies *trichotomy* on a set X iff for every $x, y \in X$ exactly one of the following three possibilities hold:

(i) $(x, y) \in R$
(ii) $x = y$, or
(iii) $(y, x) \in R$

Definition 4.5 A relation R on a set X is said to be a **partial order** or **partial ordering** iff R is reflexive, transitive, and antisymmetric. A set X, together with a partial ordering R, is called a **partially ordered set** or **poset.**

Definition 4.6 A partial ordering R is called a **linear order** or **total order** on X if $\forall\, x, y \in X$ either xRy or yRx.

Example 4.9 A relation "$\subseteq$" on a set is a partial ordering of any collection of sets since

(i) $A \subseteq A$ for any set A.
(ii) If $A \subseteq B$ and $B \subseteq A$ then $A = B$.
(iii) If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

Similarly, you can convince yourself that the relation "$\leq$" on $\mathbf{R}$ is also a partial ordering. ■

Example 4.10 Consider a power set of a set $A = \{a, b, c\}$, with a relation "$\subseteq$" on it. Figure 4.3 shows a partial ordering on $\mathcal{P}(A)$.

Example 4.11 Consider a set $X = \{2, 3, 1, 4, 12, 16\}$. Most people would consider it more natural to write this set as $X = \{1, 2, 3, 4, 12, 16\}$. Such an arrangement puts the elements of the set in a "natural" order, and, if we choose as a relation R on X to be "$\leq$," we immediately recognize a partial ordering on our set. However, there is more. Since all the elements of X are natural numbers, $\forall\, x, y \in X$, $x < y$ or $y < x$, and we have a total ordering on X (see Figure 4.4).

On the other hand, if we consider a relation R to be "$|$" (as in "x divides y") we have a partial ordering on X : $2|4$ but $2 \not| 3$, and so on. Figure 4.5 illustrates this explicitly. ■

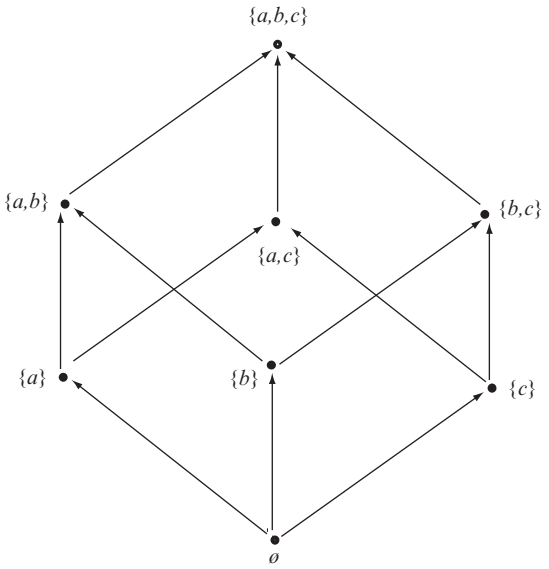


Figure 4.3

Figure 4.4

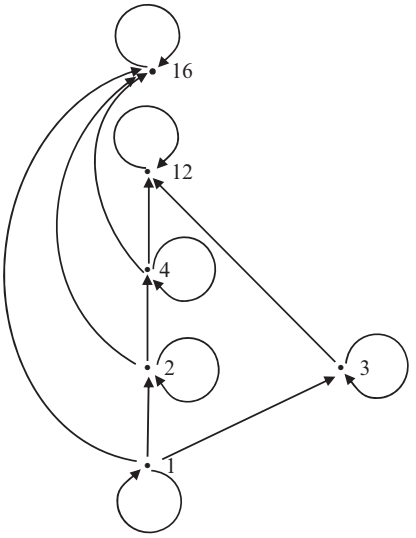


Figure 4.5

Example 4.12 Consider the relation R on $\mathbf{N}$ defined as follows: xRy iff $x \leq y$ and $x + y$ is even. Let's show that such defined R is a partial order, that is, let's show that R is reflexive, transitive, and antisymmetric.
***Reflexivity*:**

$x + x = 2x$ is even and $x \leq x \quad \forall x \in \mathbf{N}$

***Transitivity*:**

Suppose xRy and yRz. Then $x + y$ is even and $x \leq y$, and also $y + z$ is even and $y \leq z$. Now, $x + z = (x + y) + (y + z) + (-2y)$ is a sum of three even numbers, thus an even number. Finally, since "$\leq$" is transitive on $\mathbf{N}$, it follows that $x \leq z$.

***Antisymmetry*:**

Suppose xRy and yRx. Then, $x + y\, x$ is even and $x \leq y$, and also $y \leq x$. But "$\leq$" is antisymmetric on $\mathbf{N}$, and thus $x = y$.

Therefore, R is a partial order. ■

Because of a particular importance of partial ordering of a set in general, which doesn't have to be usual "$\leq$," we use the symbol "$\preccurlyeq$" and (for the lack of better terminology) read $x \preccurlyeq y$ as: "*x is less or equal to y*." Consequently, we have

Definition 4.7 A **partially ordered set** is a set X equipped with a relation "$\preccurlyeq$", which satisfies for all $x, y \in X$ the following:

(i) reflexivity : $x \preccurlyeq x$
(ii) transitivity: if $x \preccurlyeq y$ and $y \preccurlyeq z$, then $x \preccurlyeq z$
(iii) antisymmetry: if $x \preccurlyeq y$ and $y \preccurlyeq x$, then $x = y$

Definition 4.8

(i) An element x_M in a partially ordered set X is called a **maximal element** iff for all $x \in X$ if $x_M \preccurlyeq x$ then $x_M = x$.
(ii) An element x_m in a partially ordered set X is called a **minimal element** iff for all $x \in X$ if $x \preccurlyeq x_m$ then $x_m = x$.

Definition 4.9 Let "$\preccurlyeq$" be a partial order on a set X, then

(i) For any subset $S \subseteq X$, an element $u \in X$ is said to be an **upper bound** of S iff $s \preccurlyeq u, \forall\, s \in S$.

(ii) For any subset $S \subseteq X$, an element $l \in X$ is said to be a **lower bound** of S iff $l \preccurlyeq s, \forall\, s \in S$.

(iii) A subset $S \subseteq X$ is said to be **bounded** iff S has an upper bound and a lower bound.

(iv) An element $\overline{x} \in X$ is **least upper bound** of $S \subseteq X$, (l.u.b. of S), or **supremum** of S, sup(S), iff $\overline{x}$ is an upper bound of S, and for every upper bound u of S, $\overline{x} \preccurlyeq u$.

(v) An element $\underline{x} \in X$ is a **greatest lower bound** of $S \subseteq X$, (g.l.b. of S), or **infimum** of S, inf(S), iff for every lower bound l of S, $l \preccurlyeq \overline{x}$.

Example 4.13 Consider sets **N**, **Z**, **Q**, or **R** with partial order (linear order) "$\leq$." There is no maximal element for either of them. On the other hand, if we consider, say, a set $X = \{x \in \mathbf{Q} | x^2 < 2\}$ then we note that there are many upper bounds but no l.u.b. in **Q**. However, the set X has an l.u.b. in **R**, namely, $\sqrt{2}$. Similarly, the set $(x, y) \subseteq \mathbf{R}$ has neither a minimal nor a maximal element but does have l.u.b. and g.l.b. in **R**, namely, x and y, respectively. ■

Definition 4.10 We say that R is an **ordering relation** on a set X iff R is transitive an satisfies trichotomy on X.

Definition 4.11 Given a relation R from set X to set Y, we say that R^{-1} is an **inverse relation** from Y to X and we write $yR^{-1}x$ iff $xRy \;\forall\, x \in X, \forall\, y \in Y$. In other words, if R is relation from X to Y, then the inverse of R is

$$R^{-1} = \{(y, x) | (x, y) \in R\}$$

For instance, if you think of a relation R on **Z** to be "*x divides y*", then the inverse relation R^{-1} would be "*y* is a multiple of *x*."

Of course, if R is a relation, then $(R^{-1})^{-1} = R$.

Example 4.14 Let $X = \{1, 2, 3\}$ and $Y = \{a, b, c\}$. If the relation $R = \{(1, b), (1, c), (3, b)\}$ then $R^{-1} = \{(b, 1), (c, 1), (b, 3)\}$ (see Figure 4.6).

Representing relations visually can be quite messy sometimes, especially if the sets in question, although finite, are large. It can be more advantageous to portray the relations on those sets by the so-called *directed graphs* or *digraphs*.

For instance, a relation

$$R = \{(a, b), (b, b), (b, d), (c, b), (c, d), (d, a), (d, c)\}$$

on a rather small set $X = \{a, b, c, d\}$ can be pictured in Figure 4.7.

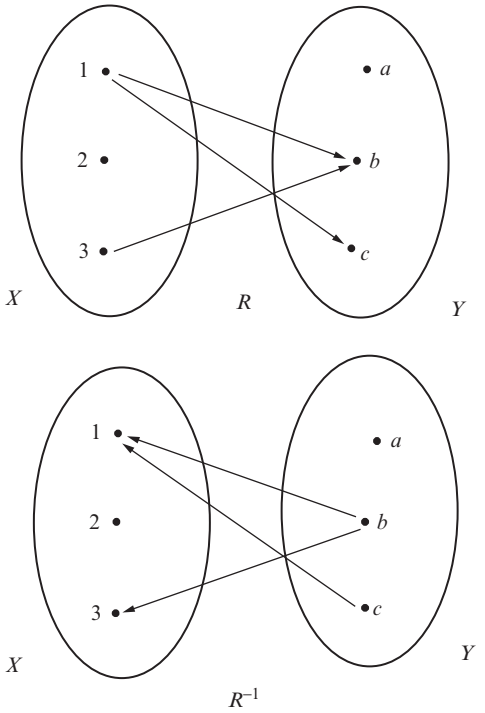


Figure 4.6

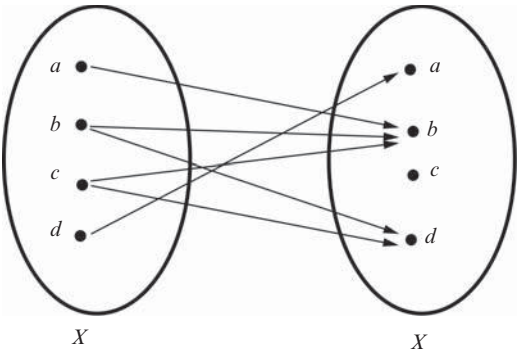


Figure 4.7

However, we can represent it more transparently the following way (Figure 4.8): ■

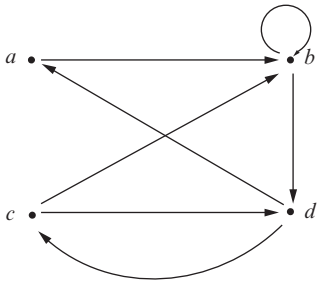


Figure 4.8

Example 4.15 Consider a set $X = \{\varnothing, \{1\}, \{3\}, \{1, 2\}\}$ with a relation R to be "$\subseteq$." Then, digraphs of R and R^{-1} are shown in Figure 4.9. ■

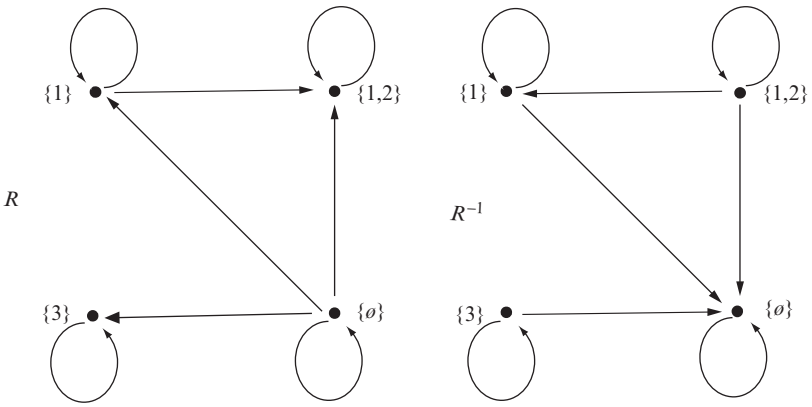


Figure 4.9

Example 4.16 Let $R =$ "divide," $S =$"$\geq$," and $T = \{(x, y)$ iff $x + y > 7\}$ be three relations on a set $X = \{2, 3, 6\}$. The corresponding digraphs are given in Figure 4.10. ■

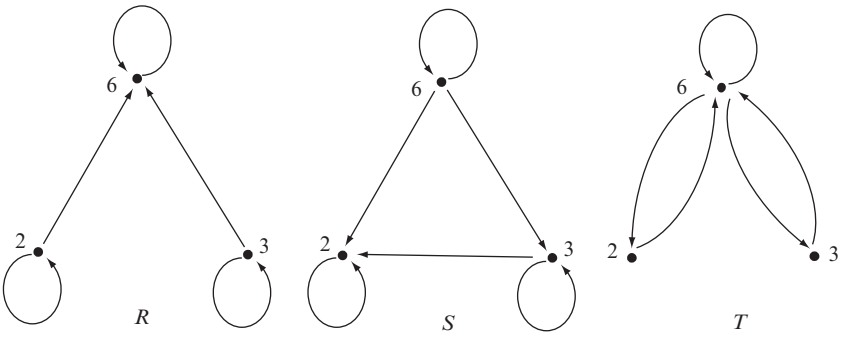


Figure 4.10

Example/Exercise 4.17 Let $X = \{1, 2, 3, 4\}$ and let R on X be given by the following digraph (Figure 4.11).

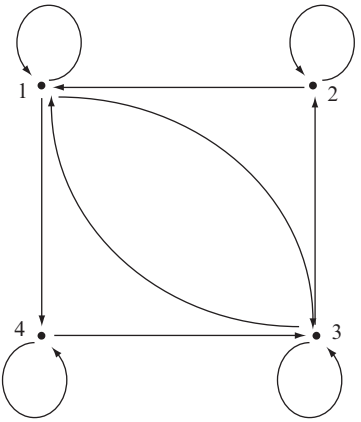


Figure 4.11

What can be said about the relation R? Is it symmetric, antisymmetric, transitive?

Example/Exercise 4.18 Let R be a relation from $\mathbf{R}$ to $\mathbf{R}$ defined by

$$\text{for all } (x, y) \in \mathbf{R} \times \mathbf{R} \quad xRy \text{ iff } y = 2 \cdot |x|$$

Then, $R = \{(x, y) | y = 2 \cdot |x|\}$ and $R^{-1} = \{(y, x) | y = 2 \cdot |x|\}$ and you can certainly recognize the respective graphs (Figure 4.12):

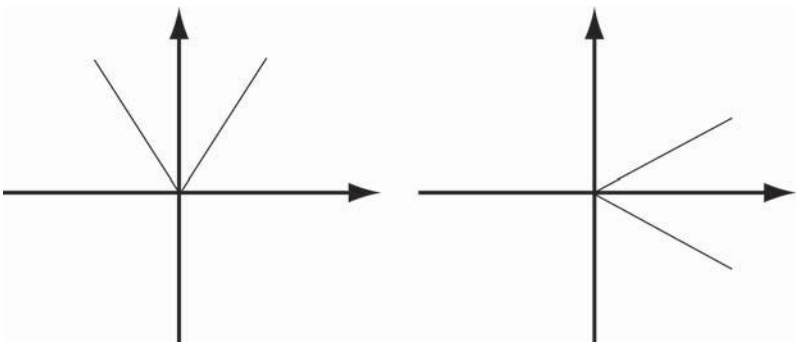


Figure 4.12

What can be said about the above relation R – symmetric, antisymmetric, transitive, and so on?

Theorem 4.1 Let R be a relation from X to Y. Then

(i) $\mathcal{D}(R^{-1}) = \mathcal{R}(R)$
(ii) $\mathcal{R}(R^{-1}) = \mathcal{D}(R)$

Proof (i) If R is a relation from X to Y, then R^{-1} is a relation from Y to X. So, let's assume that

$$(y, x) \in R^{-1}$$

that is, $(y, x) \in Y \times X$. But, since R is a relation from X to Y,

$$(x, y) \in R \subseteq X \times Y$$

where $x \in X$ and $y \in Y$.

Thus,

$$(y, x) \in Y \times X$$

and therefore

$$R^{-1} \subseteq Y \times X$$

This furthermore means that

$$y \in \mathcal{D}(R^{-1}) \text{ iff } \exists x \in X \text{ s.t. } (y, x) \in R^{-1}$$

But,

$$(y, x) \in R^{-1} \text{ iff } \exists x \in X \text{ s.t. } (x, y) \in R$$

which, in turn, is true iff $y \in \mathcal{R}(R)$.

Proof for (ii) is similar. ■

Definition 4.12 Let X be any nonempty set. We say that I_X is the **identity relation** on X if

$$I_X = \{(x, x) | x \in X\}$$

Obviously, $\mathcal{D}(I_X) = X$ and $\mathcal{R}(I_X) = X$, $\forall x \in X$. The digraph is shown in Figure 4.13.

Figure 4.13

Example 4.19 Let $X = \{1, 2, 3\}$. Then $I_X = \{(1, 1), (2, 2), (3, 3)\}$. ■

Having established the concept of a relation on a given set, we are now in the position to classify objects that are in some way "alike."

Definition 4.13 A relation R on a nonempty set X is an **equivalence relation** iff R is *reflexive, symmetric and transitive*. In the case that there is an equivalence relation on X, we write $x \sim y$ for all $x, y \in X$. In other words, R is an equivalence relation on a set X if:

(i) $x \sim x, \quad \forall\, x \in X$.
(ii) If $x \sim y$ then $y \sim x, \forall\, x, y \in X$.
(iii) If $x \sim y$ and $y \sim z$, then $x \sim z, \forall\, x, y, z \in X$.

Example 4.20 Let's verify that a relation "*has the same surname as*" is an equivalence relation. Indeed, let X be the set of all humans. Then $\forall\, x, y, z \in X$ it is true that:

(i) *x has the same surname as x, that is* $x \sim x$.
(ii) If *x has the same surname as y*, then *y has the same surname as x, if* $x \sim y$ *then* $y \sim x$.
(iii) If *x has the same surname as y, and y has the same surname as z* then *x has the same surname as z,* that is $x \sim y$ and $y \sim z$ then $x \sim z$.

Thus, "*has the same surname as*" is an equivalence relation. ■

Example/Exercise 4.21 Let R be a relation on $\mathbf{R}$ defined by: xRy iff $x^2 = y^2$. Prove that R is an equivalence relation.

Example 4.22 Let T be a set of all triangles in a plane, and let $s, t \in T$ be two triangles from T. We define the equivalence relation $s \sim t$, and say that s and t are similar, if s and t have the same corresponding angles. ■

Example 4.23 The relation "$\leq$" defined on $\mathbf{N}$ is reflexive and transitive but not symmetric, thus it is not the equivalence relation. ■

Example/Exercise 4.24 You have surely anticipated that equality is the most basic equivalence relation. Simply put $a \sim b$ iff $a = b$. Show that such defined relation is indeed an equivalence relation.

Example/Exercise 4.25 Define aRb iff for some given $n \geq 2$, a–b is divisible by n, where $a, b, n \in \mathbf{Z}$. Recall, that saying that $a - b$ is divisible by n, or that n divides $a - b$, and writing $n|a - b$, means that there exists $c \in \mathbf{Z}$, such that $a - b = nc$. Show that such defined relation is indeed an equivalence relation.

If you had at least some high school trigonometry, you may appreciate

Example 4.26 Let's define xRy for all $x, y \in \mathbf{R}$, for which $\sin^2 x + \cos^2 y = 1$. Is such defined relation R an equivalence relation?

Solution First recall the basic trigonometric identity

$$\sin^2 x + \cos^2 x = 1, \forall\, x \in \mathbf{R} \qquad (*)$$

Thus, R is obviously reflective, that is, xRx holds.

Let's check the symmetry. Suppose xRy, that is

$$\sin^2 x + \cos^2 y = 1$$

Then from (*) it follows that

$$\cos^2 y = \cos^2 x \qquad (4.1)$$

and also

$$\sin^2 y = \sin^2 x \qquad (4.2)$$

Substituting (4.1) and (4.2) into (*) we get

$$\sin^2 x + \cos^2 y = 1$$

and symmetry $xRy = yRx$ is confirmed.

For transitivity, let's take $x, y, z \in \mathbf{R}$ and suppose that xRy and yRz. From the definition of our relation, it follows that

$$\sin^2 x + \cos^2 y = 1 \qquad (4.3)$$

and

$$\sin^2 y + \sin^2 z = 1 \qquad (4.4)$$

Using the same argument to obtain (4.1) and (4.2), we get

$$\cos^2 y = \cos^2 z$$

Substituting into (4.3), we get

$$\sin^2 x + \cos^2 z = 1$$

that is, xRz. Hence, we have established that the relation R is indeed an equivalence relation. ■

Example 4.27 Let R be a relation on $\mathbf{N}^+ \times \mathbf{N}^+$ defined as follows:

$$(a, b)R(c, d) \text{ iff } ad = bc, \quad a, b, c, d \in \mathbf{N}^+$$

Show that R is an equivalence relation.

Solution We have to show (i) reflexivity, (ii) symmetry, and (iii) transitivity of R.

(i) $(a, b)R(a, b)$ since $ab = ba$, thus R is reflexive.

(ii) Suppose $(a, b) = (c, d)$. Then $ad = bc$. But $cb = da$ and therefore $(c, d) = (a, b)$, so we have proved symmetry.

(iii) Suppose $(a, b)R(c, d)$ and $(c, d)R(e, f)$. Then

$$ad = bc \text{ and } cf = de$$

Let's multiply the corresponding terms of the equations to get

$$ad \cdot cf = bc \cdot de$$

Dividing both sides by c and d we get

$$af = be$$

Thus, $(a, b)R(e, f)$, that is, R is also a transitive relation. We conclude that R is an equivalence relation. ■

Example/Exercise 4.28 Let R be a relation on $\mathbf{Z} \times \mathbf{Z}$ defined as follows:

$$(a, b)R(c, d) \text{ iff } a + b = c + d, \quad a, b, c, d \in \mathbf{Z}$$

Show that R is an equivalence relation.

Example 4.29 Let R be a relation on $\mathbf{Z}$ defined by

$$xRy \text{ iff } x + 3y = 2k, \quad \forall x, y, k \in \mathbf{Z}$$

Show that such defined R is an equivalence relation.

Solution

(i) $xRx = x + 3x = 4x = 2(2x) \in \mathbf{Z}$. Thus, R is reflexive.

(ii) Suppose xRy. Then, $x + 3y = 2k$ for some $k \in \mathbf{Z}$. Therefore, $x = 2k - 3y$ and thus, $y + 3x = y + 3(2k - 3y) = y + 6k - 9y = 6k - 8y = 2(3k - 4y) \in \mathbf{Z}$. Hence, $xRy = yRx$, that is, we have proved symmetry.

(iii) Suppose xRy and yRz. Then,

$$x + 3y = 2k, \quad \text{and} \quad y + 3z = 2l \text{ for some } k, l \in \mathbf{Z}$$

Consider

$$\begin{aligned}(x + 3y) + (y + 3z) &= 2k + 2l \\ &= x + 4y + 3z\end{aligned}$$

Therefore,

$$\begin{aligned}x + 3z &= 2k + 2l - 4z \\ &= 2(k + l - 2z) \in \mathbf{Z}\end{aligned}$$

Thus, xRz too! We have established that R is reflexive, symmetric, and transitive, thus an equivalence relation. ■

Definition 4.14 Let m and n be any two integers, and let d be a positive integer. We say that ***m* is congruent to *n* modulo *d*** iff $d|m - n$, and we write

$$m \equiv n(\text{mod } d)$$

Example 4.30

(i) 12 is congruent to 7 modulo 5, that is, $12 \equiv 7(\text{mod } 5)$, since

$$12 - 7 = 5 = 5 \cdot 1, \quad \text{that is } 5|5$$

(ii) Similarly, 6 is not congruent to (-8) modulo 4, since $6 - (-8) = 14$, and $4 \nmid 4$. ■

Theorem 4.2 For every $n \in \mathbf{N}$, congruence (mod d) is an equivalence relation on $\mathbf{Z}$.

Proof Let $d \in \mathbf{N}$, then $m \equiv n(\text{mod } d)$ iff $m, n \in \mathbf{Z}$ and $d|m - n$. We need to prove (i) reflexivity, (ii) symmetry, and (iii) transitivity:

(i) Take any $m \in \mathbf{Z}, m - m = 0$ and any d divides 0, since $d \cdot 0 = 0$. Thus, we have proved that $m \equiv m(\text{mod } d)$.

(ii) Let $m, n \in \mathbf{Z}$, and suppose $m \equiv n(\text{mod } d)$. That implies that $d|m - n$. In other words,

$$\exists x \in \mathbf{Z}, \text{ s.t. } d \cdot x = m - n$$

We need to show $n \equiv m(\text{mod } d)$, that is, we need to show that $d|n - m$. But

$$n - m = -(m - n) = -d \cdot x = d \cdot (-x)$$

Since $(-x) \in \mathbf{Z}$, d divides $n - m$, thus it follows $n \equiv m(\text{mod } d)$.

(iii) Let $m, n, p \in \mathbf{Z}$, *and suppose* $m \equiv n(\text{mod } d)$ *and* $n \equiv p(\text{mod } d)$. That means that $d|m - n$ and $d|n - p$. Furthermore, that means $\exists x, y \in \mathbf{Z}$, s.t.

$$d \cdot x = m - n \quad \text{and} \quad d \cdot y = n - p$$

We need to show that $d|m - p$, which would imply that $m \equiv p(\text{mod } d)$. The following obviously holds:

$$m - p = (m - n) + (n - p) = d \cdot x + d \cdot y = d(x + y)$$

Since $x, y \in \mathbf{Z}$, $x + y \in \mathbf{Z}$ too. Thus, $d|m - p$.

It follows that $m \equiv p(\text{mod } d)$ indeed. Thus, we have proved that congruence (mod d) is an equivalence relation. ■

Theorem 4.3 Let $a, b, c, d \in \mathbf{Z}$, and let $n \in \mathbf{N}$. If $a \equiv c(\text{mod } n)$ and $b \equiv d(\text{mod } n)$, then

(i) $a + b \equiv (c + d)(\text{mod } n)$
(ii) $a - b \equiv (c - d)(\text{mod } n)$
(iii) $ab \equiv cd(\text{mod } n)$

Proof

(i) $a \equiv c(\text{mod } n)$ means that there exists $x \in \mathbf{Z}$, s.t. $a - c = nx$. Also, $b \equiv d(\text{mod } n)$ means that there exists $y \in \mathbf{Z}$, s.t. $b - d = ny$.
Consider

$$a - c + b - d = (a + b) - (c + d) = nx + ny = n(x + y)$$

Since, $x, y \in \mathbf{Z}$, $(x + y) \in \mathbf{Z}$, too. Hence, $n \mid (a + b) - (c + d)$. Therefore

$$a + b \equiv (c + d)(\text{mod } n)$$

(ii) Similar to (i).

(iii) $a \equiv c(\text{mod } n)$ means that there exists $x \in \mathbf{Z}$, s.t. $a - c = nx$. Also, $b \equiv d(\text{mod } n)$ means that there exists $y \in \mathbf{Z}$, s.t. $b - d = ny$.
Consider

$$\begin{aligned} ab - cd &= ab - cd + bc - bc = b(a - c) + c(b - d) \\ &= b \cdot nx + c \cdot ny \\ &= n(bx + cy) \end{aligned}$$

Hence, $n|ab - cd$ and, therefore, $ab \equiv cd(\text{mod } n)$, what was to be shown. ■

Before we bring in another concept we need to recall the definition of a partition (Definition 1.19): A collection $P = \{X_i\}$ of nonempty, mutually disjoint subsets of a set X whose union is equal to X, such that $\forall\, x \in X, x$, is an element of some X_i, that is

$$\forall\, x,\ x \in \bigcup_i X_i = X \text{ and } \bigcap_i X_i = \varnothing$$

is said to be a partition of the set X (see Figure 4.14).

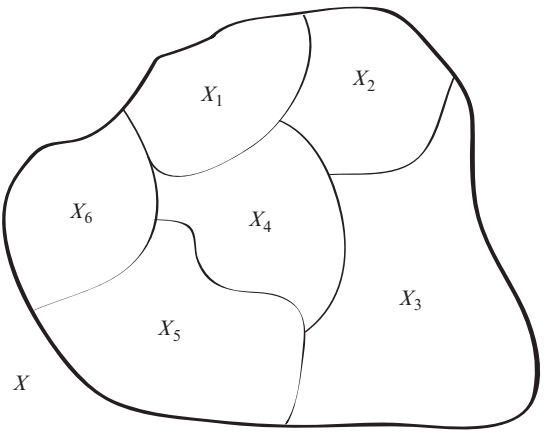


Figure 4.14 Partition $\bigcup_i X_i = X$

Example 4.31 Consider a set $X = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, and the following collection of subsets:

(i) $P_1 = \{X_1 = \{1,3,5\}, X_2 = \{2,3,7\}, X_3 = \{4,6,9\}\}$
(ii) $P_2 = \{X_1 = \{1,3,5\}, X_2 = \{2,4,6,7\}, X_3 = \{8,9\}, X_4 = \{10\}\}$

Observe that P_1 is not a partition since $X_1 \cap X_2 \neq \emptyset$, and 8 and 10 are not elements of any of the subsets. P_2, on the other hand, is a partition of X. ■

Definition 4.15 A relation R on a set X is said to be a **binary relation induced by the partition** if for all $x, y \in X$, xRy iff there is a subset X_i of the partition, such that $x, y \in X_i$.

Example 4.32 Let $X = \{0,1,2,3\}$, and let $X_1 = \{0,2\}, X_2 = \{1\}, X_3 = \{3\}$.
Obviously, X_1, X_2, X_3 form a partition of X. Does this partition induce a relation on X?

Solution Since a binary relation induced by this partition means:

$$\forall\, x, y \in X, \quad xRy \quad \text{iff} \quad x, y\ X_i$$

we get the following:

$0\,R\,0$ since both 0 and 0 are elements of X_1
$0\,R\,2$ since both 0 and 2 are elements of X_1
$2\,R\,0$ since both 2 and 0 are elements of X_1
$2\,R\,2$ since both 2 and 2 are elements of X_1

Similarly,

$3\,R\,3$ since both 3 and 3 are elements of X_2
$1\,R\,1$ since both 1 and 1 are elements of X_2

Hence,

$$R = \{(0,0),(0,2),(2,0),(1,1),(2,2),(3,3)\}$$ ■

Example 4.33 Consider the following four sets:

$$A_0 = \{n | n = 4k, k \in \mathbf{Z}\}$$
$$A_1 = \{m | m = 4k+1, k \in \mathbf{Z}\}$$
$$A_2 = \{p | p = 4k+2, k \in \mathbf{Z}\}$$
$$A_3 = \{r | r = 4k+3, k \in \mathbf{Z}\}$$

Note that the set $A = \{A_0, A_1, A_2, A_3\}$ is a partition of $\mathbf{Z}$. Next, observe that $x, y \in \mathbf{Z}$ are in the some set $A_i \subseteq A$ iff $x - y$ is a multiple of 4. Thus, the partition A induces an equivalence relation of congruence modulo 4. ■

Definition 4.16 Let X be a set and R an equivalence relation on X. If $x \in X$, we say that $\mathbf{C}(x)$ **is an equivalence class of** $\boldsymbol{x}$ (often also denoted by [x]) if

$$C(x) = [x] = \{y \in X | yRx\} = \{y \in X | y \sim x\}$$

The following theorem lists the properties of equivalence classes.

Theorem 4.4

(i) $x \in [x]$
(ii) *If* xRy, *then* $[x] = [y]$
(iii) *If* $x \not R y$, *then* $[x] \cap [y] = \varnothing$
(iv) $\cup_{x \in X}[x] = X$

Proof

(i) xRx therefore $x \in [x]$.
(ii) $xRy \Rightarrow yRx$ therefore $[x] = [y]$.
(iii) Suppose $z \in [x] \cap [y]$, that is, $z \in [x]$ and $z \in [y]$, which means that zRx and zRy.
Hence,
$$[x] \cap [y] \neq \varnothing \quad \text{iff} \quad [x] = [y]$$
(iv) Proof is left as an exercise for the reader. ■

Definition 4.17 A collection of all equivalence classes on X, denoted by X/R, is said to be the **quotient set of** $\boldsymbol{X}$ **by R**, that is

$$X/R = \{C(x) = [x] | x \in X\}$$

Theorem 4.5 Let R be an equivalence relation on a set X. Then the quotient set X/R is a partition of X.

Proof See the previous theorem. ■

Example 4.34 Consider the set $\mathbf{Z}$. We say that $x, y \in \mathbf{Z}$ have the same **parity** if both are even, or both are odd. Suppose we endow the set $\mathbf{Z}$ with a relation $R =$"*has the same parity as*." Then R is an equivalence relation. Observe that this relation induces the partition of $\mathbf{Z}$. Indeed, any element of $\mathbf{Z}$ is either an element of $\mathbf{Z}_{\text{even}}$ or $\mathbf{Z}_{\text{odd}}$, but never an element of both, that is, $\mathbf{Z}_{\text{even}} \cap \mathbf{Z}_{\text{odd}} = \varnothing$. On the other hand, $\mathbf{Z}_{\text{even}} \cup \mathbf{Z}_{\text{odd}} = \mathbf{Z}$, and thus we have an obvious partition of $\mathbf{Z}$. ■

Example 4.35 Let $X = \{1, 2, 3, 4\}$ be a set with a relation R given by

$$R = \{(1, 1), (2, 2), (1, 3), (3, 1), (3, 3), (4, 4)\}$$

(i) Show that R is an equivalence relation.
(ii) Find [1], [2], [3], [4].
(iii) Convince yourself that X/R is a partition of X.

Solution For more transparency, let's first construct the digraph of the relation (Figure 4.15):

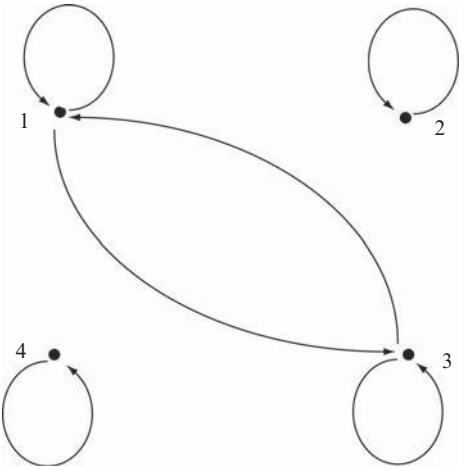


Figure 4.15

One can immediately see that R is reflexive, symmetric, and transitive, hence an equivalence relation. Also, observe that $[1] = [3] = \{1, 3\}$; $[2] = \{2\}$; $[4] = \{4\}$. Finally, without further ado, we identify $X/R = \{[1], [2], [3]\}$ as a partition of X. ■

So, the connection between an equivalence relation and a partition is twofold. Not only does an equivalence relation on a set X determine a partition of X, but also the partition can be used to establish the equivalence relation.

Theorem 4.6 Let $P(X)$ be a partition of a set X. Then there is an equivalence relation R on X such that the quotient set X/R is the same as a partition $P(X)$.

Proof Let R be a relation on X defined by: xRy iff $x, y \in X_i \subseteq X$, that is, x and y belong to the same subset of X. Let's verify that such defined R is indeed an equivalence relation. Take some $x \in X$. Since we have partitioned X, that is,

$X = \cup X_i$ and $\cap X_i = \varnothing$, x has to be an element of some X_i Obviously, xRx, that is, R is reflexive. Next, let $x, y \in X$ be such that xRy. By our definition $x, y \in X_i$, for some $X_i \subseteq X$, and thus yRx, that is, R is symmetric. Finally, take $x, y, z \in X$, such that xRy and yRz. xRy, of course, means that $x, y \in X_i \subseteq X$, and yRz means that $y, z \in X_j \subseteq X$. But, since we partitioned our set X, $X_i \cap X_j = \varnothing$, for $i \neq j$, so the element y can belong to only one subset of X, hence $X_i = X_j$. We conclude that xRz, that is, our relation R is transitive too. We established that R is an equivalence relation, and therefore, with such defined R, $P(X) = X/R$. ■

The converse is also true.

Theorem 4.7 Let R be an equivalence relation on a set X. Then the quotient set X/R is a partition of X.

Example 4.36 Let relation R on $\mathbf{Z}$ be congruence modulo 3. Find the equivalence classes with respect to R.

Solution Since the relation R is defined as

$$aRx \quad \text{iff} \quad a \equiv x (\text{mod } 3)$$

that is,

$$3 \mid a - x, \forall\, a, x \in \mathbf{Z}$$

we have

$$\begin{aligned}[a] &= \{x \in \mathbf{Z} \mid 3 \mid x - a\} \\ &= \{x \in \mathbf{Z} \mid x - a = 3k, k \in \mathbf{Z}\} \\ &= \{x \in \mathbf{Z} \mid x = 3k + a, k \in \mathbf{Z}\}\end{aligned}$$

Therefore,

$$\begin{aligned}[0] &= \{x \in \mathbf{Z} \mid x = 3k + 0\} = \{\ldots, -9, -6, -3, 0, 3, 6, 9, \ldots\} \\ [1] &= \{x \in \mathbf{Z} \mid x = 3\,k + 1\} = \{\ldots, -8, -5, -2, 1, 4, 7, \ldots\} \\ [2] &= \{x \in \mathbf{Z} \mid x = 3k + 2\} = \{\ldots, -7, -4, -1, 2, 5, 8, \ldots\}\end{aligned}$$

Notice that $[0] \cup [1] \cup [2] = \mathbf{Z}$ and $[0] \cap [1] \cap [2] = \varnothing$. Thus $\mathbf{Z}/R$ is a partition of $\mathbf{Z}$. ■

Definition 4.18 Let R be a relation from X to Y, and let S be a relation from Y to Z. We say that the **composition** of R and S is

$$\begin{aligned}R \circ S &= \{(x, z) \mid \exists\, y \in Y \text{ such that } (x, y) \in R \text{ and } (y, z) \in S\} \\ &= x(R \circ S)z\end{aligned}$$

Example 4.37 Let $X = \{1, 2, 3, 4, 5\}$, $Y = \{a, b, c, c, d\}$, and $Z = \{x, y, z, w\}$, and let

$$R = \{(1, a), (2, d), (3, a), (3, b), (3, d)\}$$
$$S = \{(b, x), (b, z), (c, y), (d, z)\}$$

Find $R \circ S$.

Solution Let's first sketch a diagram of the said relations (Figure 4.16):

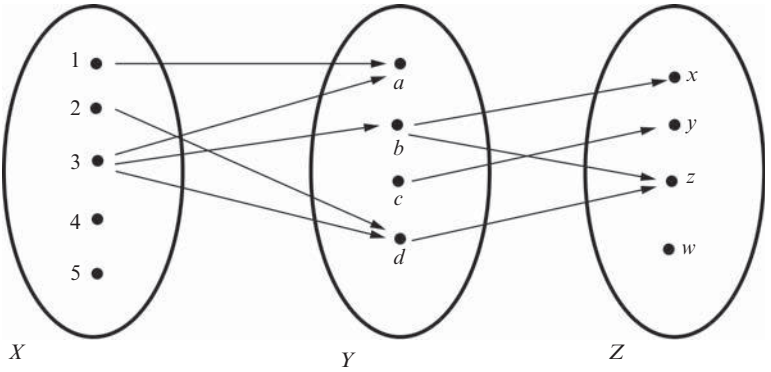


Figure 4.16

Now it's easy to see that, for example, 2 is mapped to d, which, in turn, is mapped to z, so we write

$$2Rd \quad \text{and} \quad dSz$$

In other words,

$$2(R \circ S)z$$

Similarly, we observe that $3(R \circ S)x$, and $3(R \circ S)z$. On the other hand, 1 "goes" to a, but a does not "go further," that is, $1Ra$ but a is not related to any element of Z. We evaluate the "paths" of all elements in the same way and conclude that our composition is

$$R \circ S = \{(2, z), (3, x), (3, z)\}$$ ■

You should be able to convince yourself easily that in general $R \circ S \neq S \circ R$.

Example/Exercise 4.38 Let $R = \{(x, y) \in \mathbf{R} \times \mathbf{R} | y = x + 1\}$ and $S = \{(x, y) \in \mathbf{R} \times \mathbf{R} | y = x^2\}$.

Show that $R \circ S \neq S \circ R$.

Example 4.39 Let $X = \{1, 2, 3, 4\}$ be a set with a relation

$$R = \{(1,1),\ (2,2),\ (2,3),\ (3,2), (4,2), (4,4)\}$$

Find R^2.

Solution $R^2 = R \circ R$, thus $\forall\,(x, y) \in R, \exists (y, z) \in R,$ s.t. $(x, z) \in R^2$. Thus,

$$R^2 = \{(1,1),\ (2,2),\ (2,3), (3,2), (3,3), (4,2), (4,3), (4,4)\}$$

See Figures 4.17 and 4.18. ■

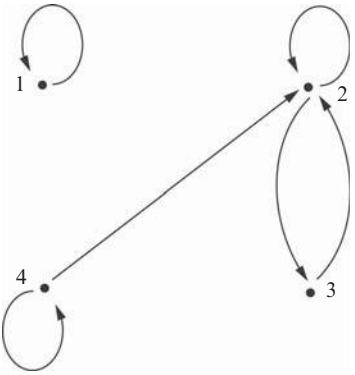


Figure 4.17 R

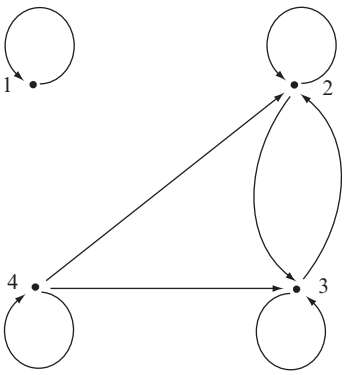


Figure 4.18 R^2.

As expected, given a relation R on an nonempty set X, we define $R^3 = R{\circ}R{\circ}R = R^2{\circ}R$, $R^4 = R{\circ}R{\circ}R{\circ}R = R^3{\circ}R$, and so on. In general,

$$R^n = R^{n-1}{\circ}R$$

Theorem 4.8 Let X, Y, Z, W be four nonempty sets, and let R be a relation from X to Y, let S be a relation from Y to Z, and finally let T be a relation from Z to W. Then,

(i) $T{\circ}(S{\circ}R) = (T{\circ}S){\circ}R$
(ii) $I_Y{\circ}R = R \quad \text{and} \quad R{\circ}I_X = R$
(iii) $(S{\circ}R)^{-1} = R^{-1}{\circ}S^{-1}$

Proof

(i) Let $(x, w) \in T{\circ}(S{\circ}R)$ for some $x \in X$ and $w \in W$. Then,

$$T{\circ}(S{\circ}R) = (T{\circ}S){\circ}R$$

iff $(\exists z \in Z)[(x, z) \in S{\circ}R, \ \textit{and} \ (z, w) \in T]$
iff $(\exists z \in Z)[(\exists y \in Y)((x, y) \in R, (y, z) \in S, \ \textit{and} \ (z, w) \in T)]$
iff $(\exists z \in Z)(\exists y \in Y)[(x, y) \in R, (y, z) \in S, \ \textit{and} \ (z, w) \in T]$
iff $(\exists y \in Y)(\exists z \in Z)[(x, y) \in R, (y, z) \in S, \ \textit{and} \ (z, w) \in T]$
iff $(\exists y \in Y)[(x, y) \in R, \ \textit{and} \ (\exists z \in Z)((y, z) \in S \ \textit{and} \ (z, w) \in T)]$
iff $(\exists y \in Y)[(x, y) \in R \ \textit{and} \ (y, w) \in T{\circ}S]$
iff $(x, w) \in (T{\circ}S){\circ}R$

(ii) Suppose $(x, y) \in I_Y{\circ}R$. Since R is a relation from X to Y, there has to be a $z \in Y$ s.t. $(x, z) \in R$ and $(z, y) \in I_Y$. But $(z, y) \in I_Y$ means that $z = y$, and thus

$$(x, z) = (x, y) \in R$$

Conversely, if $(x, y) \in R$ then, certainly, $(y, y) \in I_Y$, thus $(x, y) \in I_Y{\circ}R$, and so, indeed,

$$I_Y{\circ}R = R$$

Similarly, you can prove that $R{\circ}I_X = R$.

(iii) Having established (i) and (ii) and keeping in mind Definitions 4.11 and 4.18, we argue as follows:

$$(S \circ R)^{-1} \circ (S \circ R) = I$$

On the other hand, consider

$$\begin{aligned}(R^{-1} \circ S^{-1}) \circ (S \circ R) &= R^{-1}(S^{-1} \circ S)R \\ &= R^{-1}R = I\end{aligned}$$

Thus,

$$(S \circ R)^{-1} = R^{-1} \circ S^{-1}$$ ■

Theorem 4.9 Let R be a relation on a set X. Then

(i) R is reflexive iff $I_X \subseteq R$
(ii) R is symmetric iff $R = R^{-1}$
(iii) R is transitive iff $R \circ R \subseteq R$

Proof

(i) is obvious. Let's prove.

(ii) Suppose, first, that R symmetric. Then $(x, y) \in R$ implies that $(y, x) \in R$. But that means that $(x, y) \in R^{-1}$. Thus $R = R^{-1}$. Next, suppose $R = R^{-1}$. Then $(x, y) \in R$ implies $(x, y) \in R^{-1}$, which furthermore implies that $(y, x) \in R$. Hence, R is symmetric.

(iii) Suppose R is transitive. That means:

$$\forall\, x, y, z \in X \text{ if } (x, y) \in R \text{ and } (y, z) \in R, \text{then } (x, z) \in R$$

In other words,

$$R \circ R = \{(x, z) | \exists y \in X \text{ s.t. } (x, y) \in R \text{ and } (y, z) \in R\} \subseteq R$$

On the other hand, if $R \circ R \subseteq R$ then $\forall\, (x, z) \in R \circ R, (x, z) \in R$. But $(x, z) \in R \circ R$ means $\exists x, y \in R$ s.t. $(x, y) \in R$ and $(y, z) \in R$ and thus, R is transitive. ■

Nobody can explain what a function is, but this is what really matters in mathematics.

H. Weyl[5]

[5]H.Weyl, *Philosophy of Mathematics and Natural Science,* Princeton University Press, 1949, 2009.

4.3 FUNCTIONS

Definition 4.19 Let X and Y be two sets. W*e say that a* ***function f from set X to set Y*** *is a map that assigns to every element of X a unique element of Y. We write this as follows:*

$$f : X \to Y$$

See Figure 4.19.

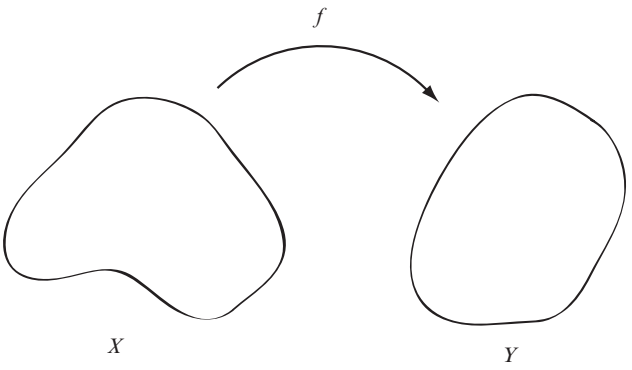


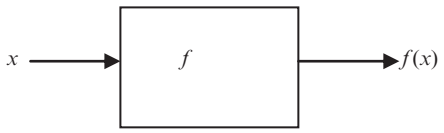
Figure 4.19 Function

If $X = Y$ we say that f is a function on X.

If there is no need to explicitly name the function, we abbreviate the notation by writing

$$x \mapsto f(x)$$

It might be convenient sometimes to think of a function f as a "black box," a "machine," into which one can feed any element $x \in X$ in order to obtain $f(x)$ emerging at the other end.



Often we simply say that f sends x to $y = f(x)$ and we write

$$X \xrightarrow{f} Y$$

Usually, the set X is called the **domain of *f***, denoted $\mathcal{D}(f)$, and Y is called the **codomain of *f*** or **range of *f***, denoted $\mathcal{R}(f)$. However, this is not exactly correct.

There is a fine, but important, subtlety that has to be addressed here. As you probably remember from your high school algebra, it may happen that when one is constructing a map $X \xrightarrow{f} Y$ from an arbitrary set X to Y, a function f simply does not "recognize" certain parts of X and cannot "carry" them over to Y. What is meant by this is that once the function is "defined," one should be careful to eliminate the parts of X (a subset of X) that would make $f(x)$ nonsensical. In other words, the *domain* of the function is the largest subset of X for which the formula describing f has a meaning. In a sense, the domain of the function is a part of the definition of the function. Analogously, we talk about the *range* of f. Thus, we have

Definition 4.20 If X and Y are two sets with a function f between them, we say that the **domain of f** is a subset of X defined by

$$\mathcal{D}(f) = \{x \in X | \exists y \in Y, \text{ s.t. } (x, y) \in f\} \subseteq X$$

Analogously, the **range of f** is a subset of Y defined by

$$\mathcal{R}(f) = \{y \in Y | \exists x \in X \text{ s.t. } (x, y) \in f\} \subseteq Y$$

Of course, it is quite possible that $\mathcal{D}(f) = X$ and $\mathcal{R}(f) = Y$, in which case we simply say that f is a function *from X to Y*, and we don't worry about excluding any $x \in X$ that would make our function nonsensical (Figure 4.20).

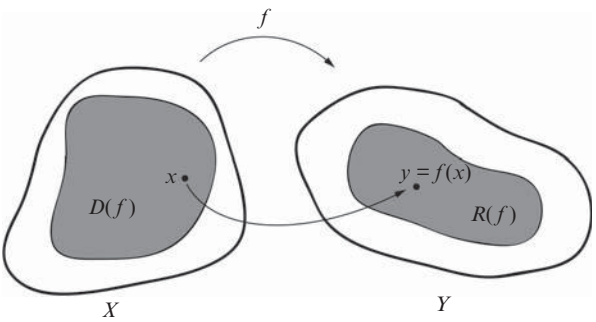


Figure 4.20 Domain and range of f

Let's look at a few examples.

Example 4.40 If $\mathbf{R}$ is the set we are working with, and we are considering a function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^2$ or, simply, by $x \mapsto x^2$, indicating the function that maps every real number to its square. We immediately see that the function is *well defined* for every $x \in \mathbf{R}$ and we say that $\mathcal{D}(f) = \mathbf{R}$.

On the other hand, if our function is defined by $f(x) = 1/x$, we have to exclude zero since the function "doesn't know what to do" with zero – *it is not defined at zero*. In this case, we say that

$$\mathcal{D}(f) = \{x \in \mathbf{R} | x \neq 0\}$$

So, our map is

$$f : \mathbf{R}\backslash\{0\} \to \mathbf{R}$$

■

Example 4.41 A function $f : \mathbf{R} \to \mathbf{R}$, defined *by* $f(x) = e^x$, is well defined for all $x \in \mathbf{R}$. We say that $\mathcal{D}(f) = \mathbf{R}$. ■

Example 4.42 Let $X = [-1, 1]$ and $Y = [-2, 3]$, and let $f = \{(x, y) | y = \sqrt{x} \in \mathbf{R}\}$. In this case, it would be **incorrect** to write

$$f : [-1, 1] \to [-2, 3]$$

but it is **correct** to write

$$f : [0, 1] \to [-2, 3]$$

■

Since the concept of a function is so important, some additional and equivalent definitions may be useful.

Definition 4.19′ A function f from a set X to a set Y is a rule that assigns to each $x \in X$ a *unique* $y = f(x) \in Y$.

Definition 4.19″ Let X and Y be two nonempty sets. A function f from set X to set Y is a subset of $X \times Y$, denoted

$$f = \{(x, y) | x \in X, y \in Y\}$$

meaning that

$$\forall\, x \in X, \quad \exists!\, y \in Y \quad \text{s.t.} \quad (x, y) \in f$$

In other words, f is *a relation with the property of being single-valued*.

We say that $y = f(x)$ is the **value of f**, or the **image of f**. The set $\boldsymbol{R(f)}$ of all values of f taken together is called the **range of f**; x is said to be the **preimage of y under f**.

Definition 4.19‴ We say that an ordered triple of sets $(X, Y; f)$ is a function, where X is a domain, Y is a codomain and f is a set of ordered pairs (x, y) such that for every $x \in X$ there is exactly one $y \in Y$.

Definition 4.19 [(iv)] Let X and Y be any two sets. A function f, from set X to set Y, is a nonempty *relation* $f \subseteq X \times Y$ such that if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$.

Example 4.43 Let $X = \{a, b, c, d, e\}$ and $Y = \{1, 2, 3\}$. Define a function from X to Y that assigns a number to every letter (see Figure 4.21). One function, say, f could be

$$f = \{(a, 1), (b, 2), (c, 3), (d, 1), (e, 2)\}$$

One simple and obvious way to represent this mapping could be

$$f = \begin{pmatrix} a & b & c & d & e \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 2 & 3 & 1 & 2 \end{pmatrix}$$

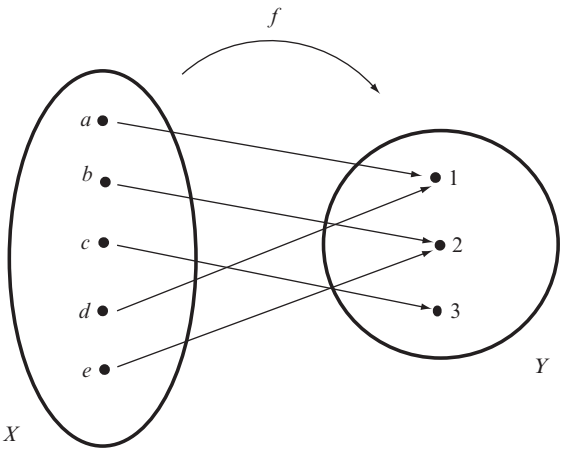


Figure 4.21 The function f

Another function, g (see Figure 4.22), could be

$$g = \{(a, 2), (b, 2), (c, 2), (d, 2), (e, 2)\}$$

Obviously, the function g sends all letters to number 2, that is

$$g = \begin{pmatrix} a & b & c & d & e \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 2 & 2 & 2 & 2 \end{pmatrix}$$

■

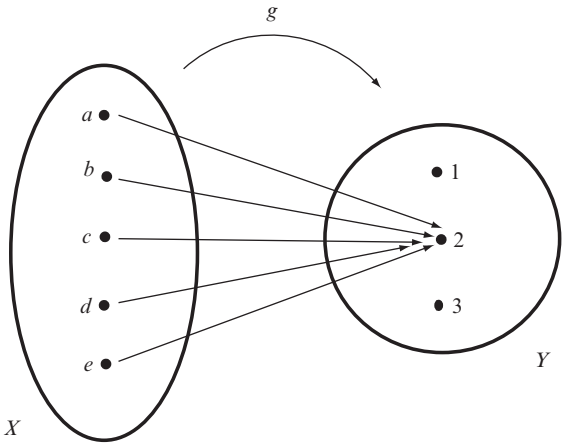


Figure 4.22 Function g

Definition 4.21 A function $f : X \to Y$ is said to be a **real-valued function**, if $\mathcal{R}(f) \subseteq \mathbf{R}$, and simply, a **real function** if both $X \subseteq \mathbf{R}$ **and** $Y \subseteq \mathbf{R}$.

Example 4.44 A function $f : \mathbf{R} \to \mathbf{R}$, defined by $f(x) = x^2$, is an example of a real function. It is worth noting that very often $f(x)$ is (wrongly) referred to as a function. Again, strictly speaking, $f(x)$ is the image of a real number x under the action of the function f. f itself is actually a set

$$f = \{(x, x^2) | x \in \mathbf{R}\}$$

■

Example 4.45 Suppose we are given the following relation:

$$R = \{(\varnothing, \{\varnothing\}), (\{\varnothing\}, \varnothing), (\varnothing, \varnothing), (\{\varnothing\}, \{\varnothing\})\}$$

Is this relation a function or not? Well, it is not! Observe that we have two cases of one element related to two, that is, we have Figure 4.23.

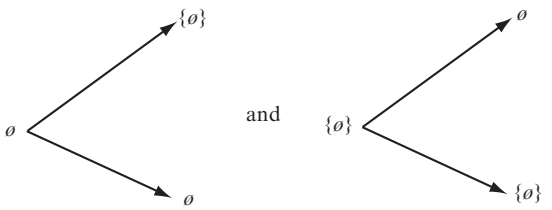


Figure 4.23

Thus, relation R is not a function. ■

Example 4.46 Consider a relation $S = \{(x, y) \in \mathbf{R} \times \mathbf{R} | x^2 + y^2 = 1\}$ with the domain $[-1, 1]$. It is easy to see that S is not a function from $[-1, 1]$ to $\mathbf{R}$ since, for example, both $((\sqrt{2}/2), (\sqrt{2}/2))$ and $((\sqrt{2}/2), -(\sqrt{2}/2))$ are the elements of S (see Figure 4.24). ■

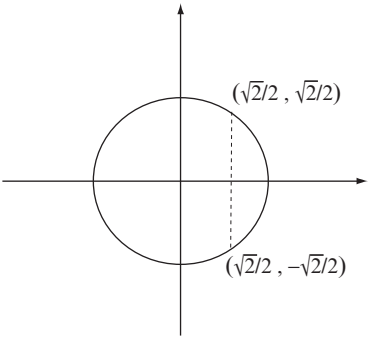


Figure 4.24

Example 4.47 Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $f(x) = x^3, \forall\, x \in \mathbf{R}$. That is, f assigns to every real number its cube. Then, for instance, $f(2) = 8$ and $f^{-1}(8) = 2, f(-3) = -27$ and $f^{-1}(-27) = -8$, and so on. ■

Example 4.48 Consider a relation R defined as follows:

$$R = \{(x, y) \in \mathbf{R} \times \mathbf{R} | y = x^2 + 2x + 1\}$$

The relation R is another example of a function from $\mathbf{R}$ to $\mathbf{R}$. ■

Another way to think about functions is to consider their *graphs*:

Definition 4.22 Let $f : X \to Y$ be a function. We say that a relation

$$G(f) = \{(x, y) | x \in X, y = f(x) \in Y\}$$

is the ***graph of f.***

Example 4.49 Let $f : X \to Y$ be a function from $X = \{1, 2, 3, 4\}$ to $Y = \{a, b, c, d\}$ defined by $f(1) = b$, $f(2) = d$, $f(3) = a$, $f(4) = b$. Then the graph of f is the following set:

$$G(f) = \{(1, b), (2, d), (3, a), (4, b)\}$$

■

Definition 4.23 Consider a universal set $\mathcal{U}$, and one of its subsets $A \subseteq \mathcal{U}$. We say that the function $\chi_A : \mathcal{U} \to \{0, 1\}$ defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

is a **characteristic function** of A.

Example 4.50 Suppose $\mathcal{U} = \{a, b, c, d, e, f, g, h\}$ and $A = \{a, d, e\}$, then the characteristic function

$$\chi_A = \{(a, 1), (b, 0), (c, 0), (d, 1), (e, 1), (f, 0), g(0), (h, 0)\}$$

■

Example 4.51 Let $\mathcal{U} = \mathbf{R}$ and $A = [1, 4)$. Then the graph of the characteristic function χ_A is shown in Figure 4.25. ■

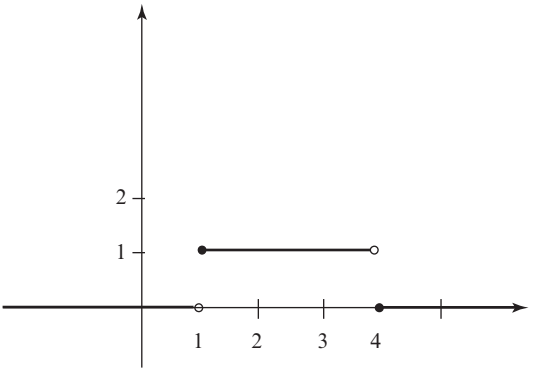


Figure 4.25

Theorem 4.10 Let $\mathcal{U}$ be a universal set, and let $A \subseteq \mathcal{U}$ and $B \subseteq \mathcal{U}$ be two of its subsets. Then

$$\chi_{A \cap B} = \chi_A \chi_B$$

Proof Let $x \in A \cap B$, that is, $x \in A$ and $x \in B$. Then, by definition of characteristic function,

$$\chi_{A \cap B}(x) = 1$$

On the other hand,

$$\chi_A(x)\chi_B(x) = 1 \cdot 1 = 1$$

Thus,

$$\chi_{A\cap B}(x) = \chi_A(x)\chi_B(x) = 1$$

Now, let's take some $y \notin A \cap B$, which would, of course, imply

$$\chi_{A\cap B}(y) = 0$$

On the other hand, if $y \notin A \cap B$ then for sure $y \in (A \cap B)^c = A^c \cup B^c$, and therefore $y \in A^c$ or $y \in B^c$. This means that

$$\chi_A(y) = 0 \quad \text{or} \quad \chi_B(y) = 0$$

In any case

$$\chi_A(y)\chi_B(y) = 0$$

So, we have

$$\chi_{A\cap B}(y) = \chi_A(y)\chi_B(y) = 0$$

Since $\chi_{A\cap B}$ and $\chi_A\chi_B$ assign the same number to each element of $\mathcal{U}$, we conclude that, indeed,

$$\chi_{A\cap B} = \chi_A\chi_B$$ ■

Here is another function you can often come across in many different branches of mathematics.

Definition 4.24 A function $d : X \times X \to \mathbf{R}$ is called a **metric** on a set X, if $\forall\, x, y, z \in X$

(i) $d(x, y) \geq 0$
(ii) $d(x, y) = 0$ iff $x = y$
(iii) $d(x, y) = d(y, x)$
(iv) $d(x, y) + d(y, z) \geq d(x, z)$

Example 4.52 Let $X = \mathbf{N}$. Then $d(x, y) = |x - y|, \forall\, x, y, z \in \mathbf{N}$ is a metric on $\mathbf{N}$.

Indeed, since $x, y, z \in \mathbf{N}$

(i) $| x - y | \geq 0 \quad \forall\, x, y \in \mathbf{N}$
(ii) $|x - y| = 0$ iff $x = y$

(iii) $d(x, y) = |x - y| = |y - x| = d(y, x)$
(iv) $|x - y| + |y - z| \geq |x - z|$ that is $d(x, y) + d(y, z) \geq d(x, z)$ ■

Similarly,

Example 4.53 Let $X = \mathbf{R} \times \mathbf{R}$. Then $d((x, y), (z, w)) = \sqrt{(x - z)^2 + (y - w)^2}$ is a metric on $\mathbf{R} \times \mathbf{R}$. ■

Definition 4.25 Suppose we are given a function $f : X \to Y$, and $A \subseteq X$. We say that a set

$$f[A] = \{f(x) | x \in A\}$$

is the **image** of A under the action of f.

Consequently, if $B \subseteq Y$ we call

$$f^{-1}[B] = \{x \in X | f(x) \in B\}$$

the **preimage** of B under the action of f (Figure 4.26).

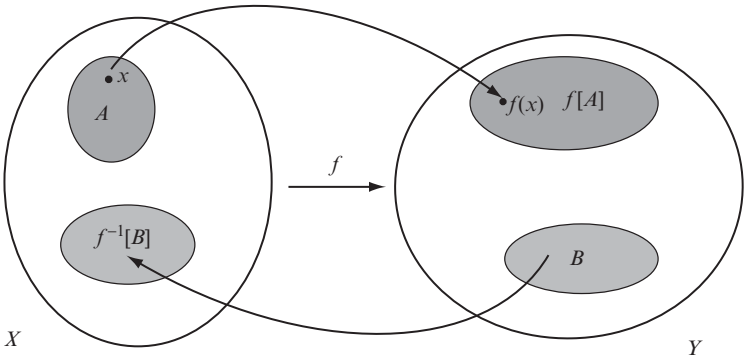


Figure 4.26 Image and preimage

Example 4.54 Let $X = \{-1, -2, -3, 0, 1, 2, 3\}$ and $Y = \{0, 1, 2, 4, 6, 9\}$, and let $f : X \to Y$ be a function from X to Y defined by

$$f(x) = x^2$$

If $A = \{-1, 3\} \subseteq X$ let's find $f(A)$ and $f^{-1}(B)$.

In this case, the analogue of Figure 4.26 is Figure 4.27.

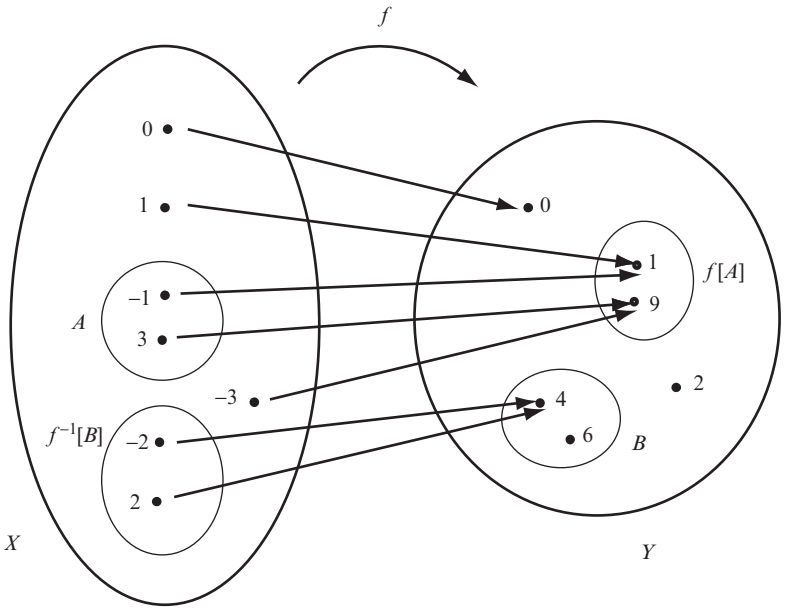


Figure 4.27

So, we have:

$$f[A] = f[\{-1, 3\}] = \{1, 9\}$$

and

$$f^{-1}[B] = f^{-1}[\{4, 6\}] = \{-2, 2\}$$
$$f^{-1}[\ \{\ 6\}] = \varnothing$$

Observe, also, that

$$f(X) = \ \{0,\ 1,\ 4,\ 9\ \}$$

and that f^{-1} is not a function. ■

Example/Exercise 4.55 Consider a function $f : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ defined by

$$f(m, n) = 2^m 3^n$$

and find $f(A \times B)$ if $A = \{1, 2, 3\}$ and $B = \{3, 4\}$.

Example 4.56 Consider again a function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^2$, and this time let $A = [\ 1, 2]$. Find $f[A]$.

Solution First, let's sketch the graph of the function f and the set A (Figure 4.28).

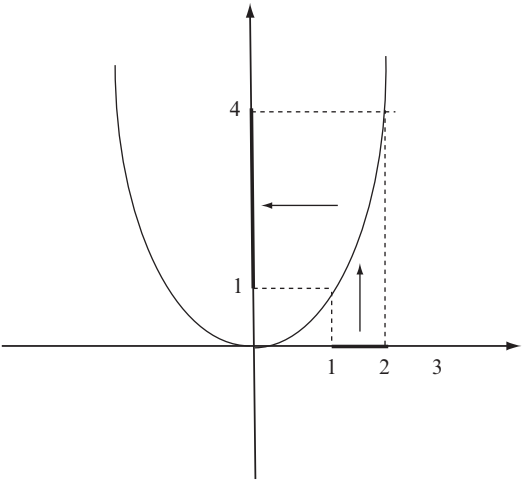


Figure 4.28 $A = [1, 2]; f[A] = [1, 4]$

Now it's easy to see from the figure that $f[A] = f([1,\ 2]) = [1,\ 4]$. ■

Definition 4.26 We say that $i : X \to X$ is the **identity function** on a set X if

$$i(x) = x, \quad \forall\, x \in X$$

In other words, the function i sends every element from X into itself (Figure 4.29).

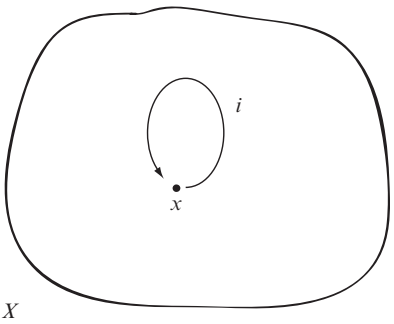


Figure 4.29 Identity function

Definition 4.27 Two functions f and g are said to be **equal** iff they have the same domain and assign the same value to every member of their common domain. Symbolically,

$$f = g \Leftrightarrow \mathcal{D}(f) = \mathcal{D}(g) \text{ and } f(x) = g(x), \quad \forall\, x \in \mathcal{D}(f)$$

As a motivation for the introduction of some more important concepts, let's mention another simple and rather common sense principle with far and wide applications. It is attributed to J.P.G.L. Dirichlet[6] who called it *Schubfach-prinzip* – a drawer/ box principle, or a Pigeonhole principle. Indeed, nothing can be more obvious: If there are more pigeons than the pigeonholes, then there must be at least one pigeonhole with at least two pigeons in it.

Before we proceed more formally, we need to recall some important concepts from Chapter 1.

Definition 4.28 We say that a function $f : X \to Y$ is ***one-to-one*** (or an ***injection***) iff

$$\forall x_1, x_2 \in X \text{ if } f(x_1) = f(x_2) \text{ then } x_1 = x_2$$

Or, equivalently,

$$\forall x_1, x_2 \in X \text{ if } x_1 \neq x_2 \text{ then } f(x_1) \neq f(x_2)$$

In other words, a function is injective if no element of Y is the image of more than one element of X (Figure 4.30).

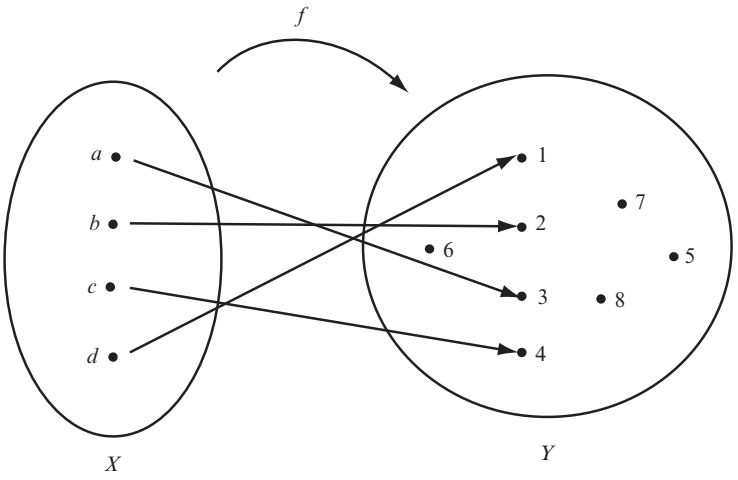


Figure 4.30 Injection

(Sometimes injections are designated with the special arrow "↣," that is, in case the function is injection we write $f : X \rightarrowtail Y$ or even simpler $X \rightarrowtail Y$.)

[6]Johann Peter Gustav Lejeune Dirichlet (1805–1859), German mathematician.

Definition 4.29 We say that a function $f : X \to Y$ is ***onto*** *(*or a ***surjection****)* if and only if

$$\forall\, y \in Y, \exists\, x \in X, \quad \text{s.t. } f(x) = y$$

(Sometimes we use "$\twoheadrightarrow$" to indicate surjection; see Figure 4.31.)

Definition 4.29′ If $f : X \to Y$ is such that $\mathcal{R}(f) = Y$, we say that f is surjection from X to Y.

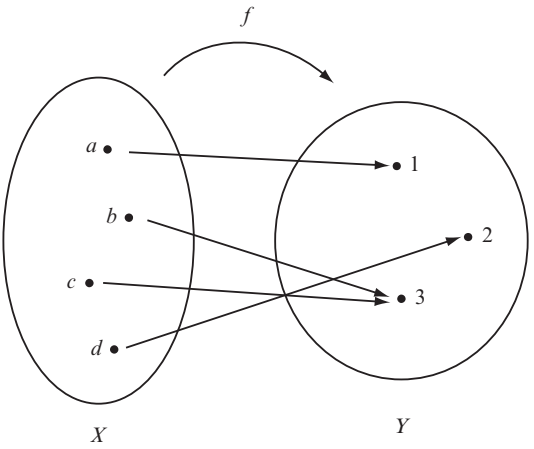


Figure 4.31 Surjection

Example 4.57 Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $f(x) = ax + b$. Show that f is one-to-one.

Solution Suppose $f(x_1) = f(x_2)$, that is, suppose

$$ax_1 + b = ax_1 + b$$

It follows immediately that $x_1 = x_2$. The converse is trivial, thus our function is one-to-one.

■

Example 4.58 Similarly, it is easy to see that $f : \mathbf{R} \to \mathbf{R}$ defined by

$$f(x) = \frac{2x}{x+1}$$

is one-to-one. ■

Example 4.59 Let $f : \mathbf{R} \to \mathbf{R}$ be defined by

$$f(x) = \frac{x}{x^2 + 1}$$

Show that f is *not* one-to-one.

Solution Suppose that f is one-to-one. Consider

$$f(x_1) = \frac{x_1}{{x_1}^2 + 1} \quad \text{and} \quad f(x_2) = \frac{x_2}{{x_2}^2 + 1}$$

If f is one-to-one then $f(x_1) = f(x_2)$ should imply that $x_1 = x_2$. Well, let's see:

$$\begin{aligned} f(x_1) = f(x_2) \Rightarrow \quad & \frac{x_1}{{x_1}^2 + 1} = \frac{x_2}{{x_2}^2 + 1} \\ \Rightarrow \quad & x_1({x_2}^2 + 1) = x_2({x_1}^2 + 1) \\ \Rightarrow \quad & x_1{x_2}^2 + x_1 = x_2{x_1}^2 + x_2 \\ \Rightarrow \quad & x_1x_2(x_2 - x_1) = x_2 - x_1 \\ \text{Therefore, } \Rightarrow \quad & x_1x_2 = 1 \\ \text{that is } x_1 = \quad & 1/x_2 \end{aligned}$$

Thus, assuming that $f(x_1) = f(x_2)$ we have obtained $x_1 \neq x_2$.

Now, take any two $x_1, x_2 \in \mathbf{R}$, such that $x_1 \neq x_2$ but $x_1x_2 = 1$. For instance, $x_1 = 2$ would imply $x_2 = 1/2$. Then, we would have

$$f(x_1) = \frac{2}{4 + 1} = \frac{2}{5}$$

$$f(x_2) = \frac{\frac{1}{2}}{\frac{1}{4} + 1} = \frac{2}{5}$$

Thus, with $x_1 \neq x_2$, we have obtained $f(x_1) = f(x_2)$.

We conclude that f is *not* one-to-one. ■

Example 4.60 Let $f : [0, \infty] \to [0, \infty]$ be a function defined by $f(x) = x^2$. Show that f is one-to-one.

Solution Suppose $f(x) = f(y), x, y \in [0, \infty] = \mathcal{D}(f)$. Then, $x^2 = y^2$. But $x, y \in \mathcal{D}(f)$ implies $x \geq 0$ and $y \geq 0$. Hence, we conclude that the only possible consequence of $x^2 = y^2$ is $x = y$. The converse is obvious.

Note, however, if the function f was a real function $f : \mathbf{R} \to \mathbf{R}$, defined as before by $f(x) = x^2$, we immediately recognize that it is not one-to-one since, for instance,

$$f(2) = f(-2) = 4$$ ■

Example 4.61 Consider a function $f : \mathbf{R} \to \mathbf{R}$ defined as follows:

$$f(x) = \begin{cases} x^2 + 2 & \text{if } x \geq 2 \\ 2 - x & \text{if } x < 2 \end{cases}$$

Observe that f is one-to-one. ■

Example/Exercise 4.62 Determine whether the function $f : \mathbf{R} \to \mathbf{R}$ defined by

$$f(x) = \begin{cases} 2 - x & \text{if } x \leq 1 \\ \dfrac{1}{x} & \text{if } x > 1 \end{cases}$$

is one-to-one or not.

Example 4.63 Consider the function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = \sin x$. Clearly, it is not one-to-one since, for instance, $\sin 0 = \sin 2\pi = 0$. ■

Example 4.64 Show that a function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = 3x - 5$ is onto.

Solution According to Definition 4.29, we need to show that every element in the co-domain is an image of some element in the domain. So, let $y \in \mathbf{R}$ be an arbitrary element from the co-domain, such that $y = f(x) = 3x - 5$. Then, $x = (y + 5)/3$ is also an element of $\mathbf{R}$, and therefore

$$f(x) = f\left(\frac{y+5}{3}\right) = 3\frac{y+5}{3} - 5 = y$$

As there are no restrictions on y, we can see that our function is indeed onto.

■

Note that, since a function is always onto its range, the question whether it is a surjection depends on the choice of the co-domain.

Now we return to the pigeonhole principle.

Definition 4.30 (Pigeonhole principle) Let X and Y be two *finite* sets with m and n elements, respectively, and let $m > n$. Then, there must be at least two elements in X that have the same image in Y. That is, a function from X to Y has

to send at least two elements from X to one element of Y. We say that $X \xrightarrow{f} Y$ *cannot be one-to-one*.

More formally, we have

Theorem 4.11 Let $n, r \in \mathbf{N}$ such that $n > r$, and let $\mathbf{N}_n = \{1, 2, 3, \ldots, n | n \in \mathbf{N}\}$. Then a function $f : \mathbf{N}_n \to \mathbf{N}_r$ is not one-to-one.

Equivalently, we could say

Theorem 4.12 For any $n \in \mathbf{N}$, if $f : \mathbf{N} \to \mathbf{N}$ is one-to-one function, then f is onto.

Example 4.65 Suppose there are 13 students attending a math seminar. Then at least 2 of them were born in the same month. ■

Example 4.66 Let's consider again the function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^2$. Observe that f is not a surjection since, for instance, -1 is an element of the codomain but not an element of the range of f. On the other hand, if we take our function to be $f : \mathbf{R} \to \mathbf{R}^+$, we claim that the surjection is recovered. Indeed, as we know, the range is a subset of the co-domain, that is, $\mathcal{R}(f) \subseteq \mathbf{R}^+$. In order to show that $\mathcal{R}(f) = \mathbf{R}^+$ we need to show that $\mathbf{R}^+ \subseteq \mathcal{R}(f)$ too. Well, if we take any $y \in \mathbf{R}^+$, then $\sqrt{y} \in \mathbf{R}$. Therefore, $\mathbf{R}^+ \subseteq \mathcal{R}(f)$. Thus, the proof. ■

Definition 4.31 A function $f : X \to Y$ that is both **one-to-one and onto** we call a **bijection** or **one-to-one correspondence** between sets X and Y (sometimes we use "$\rightarrowtail$" to indicate bijection; see Figure 4.32).

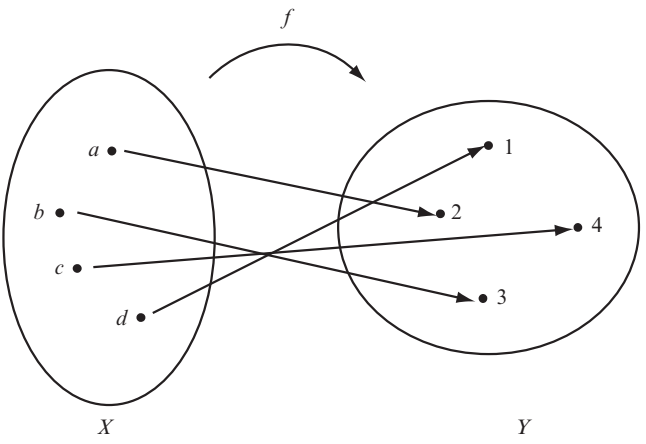


Figure 4.32 Bijection

As is evident from Figure 4.32, being one-to-one and onto means that every element of X has one and only one image in Y; and vice versa, every element in Y has one and only one preimage in X.

Example/Exercise 4.67 If X is a finite set, any surjective function $f : X \to X$ is necessarily injective. Why?

Example/Exercise 4.68 Show that the function from Example 4.64 is a bijection.

Example/Exercise 4.69

(i) Prove that $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = 3x - 5$ is a bijection.
(ii) Prove that $f : \mathbf{Z} \to \mathbf{Z}$ defined by $f(x) = 3x - 5$ is not a bijection.
(iii) Prove that $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^4 - x^2$ is not a bijection.

Example/Exercise 4.70 Let $X = Y = \mathbf{N}$. Define function $f : \mathbf{N} \to \mathbf{N}$ as follows:

$$f(n) = 2n, \quad \forall\ n \in \mathbf{N}$$

Evidently, $f(1) = 2,\ f(2) = 4,\ f(3) = 6,\ \ldots$. As you can see, our function is actually $\mathbf{N} \to \mathbf{N}_{\text{even}}$. We have already encountered this map in Chapter 1 where we have shown that $|\mathbf{N}| = |\mathbf{N}_{\text{even}}|$. I hope you can now easily show that $\mathbf{N} \to \mathbf{N}_{\text{even}}$ is a bijection.

Example 4.71 Recall that two sets are said to be equivalent iff there is a bijection between them. Let's prove that for any $a, b, c, d \in \mathbf{R}$, such that $a < b$ and $c < d$, the intervals (a, b) and (c, d) are equivalent.

Proof Consider a function $f : (a, b) \to (c, d)$ defined by

$$f(x) = \frac{d - c}{b - a}(x - a) + c, \quad \forall\, x \in (a,\ b)$$

It is fairly obvious that our function is a linear function passing through points (a, c) and (b, d). Thus, it is easily verifiable that it is a bijection. At the same time, the function is restricted to the domain (a, b) hence, $(a, b) \sim (c, d)$. ■

Definition 4.32 Let $f : X \to Y$ be a bijection. We say that

$$f^{-1} : Y \to X$$

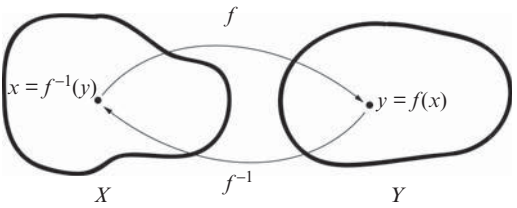


Figure 4.33 Inverse function

is the **inverse function** of f (Figure 4.33) if the following is true

$$f^{-1}(y) = x \Leftrightarrow f(x) = y$$

In other words,

$$f^{-1} = \{(y, x) \in Y \times X | (x, y) \in f\}$$

As we know by now, a relation f^{-1} does not necessarily have to be a function. Indeed, suppose there is a function $f : X \to Y$, then the relation in $Y \times X$

$$f^{-1} = \{(y, x) | (x, y) \in X, \text{ s.t. } y = f(x)\}$$

is not a function in general.

Example 4.72 Consider a mapping

$$f : [-1, 1] \to [0, 1]$$

given by

$$f(x) = x^2$$

Then, $(-1, 1) \in f^{-1}$, but also $(1, 1) \in f^{-1}$, which violates our definition of a function. ■

So, which functions have inverses and which do not? Intuitively, we feel that if $f : X \to Y$ is not injective, then there are at least two distinct elements $x_1, x_2 \in X$ with the same image $y \in Y$, and thus we would have $x_1 = f^{-1}(y)$, but also $x_2 = f^{-1}(y)$, which immediately eliminates f^{-1} as a function (Figure 4.34).

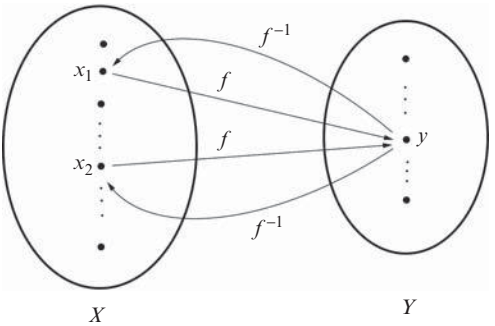


Figure 4.34

Similarly, if f is not surjective, then there is an element $y \in Y$ that is not an image of any element $x \in X$, and so $f^{-1}(y)$ does not exist (Figure 4.35).

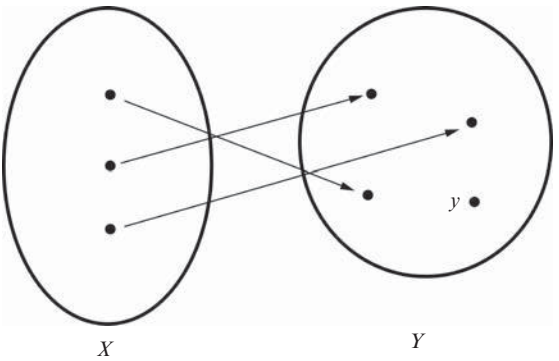


Figure 4.35

Theorem 4.13 If $f : X \to Y$ is a function, then $f^{-1} : Y \to X$ is also a function iff f is a bijection.

Proof Suppose $f : X \to Y$ is a function. Then,

(i) For f to be onto: Suppose f^{-1} is a function. Then there exists a unique $x \in X$, and $y \in Y$, such that $f^{-1}(y) = x$. But that means that $f(x) = y$ and thus, f is onto.

(ii) For f to be one-to-one: Suppose $y_1, y_2 \in Y$ such that

$$f^{-1}(y_1) = f^{-1}(y_2) = x$$

We need to show that $y_1 = y_2$. Well, since $f^{-1}(y_1) = x$ and $f^{-1}(y_2) = x$, it follows that

$$f(x) = y_1 \quad \text{and} \quad f(x) = y_2$$

which implies that $y_1 = y_2$. Thus, f is one-to-one.

So, f is onto and one-to-one, that is, f is a bijection. Hence, the assumption that f^{-1} is a function led us to the conclusion that $f : X \to Y$ is a bijection. The converse is trivial. ■

Example 4.73 Let $X = \{a, b, c, d\}$ and $Y = \{1, 2, 3, 4\}$, and let $f : X \to Y$ be a function defined by Figure 4.36.

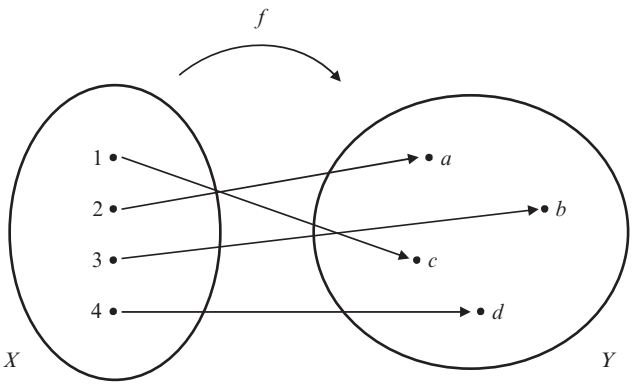


Figure 4.36

Then, as we have shown before, f can be expressed explicitly in the form of matrix

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ c & a & b & d \end{pmatrix}$$

by which we mean

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ c & a & b & d \end{pmatrix}$$

with the obvious interpretation: 1 goes (is mapped) to c, 2 goes to a, and so on. Evidently, f is a bijection and thus it has an inverse

$$f^{-1} = \begin{pmatrix} c & a & b & d \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

which, in turn, is also a bijection. ■

Definition 4.33 Let $f : X \to Y$ be a function, and let $A \subseteq X$ be a subset of X. We say that $f|_A$ is the **restriction of f to A** if

$$f|_A(a) = f(a), \quad \forall\, a \in A$$

Let's look at some examples.

Example 4.74 Let $X = \{1, 2, 3, 4, 5\}$. Consider functions $f : X \to Y$, and $f|_A : A \to Y$ defined by

$$f = \{(1, 3), (2, 6), (3, 11), (4, 18), (5, 27)\}$$

and

$$f|_A = \{(1, 3),\ (3, 11),\ (5, 27)\}$$

respectively. Evidently $f|_A$ is a restriction of f to $A = \{1, 3, 5\} \subseteq X$. ■

Example 4.75 Let $f : \mathbf{R} \to [0, \infty)$ be a function defined by $f(x) = x^2$. Then, as we know and can easily convince ourselves by inspection, f is not an injection. For instance,

$$f(2) = f(-2) = 4$$

And therefore, f does not have an inverse. How about if we restrict the domain of f to $\mathcal{D}(f) = \mathbf{R}\backslash(-\infty, 0)$? The restricted function

$$f|_+ : [0, \infty) \to [0, \infty)$$

is now injective. Indeed: for any $x, y \in [0, \infty)$, $f_+(x) = f_+(y)$, that is, $x^2 = y^2$ and therefore, $x^2 - y^2 = 0$. It follows that $(x - y)(x + y) = 0$. Since, by restricting our domain, we rule out the possibility of $x = -y$, we get that $x = y$ and thus our function is injective. It is evident that $f|_+$ is also onto, hence $f|_+$ is a bijection. ■

Example 4.76 You most likely already know that a function $\sin : \mathbf{R} \to \mathbf{R}$ is not one-to-one. However, suppose that we restrict our function to the domain $[-(\pi/2), \pi/2]$, and call

$$\sin|_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]} = \mathrm{Sin}$$

we get the function

$$\mathrm{Sin} : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to \mathbf{R}$$

which is one-to-one and onto $[-1, 1]$. ■

Example/Exercise 4.77 Show that a restriction of one-to-one function is one-to-one function.

Example/Exercise 4.78 Let $f : \mathbf{R} \to \mathbf{R}^+$ be a function defined by

$$f(x) = (x-2)^4$$

How should f be restricted in order to be a bijection?

Theorem 4.14 Let $f : X \to Y$ be an injection. Then, $\mathcal{D}(f^{-1}) = Y$ iff f is also a surjection.

Proof Suppose f is onto. Then,

$$\forall\, y \in Y, \exists x \in X \text{ s.t. } f(x) = y$$

Thus, $(x, y) \in f$ and $(y, x) \in f^{-1}$. But, since $y \in \mathcal{D}(f^{-1}), Y \subseteq \mathcal{D}(f^{-1})$. On the other hand, by definition, $\mathcal{D}(f^{-1}) \subseteq Y$. Therefore,

$$Y = \mathcal{D}(f^{-1})$$

Conversely, suppose $Y = \mathcal{D}(f^{-1})$. Take some $y \in Y$. Then, certainly, $y \in \mathcal{D}(f^{-1})$. That means that $\exists x \in X$, such that $(y, x) \in f^{-1}$, which in turn means that $(x, y) \in f$. Therefore $y = f(x)$, that is, f is a surjection. Hence the proof. ■

As a simple exercise you should now be able to prove

Theorem 4.15 Let $f : X \to Y$ be a bijection. Then $f^{-1} : Y \to X$ is a bijection too.

(Hint: See Theorem 4.13.)

Definition 4.34 We say that $h = g \circ f : X \to Z$ is a composition of two functions f and g, that is,

$$X \xrightarrow{f} Y \xrightarrow{g} Z$$

if

$$h(x) = (g \circ f)(x) = g(f(x))$$

Definition 4.34′ Given two functions $f : X \to Y$ and $g : Y \to Z$, we say that $g \circ f$ is a composition of f and g (Figure 4.37) if

$$g \circ f = \{(x, z) | \exists y \in Y, \text{ s.t. } f(x) = y, \text{ and } g(y) = z\}$$

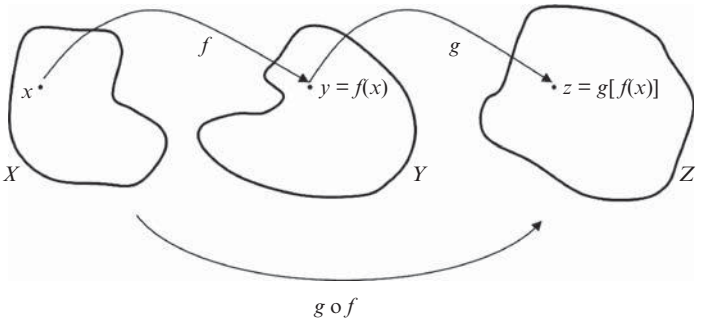


Figure 4.37 Composition of functions

Theorem 4.16 Suppose $f : X \to Y$ and $g : Y \to Z$ are two functions. Then, $g \circ f$ is a function from X to Z.

Proof First, we show that there exists a function, h, that maps X to Z.

Since f is a function, there is $x \in X$ and $y \in Y$, such that $f(x) = y$. Similarly, because g is a function, there exists $z \in Z$, such that $g(y) = z$. But we have already established the fact that $y = f(x)$. Therefore,

$$g(f(x)) = (g \circ f)(x) = h(x) = z$$

Now we show that h is *unique*.

Suppose $\exists z_1, z_2 \in Z$, such that for some $x \in X, h(x) = z_1$, and $h(x) = z_2$. However, by definition of $g \circ f, \exists y_1, y_2 \in Y$, such that $f(x) = y_1$ and $g(y_1) = z_1$. Similarly, $f(x) = y_2$ and $g(y_2) = z_2$. But, since f is a function, $f(x) = y_1 = y_2$. Also, since g is a function too, $g(y_1) = g(y_2)$. Therefore $z_1 = z_2$. Hence, h is unique. ■

Example 4.79 Let $f : \mathbf{Z} \to \mathbf{Z}$ and $g : \mathbf{Z} \to \mathbf{Z}$ be two functions defined as follows:

$$f(x) = x + 1, \quad \forall x \in \mathbf{Z}$$
$$g(x) = x^2$$

Find $f \circ g$ and $g \circ f$.

Solution

$$(f \circ g)(x) = f(g(x)) = f(x^2) = x^2 + 1$$

$$(g \circ f)(x) = g(f(x)) = g(x+1) = (x+1)^2$$

As you can see (in general) $f \circ g \neq g \circ f$. ■

Example 4.80 Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be functions defined by

$$f(x) = \sin x \quad \text{and} \quad g(x) = e^x$$

respectively. Find

(i) $f \circ g$
(ii) $g \circ f$

Solution

(i) $(f \circ g)(x) = f(g(x)) = f(e^x) = \sin(e^x)$
(ii) $(g \circ f)(x) = g(f(x)) = g(\sin x) = e^{\sin x}$ ■

Example 4.81 Let $f : X \to Y$ be a function. Observe that $f \circ f$ is defined when $\mathcal{D}(f) = \mathcal{R}(f)$, that is, when $X = Y$. ■

Example/Exercise 4.82 Let $f : \mathbf{R}\setminus\{0\} \to \mathbf{R}\setminus\{0\}$ be defined by

$$f(x) = \frac{1}{x}$$

Show that $(f \circ f)(x) = x$.

Example 4.83 Given a set $X = \{a, b, c, d\}$ with functions $f : X \to X$ and $g : X \to X$ defined by

$$f = \begin{pmatrix} a & b & c & d \\ a & c & a & c \end{pmatrix} \quad \text{and} \quad g = \begin{pmatrix} a & b & c & d \\ b & a & b & a \end{pmatrix}$$

Find $f \circ g$ and $g \circ f$.

Solution See Figure 4.38

$$f \circ g = \begin{pmatrix} a & b & c & d \\ a & c & a & c \end{pmatrix} \circ \begin{pmatrix} a & b & c & d \\ b & a & b & a \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ c & a & c & a \end{pmatrix}$$

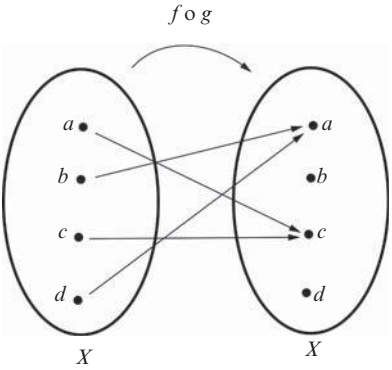


Figure 4.38 $(f \circ g) : X \to X$

$$g \circ f = \begin{pmatrix} a & b & c & d \\ b & a & b & a \end{pmatrix} \circ \begin{pmatrix} a & b & c & d \\ a & c & a & c \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ b & b & b & b \end{pmatrix}$$

See Figure 4.39

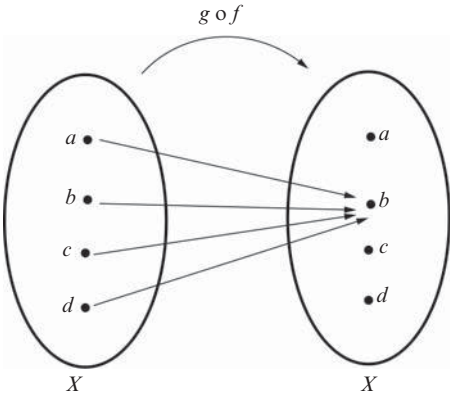


Figure 4.39 $(g \circ f) : X \to X$

■

Theorem 4.17 If $f : X \to Y$ is a bijection, and $f^{-1} : Y \to X$ is its inverse, then

(i) $f^{-1} \circ f = i_X$
and
(ii) $f \circ f^{-1} = i_Y$

Proof Suppose $f : X \to Y$ is a bijection, and let $f^{-1} : Y \to X$ be its inverse. Take some $x \in X$, and consider

$$(f^{-1} \circ f)(x) = f^{-1}[f(x)]$$

Now, since f^{-1} is the inverse function of f, $\forall\, x \in X$, and $\forall\, y \in Y$

$$f^{-1}(y) = x \quad \text{iff } f(x) = y.$$

It follows that for some $x' \in X$

$$x' = f^{-1}(y) = f^{-1}[f(x)]$$

But that means

$$f(x') = f(x)$$

And, since f is one-to-one, it follows that $x' = x$. We conclude that

$$(f^{-1} \circ f)(x) = x$$ ■

Example 4.84 Revisiting Example 4.83 one notes that

$$\begin{aligned} f^{-1} \circ f &= \begin{pmatrix} c & a & b & d \\ 1 & 2 & 3 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ c & a & b & d \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} = i_X \end{aligned}$$

and

$$\begin{aligned} f \circ f^{-1} &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ c & a & b & d \end{pmatrix} \circ \begin{pmatrix} c & a & b & d \\ 1 & 2 & 3 & 4 \end{pmatrix} \\ &= \begin{pmatrix} c & a & b & d \\ c & a & b & d \end{pmatrix} = i_Y \end{aligned}$$

which is in accord with Theorem 4.17. ■

Now we can restate our definition of the inverse function yet another way.

Definition 4.32′ A function $f : X \to Y$ is said to be invertible if an inverse relation f^{-1} is a function from Y to X, that is, $f^{-1} : Y \to X$, such that

$$f^{-1} \circ f = i_X \quad \text{and} \quad f \circ f^{-1} = i_Y$$

Example 4.85 Consider a function

$$f : \mathbf{R}\setminus\{2\} \to \mathbf{R}\setminus\{3\}$$

defined by

$$f(x) = \frac{3x}{x-2}$$

Find $f^{-1}(x)$.

Solution First, you should convince yourself that f is bijective. Then, we can argue as follows:

$$\begin{aligned}(f \circ f^{-1})(x) &= f[f^{-1}(x)] \\ &= \frac{3f^{-1}(x)}{f^{-1}(x)-2} \\ &= x\end{aligned}$$

It follows that

$$3f^{-1}(x) = xf^{-1}(x) - 2x$$

and thus,

$$f^{-1}(x)(\,x-3) = 2x$$

Therefore,

$$f^{-1}(x) = \frac{2x}{x-3}, \quad \forall\, x \in \mathbf{R}\setminus\{3\}$$

Note that, at the very beginning, in anticipation of the result we "restricted" our function's codomain to $\mathbf{R}\setminus\{3\}$. Of course, we could have said that our function maps $\mathbf{R} \setminus \{2\}$ to $\mathbf{R}$ and then, after obtaining the result, demand that our codomain be $\{x \in \mathbf{R} | x \neq 3\}$. ■

Example/Exercise 4.86 Let $f : \mathbf{R} \to (0, \infty)$ be a function defined by

$$f(x) = e^x$$

Show that the inverse exists and find the f^{-1}.

Example/Exercise 4.87 Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be two functions defined by $f(x) = 2x + 1$ and $g(x) = (x-1)/2$.

Show that both f and g are bijections and that $f^{-1} = g$.

Example 4.88 Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by

$$f(x) = \begin{cases} x^2 + 2 & \text{if } x \leq 0 \\ 2 - x^2 & \text{if } x > 0 \end{cases}$$

Find f^{-1}.

Solution Consider a function $g : \mathbf{R} \to \mathbf{R}$ given by

$$g(x) = \begin{cases} -\sqrt{x-2} & \text{if } x \geq 2 \\ \sqrt{2-x} & \text{if } x < 2 \end{cases}$$

and observe that if $x \leq 0$

$$(g \circ f)(x) = g(f(x)) = g(x^2 + 2)$$

Since $x^2 + 2 \geq 2$

$$\begin{aligned}(g \circ f)(x) &= -\sqrt{x^2 + 2 - 2} = -\sqrt{x^2} \\ &= -|x| = x\end{aligned}$$

On the other hand, if $x > 0$

$$(g \circ f)(x) = g(f(x)) = g(2 - x^2)$$

and, since $2 - x^2 < 2$,

$$(g \circ f)(x) = \sqrt{2 - 2 + x^2} = x$$

Similarly, you can check that $(f \circ g)(x) = x$, and therefore $g = f^{-1}$. ■

We expect the following theorem to hold.

Theorem 4.18 Given a function $f : X \to Y$, and i_X and i_Y identity functions on X and Y respectively, then

(i) $f \circ i_X = f$
(ii) $i_Y \circ f = f$

Proof

(i) $(f \circ i_X)(x) = f(i_X(x)) = f(x) \Rightarrow (f \circ i_X) = f$

(ii) $(i_Y \circ f)(x) = i_Y(f(x)) = f(x) \Rightarrow (i_Y \circ f) = f$ ■

The next exercise shouldn't be difficult to do.

Example/Exercise 4.89 Let $f : X \to X$ be any function. Show that $f \circ i_X = i_X \circ f = f$.

Theorem 4.19 Let $X, Y, W,$ and Z be sets, and let

$$f : X \to Y, \quad g : Y \to W, \quad h : W \to Z$$

be functions. Then,

$$h \circ (g \circ f) = (h \circ g) \circ f$$

Proof Take $x \in X$. Then,

$$(h \circ (g \circ f))(x) = h((g \circ f))(x) = h(g(f(x))) = (h \circ g)f(x)$$
$$= ((h \circ g) \circ f)(x)$$

It looks something like Figure 4.40.

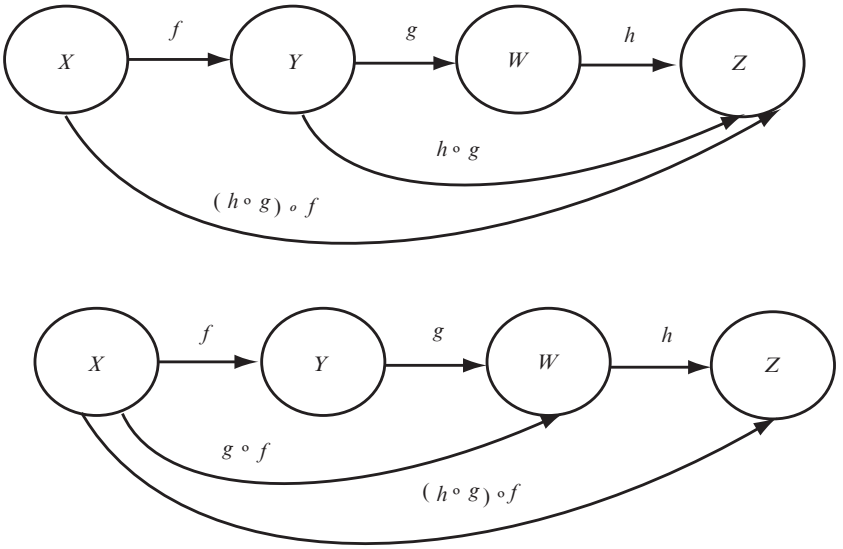


Figure 4.40 Associativity of composition ■

Theorem 4.20 If $f : X \to Y,$ and $g : Y \to Z$ are two onto functions, then $h = g \circ f$ is also an onto function.

Proof Since f and g are functions, by Theorem 4.16, $g \circ f = h$ is a function too. We need to show that $g \circ f$ is onto.

Take some $z \in Z$. We need to find $x \in X$, such that $(g \circ f)(x) = h(x) = z$. Now, since g is onto Z, there exists $y \in Y$ such that $g(y) = z$. But f is onto Y, thus there should be an $x \in X$ such that $f(x) = y$. Collecting all together

$$h(x) = (g \circ f)(x) = g[f(x)] = g(y) = z$$

Hence, $h = g \circ f$ is onto. ■

Exercise 4.90 Let $f : X \to Y$ and $g : Y \to Z$ be two functions, such that $g \circ f : X \to Z$ is onto. Must both f and g be onto?

Solution No! The diagram below convincingly shows that the fact that $g \circ f : X \to Z$ is onto doesn't necessitate that both f and g be onto (see Figure 4.41).

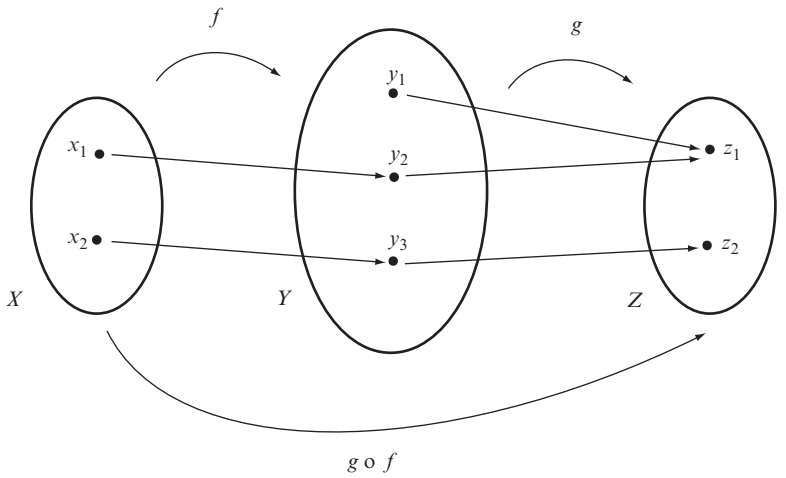


Figure 4.41 ■

Theorem 4.21 If $f : X \to Y$ and $g : Y \to Z$ are two one-to-one functions, then $h = g \circ f$ is a one-to-one function too.

Proof Let $x_1, x_2 \in X$ and suppose $(g \circ f)(x_1) = (g \circ f)(x_2)$. In other words, $g[f(x_1)] = g[f(x_2)]$. Since g is one-to-one, $f(x_1) = f(x_2)$. But f is also one-to-one, hence $x_1 = x_2$, which implies that $h = g \circ f$ is one-to-one too. ■

Example 4.91 Suppose $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ are functions, such that $g \circ f : \mathbf{R} \to \mathbf{R}$ is one-to-one. Must both f and g be one-to-one?

Solution No! Consider, for instance, $g(x) = x^2$ and $f(x) = e^x$. Obviously

$$(g \circ f)(x) = g(f(x)) = g(e^x) = e^{2x}$$

is one-to-one while g is clearly not.

■

Theorem 4.22 If $f : X \to Y$ and $g : Y \to Z$ are two bijections, then $h = g \circ f$ is a bijection too.

Proof Use proofs of Theorems 4.20 and 4.21 together. ■

Theorem 4.23 Let $f : X \to Y$ and $g : Y \to Z$ be two bijections. Then

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$$

Proof We have just seen that given two bijections, f and g, their composition $g \circ f$ is also a bijection. Thus, there exists its inverse $(g \circ f)^{-1}$.

But, remember,

$$g \circ f = \{(x, z) | \exists y \in Y, \text{ s.t. } (x, y) \in f \text{ and } (y, z) \in g\}$$

so that

$$\begin{aligned}(g \circ f)^{-1} &= \{(z, x) | \exists y \in Y, \text{ s.t. } (y, x) \in f^{-1} \text{ and } (z, y) \in g^{-1}\} \\ &= f^{-1} \circ g^{-1}\end{aligned}$$

■

Example 4.92 Let $f : X \to Y$ and $g : Y \to Z$ be two functions, and let $A \subseteq Z$. Prove that

$$(g \circ f)^{-1}[A] = f^{-1}(g^{-1}[A])$$

Proof To establish the equality, we need to prove two things:

(i) $(g \circ f)^{-1}[A] \subseteq f^{-1}(g^{-1}[A])$
(ii) $f^{-1}([A]) \subseteq (g \circ f)^{-1}[A]$

For (i): Let $x \in (g \circ f)^{-1}[A]$. That implies that

$$(g \circ f)(x) = g(f(x)) \in A$$

but, since $g[f(x)] \in A$

$$f(x) \in g^{-1}[A]$$

Therefore,

$$x \in f^{-1}(g^{-1}[A])$$

Hence,

$$(g \circ f)^{-1}[A] \subseteq f^{-1}(g^{-1}[A])$$

For (ii): Suppose that $x \in f^{-1}(g^{-1}[A])$. That implies that

$$f(x) \in g^{-1}[A]$$

which, furthermore, means that

$$(g \circ f)(x) = g(f(x)) \in A$$

It follows that

$$x \in (g \circ f)^{-1}[A]$$

Consequently,

$$f^{-1}(g^{-1}[A]) \subseteq (g \circ f)^{-1}[A]$$

We have proved (i) and (ii), thus

$$(g \circ f)^{-1}[A] = f^{-1}(g^{-1}[A])$$

See Figure 4.42.

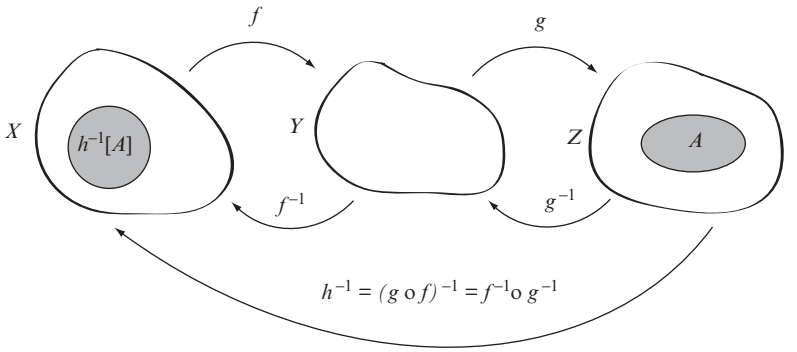


Figure 4.42 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

■

Theorem 4.24 Let $f : X \to Y$ be any function and let $A \subseteq X$ be any subset of X. Then,

$$A \subseteq f^{-1}(f[A])$$

Proof Let $a \in A \subseteq X$ be any element of A, then, $f(a) \in f[A]$. On the other hand,

$$f^{-1}(f[A]) = \{a \in A | f(a) \in f[A]\}$$

So, we have that $a \in A$ and $f(a) \in f[A]$, which implies that $a \in f^{-1}(f[A])$. Thus,

$$A \subseteq f^{-1}(f[A])$$

See Figure 4.43. ■

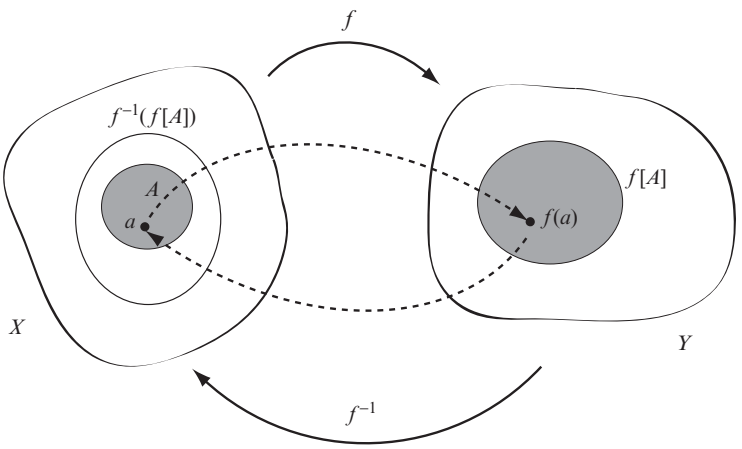


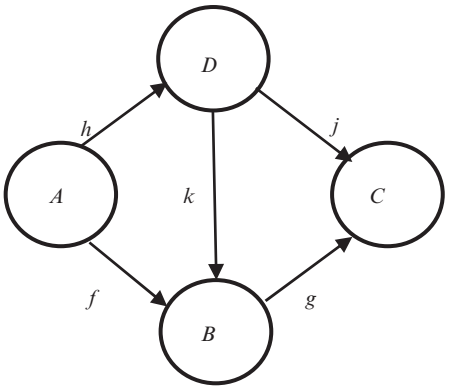
Figure 4.43 $A \subseteq f^{-1}(f[A])$

Example 4.93 Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $(x) = x^2$, and let $A = \{2\}$. Then, obviously, $f[A] = f(2) = 4$. Thus, $2 \in f^{-1}(4)$, that is, $f^{-1}(4)$ is a set of all $x \in \mathbf{R}$ such that $x^2 = 4$. Hence, $f^{-1}(4) = \{-2, 2\} \neq \{2\} = A$. So, indeed,

$$A \subseteq f^{-1}(f[A])$$ ■

Definition 4.35 A diagram (of functions) is said to be commutative if for any pair of sets X and Y (in the diagram) any two paths from X to Y are equal.

Example 4.94 Suppose the diagram below is commutative:

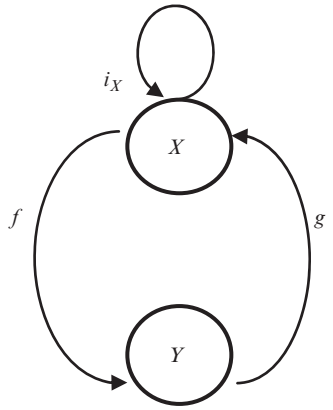


Then,

$$k \circ h = f, \quad g \circ k = j, \quad g \circ f = j \circ h = g \circ k \circ h$$

■

Example 4.95 Consider the following commutative diagram:



Observe that, since the diagram is commutative, $g \circ f = i_X$. That implies that $g \circ f$ is one-to-one and, therefore, f must be one-to-one. Furthermore, $g \circ f$ is onto, so g is also onto. On the other hand, since we do not know whether $f \circ g = i_Y$ or not, we cannot conclude that $g = f^{-1}$. ■

Theorem 4.25 Let X and Y be finite sets, such that $|X| = n$ and $|Y| = k$, then there are k^n functions $X \to Y$.

Proof Easy! (cf. Figure 4.35) Obviously, there are k choices (k possible images) for each of the n elements of X. Each arrow represents one of those k

choices. Since any image of $x_i \in X$, $(i = 1, 2, \ldots, n)$ is independent of the any image of $x_j \in X$, $(j = 1, 2, \ldots, n)$, all together we have $k \cdot k \cdot k \cdot \cdots \cdot k = k^n$ ways to choose images for all the elements of X, thus k^n possible functions from X to Y (see Figure 4.44). ■

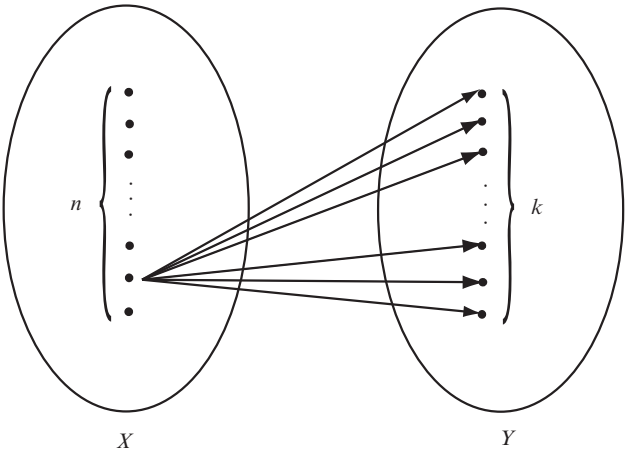


Figure 4.44

Consequently, the following theorem holds.

Theorem 4.26 Let X and Y be finite sets such that $|X| = |Y| = n$, then there are $n!$ bijections $X \to Y$.

Theorem 4.27 Let $f : X \to Y$ be a function and let $A \subseteq X$, and $B \subseteq Y$. Then,

(i) If f is one-to-one, then $A = f^{-1}(f[A])$.
(ii) $f(f^{-1}[B]) \subseteq B$.
(iii) $f(f^{-1}[B]) = B$ iff $B \subseteq R(f)$.

Proof (Hint: See Chapter 1.) ■

Theorem 4.28 Let $f : X \to Y$ be a function and let A and B be subsets of X. Then,

(i) If $A \subseteq B$ then $f[A] \subseteq f[B]$.
(ii) $f[A \cup B] = f[A] \cup f[B]$.
(iii) $f[A \cap B] \subseteq f[A] \cap f[B]$ and, if f is one-to-one, then

$$f[A \cap B] = f[A] \cap f[B]$$

Proof

(i) Take some $y \in f[A]$. Then $\exists a \in A \subseteq X$ such that $f(a) = y$. Since $A \subseteq B$, a is definitely an element of B, and thus $y = f(a) \in f[B]$. It follows that $f(A) \subseteq f(B)$ as claimed.

(ii) Suppose $y \in f[A \cup B]$. Then, $\exists x \in A \cup B$ such that $f(x) = y$. But, x being an element of $A \cup B$, it means that either

1. $x \in A$ or
2. $x \in B$

Consider (1):
$x \in A$ implies that $y = f(x) \in f[A]$ and thus, $y \in f[A] \cup f[B]$. So,

$$f[A \cup B] \subseteq f[A] \cup f[B]$$

Similarly, we approach case (2) and again get

$$f[A \cup B] \subseteq f[A] \cup f[B]$$

Next, we want to show that $f[A] \cup f[B] \subseteq f[A \cup B]$. Well, it is always true that both $A \subseteq A \cup B$ and $B \subseteq A \cup B$, which, considering (i), implies $f[A] \subseteq f[A \cup B]$ and $f[B] \subseteq f[A \cup B]$. Consequently, since any element of $f[A] \cup f[B]$ is either an element of $f[A]$ or $f[B]$, we conclude that

$$f[A] \cup f[B] \subseteq f[A \cup B]$$

So, we have

$$f[A \cup B] \subseteq f[A] \cup [B] \quad \text{and} \quad f[A] \cup f[B] \subseteq f[A \cup B]$$

and thus,

$$f[A \cup B] = f[A] \cup f[B]$$

(iii) Let $y_0 \in f[A \cap B]$. Then, $\exists x_0 \in A \cap B$ such that $y_0 = f(x_0)$. But $x_0 \in A \cap B$ implies $x_0 \in A$ and $x_0 \in B$. Therefore, $f(x_0) \in f[A]$, and $f(x_0) \in f[B]$, implying that $y_0 = f(x_0) \in f[A] \cap f[B]$. We conclude that

$$f[A \cap B] = f[A] \cap f[B]$$

Observe that the above argument would not hold in the case that our function was not one-to-one. The best we could say would be

$$f[A \cap B] \subseteq f[A] \cap f[B]$$

■

Example 4.96 Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $f(x) = x^2$, and let $A = [-3, 2]$ and $B = [1, 5]$. Let's show that $f[A \cap B] \subseteq f[A] \cap f[B]$. Certainly our function is not an injection. However, we have

$$f[A \cap B] = f[[-3, 2] \cap [1, 5]] = f[[1, 2]]$$
$$= [1, 4]$$

On the other hand,

$$f[[-3, 2]] \cap f[[1, 5]] = [9, 4] \cap [1, 25]$$
$$= [1, 9]$$

So, $f[A \cap B] \subseteq f[A] \cap f[B]$ as expected. ■

Example 4.97 Let $X = \{-3, -2, -1, 0, 1, 2, 3\}$ and $Y = \{0, 1, 2, 4, 6, 9\}$. Furthermore, let $f : X \to Y$ be a function defined by $f(x) = x^2$. If $A = \{-1, 3\} \subseteq X$ and $B = \{4, 6\} \subseteq Y$, let's examine $f[A]$ and $f^{-1}[B]$.

Figure 4.45 will make everything more transparent.

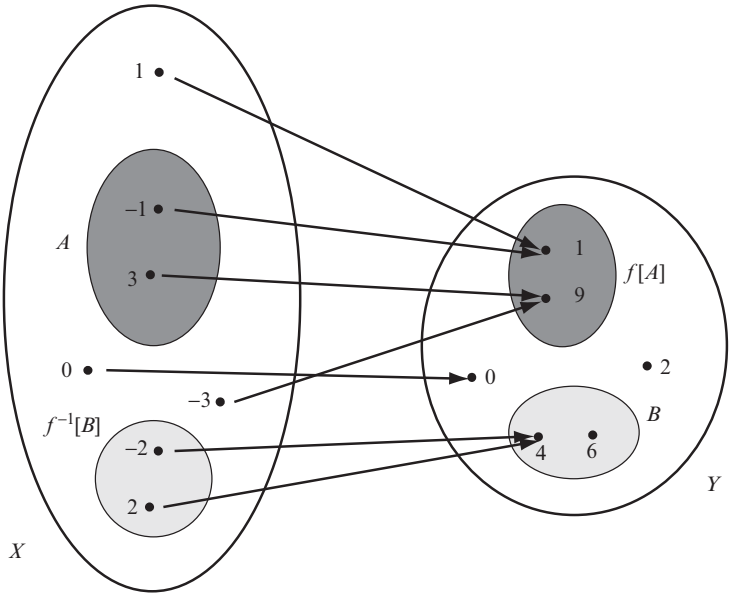


Figure 4.45

Let's first consider our function as

$$f = \{(0, 0), (-1, 1), (1, 1), (-2, 4), (2, 4), (-3, 9), (3, 9)\}$$

We see that

$$f(X) = \{0, 1, 4, 9\}$$

and

$$f[A] = f[\{-1, 3\}] = \{1, 9\}, \text{and also } f(-3) = \{9\}$$

Now, since f is not a bijection, f^{-1} is not a function from Y to X, so it would be incorrect to write, say, $f^{-1}(1)$ or $f^{-1}(9)$. We can, however, write

$$f^{-1}[\{1\}] = \{-1, 1\}$$

and

$$f^{-1}[\{9\}] = \{-3, 3\}$$

Similarly,

$$f^{-1}[B] = \{-2, 2\}$$

■

Example 4.98 Let $X = [0, 10]$ and $Y = [0, 20]$. Furthermore, let $f : X \to Y$ be a function defined by $f(x) = -x^2 + 8x$. Let $A = [0,\ 3]$ and $B = [2, 7]$, find

(i) $f[A] \cup f[B]$
(ii) $f[A \cup B]$
(iii) $f[A] \cap f[B]$
(iv) $f[A \cap B]$

Solution

(i) Since $f(x) = -x^2 + 8x$, it follows that $f[A] = [0, 15]$ and $f[B] = [7, 16]$. Therefore, $f[A] \cup f[B] = [0, 16]$. Similarly,
(ii) $A \cup B = [0, 3] \cup [2, 7] = [0, 7]$. Thus $f[A \cup B] = [0, 16]$.
(iii) and (iv) are left for the reader to do. ■

Example 4.99 Suppose we want to find $f^{-1}[\{y \in \mathbf{R} | 4 \leq y \leq 25\}]$ if the function $f : \mathbf{R} \to \mathbf{R}$ is defined by $f(x) = x^2$. By inspecting the domain of f^{-1}, we immediately note

$$f^{-1}[\{y \in \mathbf{R} | 4 \leq y \leq 25\}] = [2, 5] \cup [-5, -2]$$ ■

In Chapter 1 we discussed the methods of comparing sets, in particular, the sets of infinite cardinality. There, as you would recall, we defined two sets X and Y to be equivalent, that is, $|X| = |Y|$, iff there existed a bijection $f : X \to Y$. Also, recall that we called a set X the infinite set iff there existed at least one proper subset $A \subseteq X$ such that $|A| = |X|$. Before the next example, it may be instructive

to recall another well-known and useful function from $\mathbf{R}$ to $\mathbf{R}^+$, the **exponential function**, defined by

$$\exp_b(x) = b^x, \quad \forall\, x \in \mathbf{R}$$

and its inverse, the **logarithmic function**, defined by

$$\log_b x = y \Leftrightarrow b^y = x, \quad \forall\, x \in \mathbf{R}^+$$

Now, we can do

Example 4.100 Let's prove something that we already "know," that is, let's prove that the set $\mathbf{R}$ is equivalent to the set of all positive real numbers $\mathbf{R}^+$. The "only" thing we need to do is find a "good" function, a bijection, $f : \mathbf{R} \to \mathbf{R}^+$ and we are done. But that shouldn't be difficult since we remember from our high school math the two functions that will do exactly what we need: the exponential function,

$$f(x) = b^x, \quad b \geq 0$$

and the logarithmic function $f(x) = \log_b x$ Remember the definition of logarithmic function:

$$y = f(x) = \log_b x \Rightarrow\ b^y = b^{\log_b x} = x$$

Now we ask: Is $f(x) = b^x$ one-to-one? Well, suppose $f(x_1) = f(x_2)$. Then

$$b^{x_1} = b^{x_2}$$

Thus,

$$\log_b b^{x_1} = \log_b b^{x_2}$$

implies

$$x_1 \log_b b = x_2 \log_b b$$

that is, $x_1 = x_2$.

So, our function is one-to-one. Let's see if it is also onto.

Take some $y \in \mathbf{R}^+$. Then,

$$x = \log_b y$$

Therefore,

$$b^x = b^{\log_b y} = y = f(x)$$

Hence, $|\mathbf{R}| = |\mathbf{R}^+|$. ■

You are most likely familiar with the next two examples from your elementary math courses. Nevertheless, it is worthwhile to remind ourselves of the well-known properties of logarithms.

Example 4.101 Show that for all $A, B > 0$ and b not equal to 1

(i) $\log_b AB = \log_b A + \log_b B$
(ii) $\log_b \frac{A}{B} = \log_b A - \log_b B$
(iii) $\log_b A^p = p\log_b A$

Solution Let $\log_b A = x$ and $\log_b B = y$. Then $A = b^x$ and $B = b^y$. It follows

(i) $AB = b^x b^y = b^{x+y}$
thus
$$\log_b AB = x + y = \log_b A + \log_b B$$

(ii) $\frac{A}{B} = \frac{b^x}{b^y} = b^{x-y}$
thus
$$\log_b \frac{A}{B} = x - y = \log_b A - \log_b B$$

(iii) $A^p = (b^x)^p = b^{xp}$
thus
$$\log_b A^p = px = p\log_b A$$ ■

Example 4.102 Show that for any three positive real numbers $a, b, x \in \mathbf{R}^+$

$$\log_b x = \frac{\log_a x}{\log_a b}$$

Solution Let

$$y_1 = \log_a x, \quad \text{and} \quad y_2 = \log_b x. \tag{4.5}$$

Then, from the definition of logarithm, we have

$$a^{y_1} = x$$

and

$$b^{y_2} = x$$

Therefore,

$$a^{y_1} = b^{y_2}$$

and

$$\log_a a^{y_1} = \log_a b^{y_2}$$

So,

$$y_1 \log_a a = y_2 \log_a b \tag{4.6}$$

From (4.5) and (4.6), we have

$$\log_a x \cdot \log_a a = \log_b x \cdot \log_a b$$

that is

$$\log_a x = \log_b x \cdot \log_a b$$

or

$$\log_b x = \frac{\log_a x}{\log_a b}$$

■

Example 4.103 With logarithmic function at hand, let's show once again that the set **R** is of infinite cardinality.

As before, we would like to show that there is a bijection between the set **R** and its proper subset $\mathbf{R}^+$. So, let $f : \mathbf{R} \to \mathbf{R}^+$ be a function defined by $f(x) = 2^x$. Suppose

$$f(x_1) = f(x_2), \quad \text{that is,} \quad 2^{x_1} = 2^{x_2}$$

Then,

$$\log_2 2^{x_1} = \log_2 2^{x_2}$$

implying that $x_1 = x_2$. So, f is one-to-one. To see that it is also onto, take $y \in \mathbf{R}^+$ such that $x = \log_2 y$. It follows that

$$f(x) = 2^x = 2^{\log_2 y} = y$$

Thus, f is onto. Since there is a bijection between **R** and its proper subset $\mathbf{R}^+$, we conclude that the set **R** is infinite. ■

Theorem 4.29 Let X and Y be two nonempty finite sets such that $|X| = |Y| = n$, and let $f : X \to Y$ be a function from X to Y. Then f is an injection iff f is a surjection.

Proof Let $|X| = |Y| = n$, and let f be one-to-one. In that case there are n distinct images of n distinct elements of X. In other words, $\mathcal{R}(f) = Y$ and therefore f is onto. Conversely, let f be onto. Then each of the n elements of Y is an image of some element of X. In other words, the n elements of X have n distinct images in Y, thus f is one-to-one. ■

The astute reader might have noticed that we have emphasized in the above theorem that the sets should be finite. Otherwise the theorem wouldn't hold. Indeed, consider, for instance, the function $f : \mathbf{Z} \to \mathbf{Z}$ defined by $f(x) = 2x$. Obviously, f is one-to-one yet the range is the set of all even integers, thus f is not onto.

Theorem 4.30 Let $X \neq \varnothing$ and Y be any two sets. If there is an injection $f : X \to Y$, then there is also a surjection $g : Y \to X$.

Proof Define g as follows:

$$g(y) = \begin{cases} x & \text{if } f(x) = y \\ a & \text{if no such } x \text{ exists} \end{cases}$$

Since f is an injection, if x exists it has to be unique, thus our function is well defined. So, for any $x \in X, x = g(y) = g(f(x))$. Therefore, g is surjective. ■

The next three important theorems are not easy and may be skipped on the first reading.

Theorem 4.31 Let $f : X \to Y$ and $g : Z \to W$ be one-to-one function, and let $X \cap Z = \varnothing$. Furthermore, define a function $h : X \cup Z \to Y \cup W$ by

$$h(x) = \begin{cases} f(x) & \text{if } x \in X \\ g(x) & \text{if } x \in Z \end{cases}$$

If $Y \cap W = \varnothing$ then h is also one-to-one. If, in addition, f and g are bijective then h is also bijective.

Proof Suppose $x_1, x_2 \in X \cup Y$, such that $h(x_1) = h(x_2) = y$. It follows that $y \in Y \cup W$, that is, $y \in Y$ or $y \in W$. Let's say $y \in Y$. Since we have assumed that $Y \cap W = \varnothing$ it follows immediately that $y \notin W$. Hence $x_1, x_2 \in X$ and therefore

$$\begin{aligned} h(x_1) &= f(x_1) \\ \text{and} \qquad &= h(x_2) \\ &= f(x_2) \end{aligned}$$

But since f is one-to-one, and $f(x_1) = f(x_2)$, it follows that $x_1 = x_2$, thus h is also one-to-one.

Let's say now that $y \in W$. Reasoning as before, we conclude that $y \notin Y$. Hence $x_1, x_2 \in Z$, and therefore

$$\begin{aligned} h(x_1) &= g(x_1) \\ &= h(x_2) \\ &= g(x_2) \end{aligned}$$

By the same argument as before, we conclude that in this case h is one-to-one too. The readers should prove the bijective case on their own. ■

Theorem 4.32 Let X and Y be two nonempty sets, such that $Y \subseteq X$. If there exists an injection $X \to Y$, then there exists a bijection $X \to Y$.

Proof First, consider a simple case $X = Y$. Observe that the identity function

$$i_X : X \to Y = X$$

is obviously a bijection.

Take now a proper subset $Y \subset X$. It follows that $X \setminus Y \neq \emptyset$. Furthermore, let

$$f : X \to Y$$

be an injection. We need to prove that f is also onto. Suppose it is not. In that case $\mathcal{R}(f) \subset Y$, thus $Y \setminus \mathcal{R}(f) \neq \emptyset$. Consider the subset of Y defined the following way:

$$Y' = \{f^n(x) | x \in X \setminus Y, n \in \mathbf{N}\}$$

where by $f^n(x)$ we mean $(f \circ f \circ \cdots \circ f)(x)$.

Note that

$$Y' \subseteq R(f)$$

which implies that $\forall\, x \in X \setminus Y,\ f(x) \in Y'$. Furthermore, $\forall\, x \in X \setminus Y\ (f \circ f)(x) \in Y'$, $(f \circ f \circ f)(x) \in Y'$, and so on. Now we construct a new set

$$Z = (X \setminus Y) \cup Y'$$

and a new function

$$f_1 : Z \to Y'$$

We want to show that f_1 is onto.

Take some $x \in X \setminus Y$. There has to be $y \in Y'$ such that for some $n \in \mathbf{N}$

$$y = f^n(x)$$

This also implies that $y = f(x)$ for some $x \in X \setminus Y$, or $y = f(x)$ for some $x \in Y'$. Therefore, $y = f_1(x)$ if we take $x \in Z$. Hence, f_1 is onto. Now, remember, f is one-to-one thus f_1 is also one-to-one; hence, $f_1 : Z \to Y'$ is a bijection. We are not done yet. We need one more set, $W = Y \setminus Y'$. Recall, $Y \setminus R(f) \neq \emptyset$, and also $Y \setminus R(f) \subseteq Y \setminus Y'$, thus $W \neq \emptyset$. Observe that $W \cap Y' = \emptyset$, and $W \cap Z = \emptyset$. Also,

$$Z \cup W = X, \quad \text{and} \quad Y' \cup W = Y.$$

With this in mind, we need the previous theorem to construct the function

$$h : Z \cup W \to Y' \cup W$$

that is

$$h : X \to Y$$

as follows

$$h(x) = \begin{cases} f_1(x) & \text{if } x \in Z \\ i_W(x) & \text{if } x \in W \end{cases}$$

Since f_1 is bijective, and $i_W : W \to W$ is certainly bijective, h is bijective too. ■

This may be the right time to recall (see Chapter 1, Theorem 1.25) and finally prove

Theorem 4.33 (Schröder–Bernstein[7]) If A and B are any two sets such that $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

A couple of comments may be in order before we start with the proof. First, since we are talking about any set, thus also the sets with infinite cardinality, assiduous reader might consider the use of the symbol "$\lesssim$" more appropriate than the usual "$\leq$." However, we will continue with our accustomed sign "$\leq$." Second, and most importantly, the same reader may by now be aware of the fact that the theorem can be stated in the equivalent form:

Let A and B be any two sets. If there exist one-to-one functions $f : A \to B$ and $g : B \to A$, then there is also a bijection $\varphi : A \to B$. Thus, $|A| = |B|$.

[7]Ernst Schröder (1841–1902) German Mathematician. Felix Bernstein (1878–1956) German mathematician.

Proof The proof is not easy, so don't despair. It might take you some time to get it completely.

Obviously, in order to prove our claim, it would be sufficient to find a bijection $A \xrightarrow{\varphi} B$. Note, however, that the theorem doesn't assume that either f or g is onto.

So, consider a function $h = f^{-1} \circ g^{-1}$. Observe that h is defined on some subset of A and with the range in A, that is

$$\mathcal{D}(h) = \{x \in A | x \in \mathcal{R}(g) \text{ and } g^{-1}(x) \in \mathcal{R}(f)\}$$

Let's now define a new function

$$\varphi_1 : \mathcal{R}(g) \to B$$

by

$$\varphi_1(x) = g^{-1}(x)$$

and also

$$\varphi_n(x) = g^{-1} \circ h_n(x)$$

where

$$h_n = \underbrace{h \circ h \circ \cdots \circ h}_{n}$$

Now, observe that $\mathcal{D}(\varphi_n) \subseteq A$ and $\mathcal{R}(\varphi_n) \subseteq B$. Next, define $A_0 = \{x \in A | \exists\, n, \text{ s.t. } \varphi_n(x) \text{ is well defined and } \varphi_n(x) \notin \mathcal{R}(f)\} \subseteq A$.

It follows that for any $x \in A$, either $x \in A_0$ or $x \notin A_0$. So, we construct our function $\varphi : A \to B$ as follows:

$$\varphi(x) = \begin{cases} f(x) & x \notin A_0 \\ g^{-1}(x) & x \in A_0 \end{cases}$$

We claim that φ is a bijection. If we could prove that, then we are done. Well, let's see.

One-to-one:

Let $\varphi(x_1) = \varphi(x_2)$. We need to check whether $x_1 = x_2$. Well, if both x_1 and x_2 **are** the elements of A_0, or both x_1 and x_2 **are not** the elements of A_0 then $x_1 = x_2$. Why? Because, remember, both f and g are one-to-one. Now, how about if $x_1 \notin A_0$ and $x_2 \in A_0$? In that case

$$f(x_1) = \varphi(x_1) = \varphi(x_2) = g^{-1}(x_2)$$

But this cannot be. Since $x_2 \in A_0$, there is an n such that $\varphi_n(x_2)$ is well defined, but $\varphi_n(x_2) \notin \mathcal{R}(f)$. Let's check the case when $n = 1$. Then

$$\begin{aligned}\varphi_1(x_2) &= g^{-1}(x_2)\\ &= f(x_1) \in \mathcal{R}(f)\end{aligned}$$

For $n > 1$ we have

$$\begin{aligned}\varphi_n(x_2) &= g^{-1}\circ h_n(x_2)\\ &= g^{-1}\circ h_n\circ(g\circ f)(x_1)\\ &= g^{-1}\circ h_{n-1}(x_1)\\ &= \varphi_{n-1}(x_1)\end{aligned}$$

We know that if $\varphi_n(x_2) \notin \mathcal{R}(f)$, then $\varphi_{n-1}(x_1) \notin \mathcal{R}(f)$. That would furthermore imply that $x_1 \in A_0$, which is a contradiction. Therefore, φ is one-to-one.

Onto:

In order for φ to be onto, we need, as usual, to prove that $\forall\, b \in B$, $\exists a \in A$, such that $\varphi(a) = b$. Since g is one-to-one, let's take $a = g(b)$. Recall that either $a \in A_0$ or $a \notin A_0$. So, suppose first that $a \in A_0$. Then

$$\begin{aligned}\varphi(a) &= g^{-1}(a)\\ &= b\end{aligned}$$

and we are done.

The case when $a \notin A_0$ is a little trickier. For all, n, either $\varphi_n(a) \notin \mathcal{R}(f)$ or $\varphi_n(a)$ is not defined. If $n = 1$

$$\varphi_1(a) = g^{-1}(a)$$

is defined, hence $\varphi_1(a) \in \mathcal{R}(f)$. Let

$$\begin{aligned}a_1 &= f^{-1}(\varphi_1(a))\\ &= (f^{-1}\circ g^{-1})(a)\end{aligned}$$

We would like to prove that $a_1 \notin A_0$. Well, let's see. Suppose the opposite, that is, suppose $a_1 \in A_0$. Then, there is an n such that $\varphi_n(a_1)$ is defined, and $\varphi_n(a_1) \notin \mathcal{R}(f)$. But, remember,

$$\begin{aligned}\varphi_n(a_1) &= \varphi_n(f^{-1}\circ g^{-1})(a)\\ &= \varphi_{n+1}(a)\end{aligned}$$

This is a contradiction since we have established the fact that for all n, either $\varphi_{n+1}(a)$ is not defined, or $\varphi_{n+1}(a) \notin \mathcal{R}(f)$. Thus, $a \notin A_0$. Then, by definition,

$$\varphi(a_1) = f(a_1) = f(f^{-1} \circ g^{-1})(a) = g^{-1}(a) = b$$

Hence φ is onto and our proof is complete. ■

Example 4.104 Let's see whether we can prove that $|(0, 1)| = |[0, 1]|$.

We would like to show that $|(0, 1)| \leq |[0, 1]|$ and $|(0, 1)| \geq |[0, 1]|$, from which we could conclude that $|(0, 1)| = |[0, 1]|$.

First, observe that $(0, 1) \subseteq [0, 1]$. So, we can safely say that $|(0, 1)| \leq |[0, 1]|$. Now, let's pick an interval with a property that $[0, 1]$ is its subset, say, $(-1, 2)$. Since $[0, 1] \subseteq (-1, 2)$, it is true that $|[0, 1]| \leq |(-1, 2)|$. But we already know that $|(1, 0)| = |(-1, 2)|$, which implies $|[0, 1]| \leq |(1, 0)|$. So, we have

$$|(0, 1)| \leq |[0, 1]| \quad \text{and} \quad |(0, 1)| \geq |[0, 1]|$$

By Schröder–Bernstein theorem (S–B-theorem), it follows that $|(0, 1)| = |[0, 1]|$. ■

Exercise 4.105 Show that $|\mathcal{P}(\mathbf{N})| = |\mathbf{R}|$.

Solution We already know that $|(0, 1)| = |\mathbf{R}|$. So, let's consider a function $f : (0, 1) \to \mathcal{P}(\mathbf{N})$. We would like to prove that f is one-to-one.

Recall that any $a \in (0, 1)$ can be uniquely expressed as $a = 0.a_1a_2a_3 \cdots$ where each $a_i \in \{0, 1, 2, \ldots, 9\}$, and there is no $n_0 \in \mathbf{N}$ such that $a_n = 9$ for all $n \geq n_0$. That is, from some point on, in decimal expansion of a, we don't want all 9's.

Now we need to choose a suitable function in order to take advantage of the S–B-theorem. So, let's define our function f as follows:

$$f(a) = \left\{10^{n-1} \mid n \in \mathbf{N}\right\}$$

In other words, every $a \in (0, 1)$ is mapped to some subset $A \subseteq \mathcal{P}(\mathbf{N})$. For example, if $a = 0.123$ then $f(a) = A = \{1, 20, 300, 4000\}$. Similarly, $f\left(\frac{1}{3}\right) = f(0.333\ldots) = \{3, 30, 300, \ldots\}$. Is our function one-to-one? Suppose for some $a, b \in \mathbf{N}$, $f(a) = f(b)$, where, again, $a = 0.a_1a_2a_3\ldots$ and $b = 0.b_1b_2b_3\ldots$ (with the same constraints on their decimal expansion as mentioned before). Then,

$$A = \left\{10^{n-1}a_n \mid n \in \mathbf{N}\right\} = \left\{10^{n-1}b_n \mid n \in \mathbf{N}\right\} = B$$

Observe that, if the ith digit in the decimal expansion of a, a_i is different from zero, then $10^{i-1}a_i \in A$ is a unique number in the interval $[10^{i-1}, 9 \cdot 10^{i-1}]$. Similar arguments hold also for b_i, that is, if $b_i \neq 0, 10^{i-1}b_i \in [10^{i-1}, 9 \cdot 10^{i-1}] \subseteq B$.

But, remember, $A = B$, that is

$$10^{i-1}a_i = 10^{i-1}b_i$$

and therefore $a_i = b_i$. How about if $a_i = 0$? Well, in that case $a_i = 0 \in A$ and there is no number in the interval $[10^{i-1}, 9 \cdot 10^{i-1}]$ belonging to A. But, again, since $A = B, 0 \in B$ and there is no number in the interval $[10^{i-1}, 9 \cdot 10^{i-1}]$ belonging to B. Thus, $b_i = 0$. So, in all the cases $a_i = b_i$, $\forall\, i \in \boldsymbol{N}$, hence, $a = b$ and our function is one-to-one. Hence, $|(0, 1)| \leq |\mathcal{P}(\boldsymbol{N})|$.

Now, we need another suitable function, namely, $g : (\mathbf{N}) \to (0, 1)$ as follows: For every $S \subseteq \mathbf{N}$ we define $g(S) = 0.s_1s_2s_3 \ldots$, such that

$$s_n = \begin{cases} 1 & \text{if } n \in S \\ 2 & \text{if } n \notin S \end{cases}$$

So, $g(S) \in \mathbf{R}$ is a number whose decimal expansion consists only of 1s and 2s. In order to be able to apply the S-B-theorem, we need our function g to be one-to-one too. Let's see. Suppose $S, T \subseteq \mathbf{N}$ and suppose that $g(S) = g(T)$. Then

$$\begin{aligned} g(S) &= 0.s_1s_2s_3 \ldots = s \\ &= g(T) = 0.t_1t_2t_3 \ldots = t \end{aligned}$$

where

$$t_n = \begin{cases} 1 & \text{if } n \in T \\ 2 & \text{if } n \notin T \end{cases}$$

We need to show that $S = T$. Let $k \in S$. Then, $s_k = 1$. But $s = t$, so $t_k = 1$ too, which implies that $k \in T$ and therefore $S \subseteq T$. Similarly, we show that $T \subseteq S$, thus $S = T$ and we conclude that our function g is one-to-one. So, we have that $|\mathcal{P}(\mathbf{N})| \leq |(0,\ 1)|$. By the S-B-theorem $|\mathcal{P}(\mathbf{N})| = |(0, 1)|$. Finally, we recall that $|(0, 1)| = |\mathbf{R}|$ and the proof that $|\mathcal{P}(\mathbf{N})| = |\mathbf{R}|$ is complete.

4.4 SUPPLEMENTARY PROBLEMS

1. With relations R on $\mathbf{R}$ given below, determine and explain which one is and which one is not an equivalence relation.

(i) xRy iff $|x - y| \leq 1$

(ii) xRy iff $xy > 0$

(iii) xRy iff $x + y = 0$

2. Given a set $X = \{x, y, w, z\}$ with the following relations:

$$R_0 = \varnothing$$
$$R_1 = \{(x, w), (y, x)\}$$
$$R_2 = \{(x, x), (x, y), (y, w), (x, w), (z, z)\}$$
$$R_3 = \{(x, x), (x, y), \ (y, x), (y, y), (w, w), (z, z)\}$$

Determine which of the relations are:
(i) reflexive
(ii) symmetric
(iii) antisymmetric
(iv) transitive.

3. Let R be a relation defined on $\mathbf{Z}$ by: xRy if $11x - 5y = 2k, x, y, k \in \mathbf{Z}$. Show that R is an equivalence relation.

4. Let $f : X \to Y$ be a function. Define a relation "$\sim_f$" on X by: $x \sim_f y$ iff $f(x) = f(y)$ Show that "$\sim_f$" is an equivalence relation.

5. Let R be a relation defined on $\mathbf{Z}$ by: xRy if $x = y (\text{mod } 3)$. Determine whether R is an equivalence relation.

6. We say that a relation R on a set X is cyclic iff for all $x, y, z \in X$, if aRb and bRc then zRx. Show that
(i) If a relation R is symmetric and transitive, then it is also cyclic.
(ii) If a relation R is symmetric and cyclic, then it is also transitive.

7. Let "$\preccurlyeq$" be a relation on the set $\mathbf{R} \times \mathbf{R}$ defined by

$$(x, y) \preccurlyeq (v, w) \text{ iff } x < v \text{ or } (x = v \text{ and } y \leq w)$$

Show that "$\preccurlyeq$" is a partial order on $\mathbf{R} \times \mathbf{R}$.

8. Consider the set of linear functions

$$F = \{f : \mathbf{R} \to \mathbf{R} | f(x) = ax + b, \ a, b \in \mathbf{R}\}$$

and define a relation "$\preccurlyeq$" on $\mathbf{R}$ as follows:

$$f \preccurlyeq g \text{ iff } f(1) - f(0) \leq g(1) - g(0)$$

Is this an equivalence relation?

9. Let "$\leq$" be a relation on a set $\mathcal{F}$ of all functions from $\mathbf{R}$ to $\mathbf{R}$ defined by

$$f \leq g \text{ iff } f(x) \leq g(x), \quad \forall x \in \mathbf{R}$$

Show that "$\leq$" is a partial order on $\mathcal{F}$.

10. Let $R = \{(1,5),(2,2),(3,4),(5,2)\}$ and $S = \{(1,4),(3,5),(4,1)\}$ be two relations. Find

(i) $R \circ S$

(ii) $S \circ R$

and sketch the corresponding digraphs.

11. Let $X = \{1,2,3\}$ be a set with a relation $R = \{(1,1),(2,2),(3,3),(1,2),(2,1)\}$. Find X/R.

12. Let R be a relation on $\mathbf{Z}$ defined by xRy iff $x^2 = y^2, x, y \in \mathbf{Z}$. What is the $[0], [1]$ and $[4]$?

13. Determine for each of the following digraphs if they represent relations that are reflexive, symmetric, transitive, and antisymmetric.

14. Determine whether any of the following digraphs represent the equivalence relation.

Let R and S be relations on a set X. Determine whether the following statements are true:

(i) If R and S are reflexive, then $R \cap S$ is reflexive.

(ii) If R and S are reflexive, then $R \cup S$ is reflexive.

(iii) If R and S are symmetric, then $R \cap S$ is symmetric.

(iv) If R and S are symmetric, then $R \cup S$ is symmetric.

(v) If R and S are transitive, then $R \cap S$ is transitive.

(vi) If R and S are transitive, then $R \cup S$ is transitive.

15. Let R and S be equivalence relations on a set X. Show that $R \cap S$ is also an equivalence relation on X.

16. Let R is an equivalence relation on a set X. Determine whether $R \cup R^{-1}$ is an equivalence relation on X.

17. Let $X = \{a,b,c,d\} Y = \{1,2,3,4\}$, and let $f : X \to Y$ be given by

$$f = \begin{pmatrix} a & b & c & d \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

Find its inverse.

18. Determine whether the following relations are functions or not:

(i) $R = \{(1,1),(2,2),\ (3,3),\ (2,3),(4,3),(5,3)\}$

(ii) $S = \{(a,b),(b,c),(c,d),(d,e)\}$

19. Determine whether the following relation is a function:

$$R = \{(x,y) \in \mathbf{Z} \times \mathbf{Z} | y^2 = x\}$$

20. Let $X = \{a,b,c,d\}$, and let f and g be functions given by

$$f = \begin{pmatrix} a & b & c & d \\ a & c & a & c \end{pmatrix} \quad \text{and} \quad g = \begin{pmatrix} a & b & c & d \\ b & a & b & a \end{pmatrix}$$

Find $f \circ g$ and $g \circ f$.

21. Let $p_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$ and $p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$ be two permutations of five elements. Find

(i) $p_1 \circ p_2$

(ii) p_1^{-1} and p_2^{-1}

22. Let $X = \{1,2,3,\}$ and $Y = \{a,b,c\}$. Define functions f and g by

$$f = \{(1,a),(2,b),(3,a)\} \quad \text{and} \quad g = \{(a,\ 1),(b,3),(c,2)\}$$

Find

(i) $f \circ g$

(ii) $g \circ f$

(iii) f^{-1}

(iv) g^{-1}

23. Prove that $f : \mathbf{Z} \to \mathbf{Z}$ defined by $f(x) = 3 - x$ is a bijection.

24. Prove that $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = 3x - 7$ is a bijection; Find f^{-1} and find $f \circ f$.

25. Prove that $f : \mathbf{Z} \to \mathbf{Z}$ defined by $f(x) = 3x - 7$ is one-to-one but not onto.

26. Let $N = \{n \in \mathbf{N} | n \geq 100\}$ and define a function $f : N \to \mathbf{N}$ by $f(n) =$ the sum of digits of n. Determine whether f is one-to-one and/or onto.

27. Prove that $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^4 - x^2$ is neither one-to-one nor onto.

28. Determine whether $f : \mathbf{N} \to \mathbf{N}$ defined by $f(x) = x^2$ is one-to-one and onto.

29. Prove that $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^5 + 3$ is one-to-one and onto.

30. Consider a function $f : \mathbf{R} \times \mathbf{R} \to \mathbf{R} \times \mathbf{R}$ defined by $f((x, y)) = (x + y, xy)$. Determine whether this function is one-to-one and/or onto.

31. For which of the following functions f, the relation f^{-1} is also a function?

(i) $f(x) = \frac{x+1}{x+2}$

(ii) $f(x) = e^{x+2}$

(iii) $f(x) = \log x$

(iv) $f(x) = \sin x$

32. Let $f : \mathbf{R}\backslash\{1\} \to \mathbf{R}\backslash\{1\}$ be a function defined by

$$f(x) = \frac{x}{x-1}$$

(i) Prove that f is bijective.

(ii) Find f^{-1}.

(iii) Find $f \circ f$.

(iv) Find $f^{-1} \circ f^{-1}$.

33. Let $f : [0, \infty) \to [0, \infty)$ be a function defined by $f(x) = x^2$. Determine whether f is one-to-one or not.

34. Consider a function $f : X \times Y \to X$ defined by $f(x, y) = x$. Determine whether f is injective and/or surjective.

35. Let $f : (0, \infty) \to (0, \infty)$ be a function defined by

$$f(x) = \frac{1}{x}$$

Prove that f is a bijection and find its inverse.

36. Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $f(x) = \begin{cases} 2x & \text{if } x \in \mathbf{Q} \\ 3x & \text{if } x \in \mathbf{R} \setminus \mathbf{Q} \end{cases}$.
Prove that f is a bijection and find its inverse.

37. Let $f : \mathbf{R}\backslash\{0\} \to \mathbf{R}\backslash\{0\}$ be a function defined by

$$f(x) = 1 - \frac{1}{x}$$

Show that $f \circ f \circ f = i$, and find f^{-1}.

38. Which of the following functions is injective/surjective:

(i) $f(x) = x^3 - x + 2$

(ii) $g(x) = \frac{x^2+1}{x^2-1}$

39. Let $f : X \to Y$ be a function defined by $f(x) = x^3 - x$. Find sets X and Y for f to be a bijection.

40. Consider a set $X = \mathbf{R} \times \mathbf{R}$. Show that

$$d((x, y), (z, w)) = |x - y| + |y - w|$$

is a metric on $\mathbf{R} \times \mathbf{R}$.

41. Is there a function $f : \mathbf{R} \to \mathbf{R}$ that is onto but not one-to-one?

42. Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $f(x) = x^2$. Let $A = [-3, 2]$ and $B = [1, 5]$. Show that

(i) $f[A \cup B] = f[A] \cup f[B]$

(ii) $f[A \cap B] \subseteq f[A] \cap f[B]$

43. Let $f : \mathbf{R} \to \mathbf{R}$ be a function defined by $f(x) = 10x - x^2$. Find

$$f([1, 6]) \cap f([4, 7]).$$

44. Let $f : X \to Y$ be a one-to-one function and let $A \subseteq X$. Furthermore, define a function $F : P(X) \to P(Y)$ by $F(A) = f[A]$. Show that F is one-to-one.

45. Use Schröder–Bernstein theorem to prove that

(i) $|[0, 1)| = |(0, 1)|$

(ii) $|\mathbf{R}| = |\mathbf{R} \times \mathbf{R}|$

5

GROUP THEORY

> I seem to have been only like a boy playing on the sea shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me.
>
> I. Newton[1]

5.1 INTRODUCTION

There is, arguably, no other mathematical discipline that originated from such simple concepts, and with such profound implications to virtually all branches of mathematics, physics, and many other sciences, as Group Theory. The axiomatics of groups are so simple, so natural, and so "how-could-it-have-been-otherwise," that one wonders how something so self-evident and so ubiquitous wasn't discovered much earlier in its own right, rather than being motivated by problems in number theory, geometry, the theory of algebraic equations, and the like. And yet, how could something so "simple" could have evolved into such an elaborate, encompassing mathematical theory underpinning the fundamental laws of elementary particle physics, the Standard Model, String Theory, and so on.

[1]Brewster, D., *Life of Sir Isaac Newton,* Nabu Press, 2010.

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

It is also safe to say that groups occur abundantly in nature. When looking at natural objects that usually attract attention, say, a beautiful flower, a crystal, a geometric figure or object, or even better, a human face, a piece of art, a sculpture, a painting, or still more abstractly, just a few musical tunes from your favorite composition – what makes you say that those "things" are beautiful, harmonious, symmetric; why are they "pleasing to the senses?" Is there something that, regardless of their superficial differences, is common to all of them? It is here that mathematics, abstract algebra in particular, enters the quest. With a little consideration, one recognizes the structure of a group. Take an object, *any object X*, endowed with some structure and consider a map $\varsigma : X \to X$ that maps the object onto itself while preserving that structure – we say that there is a symmetry. Group theory formalizes the essential aspects of symmetry – we say that symmetries form a group. By focusing on the very transformations themselves, we gain an insight into the fundamental makeup of the object we are studying.

5.2 FUNDAMENTAL CONCEPTS OF GROUP THEORY

Definition 5.1 Given a set G, by a **binary operation** "$*$" on G, we mean a function $*\colon G \times G \to G$ that maps any ordered pair of elements of G to an element of G. In particular, if an ordered pair $(a, b) \in G$ is mapped into an element $c \in G$, then we write $c = a * b$.

The operation "$*$" could be our usual addition "$+$," or multiplication "$\cdot$," performed on a set of real numbers or functions, but it could also be something more abstract, say, permutations, rotations, or translations, and so on, performed on sets that are definitely not sets of numbers.

Definition 5.2 A binary operation "$*$" is said to be **well defined** on a set X if

(i) Exactly one element is assigned to each possible ordered pair of elements of X.

(ii) For each ordered pair of elements of X, the element assigned to it is again an element of X We say that X is closed under "$*$," or that **closure** is satisfied.

Example 5.1 Suppose we take the set $\mathbf{Z}$ and consider the usual addition "$+$," or multiplication "$\cdot$," as the binary operation "$*$," such that an ordered pair of two elements of $\mathbf{Z}$, say, $(2, 3)$ is mapped into $2 + 3 = 5$ by "$+$," or into $2 \cdot 3 = 6$ by "$\cdot$." Note that the "results" of our "operations" are again elements of $\mathbf{Z}$. We say that our operations are well defined since they satisfy both conditions (i) and (ii) in Definition 5.2. ■

Example 5.2 If we consider set **N** and take for our operation the usual subtraction "−," then closure is not satisfied since, for instance, $2 - 5 = -3 \notin \mathbf{N}$, and, therefore, our operation is not a well-defined binary operation. ■

Example 5.3 Consider the operation

$$a * b = \frac{a+b}{ab}$$

with $a, b \in \mathbf{Z}$. Obviously, closure is not satisfied on set **Z**, since one can easily find two integers a and b such that $a * b \notin \mathbf{Z}$. ■

Example 5.4 Consider a set $\mathcal{F}$ of all real functions $f : \mathbf{R} \to \mathbf{R}$. If the operation "$*$" is the familiar addition "+," multiplication "$\cdot$", or composition "$\circ$" of functions, then, in each of those cases, "$*$" is a well-defined binary operation. ■

Example 5.5 Consider set **Q** and the operation

$$a * b = \frac{a}{b}, \quad \forall a, b \in \mathbf{Q}$$

This is obviously not a well-defined operation on **Q** for no rational number can be assigned to the pair $(a, 0)$. ■

Example/Exercise 5.6 Determine whether the following two operations are well defined:

(i) $a * b = |a - b| a, b \in \mathbf{Z}$
(ii) $a * b = a \ln b, \ \ a, b \in \mathbf{R}$

Example 5.7 As you may recall, if X is a set with n elements, that is, $|X| = n$, there are $n!$ different rearrangements/permutations of n elements, counting no rearrangement as the 0th permutation. We say that a **permutation** of X is a bijection $p : X \to X$. Generally, we denote the set of all permutations of X by P_X. As an example, let P_3 be the set of all permutation of three elements, that is, $P_3 = \{p_0, p_1, \ldots, p_5\}$, such that

$$p_i = \begin{pmatrix} 1 & 2 & 3 \\ a_1 & a_2 & a_3 \end{pmatrix}$$

In the top row of the array (matrix), we write elements 1, 2, 3. In the second row, we write their image under the permutation p. Thus, a_i represents one of the numbers 1, 2, 3. We say $p_i(j) = a_j, \ i = 0, 1, 2, 3, 4, 5; \ a_j = 1, 2, 3$. For example:

$$p_0(1) = 1, \quad p_0(2) = 2, \quad p_0(3) = 3;$$
$$p_1(1) = 2, \quad p_1(2) = 3, \quad p_1(3) = 1; \ \text{and so on}$$

Hence,

$$p_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \text{and so on}$$

Consider now the "multiplication" of permutations defined by the following example:

$$p_1(p_2(1)) = p_1(3) = 1$$
$$p_1(p_2(2)) = p_1(2) = 3$$
$$p_1(p_2(3)) = p_1(1) = 2$$

You recognize this multiplication as the composition of functions, $p_i's$, that is

$$p_i \cdot p_j = p_i \circ p_j = p_k$$

Hence, we conclude that for the example above

$$p_1 p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = p_3$$

We can define any other "product" $p_i p_j$ in the same way. Before we continue, let's convince ourselves that $p_i p_j \neq p_j p_i$. Since we already know the result of $p_1 p_2$, it suffices to check $p_2 p_1$:

$$p_2(p_1(1)) = p_2(2) = 2$$
$$p_2(p_1(2)) = p_2(3) = 1$$
$$p_2(p_1(3)) = p_2(1) = 3$$

We have obtained another permutation, say $p_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$, which is obviously not equal to p_3. We could have recognized this after the first step, since evidently $p_1 p_2(1) \neq p_2 p_1(1)$. With a little effort, you can convince yourself that any "multiplication" always produces one of $p_0, \ldots, p_5$. So, P_3 is closed under the above-defined multiplication.

So, we claim that "multiplication" of permutations is another well-defined binary operation. ■

If you are familiar with matrices, the next example illustrates another case of a well-defined binary operation.

Example 5.8 Let $\mathcal{M}_{nn}$ be a set of all (n, n)-matrices over $\mathbf{R}$. $\mathcal{M}_{nn}$ is closed under the addition of matrices. ■

Example/Exercise 5.9 Let $\mathcal{M}_2(\mathbf{R})$ be the set of all matrices of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, a, b \in \mathbf{R}$. Determine whether $\mathcal{M}_2(\mathbf{R})$ is closed under (i) matrix addition and (ii) matrix multiplication.

We now introduce some additional features to structures that are closed with respect to a given binary operation.

Definition 5.3 Let $*\colon G \times G \to G$ be a binary operation on G. We say that "$*$" is **commutative** if

$$a * b = b * a, \quad \forall a, b \in G$$

Example 5.10 Let "$*$" be an ordinary multiplication on the set $\mathbf{R}$. Obviously, "$*$" is a well-defined commutative operation. ■

Example 5.11 Let "$*$" be a binary operation on $\mathbf{Q}$ defined by

$$a * b = \frac{ab}{2}$$

This is another well-defined commutative operation. ■

Example 5.12 Let $\mathcal{M}_{nn}$ be the set of all (n, n)-matrices over $\mathbf{R}$. If the binary operation "$*$" is regular addition of matrices, we have a well-defined commutative operation since

$$A + B = B + A, \quad \forall A, B \in \mathcal{M}_{nn}$$

On the other hand, if we take "$*$" to be regular matrix multiplication, we have a well-defined noncommutative operation since, in general,

$$AB \neq BA$$

■

Definition 5.4 Let $*\colon G \times G \to G$ be a binary operation on G. We say that "$*$" is **associative** if

$$a * (b * c) = (a * b) * c, \quad \forall a, b, c \in G$$

Example 5.13 Let "$*$" be a binary operation on $\mathbf{R}$ defined by

$$\forall a, b, c \in \mathbf{R}, \quad a * b = a + b + ab$$

Verify that

(i) $(a * b) * c = a * (b * c), \ \forall a, b, c \in \mathbf{R}$
(ii) $a * b = b * a, \ \forall a, b \in \mathbf{R}$

Solution
(i)

$$\begin{aligned}(a * b) * c &= (a + b + ab) * c \\ &= a + b + ab + c + (a + b + ab)c \\ &= a + b + c + ab + ac + bc + abc \\ &= a + (b + c + bc) + a(b + c + bc) \\ &= a * (b * c)\end{aligned}$$

(ii)

$$\begin{aligned}a * b &= a + b + ab \\ &= b + a + ba \\ &= b * a\end{aligned}$$

■

Example/Exercise 5.14 Consider the following operations and determine whether they are commutative and/or associative:

(i) $a * b = 2^{ab}, \ \text{with} \ a, b \in \mathbf{N}$
(ii) $a * b = a^b, \ \text{with} \ a, b \in \mathbf{N}$

Example 5.15 Let X be a set, and let $F = \{f | f : X \to X\}$ be the set of functions mapping X to X. In Chapter 4, we proved that

$$(f \circ (g \circ h))(x) = ((f \circ g) \circ h)(x)$$

Thus, the composition of functions is another example of an associative operation. ■

Theorem 5.1 If "$*$" is an associative operation on a set G, then

$$(a * b) * (c * d) = [a * (b * c)] * d, \quad \forall a, b, c, d \in G$$

Proof Let $g = a * b$, then

$$\begin{aligned}(a * b) * (c * d) = g * (c * d) = (g * c) * d = [(a * b) * c] * d \\ = [a * (b * c)] * d\end{aligned}$$

■

Definition 5.5 We call a nonempty set G equipped with a binary operation "$*$" and denoted by $(G; *)$ an **algebraic structure** or a **groupoid**.

A nonempty set G, equipped with a binary operation "$*$" that is associative is called a **semigroup** $(G; *)$.

Example 5.16 Suppose we are given a set $G = \{1, 2, 3\}$ with an operation "$*$" described by the following table:

$*$	1	2	3
1	3	2	1
2	3	1	3
3	2	3	3

Our algebraic structure $(G; *)$ is such that, for example, $2 * 1 = 3, 3 * 1 = 2$. What about associativity? Well, let's just check:

$$2 * (3 * 1) = 2 * 2 = 1$$

$$(2 * 3) * 1 = 3 * 1 = 2$$

So, associativity is not satisfied. ■

Example 5.17 Consider a set $G = \{a, b, c, d\}$ with an operation "$*$" described by the following table:

$*$	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

Closure is obvious. What about associativity? Well, let's see:

$$b * (c * d) = b * b = a$$

$$(b * c) * d = d * d = a$$

$$d * (b * c) = d * d = a$$

$$(d * b) * c = c * c = a$$

You can check the rest, and convince yourself that associativity holds. Hence, $(G; *)$ is a semigroup. ■

Example 5.18 Let $(G; *)$ be a semigroup. Show that $\forall a, b, c, d \in G$

$$(ab)cd = a(bc)d = ab(cd)$$

Solution First, observe that the definition of a semigroup requires associativity; thus,

$$\begin{aligned}(ab)(cd) &= [(ab)c]d \\ &= [a(bc)]d \\ &= a(bc)d\end{aligned}$$

Similarly, one can establish the second equality. ■

Definition 5.6 Let G be a semigroup, and let $a \in G$. We define $a^1 = a$ and $a^{n+1} = a * a^n$ for $n \geq 1$. Consequently, we define

$$a^n * a^m = a^{n+m} = a^{m+n} = a^m * a^n$$

and

$$(a^n)^m = (a^m)^n = a^{mn}$$

Example/Exercise 5.19 Show by induction that the equalities in Definition 5.6 indeed hold.

Definition 5.7 (A group)
Let $(G; *)$ be a set with a binary operation "*" such that the following conditions are satisfied:

G1 : $a * b \in G, \quad \forall a, b \in G$	**closure**
G2 : $a * (b * c) = (a * b) * c, \quad \forall a, b, c \in G$	**associativity**
G3 : $\exists e \in G,$ s.t. $a * e = e * a = a, \quad \forall a \in G$	**identity**
G4 : $\forall a \in G, \ \exists\, a^{-1} \in G,$ s.t. $a * a^{-1} = a^{-1} * a = e$	**inverse**

We say that $(G; *)$ is a **group**, and call G1–G4 **group axioms**.

Note: (1) Many authors call the **identity element *e*** the **neutral element** with respect to "**multiplication**" "$*$," and sometimes denote it by **1**, especially if "$*$" represents ordinary multiplication. Similarly, if the group operation is ordinary addition, the neutral element is denoted by **0** for obvious reasons. In order to avoid confusion, I think it is best to stick with ***e*** for the neutral element regardless of the group operation.

(2) Also, a^{-1} is just a symbol for the inverse element of a, and in general does not mean $1/a$.

Example 5.20 It is easy to see that $(\mathbf{Z}; +), (\mathbf{Q}^+ = \{q \in \mathbf{Q} | q > 0\}; \cdot)$ and $(\mathbf{R}^+ = \{x \in \mathbf{R} | x > 0\}; \cdot)$ satisfy G1–G4, and thus are groups.

Similarly, $(\mathbf{R}; +)$ and $(\mathbf{R}^*; \cdot) = (\mathbf{R} \setminus \{0\}; \cdot)$ are groups, while $(\mathbf{N}; +)$ and $(\mathbf{Z}; \cdot)$ are not. Why? ■

Example 5.21 The set $(\mathbf{N}; +)$ is not a group. There is no neutral element in $(\mathbf{N}; +)$. ■

Example 5.22 A set of integers modulo 6, $(\mathbf{Z}_6; +)$, forms a group whose **Cayley table** is as follows:

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

You should verify this. ■

Example 5.23 A set $\mathcal{F}(\mathbf{R}) = \{f | f : \mathbf{R} \to \mathbf{R}\}$, a set of all real-valued functions of a real variable under ordinary addition of functions is a group. Indeed,

$$\begin{aligned}(f+g)(x) &= f(x) + g(x) \in G, \quad \forall f, g \in \mathrm{F}(\mathbf{R}) \\ [f + (g+h)](x) &= f(x) + [g(x) + h(x)] \\ &= [f(x) + g(x)] + h(x) \\ &= [(f+g) + h](x), \quad \forall f, g, h \in \mathrm{F}(\mathbf{R})\end{aligned}$$

If we choose for the neutral element of $\mathcal{F}(\mathbf{R})$ the function o, defined by $o(x) = 0, \forall x \in \mathbf{R}$, then obviously,

$$(o+f)(x) = (f+o)(x) = f(x), \quad \forall f \in F(\mathbf{R})$$

and therefore for the inverse we have

$$(-f)(x) = -f(x), \forall f \in \mathcal{F}(\mathbf{R}), \quad \text{and} \quad \forall x \in \mathbf{R}.$$ ■

Definition 5.8 (Abelian[2] group)

If $(G; *)$ is a group such that $a * b = b * a$, for all $a, b \in G$, we say that G is a **commutative** or an **abelian group.**

Definition 5.9 We say that a group $(G; *)$ is a **finite group** if G is a finite set.

Even before discussing more examples, we are already in a position at this point to prove some elementary properties of groups.

Theorem 5.2 If G is a group, then the identity element e is unique.

Proof Suppose there are two identity elements $e, e' \in G$. Then,

$$e * e' = e' \text{ since } e \text{ is the identity}$$

but also

$$e * e' = e \text{ since } e'' \text{ is the identity}$$

Therefore, $e = e'$. ■

Theorem 5.3 Let G be a group, and $a \in G$, such that $a * a = a$. Then, $a = e$.

Proof Since G is a group, there should be an $a' \in G$ such that $a' * a = e$. Then, multiplying the equation $a * a = a$ on the left by a' gives

$$\begin{aligned} a' * (a * a) &= a' * a = e \\ &= (a' * a) * a = e * a = a \end{aligned}$$

and so, $a = e$. ■

Theorem 5.4 If G is a group, then every element of G has a unique inverse.

Proof Let G be a group with identity element e. Suppose that $a \in G$ has two inverses a_1^{-1} and a_2^{-1}. Consider

$$\begin{aligned} a_1^{-1} &= e * a_1^{-1} \\ &= (a_2^{-1} * a) * a_1^{-1} \\ &= a_2^{-1} * (a * a_1^{-1}) \\ &= a_2^{-1} * e \\ &= a_2^{-1} \end{aligned}$$

■

[2]Named after the Norwegian mathematician Niels Henrik Abel (1802–1829).

Corollary 5.1 If G is a group and $a \in G$ then

$$(a^{-1})^{-1} = a \qquad (*)$$

Proof Since G is a group and $a \in G$ there has to be an $a^{-1} \in G$, such that

$$a * a^{-1} = a^{-1} * a = e$$

"Multiplying" both sides of (*) by a^{-1} we obtain

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

Since the inverse is unique, it follows that

$$(a^{-1})^{-1} = a$$ ■

Now that we are familiar with the idea of a binary operation "$*$" on groups, we can drop the "$*$" and from now on, instead of writing $a * b$ we will simply write ab, and call our operation *multiplication*. We will keep in mind, however, that this multiplication could be quite different from ordinary multiplication.

Theorem 5.5 Let G be a group, and let $a \in G$ be any element of G, then

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

Proof In Definition 5.5 and Example 5.19, it was shown that if G is a semigroup, then for any $a \in G$

$$(a^m)^n = a^{mn}, \qquad \forall m, n \in \mathbf{Z}$$

So, putting $m = -1$ our statement also holds if G is a group. ■

Theorem 5.6 Let G be a group, and let $a, b \in G$. Then $(ab)^{-1} = b^{-1}a^{-1}$.

Proof Since G is a group and $a, b \in G$, then $ab \in G$ too. That furthermore implies that there should exist elements $a^{-1}, b^{-1}, (ab)^{-1} \in G$, such that

$$aa^{-1} = a^{-1}a = e$$
$$bb^{-1} = b^{-1}b = e$$

but also

$$(ab)(ab)^{-1} = (ab)^{-1}(ab) = e \qquad (*)$$

Now consider the product

$$(ab)(b^{-1}a^{-1})$$

which, because of associativity, we can write

$$\begin{aligned}(ab)(b^{-1}a^{-1}) &= a(b^{-1}b)a^{-1}\\ &= aea^{-1}\\ &= aa^{-1}\\ &= e \qquad (**)\end{aligned}$$

From (*) and (**), it follows that $(ab)^{-1} = b^{-1}a^{-1}$, which was to be shown. ■

Example/Exercise 5.24 Show that if G is a group, and $a_1, a_2, \ldots, a_n \in G$, then

$$(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}$$

Theorem 5.7 (Left and right cancellation law) For any group G and any $a, b \in G$, there are unique $x, y \in G$, such that $xa = b$ and $ay = b$.

Proof First take

$$xa = b \qquad (5.1)$$

Multiplying both sides by a^{-1}, we get

$$\begin{aligned}xaa^{-1} &= ba^{-1}\\ xe &= ba^{-1}\\ x &= ba^{-1} \qquad (5.2)\end{aligned}$$

Let's check our "solution":

$$\begin{aligned}xa &= (ba^{-1})a\\ &= b(a^{-1}a)\\ &= be\\ &= b\end{aligned}$$

So, (5.2) is indeed the solution of equation (5.1). Is it unique? Well, suppose there are two solutions s_1 and s_2, such that

$$s_1 a = b$$

and

$$s_2 a = b$$

In that case, we can write

$$s_1 a = s_2 a \tag{5.3}$$

Multiplying both sides of (5.3) by a^{-1} yields

$$s_1(aa^{-1}) = s_2(aa^{-1})$$
$$s_1 e = s_2 e$$
$$s_1 = s_2$$

The second part can be proved in a similar way. ■

Theorem 5.8 A semigroup G is a group iff for all $a, b \in G$ each of the equations $xa = b$ and $ay = b$ has a solution.

All of these elementary properties can be easily verified on the examples that follow.

Example 5.25 Consider a set $G = \{e, a, b, c\}$ with multiplication given by the following Cayley table:

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Show that G is a group.

Solution By simply inspecting the table, we see that closure is satisfied. To verify associativity, let's consider

$$a(bc) = aa$$
$$= e$$
$$= (ab)c$$

$$= cc$$

$$= e$$

Similarly, you can verify that all other "products" are associative.

Note that there exists an identity e, and finally, that every element is its own inverse. Hence, G is a group, which is known by the name of the Klein four-group (*Viergruppe*), often denoted by the symbol V. ■

Example/Exercise 5.26 Let $G = \{e, a, b, c, d\}$ be endowed with a binary operation "$*$" obeying the following "multiplication" table:

$*$	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	c	d	e	a
c	c	d	e	a	b
d	d	e	a	b	c

Show that $(G; *)$ is a group.

Example 5.27 Let $P = \{E, O\}$ be a set of two elements "(*E*)*ven*" and "(*O*)*dd*", endowed with an operation "+," such that

$$O + O = E, \quad E + E = E, \quad \text{and} \quad O + E = E + O = O$$

Then, $(P; +)$ is an abelian group called the **parity group**. This is easily verified. ■

Example 5.28 Consider $(\mathbf{Q}^+; *) = (\mathbf{Q} \setminus \{0\}; *)$, where $*$ is defined by

$$a * b = \frac{ab}{2}$$

Show that $(\mathbf{Q}^+; *)$ is a group.

Solution Closure is obvious.

Associativity:

$$\begin{aligned} a * (b * c) = a * \frac{bc}{2} &= \frac{abc}{4} \\ &= \frac{ab}{2} * c \\ &= (a * b) * c \end{aligned}$$

Identity:

$$2 * a = a * 2 = a$$

So, for all $a \in (\mathbf{Q}^+; *)$, $e = 2$.

Inverse:

Consider

$$a * \frac{4}{a} = \frac{a\frac{4}{a}}{2} = \frac{\frac{4}{a}a}{2} = 2$$

For any $a \in (\mathbf{Q}^+; *)$, $a^{-1} = \frac{4}{a}$.

Thus, $(\mathbf{Q}^+; *)$ is a group. ■

If you are unfamiliar with matrices, you can postpone the next three exercises until you have studied Chapter 6.

Example/Exercise 5.29 Show that the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

with usual matrix multiplication form the Klein four-group.

Example/Exercise 5.30 Show that the set of matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix},$$
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}$$

with usual matrix multiplication is a group.

Example 5.31 Let GL(n) be a set of all **regular** (n, n)-matrices. Endowed with usual matrix multiplication, GL(n) becomes a group called the **general linear group**. ■

Example 5.32 Let $\mathrm{SL}(2, \mathbf{R}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \,\middle|\, a, b, c, d \in \mathbf{R}, ad - bc = 1 \right\}$ be the set of all $(2, 2)$ – real matrices with determinant 1. With matrix multiplication as the binary operation, $\mathrm{SL}(2, \mathbf{R})$ is a group called **special linear group**. (You may want to verify this.) ■

Example 5.33 Consider the set $\mathbf{C} = \{ re^{i\varphi} | r, \varphi \in \mathbf{R}, \quad i = \sqrt{-1} \}$. Show that with regular multiplication as the binary operation $(\mathbf{C}; \cdot)$ is a group.

Solution The proof is almost trivial.

Closure: Let $r_1 e^{i\varphi_1}, r_2 e^{i\varphi_2} \in \mathbf{C}$ be any two elements of $\mathbf{C}$. Then,

$$\begin{aligned} r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2} &= r_1 r_2 e^{i(\varphi_1+\varphi_2)} \\ &= Re^{i\phi} \in \mathbf{C} \end{aligned}$$

where, of course, $R = r_1 r_2$ and $\Phi = \varphi_1 + \varphi_2$.

Associativity: For any $r_1 e^{i\varphi_1}, r_2 e^{i\varphi_2}, r_3 e^{i\varphi_3} \in \mathbf{C}$,

$$\begin{aligned} r_1 e^{i\varphi_1} \cdot (r_2 e^{i\varphi_2} \cdot r_3 e^{i\varphi_3}) &= r_1 r_2 r_3 e^{i(r_1+r_2+r_3)} \\ &= r_1 r_2 e^{i(r_1+r_2)} \cdot r_3 e^{i\varphi_3} \\ &= (r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2}) \cdot r_3 e^{i\varphi_3} \end{aligned}$$

Neutral element: Obviously it is 1.

Inverse: For every $re^{i\varphi} \in \mathbf{C}$, there exists $(1/r)e^{-i\varphi} \in \mathbf{C}$, such that

$$re^{i\varphi} \cdot \frac{1}{r} e^{i\varphi} = 1$$

Thus, $(\mathbf{C}; \cdot)$ is indeed a group. ■

Definition 5.10 (Symmetric group)

Let X be a nonempty set with n elements, and let P_X be a set of bijective mappings $X \to X$ (i.e., permutations), then $(P_X; \circ)$ is a group under the usual composition of functions. This group is called the **group of permutations** of X or the **symmetric group** on X, and is often denoted as S_X.

Example 5.34 Consider the set $X = \{1, 2, 3\}$, and the family of all permutations

$$P_X = \{p_i | p_i : X \to X, i = 0, \ldots, 5\}$$

such that

$$p_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \text{ and so on}$$

It is easy to check that, with the composition of permutations (see Example 5.7) as the binary operation on P_X, all group axioms G1–G4 are satisfied, and we obtain the group S_3. Let's just check if S_3 satisfies the cancellation law: If

$$p_i \circ p_j = p_i \circ p_k$$

then

$$p_j = p_k$$

Well,

$$\begin{aligned} p_j &= p_0 \circ p_j \\ &= (p_i^{-1} \circ p_i) \circ p_j \\ &= p_i^{-1} \circ (p_i \circ p_j) \\ &= p_i^{-1} \circ (p_i \circ p_k) \\ &= (p_i^{-1} \circ p_i) \circ p_k \\ &= p_0 \circ p_k \\ &= p_k \end{aligned}$$

Similarly, we can show that $p_j \circ p_i = p_k \circ p_i$ implies $p_j = p_k$. ■

Example 5.35 Recall that a point in a plane is given when its coordinates are given. In other words, a point A is uniquely represented in a plane by an ordered pair of real numbers $(x,\ y)$. Suppose now that we want to translate the point A to a new position $(x + a,\ y + b)$, that is, the point $A(x,\ y)$ becomes a new point $A'(x + a,\ y + b)$. In order to distinguish the point from its "translation," we will denote the process of translation $A \to A'$ by $[a,\ b]$. We can continue this process and move point A' to another position (point) A'' by translation $[c, d]$, and so on. So, our transformation looks like this:

$$A \xrightarrow{[a,b]} A' \xrightarrow{[c,d]} A''$$

We express our transformation symbolically as follows:

$$[a, b] * [c, d] = [a + c, b + d]$$

As before, we will drop "$*$" and write our "multiplication" simply as

$$[a,\ b][c,\ d] = [a + c,\ b + d]$$

Show that the set of transformations described above is a commutative group.

Solution Considering the fact that the coordinates of a point are real numbers, it is not difficult to verify that axioms G1–G4 are always satisfied under multiplication. Furthermore, since $x + y = y + x$ for all $x, y \in \mathbf{R}$, $[a, b][c, d] = [c, d][a, b]$, thus our group is also commutative. ■

Example 5.36 Let X be the set of all points of some geometric figure, say, an equilateral triangle. Consider a set T of all permutations $\tau : X \to X$ of this set, such that τ preserves the distance between any two points $x, y \in X$. In other words, we want $d(x, y) = d(\tau(x),\ \tau(y))$ for all $x, y \in X$, and all $\tau \in T$. Furthermore, we want the product of those permutations to preserve distances too, that is, we want

$$\begin{aligned} d[\sigma\tau(x), \sigma\tau(y)] &= d[\sigma(\tau(x), \sigma(\tau(y)] \\ &= d[\tau(x), \tau(y)] \\ &= d(x, y) \end{aligned}$$

$\forall x, y \in X$ and $\forall \sigma, \tau \in T$.

It follows that $\sigma\tau \in T$. Therefore,

$$\begin{aligned} d(x, y) &= d[\sigma(\sigma^{-1}(x), \sigma(\sigma^{-1}(y)] \\ &= d[e(x), e(y)] \\ &= d(x, y) \end{aligned}$$

So, $\sigma^{-1} \in T$, and $e \in X$ is the identity, that is, indicating "no permutation."

As an example, consider an equilateral triangle ΔABC. Recall that angles of an equilateral triangle $\angle BAC$, $\angle ABC$, and $\angle BCA$ are all equal to $\pi/3$. Observe furthermore that the bisectors of these angles, let's call them t_A, t_B, and t_C, respectively, intersect at one point, say, O. Now, consider the set of all transformations T that map the triangle to itself. What are those? Well, let's see (Figure 5.1):

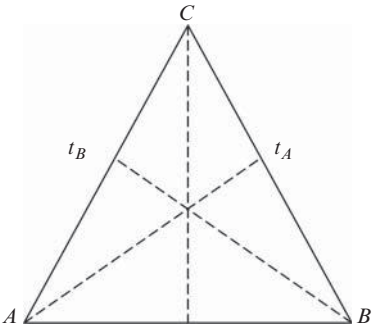


Figure 5.1

We can rotate our triangle in the positive direction (counterclockwise) through $0, 2\pi/3$, and $4\pi/3$, which results in the vertices moving in the following manner:

$$e = \begin{cases} A \to A \\ B \to B \\ C \to C \end{cases} \qquad \alpha = \begin{cases} A \to B \\ B \to C \\ C \to A \end{cases} \qquad \beta = \begin{cases} A \to C \\ B \to A \\ C \to B \end{cases}$$
$$= (A, B, C) \qquad = (B, C, A) \qquad = (C, A, B)$$

where we think of triples $(A, B, C), (B, C, A), (C, A, B)$ as group elements.

Before we continue this way, let's recognize something we know from before (cf. Example 5.34). As you have probably noticed yourself, nothing prevents us from writing the transformations above in such a way that you may find it to be more practical. Simply substituting

$$A = 1, \quad B = 2, \quad \text{and} \quad C = 3$$

we can write the above transformations in matrix form:

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \quad \alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \beta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Or even more simply, we can write

$$e = (1, 2, 3) \quad \alpha = (2, 3, 1) \quad \beta = (3, 1, 2)$$

Now, observe that $\beta = \alpha\alpha = \alpha^2$. Indeed,

$$\begin{aligned} \alpha^2 = \alpha\alpha &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \\ &= (2, 3, 1)(2, 3, 1) = (3, 1, 2) = \beta \end{aligned}$$

This is exactly how we want: two consecutive rotations through $2\pi/3$ amount to one rotation of $4\pi/3$.

The next three symmetries are reflections about t_A, t_B, and t_C, which produce the following transformation of vertices:

$$\gamma = \begin{cases} A \to A \\ B \to C \\ C \to B \end{cases} \qquad \delta = \begin{cases} A \to C \\ B \to B \\ C \to A \end{cases} \qquad \eta = \begin{cases} A \to B \\ B \to A \\ C \to C \end{cases}$$
$$= (A, C, B) \qquad = (C, B, A) \qquad = (B, A, C)$$

As before, we will write this as

$$\gamma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \delta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \quad \eta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Inspecting the transforms carefully, we note that

$$\delta = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (3,1,2)(1,3,2) = \alpha^2\gamma$$

and finally,

$$\eta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (2,3,1)(1,3,2) = \alpha\gamma$$

Hence, the set of all symmetry transformations of an equilateral triangle is

$$T = \{e, \alpha, \alpha^2, \gamma, \alpha^2\gamma, \alpha\gamma\}$$

It is now easy to verify that axioms G1–G4 are satisfied, thus, T is a group commonly known as D_3, the **dihedral group** of degree 3. An important note may be in order: one should not confuse the "degree" of the dihedral group with its "order." Dihedral groups are often denoted as D_{2n}, where n is the degree while $2n$ represents the order (cardinality)of a group (cf. Definition 5.12). ■

Example/Exercise 5.37 Convince yourself that the multiplication table for D_3 is

$\cdot$	e	α	α^2	γ	$\alpha\gamma$	$\alpha^2\gamma$
e	e	α	α^2	γ	$\alpha\gamma$	$\alpha^2\gamma$
α	α	α^2	e	$\alpha\gamma$	$\alpha^2\gamma$	γ
α^2	α^2	e	α	$\alpha^2\gamma$	γ	$\alpha\gamma$
γ	γ	$\alpha^2\gamma$	$\alpha\gamma$	e	α^2	α
$\alpha\gamma$	$\alpha\gamma$	γ	$\alpha^2\gamma$	α	e	α^2
$\alpha^2\gamma$	$\alpha^2\gamma$	$\alpha\gamma$	γ	α^2	α	e

Example 5.38 Similarly, we construct D_4. Take any square with vertices $1, 2, 3, 4$, $\boxdot_{1234}$ (Figure 5.2) and consider moving the square to make it coincide with its former self. In other words, we would like to distinguish all possible ways in which a square (any square object) can be repositioned.

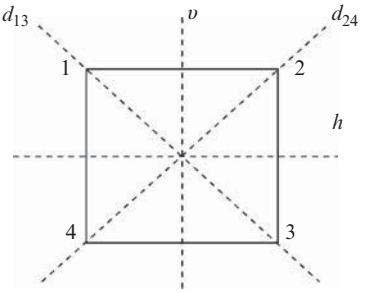


Figure 5.2

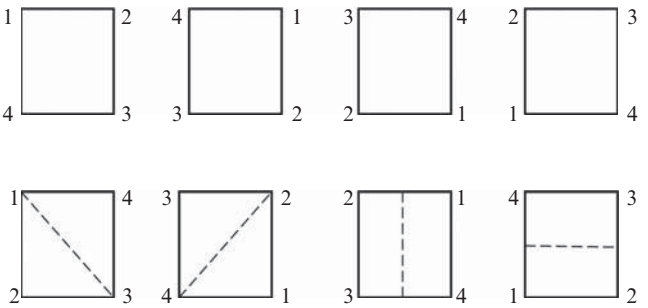


Figure 5.3

First, the obvious symmetries are obtained by rotating the square clockwise about its center through angles of $90°, 180° = 2 \times 90°, 270° = 3 \times 90°$. In other words, if our "original" square was $\boxdot_{1234}$, after a $90°$-clockwise rotation we obtain $\boxdot_{2341}$. A $180°$-rotation gives us $\boxdot_{3412}$, and so on. Of course, a rotation of $360°$ gives the original configuration – as if no transformation was performed. We will call it ρ_0. Explicitly, the elements are shown in Figure 5.3.

Writing all of this as a permutation, we have

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad \rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix},$$

$$\rho_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$$

There are four more symmetries to consider. We can flip the square about diagonals d_{13} and d_{24} and also about the horizontal and vertical axes, h and v. Let's call the flips about d_{13} and d_{24}, ρ_4 and ρ_5, respectively. Finally, the flip about h we

denote by ρ_6, and the one about v, ρ_7. This gives us four more transformations:

$$\rho_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}, \quad \rho_5 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \quad \rho_6 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix},$$

$$\text{and } \rho_7 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

Thus, set $S = \{\rho_0, \rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6, \rho_7\}$ is the set of symmetries of the square. Let's endow S with an operation "$\circ$" as the composition of transformations ρ_i's. Recall that, for instance,

$$\rho_1 \circ \rho_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \rho_7$$

One can easily convince oneself that set S is closed under such defined operations, that is

$$\rho_i \circ \rho_j \in S, \quad \forall i, j = 0, 1, \ldots, 7$$

Associativity holds, that is

$$\rho_i \circ (\rho_j \circ \rho_k) = (\rho_i \circ \rho_j) \circ \rho_k$$

ρ_0 is a neutral element, that is

$$\rho_i \circ \rho_0 = \rho_0 \circ \rho_i = \rho_i$$

Finally, for every $\rho_i \in S$ there exists an inverse $\rho_i^{-1} \in S$, such that

$$\rho_i \circ \rho_i^{-1} = \rho_i^{-1} \circ \rho_i = \rho_0$$

So, we observe that any transformation is equivalent to one of those from set S. Thus, under the composition "$\circ$" as the binary operation, the set $(S; \circ)$ becomes the *group of symmetries of the square*, often called the **octic group**. ■

Consequently, we have

Definition 5.11 The ***n*th dihedral group** is a group of symmetries of a regular n-tgon.

On the other hand, if we can consider only rotations of a square about its center, we obtain a group R_4 as described in the following

Example 5.39 Consider a set $R_4 = \{e, r_1, r_2, r_3\}$ of rotations of a square about its center, making it coincide with itself. Let's endow the set with an operation "*"

representing addition of rotations. Then $(R_4; *)$, defined by the following multiplication table, becomes a **group of rotations of the square**.

$*$	e	r_1	r_2	r_3
e	e	r_1	r_2	r_3
r_1	r_1	r_2	r_3	e
r_2	r_2	r_3	e	r_1
r_3	r_3	e	r_1	r_2

Note the difference between the group R_4 and the Klein four-group from Example 5.25. ■

Example/Exercise 5.40 Find the symmetries of the following shape, and construct the multiplication table showing that the symmetries form a group (see Figure 5.4).

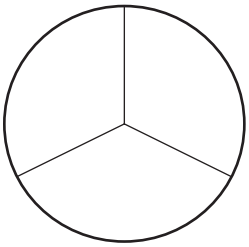


Figure 5.4

Example 5.41 Let X be a set. Then the family of all subsets of X, equipped with the symmetric difference of sets Δ (see Chapter 1) as a binary operation, forms the **Boolean group** $B(X)$. Recall that if $A, B \subseteq X$, then

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

Closure is obvious. Associativity holds (see prob. I. 23).

The neutral element is $\emptyset$. Since $A \Delta A = \emptyset$, the inverse of A is A itself. Thus, $B(X)$ is indeed an abelian group. ■

I hope you won't find proofs of the following two theorems too difficult.

Theorem 5.9 For every nonempty set X, the algebraic structure $(S_X; \circ)$ is a permutation group, with "$\circ$" denoting the composition of permutations.

Proof First, recall that permutations of set X, $X \to X$, are bijections. Furthermore, the composition of bijections is again a bijection. Thus, it has an inverse that is a bijection too. Finally, the "neutral element" is the identity map, certainly a bijection. So, indeed, $(S_X; \circ)$ is a group. ■

Example 5.42 For illustration, let's take three simple bijections from the set F of all bijections from $\mathbf{R}$ to $\mathbf{R}$:

$$f : \mathbf{R} \to \mathbf{R}, \quad g : \mathbf{R} \to \mathbf{R}, \quad \text{and} \quad h : \mathbf{R} \to \mathbf{R}$$

defined by

$$f(x) = x + 1, \quad g(x) = 2x, \quad h(x) = \frac{1}{2}x$$

All three functions are well defined and, as we recall, the composition of bijections is a bijection too. For instance,

$$(f \circ g)(x) = f[g(x)] = f[2x] = 2x + 1$$

Similarly,

$$(g \circ f)(x) = g[f(x)] = g[x + 1] = 2x + 2$$

The set F is closed with respect to composition. Let's check associativity:

$$(f \circ (g \circ h))(x) = f \circ (g[h(x)]) = f\left\{g\left[\frac{1}{2}x\right]\right\} = f\{x\} = x + 1$$

$$((f \circ g) \circ h)(x) = (f \circ g)[h(x)] = f\left\{g\left[\frac{1}{2}x\right]\right\} = f\{x\} = x + 1$$

The identity is obvious: $i(x) = x$. Indeed,

$$f(x) = (i \circ f)(x) = i[f(x)] = i(x + 1) = x + 1$$
$$= f[i(x)] = (f \circ i)(x)$$

The same is valid of any function from F. Finally, it is trivial to verify that every function from F has an inverse. Taking, for example, the same f again, we have

$$(f^{-1} \circ f)(x) = f^{-1}[f(x)] = f^{-1}(x + 1) = x + 1 - 1 = x = i(x)$$
$$(f \circ f^{-1})(x) = f[f^{-1}(x)] = f(x - 1) = x - 1 + 1 = x = i(x)$$

Thus, the set of all permutations/all bijections is a group. ■

Theorem 5.10 A group S_n has $n!$ elements.

Proof By induction. ■

Example/Exercise 5.43 Let $H = \{\rho_0, \ldots, \rho_{12}\}$ be the set of symmetries of a regular hexagon; see Figure 5.5,

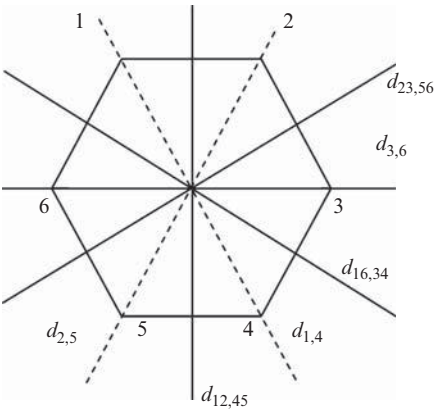


Figure 5.5

where

$$\rho_0 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}, \quad \rho_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix}, \text{ and so on}$$

With the composition of symmetries as the binary operation, construct the multiplication table for $(H; \circ)$, and show that $(H; \circ)$ is a group.

Example 5.44 Show that the set $\mathbf{Z}_n = \{0, 1, 2, \ldots, n-1\}$ is a group under addition (mod n).

Solution Theorem 4.3 shows that addition (mod n) is a well-defined binary operation. Closure is obviously satisfied and you can easily convince yourself that associativity also holds. The neutral element is 0. What about the inverse? Let's see. Take any nonzero element $x \in \mathbf{Z}_n$ and consider $n - x$. Then,

$$x + (n - x) = n = 0(mod\ n)$$

Thus, $n - x$ is the inverse of x. Of course, it follows that 0 is its own inverse. Hence, G1–G4 are satisfied and we have a group indeed. ■

Example/Exercise 5.45 Consider the set $C = \{1, -1, i, -i; \cdot\}$, where $i = \sqrt{-1}$, and "$\cdot$" indicates ordinary multiplication. Show that as defined, $(C; \cdot)$ is a group.

Example 5.46 Consider a set $K = \{-1, 1, -i, i, -j, j, -k, k\}$ with multiplication defined in the following way:

$$i^2 = j^2 = k^2 = -1$$

and

$$ij = k = -ji, \quad \text{and so on cyclically}$$

Then, K becomes a group called the **quaternion group**. ■

In case you haven't been exposed to the basic concepts of algebra of complex numbers and thus wondered about the strange number $i = \sqrt{-1}$, you may find the next two examples especially interesting.

Example 5.47 Let's consider the set of ordered pairs of numbers

$$\mathbf{C} = \{(a, b) | a, b \in \mathbf{R}; (a, b) = (c, d) \Leftrightarrow a = c \text{ and } b = d\}$$

and let's define on $\mathbf{C}$ a binary operation, the "product of pairs," as follows: For all $(a, b), (c, d) \in \mathbf{C}$

$$(a, b) \cdot (c, d) = (ac - bd, bc + ad) \tag{5.4}$$

Consequently, with a product as defined, we are naturally led to a quotient of pairs. That is, if (x, y) is a quotient of pairs (a, b) and (c, d) then, of course,

$$(x, y) \cdot (c, d) = (a, b) \tag{5.5}$$

Following (5.4) it means that

$$(xc - yd, xd + yc) = (a, b) \tag{5.6}$$

that is

$$xc - yd = a, \qquad xd + yc = b \tag{5.7}$$

Solving the system (5.7) we get

$$x = \frac{ac + bd}{c^2 + d^2}, \qquad y = \frac{bc - ad}{c^2 + d^2} \tag{5.8}$$

In other words, we define the quotient

$$\frac{(a, b)}{(c, d)} = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2}\right)$$

Obviously, the quotient is defined only if c and d are not both equal to zero.

Now we are ready to recognize that a set $\mathbf{C}^+ = \mathrm{C} \setminus \{(0,0)\}$ with multiplication defined by (5.4) satisfies axioms G1–G4 and is thus a group:

Closure is obvious. Let's check associativity. For all $(a,b),(c,d),\ (e,f) \in \mathbf{C}$,

$$\begin{aligned}(a,b)\cdot((c,d)\cdot(e,f)) &= (a,b)\cdot(ce-df,de+cf)\\ &= (a(ce-df)-b(de+cf),b(ce-df)+a(de+cf))\\ &= (ace-adf-bde-bcf,bce-bdf+ade+acf)\\ &= ((ac-bd)e-(ad+bc)f,(ac-bd)f+(ad+bc)e)\\ &= ((a,b)\cdot(c,d))\cdot(e,f)\end{aligned}$$

The neutral element is $(1,0)$. Indeed, for every $(a,b) \in \mathbf{C}^+$,

$$(1,0)\cdot(a,b) = (1\cdot a - 0\cdot b, 0\cdot a + 1\cdot b) = (a,b)$$
$$(a,b)\cdot(1,0) = (a\cdot 1 - b\cdot 0, b\cdot 1 + a\cdot 0) = (a,b)$$

Finally, for every $(a,b) \in \mathbf{C}^+$, there exists

$$(a,b)^{-1} = \left(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2}\right)$$

such that

$$(a,b)\cdot(a,b)^{-1} = (a,b)^{-1}(a,b) = (1,0)$$

Thus $(\mathbf{C}^+;\ \cdot)$ is indeed a group. ■

Example/Exercise 5.48 Endowing the same set $\mathbf{C}$ with addition defined as follows:

$$(a,b)+(c,d) = (a+c,b+d), \quad \forall\ (a,b),\ (c,d) \in \mathbf{C}$$

you can easily show that $(\mathbf{C};\ +)$ is also a group whose neutral element with respect to addition is $(0,0)$.

Now let's digress a little, and as a byproduct of the aforementioned discussion, let's finally try to "justify" the infamous $i = \sqrt{-1}$.

Suppose we wanted to solve the simple quadratic equation

$$x^2 + 1 = 0 \qquad (*)$$

You would recall from high school algebra that such an equation doesn't have a real solution, that is, we cannot find $x \in \mathbf{R}$ such that $x^2 = -1$. Obviously, our set $\mathbf{R}$ is not rich enough to accommodate even an equation as simple as (*). Clearly, we need something more.

Suppose we take our set **C** as in Examples 5.47 and 5.48 and consider

$$\begin{aligned}(0,1)^2 &= (0,1)\cdot(0,1)\\ &= (-1,0)\end{aligned}$$

So the analogue of $x^2 = -1$ would be

$$(x,\ y)^2 = (-1,\ 0) \qquad (**)$$

Evidently, the solution of equation (**) is $(0,\ 1)$. Finally, if we denote the pair $(0,\ 1)$ by i, and $(-1,\ 0)$ by -1 we get

$$i^2 = -1$$

The set **C**, introduced above, is the set of **complex numbers,** which you might have encountered before in the more familiar form of $a + ib$. Indeed, with our "pair representation"

$$\begin{aligned}(a,b) &= (a,0) + (b,0)(0,1)\\ &= a + ib\end{aligned}$$

The ordered pair representation of a complex number was introduced by Hamilton[3] in 1831 as a more logically satisfying expression that, in particular, evades the unattractive term "imaginary" number.

Example 5.49 Let $\mathbf{R}^n = \mathbf{R}\times\mathbf{R}\times\cdots\times\mathbf{R} = \{(x_1, x_2, \ldots, x_n) | x_i \in \mathbf{R}\}$ be the set of all ordered real n-tuples. If we supply set $\mathbf{R}^n$ with ordinary coordinate addition

$$\begin{aligned}x + y &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n)\\ &= (x_1 + y_1, \ldots, x_n + y_n)\end{aligned}$$

then $\mathbf{R}^n$ becomes an abelian group. ■

Example 5.50 Consider the set $F = \{f_1, \ldots, f_6\}$ such that

$$f_1 = x,\quad f_2 = 1 - x,\quad f_3 = \frac{1}{x},\quad f_4 = \frac{x-1}{x},\quad f_5 = \frac{1}{1-x},\quad f_6 = \frac{x}{x-1}$$

With the composition of functions as the binary operation, F becomes a non-abelian group. You should verify this. ■

[3]Hamilton, W. R. (1805–1865), Irish mathematician and physicist.

Example/Exercise 5.51 Let $p : \mathbf{R} \to \mathbf{R}$ be a polynomial in one variable with real coefficients, that is

$$p(x) = \sum_{i}^{n} a_i x^i, \quad a_i \in \mathbf{R}$$

and let

$$P_n = \{p | p, n < m \in \mathbf{N}\} \cup \{0\}$$

be the set of all polynomials whose degree is less than $m \in \mathbf{N}$. Show that with ordinary addition of polynomials, P_n is a group.

Example 5.52 Consider the set of all polynomials of first degree, that is, all functions $f : \mathbf{R} \to \mathbf{R}$ defined by

$$f_{a,b}(x) = ax + b$$

where $a, b \in \mathbf{R}$ with $a \neq 0$. This set equipped with the composition of functions as the binary operation forms an **Affine group** $\mathrm{Aff}(1, \mathbf{R})$.

Indeed, let's take another function, $f_{c,d}(x) = cx + d, \quad c \neq 0$, and check closure:

$$\begin{aligned}(f_{a,b} \circ f_{c,d})(x) &= f_{a,b}(cx + d) = a(cx + d) + b \\ &= acx + ad + b \\ &= f_{ac,ad+b}(x)\end{aligned}$$

Since $ac \neq 0, \quad f_{ac,ad+b} \in \mathrm{Aff}(1, \mathbf{R})$.

Associativity is obvious since the composition of functions is associative.

The neutral element (identity) is $f_{1,0} : \mathbf{R} \to \mathbf{R}$:

$$\begin{aligned}(f_{a,b} \circ f_{1,0})(x) &= f_{a,b}(1 \cdot x + 0) = ax + b = f_{a,b} \\ (f_{1,0} \circ f_{a,b})(x) &= f_{1,0}(ax + b) = 1 \cdot (ax + b) + 0 = f_{a,b}\end{aligned}$$

Finally, let's show that the inverse is $f_{a^{-1},-a^{-1}b} : \mathbf{R} \to \mathbf{R}$

$$\begin{aligned}(f_{a,b} \circ f_{a^{-1},-a^{-1}b})(x) &= f_{a,b}(a^{-1}x - a^{-1}b) = a(a^{-1}x - a^{-1}b) + b \\ &= x = f_{1,0} \\ (f_{a^{-1},-a^{-1}b} \circ f_{a,b})(x) &= f_{a^{-1},-a^{-1}b}(ax + b) = a^{-1}(ax + b) - a^{-1}b \\ &= x = f_{1,0}\end{aligned}$$

■

In general, we conclude that if $F_{\mathbf{R}}$ is the set of all functions $\mathbf{R} \to \mathbf{R}$, then $(F_{\mathbf{R}}; \circ)$ is a groupoid, with "$\circ$" denoting the usual composition of functions. If we choose the set of all bijections $P_{\mathbf{R}} \subseteq F_{\mathbf{R}}$ with the same operation "$\circ$," then $(P_{\mathbf{R}}; \circ)$ is a nonabelian group.

Theorem 5.11 Let G be a group, and let $a, b \in G$. Then, $(ab)^{-1} = b^{-1}a^{-1}$.

Proof Since G is a group, and $a, b \in G$ then $ab \in G$ too. That furthermore implies that there should exist a^{-1}, b^{-1}, $(ab)^{-1} \in G$ such that

$$aa^{-1} = a^{-1}a = e$$
$$bb^{-1} = b^{-1}b = e$$

but also

$$(ab)(ab)^{-1} = (ab)^{-1}(ab) = e \qquad (*)$$

Consider the product

$$(ab)(b^{-1}a^{-1})$$

which, because of associativity, we can write

$$\begin{aligned}(ab)(b^{-1}a^{-1}) &= a(b^{-1}b)a^{-1} \\ &= aea^{-1} \\ &= aa^{-1} \\ &= e \qquad (**)\end{aligned}$$

From (*) and (**), it follows that $(ab)^{-1} = b^{-1}a^{-1}$, which was to be shown. ■

Example/Exercise 5.53 Let G be a group, and let $a_1, a_2, \ldots, a_n \in G$, then

$$(a_1 a_2 \cdots a_n)^{-1} = a_n^{-1} a_{n-1}^{-1} \cdots a_1^{-1}$$

5.3 SUBGROUPS

Definition 5.12 The number of elements of a group G (finite or infinite) is said to be the **order of the group** or the **cardinality of the group**, denoted $|G|$.

Example 5.54

(i) The group $(\mathbf{Z}_n; +)$ is a group of order n.
(ii) The group from Example 5.25 is a group of order 4, as is the group $(C; \cdot)$ from Example 5.44. The group from Example 5.34 is of order 6 (note, however, that its **degree** is 3). ■

Example 5.55

(i) The group $(\mathbf{Z}; +)$ is a group of infinite order.
(ii) The group $(\mathbf{R}^*; \cdot)$ of nonzero real numbers is a group of infinite order. ■

Definition 5.13 Let G be a group, and let $a \in G$. We define the **order of *a*** to be the smallest natural number n, such that $a^n = e$ and we denote that number by $|a|$. If no such number exists, we say that a is of **infinite order**.

Example 5.56 An element x of a group has order 1 iff $x = e$. ■

Example 5.57 Consider $(\mathbf{Z}_{10}; +)$ and find $|2|$ and $|7|$.

Solution

$$1 \cdot 2 = 2,\ \ 2 \cdot 2 = 4,\ \ 3 \cdot 2 = 6,\ \ 4 \cdot 2 = 8,\ \ 5 \cdot 2 = 0 = e, \quad \text{thus } |2| = 5$$

$$1 \cdot 7 = 7,\ \ 2 \cdot 7 = 4,\ \ 3 \cdot 7 = 1,\ \ 4 \cdot 7 = 8,\ \ 5 \cdot 7 = 5,\ \ 6 \cdot 7 = 2,\ \ 7 \cdot 7 = 9,$$

$$8 \cdot 7 = 6,\ \ 9 \cdot 7 = 3,\ \ 10 \cdot 7 = 0 = e, \quad \text{thus } |7| = 10$$ ■

Example/Exercise 5.58 Consider $(\mathbf{Z}_6; +)$ and find $|3|$ and $|5|$.

Theorem 5.12 If there exists a nonzero $m \in \mathbf{Z}$ such that $a^m = e$, then there exists an $n \in \mathbf{N}$ such that $a^n = e$.

Proof Let $m < 0$ be an integer. Then, of course, $-m > 0$ and

$$a^{-m} = (a^m)^{-1} = e^{-1} = e$$ ■

Certainly, there are groups whose every element is of infinite order.

Example 5.59 Groups $(\mathbf{Z}; +)$, $(\mathbf{Q}; +)$, and $(\mathbf{R}; +)$ are examples of groups whose every nonzero element is of infinite order. ■

Theorem 5.13 Let G be a group, and let $a \in G$, such that $|a| = n$. Then, there are exactly n *different* powers of a.

Proof Let a^m be any power of a, that is, $a^m \in \{a^0, a, a^2, \ldots, a^{n-1}\}$. Then, we can express m as

$$m = nq + r, \quad 0 \leq r \leq n, \quad n, q, r \in \mathbf{N}$$

and write

$$\begin{aligned} a^m &= a^{nq+r} = a^{nq}a^r \\ &= (a^n)^q a^r = e^q a^r \\ &= a^r \end{aligned}$$

where $r \in \{0, 1, 2, \ldots, n-1\}$. Next, we just need to prove that $a^r \neq a^s$ for $r \neq s$. Well, suppose not. Suppose there are two *different* integers $r, s \in \{0, 1, 2, \ldots, n-1\}$, such that $a^r = a^s$. Then, either $r < s$ or $s < r$. Let's consider the case $0 \leq s < r < n$. It follows that

$$0 < r - s < n \in \mathbf{N}$$

But since $a^r = a^s$ we have

$$\begin{aligned} a^{r-s} &= a^r(a^s)^{-1} \\ &= a^s(a^s)^{-1} \\ &= e \end{aligned}$$

However, this cannot be since $r - s < n \in \mathbf{N}$, and n is the order of a, that is, the smallest natural number for which $a^n = e$. Thus, we cannot have $a^r = a^s$ with $r \neq s$, that is, the powers $a^0, \ldots, a^{n-1}$ are all different. ■

Corollary 5.2 Let G be a group and $a \in G$ such that $|a| = n$. Then, $a^p = e$ iff p is a multiple of n, that is, $p = nq, q \in \mathbf{Z}$.

Example 5.60 Let G be a group and $a, b \in G$, such that $|a| = n \in \mathbf{N}$. Show that

$$|a| = |bab^{-1}|$$

Solution $|a| = n$ means that $a^n = e$. Thus,

$$\begin{aligned}(bab^{-1})^n &= b^{-n}a^n b^n \\ &= b^{-n}eb^n \\ &= (bb^{-1})^n \\ &= e^n \\ &= e = a^n\end{aligned}$$

The converse can be proved similarly. ■

Example/Exercise 5.61 Let G be a group and $a \in G$, such that $|a| = n$. Show that $a^{n-r} = (a^r)^{-1}$.

Definition 5.14 A nonempty subset H of a group G is a **subgroup of G**, denoted $H \leq G$, iff H is a group with the same operation as G.

Definition 5.14′ A subset H of a group G is a **subgroup of G** if

(i) $e \in H$
(ii) if $a, b \in H$, then $ab \in H$
(iii) if $a \in H$, then $a^{-1} \in H$

In other words, if H is a subgroup of G, then H is a group itself. Why is that? Well, the binary operation under which G becomes a group is certainly the same operation if restricted to $H \subseteq G$. It has to be associative since it is associative for all the elements of G, thus in particular for those belonging to H. The question of the neutral element is slightly more intriguing. First, the assumption was that $H \neq \emptyset$, so H has to contain at least one element, say, a. But H is closed with respect to the inverse, that is, if $a \in H$, then $a^{-1} \in H$, and therefore $aa^{-1} = e \in H$. Hence, H is a group indeed.

Example 5.62 $\mathbf{Q}^+$ under multiplication is the subgroup of $\mathbf{R}^+$ under multiplication, that is, $(\mathbf{Q}^+; \cdot) \leq (\mathbf{R}^+; \cdot)$. ■

The question that naturally occurs is : Is there a way to recognize whether a certain subset of a group G is itself a group without going through G1–G4? Fortunately, we are in possession of some relatively simple tools for determining whether or not $H \subseteq G$ is the subgroup of G. It suffices to perform one of the following tests to get a definite answer.

Theorem 5.14 A nonempty subset H of a group G is a subgroup of G iff the following two conditions are satisfied:

(i) $\forall a, b \in H, \quad ab \in H$
(ii) $\forall a \in H, \quad a^{-1} \in H$

Proof Suppose, first, that $H \leq G$, and let's show that (i) and (ii) are satisfied.

Since $H \leq G$, that is, H is a group itself, closure is automatically satisfied. We need to show that $e \in G$ is also the identity for H. Well, suppose there is an identity

$e' \in H$, such that $e'a = ae' = a, \forall a \in H$. In particular,

$$e'e' = e'$$

At the same time, since $e \in G$ is the identity for the whole group G,

$$ee' = e'e = e'$$

Thus, by the left cancellation law (LCL),

$$e = e'$$

Now, consider some $a \in H$. Since $H \subseteq G$ $a \in G$ too. That furthermore implies that $a^{-1} \in G$ and so $a^{-1}a = e \in G$. It remains to be shown that there is a unique $a^{-1} \in H$. Suppose not. Suppose there is $a' \in H$, thus $a' \in G$ too, such that $aa' = e$. But that would mean that

$$e = aa' = aa^{-1}$$

By the LCL $a' = a^{-1}$.

Finally, we need to prove the converse: if $H \subseteq G$, satisfying (i) and (ii), then H is a subgroup of G.

Let $a, b, c \in H$. a, b, c are also elements of G, since $H \subseteq G$. Since G is a group

$$a(bc) = (ab)c, \quad \forall a, b, c \in G$$

With this, and the assumption (i), we see that associativity holds in H as well. Furthermore, (ii) implies that every element of H has an inverse. What about the identity? Well, since $H \neq \emptyset$ there has to exist $a \in H$, but (ii) says that $\forall a \in H$, $\exists a^{-1} \in H$. On the other hand, by (i) we know that H is closed under multiplication, thus

$$aa^{-1} = e \in H$$

We see that axioms G1–G4 are satisfied, thus we proved that $H \leq G$. ■

Theorem 5.15 A nonempty subset H of a group G is a subgroup iff

$$\forall a, b \in H, \quad ab^{-1} \in H$$

Proof If H is a subgroup, that is, a group itself, there is nothing to prove, that is, $ab^{-1} \in H$ (has to be) for all $a, b \in H$. It remains to be shown that if $ab^{-1} \in H$, $\forall a, b \in H$ then $H \leq G$.

Let $a, b \in H$, and consider the case where $b = a$. Then,

$$ab^{-1} = aa^{-1} = e \in H$$

So, H contains the identity. Now, since $a, e \in H$, it follows that

$$ea^{-1} = a^{-1} \in H$$

Thus, we have proved that if $a, b \in H$, then $b^{-1} \in H$ too. But that also implies

$$a(b^{-1})^{-1} = ab \in H$$

Finally, the associativity of products in H comes from the associativity in G, thus H is a group, that is, a subgroup of G. ■

Corollary 5.3 Let $H \subseteq G$ be a nonempty subset of a group G. Then, H is a subgroup of G iff

$$HH^{-1} \subseteq H$$

Example 5.63 Let G be an *abelian* group. Show that $H = \{x \in G | x^2 = e\}$ is a subgroup of G.

Solution Obviously, $H \neq \emptyset$. Suppose $a, b \in H$. We would like to show that $ab^{-1} \in H$ too. Since $a, b \in H$, it follows that $a^2 = e$ and $b^2 = e$. Thus, we need to show that $(ab^{-1})^2 = e$. Well,

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

But G is abelian, so

$$\begin{aligned}&= a(ab^{-1})b^{-1} \\ &= a^2(b^{-1})^2 \\ &= a^2(b^2)^{-1} \\ &= ee^{-1} = e\end{aligned}$$

Hence, $H \leq G$. ■

Example 5.64 Suppose G is a group, such that $a, b, c \in G$, and let $H \subseteq G$, such that all a, b and c (in any order), as well as their inverses, are elements of H. Show that $H \leq G$.

Solution From the description of H, it follows that if we randomly pick an element from H, say, $abac^{-1} \in H$, then

$$\begin{aligned} abac^{-1}(abac^{-1})^{-1} &= abac^{-1}c(aba)^{-1} \\ &= abae(aba)^{-1} \\ &= aba(aba)^{-1} \\ &= e \in H \end{aligned}$$

Obviously, similar reasoning could have been used with any combination of elements and their inverses from H. Thus, H is indeed a subgroup of G. ■

Example 5.65 Let G be a group, and let $H = \{a \in G | (ax)^2 = (xa)^2, \forall x \in G\}$. Show that $H \leq G$.

Solution Obviously, $H \neq \varnothing$. Let $a, b \in H$. In order to show that $H \leq G$, we need to show that $ab^{-1} \in H$, that is, $(ab^{-1}x)^2 = (xab^{-1})^2 \ \ \forall x \in G$. To make the calculation easier and more transparent, let's put $ab^{-1} = AS$, so we have

$$\begin{aligned} (ab^{-1}x)^2 &= (Ax)^2 \\ &= (Ax)^2(xA)^{-2}(xA)^2 \\ &= (x^2A^2)(A^{-2}x^{-2}A^2x^2) \\ &= x^2(A^2A^{-2})(x^{-2}A^2x^2) \\ &= x^2x^{-2}(A^2x^2) \\ &= A^2x^2 \\ &= (xA)^2 \\ &= (xab^{-1})^2 \end{aligned}$$ ■

Example 5.66 Let G be a group, and $a \in G$. Then, $\langle a \rangle = \{a^n | n \in \mathbf{Z}\}$ is a subgroup of G. (We will have to say more about these in Section 5.4.)

Note that $\langle a \rangle \neq \varnothing$ since $a^1 = a \in G$. Suppose that $x, y \in \langle a \rangle$ such that $x = a^m$ and $y = a^n$, where $m, n \in \mathbf{Z}$. We need to check whether xy^{-1} is an element of $\langle a \rangle$:

$$xy^{-1} = a^m(a^n)^{-1} = a^m a^{-n} = a^{m-n}$$

Since $-n \in \mathbf{Z}$, $a^{m-n} = xy^{-1} \in \langle a \rangle$ and thus $\langle a \rangle \leq G$. ■

Theorem 5.16 A nonempty subset $H \subseteq G$ is a subgroup of a *finite* group G iff H is closed under the operation of G, that is,

$$\forall a, b \in H, \quad ab \in H$$

Proof $ab \in H$, $\forall a, b \in H$ means that H is closed under the binary operation, thus H is a finite semigroup. By Theorem 5.7, LCL and RCL hold for the whole group G; therefore, they must hold for all elements of $H \subseteq G$. Thus, $H \leq G$. The converse is obvious. ■

Example 5.67 Show that the set $E = \{x \in \mathbf{Z} | x = 2k, k \in \mathbf{Z}\}$ of even numbers is a subgroup of $(\mathbf{Z}; +)$.

Solution First, it's trivially obvious that $E \neq \varnothing$ and, therefore, E is closed under addition. Second, addition is associative in $\mathbf{Z}$. Third, $0 \in \mathbf{Z}$ is a neutral element, and finally, for every $x \in E$, there exists $(-x) \in E$, such that $x + (-x) = (-x) + x = 0$, and thus $(E; +)$ is indeed a subgroup of $\mathbf{Z}$.

Also, more generally, if we construct a set ${}_{t}^{\mathbf{N}}\mathbf{Z} = \{ \ldots, -3t, -2t, -t, 0, t, 2t, 3t, \ldots \}$ of all of the multiples of $t \in \mathbf{Z}$, then $({}_{t}^{\mathbf{N}}\mathbf{Z}; +)$ is also a subgroup of $\mathbf{Z}$. ■

Example 5.68 Let H be a subgroup of an *abelian* group G and let $K = \{a \in G | a^n \in H, n \in \mathbf{Z}\}$. Show that $K \leq G$.

Solution Let $a, b \in K$. Then, by the construction of K, $a^n, b^m \in H$ for some $n, m \in \mathbf{Z}$. But since $H \leq G$, $a^{nm}, b^{nm} \in H$ too. Now observe that, since G is abelian,

$$\begin{aligned} (ab)^{nm} &= a^{nm} b^{nm} \\ &= (a^n)^m (b^m)^n \in H \end{aligned}$$

So, $ab \in K$.

We can go on and consider

$$\begin{aligned} (ab)^{nm} (ab)^{-nm} &= a^{nm} b^{nm} b^{-nm} a^{-nm} \\ &= a^{nm} e a^{-nm} \\ &= e \in K \end{aligned}$$

Thus, $K \leq G$. ■

Example 5.69 Consider the group $(\mathbf{R}^*; \cdot)$ of nonzero real numbers. Then, $H = \{2^n | n \in \mathbf{Z}\}$ under the same multiplication is a subgroup of $(\mathbf{R}^*; \cdot)$. Indeed,

If $2^n, 2^m \in H$, then obviously $2^n \cdot 2^m = 2^{n+m} \in H$, since $n + m \in \mathbf{Z}$. Associativity immediately follows. The neutral element is 2^0, which furthermore implies that the inverse for 2^n is 2^{-n}. Thus, $(H; \cdot) \leq (\mathbf{R}^*; \cdot)$. However, we could have used Theorem 5.15 and simply argued:

If $2^n, 2^m \in H$, then $2^n 2^{-m} = 2^{n-m} \in H$, since $n - m \in \mathbf{Z}$. Therefore, $(H; \cdot) \leq (\mathbf{R}^*; \cdot)$, as claimed. ■

Example 5.70 Let H be a subgroup of G, and let $K = \{x \in G | xax^{-1} \in H$, iff $a \in H\}$. Show that $K \leq G$.

Solution In order to show that $K \leq G$, we need to show that $xy^{-1} \in K$, $\forall x, y \in K$.

So, let $x, y \in K$. Then, if $a \in H$, $xax^{-1} \in H$ and $yay^{-1} \in H$. Conversely, if $xbx^{-1} \in H$, then $b \in H$. Now, for some $a \in H$,

$$\begin{aligned}(xy^{-1})a(xy^{-1})^{-1} &= (xy^{-1})(yy^{-1})a(yy^{-1})\,(xy^{-1})^{-1} \\ &= x(yy^{-1})(y^{-1}ay)(y^{-1}y)x^{-1} \\ &= x(y^{-1}ay)x^{-1}\end{aligned}$$

Thus, with $y^{-1}ay = b$, we see that $xbx^{-1} \in H$. In other words, $xy^{-1} \in K$ and so $K \leq G$. ■

Example/Exercise 5.71 With G, H, and K as in the previous exercise, show that $H \leq K$.

Example 5.72 Show that $H = \{f \in \mathcal{F}(\mathbf{R}) | f(-x) = -f(x)\}$ is a subgroup of $G(\mathcal{F}(\mathbf{R}); +)$, with $f(0) = 0$.

Solution Let $f, g \in H$. We need to show that $f + g^{-1} \in H$.

f and g, being the elements of $H \subseteq G$, imply first that

$$f(0) = 0 \quad \text{and} \quad g(0) = 0$$

So,

$$(f + g)(0) = f(0) + g(0) = 0 + 0 = 0$$

Second,

$$f(-x) = -f(x) \quad \text{and} \quad g(-x) = -g(x)$$

On the other hand, since the binary operation in G is "+," we have

$$g^{-1} = -g$$

Indeed,

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

So,

$$\begin{aligned}(f + g^{-1})(x) &= f(x) + g^{-1}(x) \\ &= f(x) - g(x) = f(x) + g(-x) \in H\end{aligned}$$

■

If you are unfamiliar with matrices, you can postpone the next two exercises until you have studied Chapter 6.

Example/Exercise 5.73 Let $(\mathcal{M}(\mathbf{R}); +)$ be a group of all real matrices. If

$$H = \left\{ \begin{bmatrix} a & b \\ c & 0 \end{bmatrix} \mid a, b, c \in \mathbf{R} \right\}$$

Show that $(H; +)$ is a subgroup of $(\mathcal{M}(\mathbf{R}); +)$.

Example 5.74 Consider a set $H = \left\{ \begin{bmatrix} a & b \\ 0 & 1/a \end{bmatrix} \mid a, b \in \mathbf{R}, \quad a \neq 0 \right\}$. Set H with matrix multiplication as the binary operation is a subgroup of $\mathrm{SL}(2, \mathbf{R})$ (see Exercise 5.32). ■

Example 5.75 We call the function $f : \mathbf{R}^2 \to \mathbf{R}^2$ an **isometry** of the plane if for all points $P(a, b), Q(c, d) \in \mathbf{R}^2$ in the plane, the distance

$$d(P, Q) = d(f(P), f(Q))$$

(cf. Examples 5.34–5.37). You should convince yourself that the set $(\mathbf{R}^2)$ of all isometries of the plane is a group under the composition of functions. The following are subgroups of $(\mathbf{R}^2)$:

(i) The set $T(\mathbf{R}^2)$ of all translations of the plane.
(ii) The set $R_P(\mathbf{R}^2)$ of all rotations of the plane about a given point P.
(iii) The set $\mathfrak{I}(\mathbf{R}^2) = T(\mathbf{R}^2) \circ R_P(\mathbf{R}^2) = \{\tau \circ \rho | \tau \in T, \rho \in I\}$ of all transformations of the plane.

■

Theorem 5.17 The set $\mathrm{Sym}(X)$ of all symmetries on a set X is a subgroup of S_X.

Proof Closure: Let σ and τ be any two symmetries of X. Then, $\sigma\circ\tau$ is certainly some permutation of X. Does it preserve the distance? Well, let's see.

Take $a, b \in X$ and consider

$$\begin{aligned} d(\sigma\circ\tau(a), \sigma\circ\tau(b)) &= d(\sigma(\tau(a), \sigma(\tau(b)) \\ &= d(\tau(a), \tau(b)) \\ &= d(a, b) \end{aligned}$$

Associativity: Obvious.

Neutral element: Consider the identity permutation $\imath(a) = a$. Then, as expected,

$$d(\imath(a), \imath(b)) = d(a, b)$$

Inverse: Since σ is a symmetry, there exists (a permutation) σ^{-1}. We would like to see whether the distance is preserved under the action of σ^{-1}. So, keeping in mind that σ is a symmetry, let's check

$$\begin{aligned} d(\sigma^{-1}(a), \sigma^{-1}(b)) &= d(\sigma(\ \sigma^{-1}(a), \sigma(\sigma^{-1}(b))) \\ &= d(\imath(a), \imath(b)) \\ &= d(a, b) \end{aligned}$$

So, σ^{-1} is also a symmetry. Thus the proof. ■

Example 5.76 Take the group D_4 (cf. Example 5.38) and consider the following subsets of it:

$$A = \{\rho_0,\ \rho_2,\ \rho_6, \rho_7\}, B = \{\rho_0, \rho_1, \rho_2, \rho_3\}, C = \{\rho_0, \rho_2, \rho_4, \rho_5\}, D = \{\rho_0, \rho_7\}$$
$$F = \{\rho_0, \rho_6\}, G = \{\rho_0, \rho_2\}, H = \{\rho_0, \rho_5\}, K = \{\rho_0, \rho_4\}, E = \{\rho_0\}$$

It is easy to verify that all of the above sets are subgroups of D_4. The following subgroup diagram (Figure 5.6) shows the remarkable symmetries. ■

Example 5.77 Consider $H = \{0, 3\} \subseteq (\mathbf{Z}_6; +)$ and $K = \{0, 2, 4\} \subseteq (\mathbf{Z}_6; +)$ and show that $H, K \leq (\mathbf{Z}_6; +)$.

Solution It is sufficient to inspect the Cayley table for $(\mathbf{Z}_6; +)$ and verify the claim. The elements of H are bold faced.

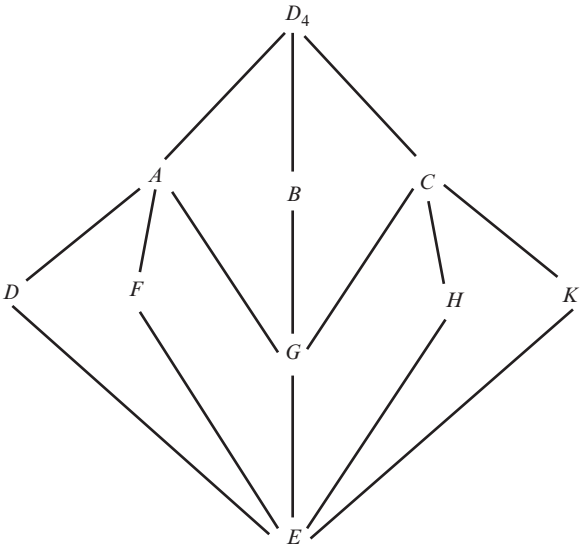


Figure 5.6 Subgroup diagram for D_4

+	0	1	2	3	4	5
0	**0**	1	2	**3**	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	**3**	4	5	**0**	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

Similarly for K:

+	0	1	2	3	4	5
0	**0**	1	**2**	3	**4**	5
1	1	2	3	4	5	0
2	**2**	3	**4**	5	**0**	1
3	3	4	5	0	1	2
4	**4**	5	**0**	1	**2**	3
5	5	0	1	2	3	4

That is, $H = \{0, 3\}$ and $K = \{0, 2, 4\}$.
We see that $H, K \leq (\mathbf{Z}_6; +)$ indeed. ■

Definition 5.15 If G is a group, then G itself is considered its own subgroup. It is obvious that $\{e\} \leq G$. We call these **trivial subgroups**. Any subgroup other than G itself and $\{e\}$ is called a **proper subgroup**.

Theorem 5.18 The union of two subgroups H and K of a group G is a subgroup, iff $H \subseteq K$ or $K \subseteq H$.

Proof Let H and K be two subgroups of a group G, such that $H \nsubseteq K$, and $K \nsubseteq H$, and let's suppose that $H \cup K \leq G$. Then there exist[4]

$$a \in H - K, \quad \text{and} \quad b \in K - H$$

On the other hand,

$$a, b \in H \cup K$$

which would imply that

$$ab \in H \cup K$$

that is,

$$ab \in H$$

But, if $ab \in H$, then

$$b = a^{-1}ab \in H$$

which is a contradiction.
Similarly,

$$ab \in K$$

implies

$$a = abb^{-1} \in K$$

a contradiction again.
We conclude that the union of two subgroups of a group G is a subgroup itself iff $H \subseteq K$ or $K \subseteq H$. ■

Theorem 5.19 The intersection of any family of subgroups of a group G is a subgroup of G.

[4]In order to avoid confusion of notation here we use "−" for the difference of sets. We will use "\" later for quotient groups.

Proof Let $\{H_i | i = 1, 2, \ldots, n\}$ be a family of subgroups of G. In that case, $e \in H_i$ for every i, that is, $e \in \cap H_i$. Suppose further that

$$a, b \in \cap H_i$$

Then, since every H_i is a subgroup of G,

$$ab^{-1} \in \cap H_i$$

Thus, $\cap H_i \leq G$, which was to be shown. ■

To summarize the conclusions of the previous two theorems: The intersection of two subgroups of a group G is a subgroup of G. However, the union of two subgroups does not necessarily have to be a subgroup of G.

Consequently, we can reason as follows: Suppose $S \subseteq G$ is some subset of a group G. Then, S is a subset of some subgroup of G (if no other but G itself). Let $H = \cap_i H_i$ be the intersection of all subgroups $H_i \leq G$ containing S. We already know from the previous theorem that H is a subgroup of G, and we say that H is a subgroup generated by S. We write $H = [S]$. Of course, if $S \leq G$ then $S = [S]$. Also, H is the smallest subgroup of G (Figure 5.7).

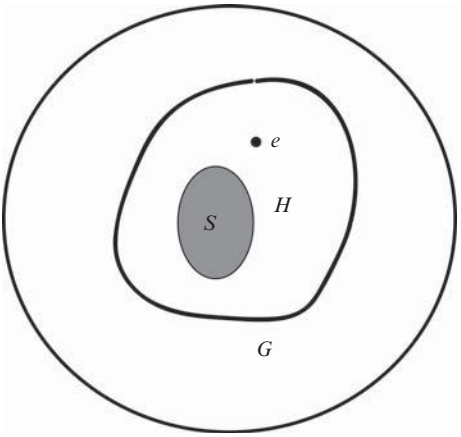


Figure 5.7

More formally, we have

Theorem 5.20 Let G be a group, and let S be a subset of G; then, there is a smallest subgroup $H \leq G$ containing S.

Proof We can always find a subgroup of G that contains S, if no other than G itself. Let H be the intersection of all subgroups of G that contain S. By the previous theorem, H itself is a subgroup of G. Thus, $S \subseteq H$. Now, suppose there is a $K \leq G$, such that $S \subseteq K$. In that case, K is one of the subgroups of G. Hence $H \leq K$, which makes it the smallest subgroup containing S. ■

Definition 5.16 If S is a subset of a group G, then the smallest subgroup of G containing S and denoted by $[X]$ is called the **subgroup generated by *S***. In particular, if $H, K \leq G$ are two subgroups of G, then the smallest subgroup containing H and K, that is, the subgroup generated by $H \cup K$, is a subgroup denoted by $[H \cup K]$ or $H \vee K$.

In case that $S \subseteq G$ consists of a single element a, that is, $S = \{a\}$, then $[S] = \langle a \rangle$, which we call a cyclic subgroup generated by a. If $[\ S] = G$, we say that S is the **set of generators** of G.

Example 5.78 The group $(\mathbf{Z}; +)$ is generated by the set $S = \{1\}$. ■

Theorem 5.21 Let G be a group, and let $S \subseteq G$ be any nonempty subset of G. Then, the subgroup generated by S is the set H of all finite products $x_1 x_2 \cdots x_n$, such that $\forall i,\ \ x_i \in S \ \text{ or } \ x_i^{-1} \in S$.

Proof First note that $S \subseteq H$.

Now, let $a, b \in H$, such that $a = x_1 x_2 \cdots x_n$ and $b = y_1 y_2 \cdots y_n$, then

$$\begin{aligned} ab^{-1} &= (x_1 x_2 \cdots x_n)(y_1 y_2 \cdots y_n)^{-1} \\ &= x_1 x_2 \cdots x_n y_n^{-1} y_{n-1}^{-1} \cdots y_1^{-1} \in H \end{aligned}$$

So $H \leq G$. Suppose there is H', another subgroup of G containing S. That implies that $\forall x \in S$, $x \in H'$. Then $x^{-1} \in H'$ too. Therefore, H' contains all finite products $x_1 x_2 \cdots x_n$, such that either $x_i \in S$ or $x_i^{-1} \in S$, $i = 1, \ \ldots \ , n$. Thus, $H \subseteq H'$. In other words, H is the smallest subgroup containing S and, therefore, the subgroup generated by S. ■

Example/Exercise 5.79 Let G be a group. For all ordered pairs $x, y \in G$, define a product called the **commutator**

$$[x, y] = xyx^{-1}y^{-1}$$

Let $[C] = \{[x, y] | x, y \in G\}$. Show that $[C]$ is a subgroup of G, called the **commutator subgroup** of G. What happens if G is an abelian group?

A few additional comments regarding symmetric groups may be in order for further examples. First, let's introduce two new concepts.

Definition 5.17 Let $p \in S_n$ be a given permutation. We say that we have an **inversion** in the permutation p if for some $i < j, p(i) > p(j)$.

Definition 5.18 Let $I(p)$ be the number of inversions of the permutation p. We call the map

$$\text{sign} : S_n \to \{1, -1\}$$

defined by $\text{sign}(p) = (-1)^{I(p)}$ the **parity** of permutation p, so that if $\text{sign}(p) = 1$, we say that the permutation p is **even**, and conversely, if $\text{sign}(p) = -1$, we say that the permutation is **odd**. For instance, for the identity permutation

$$p_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$$

$\text{sign}(p_0) = 1$, since there are no inversions, that is,

$$\text{sign}(p_0) = (-1)^{I(p_0)} = (-1)^0 = 1$$

Thus, p_0 is an even permutation. Similarly,

$$p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

is an even permutation since $I(p_1) = 2$, so $\text{sign}(p_1) = 1$, and we again have an even permutation. On the other hand, for

$$p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

$I(p_2) = 3$, thus $\text{sign}(p_2) = -1$, and p_2 is an odd permutation.

Without proof, we state the following properties of the map sign:

Theorem 5.22 For all $p, q \in S_n$,

(i) $\text{sign}(p \circ q) = \text{sign}(p) \cdot \text{sign}(q)$
(ii) $\text{sign}(p^{-1}) = \text{sign}(p)$

Now we are ready for

Theorem 5.23 Let $A_n \subseteq S_n$ be the set of all even permutations in S_n. Then, A_n is a subgroup of S_n called the **alternating group**.

Proof The proof immediately follows from the properties of the map sign and Theorem 5.22. ■

Theorem 5.24 If $n > 1$, there is the same number of even and odd permutations in S_n.

Proof Consider one permutation, say,

$$q = \begin{pmatrix} 1 & 2 & 3 \ldots & n \\ 2 & 1 & 3 & \ldots n \end{pmatrix}$$

that is, $q(1) = 2, \quad q(2) = 1$, and the rest of the elements stay unpermuted, that is, for all $i > 2, q(i) = i$. Since there is only one inversion in the permutation q, it is an odd permutation. Let $B_n \subseteq S_n$ be the set of all odd permutations in S_n, and let

$$f : A_n \to B_n$$

be the function defined by

$$f(p) = q \circ p, \quad \forall p \in A_n$$

According to the properties of the map sign, sign, $f(p)$ is an odd permutation, hence a well-defined function. We would like it to be a bijection. Well, let's see. Observe that if

$$f(p_1) = f(p_2)$$

then

$$q \circ p_1 = q \circ p_2$$

But q, p_1, p_2 are elements of the group S_n, so the LCL applies, and we have

$$p_1 = p_2$$

Thus f is an injection. Next, let's take some $r \in B_n$, and recall that q^{-1} is an odd permutation too, which makes

$$q^{-1} \circ r$$

an even permutation. Consider, now,

$$\begin{aligned} f(q^{-1}\circ r) &= q\circ(q^{-1}\circ r)\\ &= (q\circ q^{-1})\circ r\\ &= r \end{aligned}$$

Hence our function f is also surjective, which completes our proof. ■

Consequently, we have

Corollary 5.4 If $n > 1$, the alternating group A_n has $n!/2$ elements.

Finally, I would like to end this section with a subsection introducing another useful concept, the concept of

Cycles

Like many other mathematicians, after getting familiar with permutations via a two-rowed notation, you have probably felt that such a notation is rather cumbersome, especially when one is primarily interested in the group structure of permutations. Something more manageable would be welcome. So, let's briefly discuss another way of treating permutations. Consider a permutation $p \in S_n$, such that for d distinct natural numbers

$$i_1 i_2, \ldots, i_d \in \{1, 2, 3, \ldots, n\}$$

$$p(i_1) = i_2, p(i_2) = i_3, \ldots, p(i_{d-1}) = i_d, p(i_d) = i_1$$

and the rest of the elements stay unpermuted, that is, $p(j) = j$ for all $j \neq i_1, \ldots, i_d$. In other words,

$$p = \begin{pmatrix} i_1 & i_2 & \cdots & i_{d-1} & i_d & i_{d+1} & \cdots & i_n \\ i_2 & i_3 & \cdots & i_d & i_1 & i_{d+1} & \cdots & i_n \end{pmatrix}$$

We call p a **cyclic permutation** or a **cycle of length** d, or a ***d*-cycle**, and we write

$$p = (i_1, i_2, \ldots, i_d)$$

Of course, every cycle of length 1 is the identity permutation. A cycle whose length is 2, that is, a cycle that merely interchanges a pair of elements, is called a **transposition**.

More formally:

Definition 5.19 If $p \in S_n$ and $i \in \{1, 2, \ldots, n\}$, then p **fixes** i if $p(i) = i$, and p **moves** i if $p(i) \neq i$.

Definition 5.20 Let $i_1, i_2, \ldots, i_d \in \{1, 2, 3, \ldots, n\}$ be n distinct integers, and if $p \in S_n$ is such that

$$p(i_1) = i_2, p(i_2) = i_3, \ldots, p(i_{d-1}) = i_d, p(i_d) = i_1$$

and p fixes the others (if any), then we say that p is a cycle of length d, or that p is a ***d*-cycle.**

Example 5.80 A permutation from S_6

$$p = (1\ 3\ 5\ 6) = \begin{pmatrix} 1 & 3 & 5 & 6 & 2 & 4 \\ 3 & 5 & 6 & 1 & 2 & 4 \end{pmatrix}$$

which, rearranged and written equivalently and in a more common way,

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 5 & 4 & 6 & 1 \end{pmatrix}$$

is a cycle of length 4. ■

Example 5.81 A permutation from S_5

$$q = (1\ 5\ 3\ 4\ 2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 2 & 3 \end{pmatrix}$$

is a cycle of length 5. ■

Example 5.82 A permutation

$$t = (2\ 4) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 3 & 2 & 5 & 6 \end{pmatrix}$$

is an example of transposition in S_6. ■

We can now restate Definition 5.18 and say

Definition 5.21 A permutation $p \in S_n$ is even if it is the product of an even number of transpositions. Otherwise, it is odd.

Also, we can add to Theorem 5.22 the following

Lemma 5.1 If $p \in S_n$, and $t \in S_n$ is a transposition, then

$$\text{sign}(tp) = -\text{sign}(p)$$

Of course, a permutation does not have to be a cycle, but we can always write it as a product of cycles (as the following examples illustrate).

Example 5.83 Consider the permutation in S_5

$$\begin{aligned} p &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} (4)(5) \\ &= \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \end{aligned}$$

Here is what we did:

We started by noting that $1 \to 2, 2 \to 3,$ and $3 \to 1$, and we wrote this as $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$. Next, $4 \to 4,$ and $5 \to 5$, so we wrote: (4)(5). Collecting all this together we get:

$$p = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} (4)(5)$$ ■

Similarly,

Example 5.84 A permutation in S_6

$$\begin{aligned} q &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 6 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 3 & 5 & 6 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \end{aligned}$$ ■

Example 5.85 A permutation in S_9

$$\begin{aligned} r &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 6 & 4 & 7 & 2 & 5 & 1 & 8 & 9 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 6 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \begin{pmatrix} 3 & 7 & 8 & 9 \end{pmatrix} (5) \end{aligned}$$ ■

Example/Exercise 5.86 Write the following permutation

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

as a product of cycles.

This brings us to thinking about cycles in yet another way.

Remember that multiplication in S_n is a composition of functions, so for any $i \in \{1, \ldots, 9\}$ in the previous example

$$r(i) = p\{q[s(i)]\}$$

where $p = \begin{pmatrix} 1 & 6 \end{pmatrix}$, $q = \begin{pmatrix} 2 & 4 \end{pmatrix}$, and $s = \begin{pmatrix} 3 & 7 & 8 & 9 \end{pmatrix}$. We ignore the 1-cycle (5) since it is just the identity function. Thus for, say, $i = 1$ we have

$$\begin{aligned} r(1) &= p\{q[s(1)]\} \\ &= p[q(1)] \quad \text{since } s \text{ fixes 1} \\ &= p(1) \quad \text{since } q \text{ fixes 1} \\ &= 6 \end{aligned}$$

Similarly,

$$\begin{aligned} r(2) &= p\{q[s(2)]\} \\ &= p[q(2)] \quad \text{Why?} \\ &= p(4) \quad \text{Why?} \\ &= 4 \quad \text{Why?} \end{aligned}$$

and so on. ■

We need a few more things before we prove these statements in general.

Definition 5.22 Two cycles are said to be **disjoint** if they have no elements in common. For example, $\begin{pmatrix} 1 & 3 & 2 \end{pmatrix}$ and $\begin{pmatrix} 4 & 5 & 6 \end{pmatrix}$ are disjoint cycles, whereas $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$ and $\begin{pmatrix} 4 & 5 & 3 \end{pmatrix}$ are not.

Hence,

Definition 5.23 Two permutations $p, r \in S_n$ are **disjoint** if every i moved by p is fixed by r.

Theorem 5.25 Disjoint cycles/permutations commute.

Proof Intuitively this is clear. If $p, r \in S_n$, such that p moves i's but not j's, and vice versa, r moves j's but not i's, then it is reasonable to expect that $pr = rp$. Suppose

$$p = (i_1 \ \ldots \ i_k) \text{ and } r = (j_1 \ \ldots \ j_l), \qquad i, j \in \{1, 2, \ldots, n\}$$

Then,

$$\begin{aligned} pr &= (i_1 \ \dots \ i_k)\,(j_1 \ \dots \ j_l) \\ &= (j_1 \ \dots \ j_l)\,(i_1 \ \dots \ i_k) \\ &= rp \end{aligned}$$

■

Example 5.87 Consider $p, r \in S_6$ mentioned in Definition 5.22:

$$p = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 5 & 6 \end{pmatrix}$$

$$r = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 6 & 4 \end{pmatrix}$$

p and r are obviously disjoint. Let's see whether they commute.

$$\begin{aligned} pr &= \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 6 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 5 & 6 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 5 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \\ &= rp \end{aligned}$$

■

Theorem 5.26 Every permutation $p \in S_n$ is either a cycle or a product of disjoint cycles.

Proof We prove this by induction on a number $k \geq 0$ of elements moved by p. Obviously, if $n = 1$, we are dealing with S_1 and the only permutation is $p = (1)$, so our claim is evidently true. As always, assuming that the claim is true for all $n < k$, let's prove the case for $n = k$.

Let $p \in S_k$ be any permutation. Consider the following sequence of natural numbers

$$1, p(1), p^2(1), \ \dots \ , p^k(1) \ \in \{\ 1, 2, \ \dots \ , k\}$$

Due to the fact that there are only k elements in the set above, there exist two numbers r and s, such that

$$p^r(1) = p^s(1)$$

where $0 \leq r < s \leq k$, and we take $p^0(1) = 1$. Now, p, being a bijection, has an inverse, so we can write

$$p^{s-r}(1) = 1$$

Suppose $d \in \mathbf{N}$, $1 \leq d \leq k$, is the smallest number such that $p^d(1) = 1$, and let's construct a sequence of different numbers

$$i_1 = 1, i_2 = p(1), i_3 = p^2(1), \ \ldots \ , i_d = p^{d-1}(1)$$

Observe that

$$\begin{aligned} p(i_1) &= p(1) = i_2 \\ p(i_2) &= p^2(1) = i_3 \\ p(i_3) &= p^3(1) = i_4 \\ &\vdots \\ p(i_{d-1}) &= p^{d-1}(1) = i_d \\ p(i_d) &= p^d(1) = 1 = i_1 \end{aligned}$$

We obtain a cycle

$$q_1 = (\, i_1 \ i_2 \ \ldots \ i_d \,) \in S_k$$

So that the permutation p can be expressed as

$$p = q_1 \circ q_2$$

where q_2 is a permutation of the remaining $k - d$ numbers. Since, by our inductive hypothesis, q_2 is expressible as a product of disjoint cycles of $k - d < k$ numbers, we have our proof. ■

Theorem 5.27 Every permutation in $S_n, n > 1$, is the product of 2-cycles.

Proof We know from the previous theorem that every permutation can be written as

$$(i_1 i_2 \cdots i_a)(j_1 \, j_2 \cdots j_b) \cdots (k_1 \, k_2 \cdots k_c)$$

A direct computation shows that this is the same as

$$(i_1 \ i_a)(i_1 \ i_{a-1}) \cdots (i_1 \ i_2)(j_1 \, j_b)(j_1 \, j_{b-1}) \cdots (j_1 \, j_2) \cdots (k_1 \, k_c)(k_1 \, k_{c-1}) \cdots (k_1 \, k_2)$$

Thus the proof. ■

Example 5.88 If $p \in S_5$ is $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix}$, then we can write

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} \quad \blacksquare$$

Example/Exercise 5.89 Convince yourself that the identity permutation in S_n can be expressed as $\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix}$.

Example/Exercise 5.90 Convince yourself that

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \end{pmatrix} &= \begin{pmatrix} 5 & 4 \end{pmatrix} \begin{pmatrix} 5 & 3 \end{pmatrix} \begin{pmatrix} 5 & 2 \end{pmatrix} \begin{pmatrix} 5 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 5 & 4 \end{pmatrix} \begin{pmatrix} 5 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 5 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \end{pmatrix} \end{aligned}$$

Definition 5.24 A **complete factorization** of a permutation p is a factorization into disjoint cycles that contains one 1-cycle (i) for every i fixed by p.

Example 5.91 Let

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix}$$

then we can write

$$p = (1) \begin{pmatrix} 2 & 3 & 4 \end{pmatrix} (5)$$

which is a complete factorization. The factorization of the same p written as

$$p = \begin{pmatrix} 2 & 3 & 4 \end{pmatrix} = (1) \begin{pmatrix} 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 4 \end{pmatrix} (5)$$

is not considered a complete factorization. $\blacksquare$

Definition 5.25 Let $p \in S_n$ be a permutation, and let $p = q_1 q_2 \cdots q_r$ be a complete factorization of p into disjoint cycles, then

$$\text{sign}(p) = (-1)^{n-r}$$

Consequently, the theorems below are obviously true.

Theorem 5.28 If $d \geq 2$, every d-cycle can be written as a product of $d - 1$ transpositions.

Theorem 5.29 Every permutation $p \in S_n$ is a product of transpositions.

And finally, let's prove.

Theorem 5.30 A permutation $p \in S_n$ is odd iff it is a product of an odd number of transpositions t.

Proof Recall that $\text{sign}(t) = -1$, since a transposition $t \in S_n$ moves two numbers, say, i and j and keeps the rest of $n-2$ fixed. Indeed, according to Definition 5.25,

$$\begin{aligned}\text{sign}(t) &= (-1)^{n-r} \\ &= (-1)^{n-(n-1)} \\ &= -1\end{aligned}$$

Now, if p is odd, then it cannot have a factorization into an even number of transpositions; thus, it has an even number of transpositions. On the other hand, if $p = t_1 t_2 \cdots t_r$, and r is odd, then

$$\begin{aligned}\text{sign}(p) &= \text{sign}(t_1 \ \ldots \ t_r) \\ &= \text{sign}(t_1) \ \ldots \ \text{sign}(t_r) \\ &= (-1)^r \\ &= (-1)\end{aligned}$$

Thus, p is an odd permutation. ■

Theorem 5.31 The inverse of a cycle $(i_1 \ i_2 \cdots i_r)$ is a cycle $(i_r i_{r-1} \ \cdots i_1)$, that is,

$$(i_1 i_2 \cdots i_r)^{-1} = (i_r i_{r-1} \ \cdots i_1)$$

Proof If $p = (i_1 i_2 \cdots i_r) \in S_n$ and $p^{-1} = (i_r i_{r-1} \ \cdots i_1) \in S_n$, we expect that

$$\begin{aligned}pp^{-1} &= (i_1 i_2 \cdots i_r)(i_1 i_2 \cdots i_r)^{-1} \\ &= (i_1 i_2 \cdots i_r)(i_r i_{r-1} \ \cdots i_1) \\ &= (i_r i_{r-1} \ \cdots i_1)(i_1 i_2 \cdots i_r) \\ &= (i_1 i_2 \cdots i_r)^{-1}(i_1 i_2 \cdots i_r) \\ &= p^{-1}p = p^0 = (1)\end{aligned}$$

Well, let's see. First, note that the product $(i_1 i_2 \cdots i_r)(i_r i_{r-1} \ \cdots i_1)$ fixes each number between 1 and n, other than $i_1 i_2 \cdots i_r$. Second, the product sends $i_1 \to i_r \to i_1$. Also, for every $j \geq 2$, the composition acts on i_j in the following way: $i_j \to i_{j-1} \to i_j$.

In other words, each number between 1 and n is fixed by the composition and hence is equal to (1) Thus the proof. ■

Theorem 5.32 If $p = (i_1 \ \dots \ i_r) \in S_n$, then $p^{-1} = (i_1 \ \dots \ i_r)^{-1} = (i_r^{-1} \ \dots \ i_1^{-1})$.

Proof We will prove this by induction on $r \geq 2$. The claim for $r = 2$ is evidently true:

$$\begin{aligned} pp^{-1} &= (i_1 i_2)(i_2^{-1} i_1^{-1}) = i_1(i_2 i_2^{-1}) i_1^{-1} \\ &= i_1 i_1^{-1} = (1) \end{aligned}$$

So, $p^{-1} = (i_1 i_2)^{-1} = (p_2^{-1} p_1^{-1})$. We could get the same result by evaluating the product $p^{-1}p$.

Assuming that the claim is true for $r = k$, we check the assertion for $r = k + 1$:

$$\begin{aligned} (i_1 \cdots i_k i_{k+1})^{-1} &= (I_k i_{k+1})^{-1} \\ &= (i_{k+1}^{-1} I_k^{-1}) \\ &= (i_{k+1}^{-1} i_k^{-1} \cdots \ i_1^{-1}) \end{aligned}$$

where, of course, we substituted I_k for $(i_1 \ \dots \ i_k)$. Thus, our claim is valid for all $r \geq 2$. ■

Example/Exercise 5.92 You should convince yourself that, for instance, $\begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & 3 & 2 & 1 \end{pmatrix}$.

Finally, let's just mention an expected consequence of the previous theorem, namely, the case of the inverse of a disjoint cycle. Here is an illustrative.

Example 5.93 Consider the following permutation:

$$p = \begin{pmatrix} 1 & 6 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \begin{pmatrix} 3 & 7 & 8 & 9 \end{pmatrix} (5)$$

then,

$$p^{-1} = (5) \begin{pmatrix} 9 & 8 & 7 & 3 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \begin{pmatrix} 6 & 1 \end{pmatrix}$$

However, since disjoint cycles commute (see Theorem 5.25), the reversal of the order of factors is unnecessary, so

$$p^{-1} = (5) \begin{pmatrix} 9 & 8 & 7 & 3 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \begin{pmatrix} 6 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 6 \end{pmatrix} \begin{pmatrix} 2 & 4 \end{pmatrix} \begin{pmatrix} 3 & 7 & 8 & 9 \end{pmatrix} \ ■$$

5.4 CYCLIC GROUPS

Definition 5.26 A group G is called a **cyclic group** if it can be generated by a single element, *the generator of* G, that is there is an element $a \in G$, such that $G = \{a^n | n \in \mathbf{Z}\}$. We denote the cyclic group $G = \langle a \rangle$.

Definition 5.27 If $G = \langle a \rangle$ is a cyclic group, then the order of a, $|\langle a \rangle|$, is the number of elements in $\langle a \rangle$.

Example 5.94 Consider an abstract group $G = \{e, a, b\}$ of order 3, with $a \neq b \neq e$. The product ab cannot be equal to a or b since this would imply that $b = e$ or $a = e$. Thus, in order to satisfy closure, $ab = e$. What about a^2? Well, the only possibility is that $a^2 = b$. Now, what about b^2? Let's see:

$$b^2 = bb = a^2a^2 = a^4 = a^3a = ea = a$$

So, the multiplication table for G is

$\cdot$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

This is an example of a cyclic group that consists of powers of a single element, that is,

$$G = \{a, a^2, a^3 = e\} = \langle a \rangle$$

Recall the transformations of an equilateral triangle $\triangle ABC$ from Example 5.36, and consider only rotations in a plane that brings the triangle into coincidence with itself. These transformations are indeed the realization of our cyclic group G. You can easily convince yourself that the abstract multiplication table for G indeed represents the group of rotations of an equilateral triangle.

Note that the table is symmetric about the main diagonal, which is the case for abelian groups. Finally, observe that the (3×3) part of the table in Example 5.36/37 is exactly the table in Example 5.94 (of course, $\alpha = a, \beta = b$), hence $G \leq D_3$ as we expected. ■

Example 5.95 Consider the abstract group $G = \{e, a, b, c\}$ of order 4, with the following multiplication table:

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

From the table, we see that

$$\begin{aligned} a^2 &= b \\ ab &= aa^2 = a^3 \\ &= c \\ a^4 &= a^2a^2 \\ &= bb = b^2 \\ &= e \end{aligned}$$

that is, our group $G = \{e, a, a^2, a^3\} = \langle a \rangle$.

Geometrically, this group can be interpreted as a group of rotations of a square around its center which brings it into coincidence with itself (cf. Example 5.37). We could have designed our multiplication differently:

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	e	b	a	e

And we have

$$\begin{aligned} a^2 = b^2 = c^2 &= e \\ ab &= c \\ ac &= b \\ bc &= a \end{aligned}$$

We recognize this group as the Klein four-group V. (cf. Example 5.25). From the symmetry of tables, we see that both groups are abelian. ■

Example/Exercise 5.96 Show that the groups in Exercise 5.95 are the only possible groups of order 4.

Example/Exercise 5.97 Show that a group of order 4 must be abelian.

If you revisit Example 5.66, the following should be quite familiar to you.

Example 5.98 Let $G = (\mathbf{Z}; +)$, then $G = \langle 1 \rangle$. Since the binary operation in G is "+," every element of $\langle 1 \rangle$ can be written in the form $n \cdot 1$ for some $n \in \mathbf{Z}$. Another way to look at this is to interpret 1^n.

$$\underbrace{1 + 1 + \cdots + 1}_{n \text{ terms}}$$

Of course, $|\langle 1 \rangle| = |1| = \infty$. ■

Example/Exercise 5.99 Convince yourself that the same group, $(\mathbf{Z}; +)$, could also have been written as $\langle -1 \rangle$.

Hint: If n is negative, we consider (-1) as the generator and we have

$$\underbrace{(-1) + (-1) + \cdots + (-1)}_{|n| \text{terms}}$$

Example 5.100 Let $\langle a \rangle$ be a cyclic group of order n. From Theorem 5.13, we know that

$$\langle a \rangle = \{e, a, a^2, \ \ldots \ , a^{n-1}\}$$

Thus, there is the same number of elements in $\langle a \rangle$ as in $\mathbf{Z}_n$, that is, there is an obvious bijection between a and $\mathbf{Z}_n$. Indeed, observe the following "mapping":

$$\begin{array}{rcccccccc} \langle a \rangle = \{e, & a, & a^2, & \ldots, & a^{n-1}\} \\ & \updownarrow & \updownarrow & \updownarrow & & \updownarrow \\ \mathbf{Z}_n = \{0, & 1, & 2, & \ldots, & n-1\} \end{array}$$

In view of this and the two previous Example/Exercises 5.98, 5.99, it is natural to contemplate the case where $|\langle a \rangle| = \infty$ and consider the correspondence

$$\begin{array}{rccccccc} \langle a \rangle = \{\ldots, & a^{-2}, & a^{-1}, & e, & a, & a^2, \ldots\} \\ & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ \mathbf{Z} = \{\ldots, & -2, & -1, & 0, & 1, & 2, \ldots\} \end{array}$$

which is obviously a bijection $\mathbf{Z} \to \langle a \rangle$ defined by $f(n) = a^n$. ■

5.5 HOMOMORPHISMS AND ISOMORPHISMS

As you might have anticipated by now, regardless of their general elegance and apparent "simplicity," groups and group structures can be very complicated. Also, once groups are discovered, one recognizes groups and group structures in many different fields of mathematics, physics, and other sciences. It is not surprising then that a mathematician searches for tools that will hopefully enable us that, instead of studying the "original" group G, to look for a less complicated structure H, which faithfully represents a "complicated" G. That's why we invoke the concepts and techniques of functions, and look for a map $G \xrightarrow{\varphi} H$, which will "bring" us to a more manageable H, while still preserving the structure of G. So we have

Definition 5.28 Let $(G; *)$ and $(H; \circ)$ be groups. A map $\varphi : G \to H$ is called a **homomorphism** if for all $a, b \in G$

$$\varphi(a * b) = \varphi(a) \circ \varphi(b)$$

We also say that H is a **homomorphic image** of G (Figure 5.8).

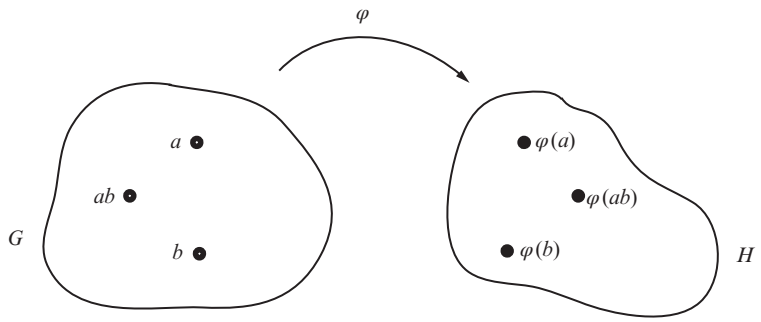


Figure 5.8

Note that homomorphisms preserve group operations, that is, the product on the left is computed in G and the product on the right is computed in H. When group operations for G and H are not explicitly given, we simply write

$$\varphi(ab) = \varphi(a)\varphi(b)$$

keeping in mind that φ respects group structure in both G and H. The set of all homomorphisms $G \to H$ is denoted by $\mathbf{Hom}(\boldsymbol{G}, \boldsymbol{H})$.

Example 5.101 Recall the parity group P from Example 5.27, and consider a function $\varphi : (\mathbf{Z}; +) \to P$, which carries every even integer to $E \in P$ and every

odd integer to $O \in P$. It is easy to see that φ is a homomorphism. Indeed, let $x, y \in \mathbf{Z}$ be any two integers. Then, either both x and y are even or both are odd, or one is even and the other is odd. In the case where both are odd or even, their sum is even, thus

$$\varphi(x + y) = \varphi(x) + \varphi(y) = O + O = E$$

or

$$\varphi(x + y) = \varphi(x) + \varphi(y) = E + E = E$$

Similarly, if, say, x is even and y is odd, their sum is odd and we have

$$\varphi(x + y) = \varphi(x) + \varphi(y) = E + O = O$$

Hence, φ is a homomorphism. ■

Example 5.102 Let G and H be two groups, such that $e' \in H$ is the identity of H. Define a map

$$\varphi \ : G \to H$$

by $\varphi(a) = e', \ \forall a \in G$.

It's easy to see that φ is a homomorphism. ■

Example 5.103 Consider a group $(\mathbf{R}; +)$ and a group of all positive real numbers $(\mathbf{R}^+; \cdot)$ with multiplication as the binary operation. Then the following function

$$\varphi \ : (\mathbf{R}; +) \to (\mathbf{R}^+; \cdot)$$

defined by $\varphi(x) = e^x$ is a homomorphism. Indeed, let $x, y \in \mathbf{R}$, then

$$\varphi(x + y) = e^{x+y} = e^x e^y = \varphi(x)\varphi(y), \forall x, y \in \mathbf{R}$$ ■

Example/Exercise 5.104 Show that a function $\varphi \ : (\mathbf{R} \times \mathbf{R}; +) \to (\mathbf{R}; +)$ given by $\varphi(x, y) = x + y$ is a homomorphism.

Example 5.105 Let $\varphi : (\mathbf{R}^*; \cdot) \to (\mathbf{R}^*; \cdot)$ be a mapping defined by $\varphi(x) = x^2$. Then, φ is a homomorphism of the group $(\mathbf{R}^*; \cdot)$ to itself. Indeed, for every $x, y \in (\mathbf{R}^*; \cdot)$

$$\varphi(xy) = (xy)^2 = x^2y^2 = \varphi(x)\varphi(y)$$ ■

Example 5.106 A function $\varphi \ : (\mathbf{R}; +) \to (\mathbf{R}; +)$ given by $\varphi(x) = x^2$ *is not* a homomorphism, since

$$\varphi(x + y) = (x + y)^2 \neq \varphi(x) + \varphi(y) = x^2 + y^2$$ ■

Theorem 5.33 Let G and H be two groups with respective identities e and e', and let $\varphi \ : G \to H$ be a homomorphism. Then,

(i) $\varphi(e) = e'$
(ii) $\varphi(x^{-1}) = (\varphi(x))^{-1} \quad \forall x \in G$

Proof

(i) Let $x \in G$ be any element of G. Then,

$$\begin{aligned} e'\varphi(x) &= \varphi(x) = \varphi(ex) \\ &= \varphi(e)\ \varphi(x) \in H \end{aligned}$$

Since H is a group, by the cancellation law

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

(ii) From (i)

$$\begin{aligned} e' = \varphi(e) &= \varphi(xx^{-1}) \\ &= \varphi(x)\varphi(x^{-1}) \end{aligned}$$

Multiplying by $(\varphi(x))^{-1}$ from the left we get

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

and therefore

$$(\varphi(x))^{-1} = \ \varphi(x^{-1})$$ ■

Example/Exercise 5.107 Show that, if $\varphi : G \to H$ is a homomorphism, then $\varphi(x^n) = (\varphi(x))^n$, $\forall n \in \mathbf{Z}$. (Hint: Use induction.)

Definition 5.29 A homomorphism φ, which is also surjective, is called an **epimorphism**.

Definition 5.30 A homomorphism of G into itself is called an **endomorphism**. The set of all endomorphisms of a group G is denoted by **End** $(\boldsymbol{G})$.

Definition 5.31 An endomorphism that is also one-to-one and onto is called an **automorphism**. The set of all automorphisms of a group G is denoted by **Aut** $(\boldsymbol{G})$.

Definition 5.32 A homomorphism φ that is also an injection is called a **monomorphism**.

Example 5.108 For any group G, G is automorphic to itself. Indeed, the identity map $i : G \to G$ is an obvious automorphism of G. ■

Definition 5.33 Let $a \in G$ be some element of a group G. A function $\varphi_a : G \to G$ defined by

$$\varphi_a(x) = axa^{-1}, \quad \forall x \in G$$

is called the **inner automorphism** of G. The set of all inner automorphisms of a group G is denoted by $\mathbf{Int}(\boldsymbol{G})$.

Example/Exercise 5.109 Show that $(\mathrm{Int}(G); \circ)$ is a group.

Definition 5.34 A homomorphism $\varphi : G \to H$ is called an **isomorphism** if φ is also a bijection. We say that G is isomorphic to H, and we write $G \cong H$.

This tells us that two isomorphic groups, with possibly different elements and different operations, are essentially (fundamentally) the same group. In other words, any property that G has, H has as well (see Figure 5.9).

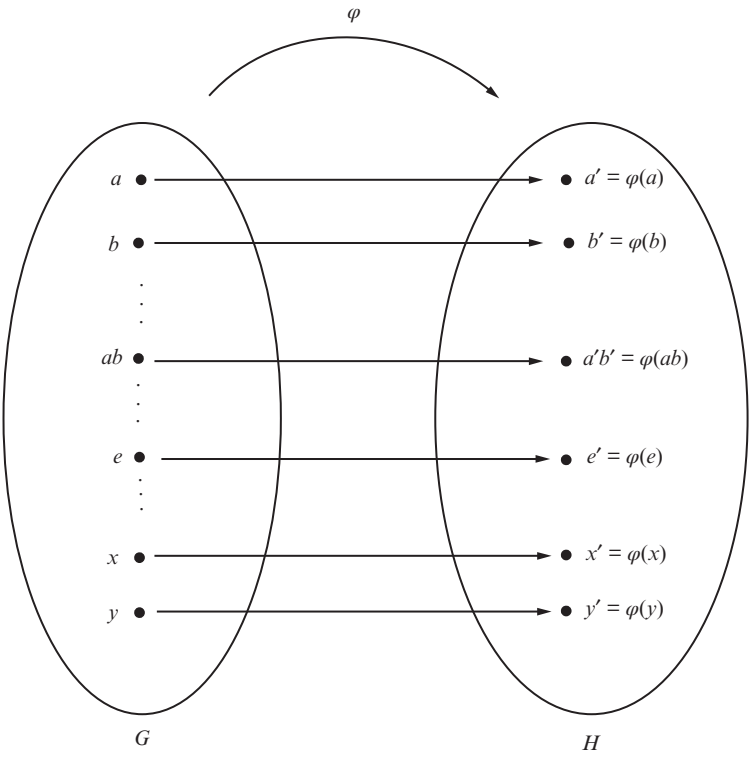


Figure 5.9 Isomorphism $\varphi : G \to H$

Example 5.110 Let G and H be two groups described by the following multiplication tables:

G

$*$	1	1	2
0	0	1	2
1	1	2	0
2	2	0	1

H

$\circ$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Whatever the objects $0, 1, 2$ and e, a, b, and the respective binary operation "$*$" and "$\circ$" might be, the one-to-one correspondence

$$\varphi : 0 \to e, 1 \to a, 2 \to b$$

is evidently an isomorphism. ■

Example 5.111 Let G be a group. Given some $a \in G$, define a map

$$\theta_a : G \to G$$

by $\theta_a(x) = axa^{-1}, \ \forall x \in G$

Show that θ_a is an automorphism.

Solution Consider

$$\begin{aligned}\theta_a(xy) &= axya^{-1} \\ &= (axa^{-1})(aya^{-1}) \\ &= \theta_a(x)\theta_a(y)\end{aligned}$$

Thus, θ_a is a homomorphism.

Next we need to check whether θ_a is injective and surjective. Consider

$$\theta_a(x) = \theta_a(y)$$

Then,

$$axa^{-1} = aya^{-1}$$

And, by applying cancellation laws, we get

$$x = y$$

Thus, θ_a is injective.

Finally, take any $x \in G$ and consider

$$\begin{aligned} x &= (aa^{-1})x(aa^{-1}) \\ &= a(a^{-1}xa)a^{-1} \\ &= \theta_a(a^{-1}xa) \end{aligned}$$

Hence, θ_a is surjective too. Consequently, θ_a is an automorphism of G. ■

Example 5.112 Show that a map $\exp : \mathbf{R} \to \mathbf{R}^+$ defined by $\exp(x) = e^x$ is an isomorphism from $(\mathbf{R}\,; +)$ to $(\mathbf{R}^+; \cdot)$ (cf. Example 5.103).

Solution We have seen that exp preserves group operations since $e^{x+y} = e^x e^y$, although in this example the groups differ in both elements and operations. Furthermore, exp has an inverse function $\ln x$, thus exp is a bijection. So exp is indeed an isomorphism from $(\mathbf{R}\,; +)$ to $(\mathbf{R}^+; \cdot)$. ■

Example 5.113 Show that groups $(\boldsymbol{Z}; +)$ and $(\boldsymbol{Q}; +)$ are not isomorphic.

Solution Suppose the contrary, that is, suppose $(\mathbf{Z}; +) \cong (\mathbf{Q}; +)$. Then, there exists an isomorphism

$$\varphi \ : \mathbf{Z} \to \mathbf{Q}$$

such that for some $q \in \mathbf{Q}$, $\varphi(1) = q$. This q has to be different from zero, since $\varphi(0) = 0$. But then $q/2 \in \mathbf{Q}$, and of course, $q/2 \neq 0$ too. Next, being an isomorphism, φ is onto, thus there exists an integer $n \neq 0$, such that $\varphi(n) = q/2$. Therefore, we have

$$\begin{aligned} \varphi(2n) &= \varphi(n + n) \\ &= \varphi(n) + \varphi(n) \\ &= \frac{q}{2} + \frac{q}{2} \\ &= q \end{aligned}$$

But φ is also one-to-one, thus

$$\begin{aligned} \varphi(1) &= q \\ &= \varphi(2n) \end{aligned}$$

implies $1 = 2n$, and therefore

$$n = \frac{1}{2} \notin \mathbf{Z}$$

which is a contradiction. We conclude that $(\mathbf{Z};+) \ncong (\mathbf{Q};+)$. ■

Example/Exercise 5.114 Prove that $(2\mathbf{Z};+) \cong (\mathbf{Z};+)$.

We have already introduced in a different context the following two definitions.

Definition 5.35 We say that the **kernel** of a homomorphism φ from a group G to a group H is the set (see Figure 5.10)

$$\mathrm{Ker}\varphi = \{x \in G | \varphi(x) = e' \in H\}$$

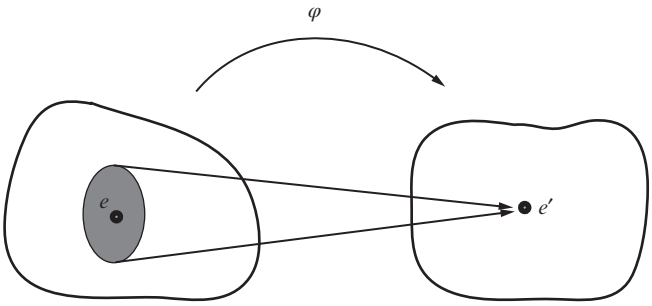


Figure 5.10 $\mathrm{Ker}(\varphi)$.

Definition 5.36 We say that the **image** of a homomorphism φ from a group G to a group H is the set (see Figure 5.11)

$$\mathrm{Im}(\varphi) = \{\varphi(x) | x \in G\} \subseteq H$$

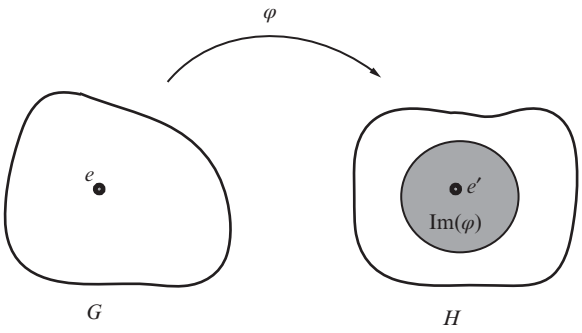


Figure 5.11 $\mathrm{Im}(\varphi)$.

Theorem 5.34 Let $\varphi : G \to H$ be a homomorphism of groups. Then, $\operatorname{Ker}\varphi$ is a subgroup of G and $\operatorname{Im}\varphi$ is a subgroup of H.

Proof First note that $\operatorname{Ker}\varphi$ and $\operatorname{Im}\varphi$ are both nonempty, since at least

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

Now, suppose $a, b \in \operatorname{Ker}\varphi$, then

$$\varphi(a) = e' = \varphi(b)$$

and so

$$\begin{aligned}\varphi(ab^{-1}) &= \varphi(a)\varphi(b^{-1}) \\ &= \varphi(a)\varphi(b)^{-1} \\ &= e'e' \\ &= e'\end{aligned}$$

Hence, $ab^{-1} \in \operatorname{Ker}\varphi$, and, therefore, $\operatorname{Ker}\varphi \leq G$.

Next, take $x, y \in \operatorname{Im}\varphi$, such that $x = \varphi(a)$ and $y = \varphi(b)$ for some $a, b \in G$. But then

$$\begin{aligned}xy^{-1} &= \varphi(a)\varphi(b)^{-1} \\ &= \varphi(ab^{-1}) \in \operatorname{Im}\varphi\end{aligned}$$

which shows that $\operatorname{Im}\varphi \leq H$. ■

In particular, if we consider φ acting only on the subgroup $A \leq G$, we have

Theorem 5.35 Let $\varphi : G \to H$ be a homomorphism of groups, and let $A \leq G$ be a subgroup of G. Then,

$$\operatorname{Im}(\varphi(A)) \leq H$$

Proof Let $a, b \in \varphi(A)$. We would like to show that $ab^{-1} \in \varphi(A)$ too. Well, since $a, b \in \varphi(A)$ there exist $x, y \in A$, such that $\varphi(x) = a$ and $\varphi(y) = b$. From the previous theorem, it follows that

$$\begin{aligned}ab^{-1} &= \varphi(x)[\varphi(y)]^{-1} \\ &= \varphi(x)\varphi(y^{-1}) = \varphi(xy^{-1})\end{aligned}$$

Since $x, y \in A \leq G$, $xy^{-1} \in A$ too. Therefore, $\varphi(xy^{-1}) \in \varphi(G)$, and thus $ab^{-1} \in \varphi(A)$, that is, $\mathrm{Im}(\varphi(A)) \leq H$, as claimed. ■

Sure enough the following is also true.

Theorem 5.36 Let $\varphi : G \to H$ be a group homomorphism, and let $K \leq H$ be some subgroup of H. Then, the *preimage* $\varphi^{-1}(K) \subseteq G$ is a subgroup of G, that is, $\varphi^{-1}(K) \leq G$.

Proof Let $x, y \in \varphi^{-1}(K)$, then $\varphi(x), \varphi(y) \in K$. But, K is a subgroup and therefore, $\varphi(x)(\varphi(y))^{-1} \in K$. Then, since φ is a homomorphism,

$$\varphi(x)(\varphi(y))^{-1} = \varphi(xy^{-1}) \in K$$

Thus, $xy^{-1} \in \varphi^{-1}(K)$, and so $\varphi^{-1}(K) \leq G$. ■

As a simple exercise you can now prove

Theorem 5.37 Let $\varphi : G \to H$ be a group homomorphism, and let A be an abelian subgroup of G. Then $\mathrm{Im}(\varphi(A))$ is an abelian subgroup of H.

Theorem 5.38 Let G, H, and K be groups, and let $\varphi : G \to H$ and $\theta : H \to K$ be homomorphisms. Then, the composition $\psi = (\theta \circ \varphi)$ is also a homomorphism.

Proof Take $x, y \in G$ and consider

$$\begin{aligned}
\psi(xy) &= (\theta \circ \varphi)(xy) = \theta(\varphi(xy)) \\
&= \theta(\varphi(x)\varphi(y)) \\
&= \theta(\varphi(x))\theta(\varphi(y)) \\
&= (\theta \circ \varphi)(x)(\theta \circ \varphi)(y) \\
&= \psi(x)\psi(y)
\end{aligned}$$

■

Theorem 5.39 Let $\varphi : G \to H$ be a homomorphism of groups. Then,

(i) φ is an epimorphism iff $\mathrm{Im}(\varphi) = H$.
(ii) φ is a monomorphism iff $\mathrm{Ker}(\varphi) = \{e\}$.

Proof Statement (i) is obvious, so let's prove (ii).

Suppose, first, that φ is a monomorphism. From the previous theorem, we know that $\varphi(e) = e' \in H$, therefore,

$$e = \varphi^{-1}(e') \in \text{Ker}(\varphi)$$

But φ is an injection by assumption, thus $\text{Ker}(\varphi)$ contains only one element, that is, $\text{Ker}(\varphi) = \{e\}$.

Conversely, suppose $(\varphi) = \{e\}$. We need to prove that φ is an injection. Well, if we choose $x, y \in G$ such that $\varphi(x) = \varphi(y)$ then, since φ is a homomorphism,

$$\begin{aligned} \varphi(x)(\varphi(y))^{-1} &= e' \\ &= \varphi(xy^{-1}) \end{aligned}$$

Hence, $xy^{-1} \in \text{Ker}(\varphi) = \{e\}$, that is

$$xy^{-1} = e$$

and therefore

$$x = y$$

Thus, φ is a monomorphism, and our proof is complete. ■

Recall that a homomorphism that is also a bijection is called an isomorphism. Consequently, we have

Theorem 5.40 Let G and H be two groups. Then, the homomorphism $\varphi : G \to H$ is an isomorphism iff φ is a surjection and $\text{Ker}(\varphi) = \{e\}$.

Theorem 5.41 Let $\varphi : G \to H$ be a group isomorphism. Then, $\varphi^{-1} : H \to G$ is also an isomorphism.

Proof First, note that φ^{-1} exists and is a bijection since φ is a bijection. We need to show that it is a homomorphism.

Let $x, y \in H$ be any two elements. Then, there exist $a, b \in G$ such that

$$\varphi(a) = x \text{ and } \varphi(b) = y$$

Consider now

$$\begin{aligned}\varphi^{-1}(xy) &= \varphi^{-1}(\varphi(a)\varphi(b)) \\ &= \varphi^{-1}(\varphi(ab)) \\ &= ab \\ &= \varphi^{-1}(x)\varphi^{-1}(y)\end{aligned}$$

which proves that φ^{-1} is indeed an isomorphism. ■

Definition 5.37 Let G be a group and $g \in G$. We say that

$$\gamma_g : G \to G$$

is a **conjugation** if for all $a \in G$, $\gamma_g(a) = gag^{-1}$.

Consequently, we have

Theorem 5.42 If G is a group, then the conjugation $\gamma_g : G \to G$ is an isomorphism.

Proof Let $g, h \in G$ be two elements of G, then

$$\begin{aligned}(\gamma_g \circ \gamma_h)(a) &= \gamma_g(hah^{-1}) \\ &= g(hah^{-1})g^{-1} \\ &= (gh)a(h^{-1}g^{-1}) \\ &= (gh)a(gh)^{-1} \\ &= \gamma_{gh}(a)\end{aligned}$$

Now, take $a, b \in G$ and consider

$$\begin{aligned}\gamma_g(ab) &= g(ab)g^{-1} \\ &= (gag^{-1})(gbg^{-1}) \\ &= \gamma_g(a)\gamma_g(b)\end{aligned}$$

Finally, observe that γ_g is also a bijection, that is,

$$\begin{aligned}\gamma_g \circ \gamma_{g^{-1}} &= \gamma_{gg^{-1}} \\ &= \gamma_e = e \\ &= \gamma_{g^{-1}g} = \gamma_{g^{-1}} \circ \gamma_g\end{aligned}$$

Thus, γ_g is an isomorphism. ■

Again, recalling that the composition of bijections is also a bijection, you should be able to prove the next theorem without difficulty.

Theorem 5.43 If $\varphi : G \to H$ and $\theta : H \to K$ are two group isomorphisms, then the composition $\psi = \theta \circ \varphi$ is an isomorphism too.

Although the following examples may look obvious to you, you should convince yourself that the claims therein are indeed true.

Example 5.115 All trivial groups of first order are (obviously) isomorphic. ■

Example 5.116 All groups of second order are isomorphic. Indeed, let $G = \{e, a\}$ be a group, where e is a neutral element. Then, of course,

$$e \cdot e = e \text{ and } a \cdot e = e \cdot a = e$$

We could have considered the possibility that $a \cdot a = a$, but that doesn't hold, since

$$a \cdot a = a$$

and

$$a \cdot e = a$$

would imply that $a = e$, and G would have only one element, and the only possible multiplication table is

$\cdot$	e	a
e	e	a
a	a	e

■

Example 5.117 All groups of third order are isomorphic. Well, let's consider designing a multiplication table for $G = \{e, a, b\}$. (cf. Examples 5.94 and 5.110)

$$e \cdot e = e$$

$$a \cdot e = e \cdot a = a$$

$$b \cdot e = e \cdot b = b$$

Suppose that $a \cdot b = a$. That would imply that $b = e$. On the other hand, if $a \cdot b = b$, then we would have $a = e$. None of those is possible. We are left with the only possibility:

$$a \cdot b = e$$

Similar arguments lead to

$$b \cdot a = e$$

Finally, let's consider the case that $a \cdot a = e$. But this is also impossible, since we would then have $a \cdot a = a \cdot b$, which would imply that $a = b$, and our group wouldn't be of third order as assumed. Similarly, $a \cdot a = a$ would imply that $a = e$, which we cannot accept, and the only reasonable case remaining is

$$a \cdot a = b$$

Similar arguments lead to

$$b \cdot b = a$$

Thus, the multiplication table for G is

$\cdot$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

■

In the case of cyclic groups, another remarkable feature emerges:

Theorem 5.44 Any two cyclic groups of the same order are isomorphic.

Proof Let $\langle x \rangle$ and $\langle y \rangle$ be two cyclic groups such that $|\langle x \rangle| = |\langle y \rangle| = n$, and let

$$\varphi : \langle x \rangle \to \langle y \rangle$$

be a map defined by $\varphi(x^k) = y^k$. Our function is well defined if

$$x^r = x^s \qquad (*)$$

implies

$$\varphi(x^r) = \varphi(x^s)$$

for some $r, s \in \mathbf{Z}$.

From (*), it follows that

$$x^{r-s} = e$$

Thus $n|r-s$, and putting $r = tn + s$ we get

$$\begin{aligned}\varphi(x^r) &= \varphi(x^{tn+s}) \\ &= y^{tn+s} \\ &= (y^n)^t y^s \\ &= y^s \\ &= \varphi(x^s)\end{aligned}$$

This proves that φ is well defined. Is it a homomorphism? The way we defined our map implies

$$\begin{aligned}\varphi(x^a x^b) &= \varphi(x^{a+b}) \\ &= y^{a+b} \\ &= y^a y^b \\ &= \varphi(x^a)\varphi(x^b)\end{aligned}$$

Hence, φ is a homomorphism. The definition of φ tells us also that $y^k \in \langle y \rangle$ is the image of $x^k \in \langle x \rangle$ and thus φ is surjective. Finally, since $|\langle x \rangle| = |\langle y \rangle| = n$, φ is also bijective. We conclude that

$$\langle x \rangle \cong \langle y \rangle$$ ■

Finally, let's formally recognize another important fact about isomorphisms:

Theorem 5.45 The relation "*to be isomorphic*" is an equivalence relation.

Proof As you certainly recall, in order to be an equivalence relation, "≅" needs to be (i) reflexive, (ii) symmetric, and (iii) transitive.

(i) That $G \cong G$ is easy to see since the identity map $e : G \to G$ is obviously an isomorphism.

(ii) For symmetry, if G and H are two groups such that $G \cong H$, then there exists an isomorphism $\varphi : G \to H$. But Theorem 5.41 tells us that in that case $\varphi^{-1} : H \to G$ is also an isomorphism, therefore $H \cong G$.

(iii) Let $G, H,$ and K be three groups such that $G \cong H$ and $H \cong K$, then there exist two isomorphisms $\alpha : G \to H$ and $\beta : H \to K$. By Theorem 5.38, the composition $\beta \circ \alpha : G \to K$ is also an isomorphism, hence $G \cong K$, which completes our proof. ■

With the equivalence relation, described by the above theorem, we can classify all groups in distinct classes of isomorphic groups, that is, each class is made of mutually isomorphic groups. In other words, at an abstract level, we can consider two groups to be the same if they belong to the same class. Thus, we can study properties of the whole collection of groups considering only one *abstract group* as a representation of the whole class. Obviously, the question now is: How does one recognize whether two groups are isomorphic or not? Of course, if the groups are of different order, or one is abelian and the other is not, one can immediately conclude that they are not isomorphic. Similarly, if one can prove that there is no one-to-one correspondence between the groups, manifestly they are not isomorphic. For instance, each group of fourth order is either isomorphic to the Klein group V or to the group of square rotation R_4, while $V \ncong R_4$.

Example 5.118 Let's show that $V \ncong R_4$. Suppose they are, that is, suppose there is an isomorphism

$$\varphi : V \to R_4$$

such that

$$\varphi(e_V) = e_{R_4}$$

where e_V and e_{R_4} are the neutral elements of the respective groups. Then for some $i \in \{1, 2, 3\}$

$$\varphi(v_i) = r_1, \quad v_i \in V,\ r_1 \in R_4$$

We would also have

$$\begin{aligned} \varphi(e_V) &= \varphi(v_i v_i) = \varphi(v_i)\varphi(v_i) \\ &= r_1 r_1 = r_2 \neq e_{R_4} \end{aligned}$$

This contradicts our assumption, and thus $V \ncong R_4$ as claimed. ■

However, the following classic theorem is true.

Theorem 5.46 (Cayley) Every group is isomorphic to some permutation group.

Proof Let G be a group. For any given $a \in G$ define a map

$$\tau_a : G \to G$$

called a *left translation* by

$$\tau_a(x) = ax, \quad \forall x \in G$$

τ_a is obviously a bijection since

$$\tau_a(x_1) = \tau_a(x_2)$$

immediately implies $x_1 = x_2$. On the other hand, let $y \in G$ be any element of G such that

$$x = a^{-1}y$$

Then,

$$\begin{aligned} \tau_a(x) &= \tau_a(a^{-1}y) \\ &= a(a^{-1}y) \\ &= y \end{aligned}$$

Thus, τ_a is a surjection too. Now consider a map

$$\varphi : G \to S_G$$

defined by

$$\varphi(a) = \tau_a, \quad \forall a \in G$$

where S_G is a symmetric group on set G.

We claim that φ is a monomorphism. Let's first check whether it is a homomorphism.

$$\begin{aligned} [\varphi(ab)](x) &= \tau_{ab}(x) = abx \\ &= a(bx) = \tau_a(bx) \\ &= \tau_a[\tau_b(x)] = (\tau_a \circ \tau_b)(x) \\ &= [\varphi(a) \circ \varphi(b)](x) \end{aligned}$$

Thus, $\varphi(ab) = \varphi(a)\varphi(b)$ so it is a homomorphism. It remains to be shown that φ is an injection too. Let

$$\varphi(a) = \varphi(b)$$

then

$$\tau_a(x) = \tau_b(x) \quad \forall x \in G$$

That, on the other hand, means that

$$ax = bx$$

that is

$$a = b$$

telling us that φ is indeed an injection. We conclude that G is isomorphic to some subgroup of S_G. ■

Note the significance of Cayley's discovery: If every group is isomorphic to a certain group of permutations then, fundamentally, permutation groups are the reflections of all groups, that is, the only groups there are. Consequently, we have

Corollary 5.5 Any group of order n is isomorphic to some subgroup of S_n.

Example/Exercise 5.119 Convince yourself that the set

$$V = \left\{(1), (12)(3,4), (13)(24), (14)(23)\right\}$$

with the usual composition of permutation is the familiar Klein four-group.

Definition 5.38 We say that $Z(G)$ is the **center of a group** G, if $Z(G)$ is the set of all those elements of G that commute with *every* element of G, that is

$$Z(G) = \{a \in G | ax = xa \quad \forall x \in G\}$$

Theorem 5.47 The center of a group G is a subgroup of G.

Proof First, note that $e \in Z(G)$ since

$$ex = xe = x \quad \forall x \in G$$

Next, suppose $a, b \in Z(G)$, then

$$\begin{aligned} ab^{-1}x &= ab^{-1}xe \\ &= ab^{-1}xbb^{-1} \\ &= ab^{-1}bxb^{-1} \\ &= aexb^{-1} \\ &= axb^{-1} \\ &= xab^{-1} \end{aligned}$$

Thus $ab^{-1} \in Z(G)$, and therefore $Z(G) \leq G$. ■

Theorem 5.48 A group G is abelian iff $Z(G) = G$.

Proof Trivial. ■

Everything that has been said above naturally leads us to inquire how to construct more subgroups of the group G if one or more of the subgroups are known. Here is what we have learned so far:

If H and K are two subgroups of the group G, then $e \in H$ and $e \in K$, thus $e \in H \cap K$, and therefore $H \cap K \neq \emptyset$. If we can find two elements $a, b \in H \cap K$, then $ab^{-1} \in H$ and $ab^{-1} \in K$, hence $ab^{-1} \in H \cap K$, which proves that $H \cap K \leq G$. And, of course, we can extend the argument to prove that the intersection of any number of subgroups of the group G is again a subgroup of G. We have already addressed this in Theorem 5.19. Another, similarly intriguing, question on the nature of the union of two subgroups was answered in Theorem 5.16.

Further on we need the following:

Definition 5.39 Let $H, K \subseteq G$ be two nonempty subsets of a group G, we define the "product" HK as follows:

$$HK = \{hk \in G | h \in H, k \in K\}$$

Theorem 5.49 Let H and K be two subgroups of a group G. Then HK is a subgroup of G iff

$$HK = KH$$

Proof First, we recall that

$$HK = \{hk \in G | h \in H, k \in K\}, \quad \forall H, K \subseteq G$$

Now, let $HK = KH$.

$HK \neq \varnothing$ since $e = ee \in HK$. Let $a, b \in HK$. Then, according to the previous definition, $a = h_1k_1$ and $b = h_2k_2$, where $h_1, h_2 \in H$, and $k_1, k_2 \in K$.

Consider now

$$\begin{aligned} ab^{-1} &= h_1k_1(h_2k_2)^{-1} \\ &= h_1k_1k_2^{-1}h_2^{-1} \end{aligned}$$

with $k_1k_2^{-1} = k_3$,

$$= h_1k_3h_2^{-1}$$

Observe that $k_3h_2^{-1} \in KH = HK$, that is, we can say $k_3h_2^{-1} = h_3k_4 \in HK$, where, as before, we assume $h_3 \in H$ and $k_4 \in K$. Collecting all of this together, and calling $h_1h_3 = h_4$, we obtain

$$\begin{aligned} ab^{-1} &= h_1h_3k_4 \\ &= h_4k_4 \in HK \end{aligned}$$

We proved that $HK \leq G$.

Conversely, suppose that $HK \leq G$ and take $a \in KH$ so that $a = kh$ for some $k \in K$ and $h \in H$. Consider

$$\begin{aligned} a^{-1} &= (kh)^{-1} \\ &= h^{-1}k^{-1} \in HK \end{aligned}$$

Thus $a \in HK$, and therefore $KH \subseteq HK$.

Take some $b \in HK$. Since $HK \leq G, b^{-1} \in HK$, we can express it as

$$b^{-1} = h'k'$$

where, of course, $h' \in H$ and $k' \in K$. It follows that

$$\begin{aligned} b &= (b^{-1})^{-1} \\ &= (h'k')^{-1} \\ &= k'^{-1}h'^{-1} \in KH \end{aligned}$$

Hence, $HK \subseteq KH$.

So, we have $KH \subseteq HK$ and $HK \subseteq KH$, thus $HK = KH$, which completes the proof of the theorem. ■

As an easy exercise you should do the following:

Example/Exercise 5.120 Let G be an abelian group and let H and K be any two subgroups of G. Show that HK is a subgroup of G.

5.6 NORMAL SUBGROUPS

Let's start by recalling (cf. Definition 5.39) that multiplication in a group G induces a product of any two subsets H and K. Suppose that H or K is a singleton, then we have

Definition 5.40 If $H \leq G$ is any subgroup of a group G, and $x \in G$, then we say that

$$Hx = \{hx | h \in H\}$$

is the **right coset** of H in G. Similarly, we define the **left coset** of H in G as

$$xH = \{xh | h \in H\}$$

Theorem 5.50 Let $H \leq G$ be a subgroup of G. Then, for any $x, y \in G$

(i) $x \in xH$
(ii) $xH = H$ iff $x \in H$
(iii) $xH = yH$ iff $x \in yH$
(iv) $xH = yH$ or $xH \cap yH = \emptyset$
(v) $xH = yH$ iff $x^{-1}y \in H$
(vi) $|xH| = |yH|$
(vii) $xH = Hy$ iff $H = xHx^{-1}$
(viii) $xH \leq G$ iff $x \in H$

Proof

(i) If $x \in G$, then $x = xe \in xH$
(ii) Suppose that $xH = H$. Then,

$$x = xe \in xH = H$$

Conversely, if we take $x \in H$ then clearly $xH \subseteq H$. On the other hand, suppose $h \in H$, then $x^{-1}h \in H$ too. Hence,

$$\begin{aligned} h &= eh \\ &= (xx^{-1})h \\ &= x(x^{-1}h) \in xH \end{aligned}$$

So we have $xH \subseteq H$ and $H \subseteq xH$, and therefore $xH = H$.

(iii) Suppose that $xH = yH$, then $x = xe \in xH = yH$. Conversely, if $x \in yH$, then $\exists h \in H$, such that $x = yh$. So we have

$$\begin{aligned} xH &= (yh)H \\ &= y(hH) \\ &= yH \end{aligned}$$

(iv) Suppose $xH \cap yH \neq \emptyset$, that is, $\exists c \in xH \cap yH$, such that, by (iii), $cH = xH$ and $cH = yH$, and therefore $xH = yH$.

(v) Suppose $xH = yH$, then obviously $H = x^{-1}yH$. Now we can use (ii) and the result immediately follows.

(vi) The easy way to prove this claim is to establish a bijection $\varphi : xH \to yH$. The map $xh \mapsto yh$ is obviously onto, and because of the cancellation law it is also one-to-one, thus φ is a bijection. Hence, $|xH| = |yH|$ indeed.

(vii) Observe that $xH = Hx$ implies

$$\begin{aligned} xHx^{-1} &= (Hx)x^{-1} \\ &= H(xx^{-1}) \\ &= He \\ &= H \end{aligned}$$

The converse can be proved similarly.

(viii) Suppose $xH \leq G$, then $e \in xH$ and therefore $xH \cap eH = xH \cap H \neq \emptyset$. From (iv), it follows that $xH = H$ and thus, by (ii), we conclude that $x \in H$. Conversely, if $x \in H$, (ii) will again imply that $xH = H$. ■

Theorem 5.51 Let G and H be two groups with $x, y \in G$. If $\varphi : G \to H$ is a homomorphism, then

$$\varphi(x) = \varphi(y) \text{ iff } \operatorname{Ker}(\varphi)x = \operatorname{Ker}(\varphi)y$$

Proof Observe that:

$$\varphi(x) = \varphi(y) \text{ iff } \varphi(x)(\varphi(y))^{-1} = e$$

But φ is a homomorphism, so we have

$$\varphi(xy^{-1}) = e$$

and this is true iff $xy^{-1} \in \text{Ker}(\varphi)$ which, furthermore, implies that

$$\text{Ker}(\varphi)x = \text{Ker}(\varphi)y$$

The converse is obvious. ■

What did we learn from this? Well, the theorem tells us that the elements of a group G that have the same image under the action of a homomorphism $\varphi : G \to H$ are in the same coset of $\text{Ker}(\varphi)$. Or conversely, if the elements are in the same coset of $\text{Ker}(\varphi)$, then they have the same image in H. As you will see shortly, this simple theorem has significant consequences.

Definition 5.41 Let G be a group, and let N be a subgroup of G. We say that N is a **normal subgroup** of G, and we write $N \lhd G$, if

$$xNx^{-1} \subseteq N, \quad \forall x \in G \tag{*}$$

where

$$xNx^{-1} = \{xnx^{-1} \mid x \in G, n \in N\}$$

The element xnx^{-1} is called the **conjugate of** $n \in N$ by x, and the set xNx^{-1} is called the **conjugate of** N by x.

We conclude that $N \lhd G$ iff $xnx^{-1} \in N, \ \forall x \in G \, n \in N$. Observe that the definition (*) could have been equivalently stated as follows: If we put $y = x^{-1}$, then

$$y^{-1}Ny \subseteq N$$

Thus,

$$N \subseteq yNy^{-1}$$

that is,

$$N \subseteq x^{-1}Nx \tag{**}$$

From (*) and (**) immediately follows

$$x^{-1}Nx = N$$

Or, even more symmetric,

$$Nx = xN$$

In short, $N \leq G$ is a normal subgroup if it is closed with respect to conjugates (see Figure 5.12).

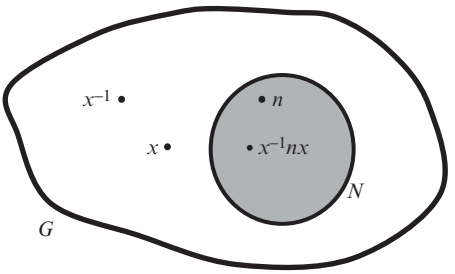


Figure 5.12 Normal subgroup

Thus,

Theorem 5.52 A subgroup $N \leq G$ is normal iff $xNx^{-1} \subseteq N, \quad \forall x \in G$.

Theorem 5.53 Let N be a subgroup of G. Then, the following are equivalent:

(i) $N \lhd G$
(ii) $x^{-1}Nx = N, \quad \forall x \in G$
(iii) $Nx = xN, \quad \forall x \in G$
(iv) $(xN)(yN) = (xy)N, \quad \forall x, y \in G$

Proof (i) $\Leftrightarrow$ (ii) and (ii) $\Leftrightarrow$ (iii) are obvious from Definition 5.41. Let's check (iii) $\Rightarrow$ (iv):

$$\begin{aligned}(xN)(yN) &= x(Ny)N \\ &= x(yN)N \\ &= (xy)(NN)\end{aligned}$$

What about NN? Well, $NN \subseteq N$, since N is closed under multiplication. Furthermore, $N = eN \subseteq NN$ thus, $NN = N$. Therefore,

$$(xy)(NN) = (xy)N$$

what was to be shown.

Similarly, we can prove (iv) ⇒ (iii).

Finally, let's check

(iv) ⇒ (i): For N to be a normal subgroup of G, we need that $xNx^{-1} = N$. Let's consider

$$\begin{aligned} xNx^{-1} &= xNx^{-1}e \subseteq xNx^{-1}N \\ &= (xN)(x^{-1}N) \\ &= (xx^{-1})N \\ &= eN = N \end{aligned}$$

Hence, $N \lhd G$. ■

Before we offer a few additional theorems, let's state some facts that the reader has probably anticipated:

Every group G has at least two normal subgroups – $\{e\}$ and G itself.

On the other hand, if the group is abelian, all subgroups are normal. Indeed, let $N \leq G$ be any subgroup of the abelian group G. Then,

$$xn = nx, \quad \forall x \in G \text{ and } \forall\, n \in N$$

Thus, $N \lhd G$.

Similarly, you should convince yourself of the validity of the claim in the next

Example/Exercise 5.121 The center of a group $Z(G)$ is a normal subgroup of G (cf. Theorem 5.47).

Theorem 5.54 Let $N \lhd G$ be a normal subgroup of G. If $xN = wN$ and $yN = zN$, then

$$xyN = wzN$$

Proof Observe that if $xN = wN$, then $x \in yN$, and we can write $x = wn_1$ for some $n_1 \in N$. Similarly, if $yN = zN$, then $y \in zN$, and thus we write $y = zn_2$ for some $n_2 \in N$. So, we have

$$\begin{aligned} xy &= (wn_1)(zn_2) \\ &= w(n_1 z)n_2 \end{aligned}$$

■

Definition 5.42 If a group G has no other normal subgroups, except the trivial ones $\{e\}$ and G itself, we say that the group G is a **simple group**.

Example 5.122 Consider the alternating subgroup A_3 of the symmetric group S_3

$$A_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$

You can easily convince yourself that $A_3 \lhd S_3$. Take, for instance,

$$a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in A_3$$

and, say,

$$x = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \in S_3$$

and consider

$$\begin{aligned} xax^{-1} &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \in A_3 \end{aligned}$$

The rest is similar.

On the other hand, take a subset of S_3, $H = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\}$. $H \leq S_3$ (you should check this), but it is not a normal subgroup. Indeed, let $x \in S_3$ be

$$x = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

and consider

$$\begin{aligned} xhx^{-1} &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \notin H \end{aligned}$$

Hence, $H \not\lhd S_3$. ■

Theorem 5.55 Let $M, N \lhd G$ be two normal subgroups of G. Then, $M \cap N$ is a normal subgroup of G.

Proof Let $M, N \leq G$ be two normal subgroups of G. We proved in Theorem 5.19 that the intersection of any family of subgroups of a group G is a subgroup of G thus, in particular, $M \cap N \leq G$. Now, let $x \in G$ and $a \in M \cap N$. We need to show that

$$xax^{-1} \in M \cap N$$

Since $a \in M \cap N$, $a \in M$ and $a \in N$. But both $M \lhd G$ and $N \lhd G$, therefore $xax^{-1} \in M$ and $xax^{-1} \in N$, hence $xax^{-1} \in M \cap N$ and $M \cap N \lhd G$. ■

Theorem 5.56 Let $\eta : G \to H$ be an epimorphism of groups and, furthermore, let $N \lhd G$ be a normal subgroup of G. Then, $\eta(N) \lhd H$ is a normal subgroup of H.

Proof Certainly, $\eta(N)$ is a subgroup of H. We need to show that it is a normal subgroup. Let $h \in H$ be some element of H and, since η is an epimorphism, that is, a surjection, there exists an $x \in G$ such that $\eta(x) = h$. Also, there exists $n \in N$ such that $\eta(n) = y \in \eta(N)$. But $N \lhd G$, therefore,

$$xnx^{-1} \in N$$

Thus,

$$\eta(xnx^{-1}) \in \eta(N)$$

that is,

$$\eta(x)\eta(n)[\eta(x)]^{-1} \in \eta(N)$$

and so,

$$hyh^{-1} \in \eta(N)$$

Hence, $\eta(N) \lhd H$. ■

Theorem 5.57 Let $\varphi : G \to H$ be a homomorphism of groups and let $M \lhd H$ be a normal subgroup of H. Then, the preimage $\varphi^{-1}(M) \leq G$ is a normal subgroup of G.

Proof Observe that, since G and H are groups and φ is a homomorphism, $\varphi^{-1}(M) \leq G$. Take $x \in G$, and $a \in \varphi^{-1}(M)$ and note that, since $M \lhd H$,

$$\varphi(x)\varphi(a)(\varphi(x))^{-1} = \varphi(xax^{-1}) \in M$$

and therefore,

$$xax^{-1} \in \varphi^{-1}(M)$$

We conclude that $\varphi^{-1}(M) \lhd G$. ■

You should be now able to do the following exercise with no difficulty.

Example/Exercise 5.123 Let $M, N \leq G$ be two subgroups of G such that N is also a normal subgroup. Show that

(i) $MN \subseteq G$ is a subgroup of G and
(ii) if $M \lhd G$, then $MN \lhd G$.

Theorem 5.58 Let $\varphi : G \to H$ be a homomorphism of groups. Then, $\text{Ker}(\varphi) \lhd G$.

Proof Let $a, b \in \text{Ker}(\varphi)$. This means that $\varphi(a) = e$ and $\varphi(b) = e$. Also, since φ is a homomorphism,

$$\varphi(ab) = \varphi(a)\varphi(b) = ee = e$$

Thus, $ab \in \text{Ker}(\varphi)$.

If $a \in \text{Ker}(\varphi)$ then $\varphi(a) = e$, implies

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

and so, $a^{-1} \in \text{Ker}(\varphi)$.

Finally, with $a \in \text{Ker}(\varphi)$ and some $x \in G$, let's consider

$$\begin{aligned} \varphi(xax^{-1}) &= \varphi(x)\varphi(a)\varphi(x^{-1}) \\ &= \varphi(x)\varphi(a)(\varphi(x))^{-1} \\ &= \varphi(x)e(\varphi(x))^{-1} \\ &= e \end{aligned}$$

Hence, $xax^{-1} \in \text{Ker}(\varphi)$, and therefore $\text{Ker}(\varphi) \lhd G$. ■

5.7 CENTRALIZER, NORMALIZER, STABILIZER

Definition 5.43 Let $A \subseteq G$ be any nonempty subset of G. We say that a set

$$C_G(A) = \{x \in G | xax^{-1} = a, \quad \forall a \in A\}$$

is the **centralizer of** A in G. In other words, $C_G(A)$ is the set of all elements of G that commute with every element of A.

We say that a subgroup $H \leq G$ **centralizes** A if $H \leq C_G(A)$.

Theorem 5.59 The centralizer $C_G(A)$ is a subgroup of G.

Proof First note that $C_G(A) \neq \emptyset$, since, certainly, $e \in C_G(A)$. Suppose $x, y \in C_G(A)$, that is,

$$xax^{-1} = a \text{ and } yay^{-1} = a, \quad \forall a \in A$$

Consider now

$$\begin{aligned}(xy)a(xy)^{-1} &= (xy)a(y^{-1}x^{-1}) \\ &= x(yay^{-1})x^{-1} \\ &= xax^{-1} \\ &= a\end{aligned}$$

Thus, $xy \in C_G(A)$ and $C_G(A)$ is closed under multiplication, hence $C_G(A) \leq G$. ■

Definition 5.44 Let $A \subseteq G$ be any nonempty subset of G. Defining

$$xAx^{-1} = \{xax^{-1} | x \in G, a \in A\}$$

we say that

$$N_G(A) = \{x \in G | xAx^{-1} = A\}$$

is the **normalizer** of A in G.

Example 5.124 Show that

(i) $N_G(A)$ is a subgroup of G, and
(ii) $C_G(A)$ is a subgroup of $N_G(A)$.

Solution Let $x, y \in N_G(A)$.

(i) Consider

$$\begin{aligned}(xy)A(xy)^{-1} &= (xy)A(y^{-1}x^{-1}) \\ &= x(yAy^{-1})x^{-1} \\ &= xAx^{-1} \\ &= A\end{aligned}$$

So, $xy \in N_G(A)$ and $N_G(A)$ is closed under multiplication, thus $N_G(A) \leq G$.

(ii) If $x \in C_G(A)$, then $xax^{-1} = a, \quad \forall a \in A$, therefore, $C_G(A) \leq N_G(A)$. ■

We briefly mentioned the concept of a commutator product and commutator group in Example 5.79. Let's now define it formally.

Definition 5.45 Let $H, K \leq G$ be two subgroups of G. We say that

$$[C] = [H, K] = [[h, k] = hkh^{-1}k^{-1}; h \in H, k \in K]$$

is a **commutator subgroup** of G. Note that the set of all commutators need not be a subgroup, that is, $[H, K]$ is a subgroup only if it is generated by the indicated commutators.

Example/Exercise 5.125 Show that $[H, K] = [K, H]$.

Definition 5.46 Let G be a group and X be a set. We say that X is the ***G*-set** if there is a mapping

$$\alpha : G \times X \to X$$

called the **action** (of G on X), defined by

$$\alpha(g, x) = gx$$

such that for all $g, h \in G$ and $x \in X$

(i) $\alpha(g, x)\alpha(h, x) = \alpha(gh, x) = (gh)x$
(ii) $\alpha(ex) = ex = x$

If G acts on X, we usually write gx instead of $\alpha(g, x)$, and axioms (i) and (ii) read

(i′) $(gx)(hx) = g(hx) = ghx$
(ii′) $ex = x$

Example 5.126 Suppose $X = G$ and G acts on itself by conjugation.

$$gx = gxg^{-1}$$

Then, the G-set is G itself. Indeed, take $g, h, x \in G$, and consider

$$\begin{aligned}(gh)x &= (gh)x(gh)^{-1} \\ &= (gh)x(h^{-1}g^{-1}) \\ &= g(hxh^{-1})g^{-1} \\ &= g(hx)g^{-1} \\ &= g(hx)\end{aligned}$$

And, of course, $ex = x$. Thus, G is the G-set itself. ■

Example 5.127 Every group acts on the family of its subgroups by conjugation. ■

Example 5.128 Every group G, acting on itself by left multiplication, α_l, is also a G-set. Indeed, let $g_1, g_2 \in G$ be two elements from G, then g_1 acts on g_2 by $\alpha_l(g_1, g_2) = g_1g_2$. Similarly, if $H \leq G$ we can regard G as an H-set where the action α of $h \in H$ on $g \in G$ is given by $\alpha(h, g) = hg$. ■

Definition 5.47 A group G is said to act **transitively** on a G-set X if for every $x, y \in X$, there exists $g \in G$ such that $gx = y$.

Theorem 5.60 Let X be a G-set, and let $x, y \in X$. We say that $x \sim y$ iff there exists $g \in G$ such that $gx = y$. Then, "$\sim$" is an equivalence relation.

Proof To check reflexivity, we just note that for each $x \in X$, $ex = x$ so, $x \sim x$.

For symmetry, suppose $x \sim y$, that is, there exists $g \in G$ such that $gx = y$. Well, consider then

$$g^{-1}y = g^{-1}(gx) = ex = x$$

So, $y \sim x$.

Finally, let $x \sim y$ and $y \sim z$. Then there exist $g_1, g_2 \in G$ such that

$$g_1x = y \text{ and } g_2y = z$$

So we have

$$(g_2g_1)x = g_2(g_1x) = g_2y = z$$

and therefore, $x \sim z$. Thus, "$\sim$" is an equivalence relation indeed. ■

Theorem 5.61 Let X be a G-set with action α.

(i) Then there exists a homomorphism

$$\Phi : G \to S_X$$

defined by

$$(\Phi(g))(x) = gx = \alpha(g, x), g \in G, x \in X$$

(ii) Conversely, every homomorphism $\alpha : G \to S_X$ defines an action given by

$$gx = \alpha(g)x$$

which makes X a G-set

Proof

(i) Let $g, h \in G$ be two elements of G, and let $x \in X$ Then,

$$\begin{aligned}(\Phi(gh))(x) &= (gh)(x)\\ &= g(hx)\\ &= g(\alpha(h)(x))\\ &= g\alpha(h, x)\\ &= \alpha(g)(\alpha(h, x))\\ &= \alpha(g, (\alpha(h, x)))\\ &= \alpha(g)((\Phi(h))(x))\\ &= (\Phi(g))(\Phi(h))(x)\end{aligned}$$

(ii) Since

$$gx = \alpha(g)x$$

consider

$$\begin{aligned}(gh)x &= \alpha(gh)(x)\\ &= (\alpha(g)\alpha(h))(x)\\ &= \alpha(g)(\alpha(h)(x))\\ &= \alpha(g)(hx)\\ &= g(hx)\end{aligned}$$

And finally,

$$\begin{aligned} ex &= (\alpha(e))(x) \\ &= (\alpha(gg^{-1}))(x) \\ &= (\alpha(g)\alpha(g^{-1}))(x) \\ &= \alpha(g)(\alpha(g^{-1})(x)) \\ &= \alpha(g)(g^{-1}x) \\ &= g(g^{-1}x) \\ &= (gg^{-1})(x) \\ &= x \end{aligned}$$

■

Definition 5.48 Let G be a group acting on a set X, and let $x \in X$. We call the set

$$G_x = \{\ g \in G |\ gx = x\ \}$$

the **stabilizer** of x in G.

Theorem 5.62 The stabilizer G_x is a subgroup of G.

Proof Note that, obviously, $G_x \neq \emptyset$, since at least $e \in G_x$. Suppose $g \in G_x$, then

$$\begin{aligned} x = ex &= (g^{-1}g)x \\ &= g^{-1}(gx) \\ &= g^{-1}x \end{aligned}$$

Therefore, $g^{-1} \in G_x$.

Now, take $g, h \in G_x$, then

$$\begin{aligned} (gh)x &= g(hx) \\ &= gx \\ &= x \end{aligned}$$

We conclude that $G_x \leq G$. ■

Definition 5.49 Let G be a group, and let X be a G-set. We say that the set

$$\mathrm{O}(x) = \{gx |\ g \in G, x \in X\} \subseteq X$$

that is, the set of all elements of X, to which x can be moved by the elements of G, is the ***G*-orbit**, or just the **orbit of *x* in *G***.

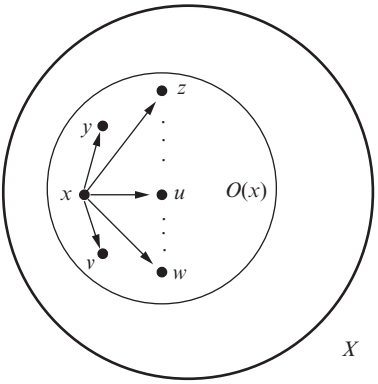


Figure 5.13 Orbit of $x \in X$ in G

We can now restate Definition 5.47 as

Definition 5.50 A group G **acts transitively** on a set X when there is an $x \in X$, such that $\mathrm{O}(x) = X$.

Example 5.129 Let's find all orbits of $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$ under the action of

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 6 & 7 & 4 & 1 & 5 & 2 \end{pmatrix} \in S_8$$

For $O\,(1)$, we note that

$$1 \xrightarrow{\pi} 3 \xrightarrow{\pi} 6 \xrightarrow{\pi} 1 \xrightarrow{\pi} \cdots$$

So, the orbit of 1 is

$$O\,(1) = \{1, 3, 6\}$$

Next, we choose an element from X that is not in $O\,(1)$, say 2, and, as before, we observe that

$$2 \xrightarrow{\pi} 8 \xrightarrow{\pi} 2 \xrightarrow{\pi} \cdots$$

and the orbit of 2 is

$$O\,(2) = \{2, 8\}$$

Finally, to find the orbit of an element from X that is in neither $O(1)$ nor $O(2)$ we note that

$$4 \xrightarrow{\pi} 7 \xrightarrow{\pi} 5 \xrightarrow{\pi} 4 \xrightarrow{\pi} \cdots$$

So,

$$O(4) = \{4, 5, 7\}$$

Thus, we have obtained the list of all orbits of elements from X. We can find the orbits under the action of any element of S_8 in the same way. ■

It may be instructive at this point to revisit Theorem 5.60 and realize that, after choosing a permutation $\pi \in S_n$ acting on a set X, one can construct equivalence classes in X, which are exactly the orbits obtained by the action of π.

Example 5.130 Let $X = \{1, 2, 3, \ldots, n\}$ and let $e \in S_n$ be the identity permutation of X. Then, the orbits under the action of e are one-element subsets of X. ■

Example 5.131 Remember Cayley's theorem (Theorem 5.46). We saw there that a group G acts on itself by (left) translation $\tau_a : G \to G$, defined by

$$\tau_a(x) = ax, \quad a, x \in G$$

If G acts on itself, that is, G is a G-set itself, then

$$\mathrm{O}(x) = \{gx \mid g, x \in G\} = G$$

For, if $g \in G$ then $g = (gx)x^{-1}$.

On the other hand, if G acts on itself by conjugation, then $\mathrm{O}(x) = G_x$. ■

Example 5.132 If G acts by conjugation on the family of all its subgroups, then

$$\begin{aligned} O(H) &= \{gHg^{-1} \mid g \in G, h \in H, H \leq G\} \\ &= \text{set of all conjugates of } H \end{aligned}$$

Also note that $G_H = N_G(H)$. ■

Example 5.133 Let G be a dihedral group D_4 acting on the set $X = \{1, 2, 3, 4\}$ of vertices of a square (cf. Figure 5.3).

To make things even more transparent, this time we write our dihedral group as

$$\begin{aligned} G &= G_{\text{Rot}} \cup G_{\text{Ref}} \\ &= \left\{(e), \begin{pmatrix}1 & 2 & 3 & 4\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 4\end{pmatrix}, \begin{pmatrix}1 & 4 & 3 & 2\end{pmatrix}\right\} \\ &\quad \cup \left\{\begin{pmatrix}2 & 4\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}, \begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}3 & 4\end{pmatrix}, \begin{pmatrix}1 & 4\end{pmatrix}\begin{pmatrix}2 & 3\end{pmatrix}\right\} \end{aligned}$$

Observe that for each vertex $i \in X$ there is some $g \in G$, such that $g1 = i$, that is, D_4 acts transitively. Thus, $O(1) = X$. ■

5.8 QUOTIENT GROUP

As we have seen thus far, in order to further our study of the structure of the group G, we often search for a smaller group related to G hoping that the smaller group will shed light on the structure of the more complicated G. So, before we proceed, recall Definition 5.40 where, given $H \leq G$ and $x \in G$, we defined the left and right cosets, respectively, as

$$xH = \{xh \mid h \in H\}$$

$$Hx = \{hx \mid h \in H\}$$

Let's now consider the set Q_L of all (different) left cosets of $H \leq G$, that is,

$$Q_L = G/H = \{xH \mid x \in G; H \leq G\}$$

Analogously, we can have the set Q_R of all right cosets of $H \leq G$, that is,

$$Q_R = H/G = \{Hx \mid x \in G; H \leq G\}$$

Note that in general

$$G/H \neq H/G$$

This is understandable since H is not necessarily a normal subgroup. It follows that $xH \neq Hx$.

However, we have

Theorem 5.63 G/H is a homomorphic image of G.

Proof Let G be a group and let $H \leq G$. Consider the function $\varphi : G \to G/H$, defined by $\varphi(x) = xH$, that is, φ maps every element of G to its own coset. Then,

$$\varphi(xy) = xyH = xHyH = \varphi(x)\varphi(y)$$

So, φ is a homomorphism, called the **natural homomorphism** of G to G/H, and G/H is said to be a **homomorphic image of *G***. ■

Theorem 5.64 Let G be a group and $H \leq G$. Then, for every $x \in G$, $|xH| = |H|$, that is, xH and H have the same number of elements. (The same holds for Hx.)

Proof Consider a map

$$\varphi : H \to xH$$

defined by

$$\varphi(h) = xh, \quad x \in G, h \in H$$

Obviously, φ is a surjection. But also, since

$$\varphi(h_1) = \varphi(h_2), \quad h_1, h_2 \in H$$

implies

$$xh_1 = xh_2$$

thus,

$$h_1 = h_2$$

Therefore, φ is an injection too. Hence, φ is a bijection and we conclude that H and xH have the same number of elements. ■

Since any coset xH has the same number of elements as H, it follows that all cosets of H have the same number of elements. Thus,

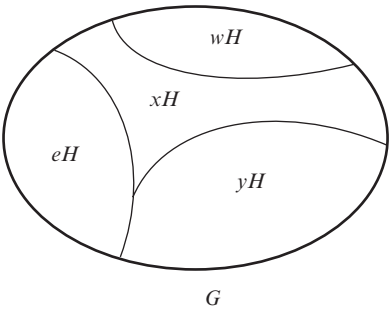


Figure 5.14

Theorem 5.65 Let G be a group, and $H \leq G$. Then, the family of all cosets xH, where $x \in G$, forms a partition of G.

Proof Recall Theorem 5.50(iv) where we proved that xH and yH are either disjoint or equal. If they are disjoint, the proof of our theorem is done. Suppose $xH \cap yH \neq \varnothing$ and let $a \in xH \cap yH$. Being in xH, $a = xh_1$ for some $h_1 \in H$. But, a is also in yH, so we can write $a = yh_2$ for some $h_2 \in H$. Hence, we have

$$xh_1 = yh_2$$
$$x = yh_2h_1^{-1}$$

So, $x \in yH$ and therefore, $xH = yH$.

Finally, we need to show that every element $g \in G$ is in one of the cosets. Well, since H is a subgroup, $e \in H$, but then for any $g \in G$,

$$g = ge \in gH$$

Thus, the family of all cosets of H indeed forms a partition of G. ■

As a simple exercise you may want to prove the following:

Theorem 5.66 Sets G/H and H/G are equipotent (have the same number of elements).

(Hint: If you consider that $(xH)^{-1} = Hx^{-1}$, then it's easy to show that $xH \mapsto (xH)^{-1}$ is a bijection.)

Definition 5.51 Let $H \leq G$ be any subgroup of G. The cardinal number of the left (right) cosets of H is called the **index of a subgroup** $\boldsymbol{H}$, $\boldsymbol{\mu}$, and is denoted by $[G : H]$, that is, $\mu = |G/H| = [G : H]$. If μ is finite, we say H is a subgroup of finite index.

In other words, if μ is finite, that is, G is finite, then

$$\mu = [G : H] = \frac{|G|}{|H|}$$

Recall (cf. Definition 5.39) that given two subsets $H, K \subseteq G$ we can always design a product

$$HK = \{hk \mid h \in H, k \in K\}$$

so it is tempting to take $H \leq G$ and consider the products $(Hx)(Hy)$, hoping to get a group structure on the family of right cosets (similarly with the family of left cosets). However, even if $H \leq G$ is the family of all, say, right cosets, H/G need not be closed under multiplication, thus H/G is not necessarily a group. (The same is true for G/H.) On the other hand, things are quite different if N is a normal subgroup of G.

We are now ready to prove a theorem of fundamental importance in the theory of finite groups.

Theorem 5.67 (Lagrange) Let G be a finite group and $H \leq G$ any of its subgroups. Then, the order of H divides the order of G.

Proof Let $x_1H,\ x_2H,\ \ldots\ , x_rH$ be a set of r distinct cosets of H in G. Then, for every $x \in G$, we can find some $i \in \{1, 2, \ldots\ , r\}$, such that

$$xH = x_iH$$

From Theorem 5.50(i) we know that $x \in xH$. Thus, every element of G belongs to one of the cosets $x_1H, x_2H, \ldots, x_rH$, that is,

$$G = x_1H \cup x_2H \cup \cdots \cup x_rH$$

Theorem 5.65(iv) tells us that $x_1H \cap x_2H \cap \cdots \cap x_rH = \varnothing$, so we can write

$$|G| = |x_1H| + |x_2H| + \cdots + |x_rH|$$

However, $|x_iH| = |H|$ for all i, so it follows

$$|G| = r|H|$$

And that is what was to be shown. ■

Corollary 5.6 If G is a finite group and $g \in G$, then the order of g divides the order of G.

Proof The order of an element g is by definition the order of the subgroup generated by that element, so the result follows at once from Theorem 5.67. ■

Corollary 5.7 A group of prime order is cyclic.

Proof Let p be a prime, and let G be a group such that $|G| = p$. Take $g \in G$, $g \neq e$. Then, the cyclic group $\langle g \rangle$ is not trivial, that is, it has more than one element. Hence, $|\langle g \rangle| \neq 1$ and should divide $|G| = p$. But p is prime, so $|\langle g \rangle| = |G|$ and therefore, $\langle g \rangle = G$. ■

Corollary 5.8 If G is a finite group and $g \in G$, then $g^{|G|} = e$.

Proof We know from Corollary 5.6 that the order of g divides the order of G. In other words, $|G| = k|g|, k \in \mathbf{N}$. Therefore,

$$\begin{aligned} g^{|G|} &= g^{k|g|} \\ &= e^k \\ &= e \end{aligned}$$

■

Concluding our discussion of Lagrange's theorem, let's note the important and not immediately obvious fact: *The converse of Lagrange's theorem is not true.*

Theorem 5.68 Let $N \lhd G$ be a normal subgroup of G. Then G/N is a group, called the **quotient group** or the **factor group** (with the natural multiplication of classes, which for simplicity from now on, we denote by $[x][y] = (xN)(yN),\ \ x, y \in G$).

Proof First, note that $N \lhd G$ implies

$$xN = Nx$$

thus,

$$G/N = N/G$$

So, we simply talk about the quotient sets

$$G/N = \{[x] = xN | x \in G\}$$

Let's now check whether multiplication is well defined:

$$\begin{aligned}[x][y] &= (xN)(yN) \\ &= x(Ny)N \\ &= x(yN)N \\ &= (xy)(NN) \\ &= (xy)N \\ &= [xy] \in G/N\end{aligned}$$

Associativity immediately follows:

$$\begin{aligned}[x]([y][z]) &= [x][yz] \\ &= [xyz] \\ &= [xy][z] \\ &= ([x][y])[z]\end{aligned}$$

Next, the identity:

$$\begin{aligned}[e] &= eN \\ &= N\end{aligned}$$

Indeed,

$$\begin{aligned}[e][x] &= [ex] \\ &= [x]\end{aligned}$$

And, finally, the inverse:

$$\begin{aligned}[x]^{-1}[x] &= [x^{-1}x] \\ &= [e]\end{aligned}$$

So, G/N is a group indeed. ■

Example 5.134 Consider the group $(\mathbf{Z}; +) = \mathbf{Z}$, and let $m \in \mathbf{Z}$ be some integer. Then $m\mathbf{Z} \leq \mathbf{Z}$ and, furthermore, $m\mathbf{Z} \lhd \mathbf{Z}$ since $\mathbf{Z}$ is a commutative group. Let's construct the quotient group

$$\mathbf{Z}/m\mathbf{Z} = \{[k] = k + m\mathbf{Z} | k \in \mathbf{Z}\}$$

Clearly, the elements of this group are

$$\begin{aligned}[0] &= m\mathbf{Z} \\ [1] &= 1 + m\mathbf{Z} \\ &\vdots \\ [m-1] &= (m-1) + m\mathbf{Z} \\ [m] &= m + m\mathbf{Z} \\ &= m(1 + \mathbf{Z}) \\ &= m\mathbf{Z} \\ &= [0] \\ [m] &= m + m\mathbf{Z} \\ &= m(1 + \mathbf{Z}) \\ &= m\mathbf{Z} \\ &= [0]\end{aligned}$$

and so on.

Similarly,

$$\begin{aligned}[-1] &= -1 + m\mathbf{Z} \\ &= (m-1) - m + m\mathbf{Z} \\ &= (m-1) + m(-1 + \mathbf{Z}) \\ &= (m-1) + m\mathbf{Z} \\ &= [m-1]\end{aligned}$$

and so on. Defining the group operation as

$$[k] + [l] = [k + l]$$

it is easy to recognize $\mathbf{Z}/m\mathbf{Z}$ as a group of order m. ■

As you can see, the construction of a quotient group $\mathbf{Z}/m\mathbf{Z}$ is a generalization of the construction of $\mathbf{Z}_m$ from $\mathbf{Z}$. Indeed, for any given $m, k \in \mathbf{Z}$, we consider the congruence class $[k]$ of $k(\text{mod } m)$ as the coset $k + \langle m \rangle$. But, remember, $\mathbf{Z}$ is abelian, so $\langle m \rangle \lhd \mathbf{Z}$, and therefore $\mathbf{Z}/\langle m \rangle$ is a quotient group whose elements are cosets $k + \langle m \rangle$. Hence, as you would expect, the binary operation is

$$(k + \langle m \rangle) + (l + \langle m \rangle) = k + l + \langle m \rangle$$

which we can write in the congruence class notation as above

$$[k] + [l] = [k + l]$$

Thus, $\mathbf{Z}/\langle m \rangle \cong \mathbf{Z}_m \cong \mathbf{Z}/m\mathbf{Z}$

Theorem 5.69 Let $N \lhd G$ be a normal subgroup of G. Then a function

$$\pi : G \to G/N$$

called the **natural map** or the **natural projection**, and defined by $\pi(g) = [g]$, is an epimorphism with kernel N.

Proof First, we see that π is a homomorphism:

$$\begin{aligned}\pi(g)\pi(h) &= (gN)(hN) \\ &= g(Nh)N \\ &= (gh)(NN) \\ &= (gh)N \\ &= \pi(gh)\end{aligned}$$

Next, consider

$$\begin{aligned}\mathrm{Ker}(\pi) &= \pi^{-1}([e]) \\ &= \pi^{-1}(eN) \\ &= \pi^{-1}(N) \\ &= N\end{aligned}$$

So, π is surjective too, hence π is an epimorphism. ■

Corollary 5.9 Every normal subgroup is the kernel of some homomorphism.

For the next theorem, we need to recall the concept of a commutator and commutator group (cf. Example 5.79 and Definition 5.45):

If $a, b \in G$ are two elements of a group G, we said that the commutator of a and b, denoted by $[a, b]$, is

$$[a, b] = aba^{-1}b^{-1}$$

Furthermore, we said that the subgroup G' generated by the set of all commutators in G is called the commutator subgroup of G. So, we have

Theorem 5.70 The commutator subgroup $G' \leq G$ is a normal subgroup of G. Also, if $H \lhd G$ is a normal subgroup of G, then G/H is abelian iff $G' \leq H$.

Proof Let $x \in G'$ be any element of G', that is, $x = [a, b] = aba^{-1}b^{-1}$, $a, b \in G$. Then,

$$\begin{aligned}x^{-1} &= (aba^{-1}b^{-1})^{-1} \\ &= bab^{-1}a^{-1} \\ &= [a, b]^{-1} \in G'\end{aligned}$$

Now, take some $g \in G$ and consider

$$\begin{aligned}gxg^{-1} &= g[a, b]g^{-1} \\ &= g(aba^{-1}b^{-1})g^{-1} \\ &= ga(g^{-1}g)b(g^{-1}g)a^{-1}(g^{-1}g)b^{-1}g^{-1} \\ &= (gag^{-1})(gbg^{-1})(ga^{-1}g^{-1})(gb^{-1}g^{-1}) \\ &= (gag^{-1})(gbg^{-1})(gag^{-1})^{-1}(gbg^{-1})^{-1} \in G'\end{aligned}$$

Clearly, any element $y \in G'$ can be represented as a product of a finite number of commutators, that is,

$$y = x_1 x_2 \cdots x_n$$

And therefore, for any $g \in G$, we can write

$$\begin{aligned} gyg^{-1} &= g(x_1 \ \cdots x_n)g^{-1} \\ &= g(x_1(g^{-1}g)x_2(g^{-1}g)\cdots(g^{-1}g)x_n)g^{-1} \\ &= (gx_1g^{-1})\cdots(gx_ng^{-1}) \in G' \end{aligned}$$

We conclude that $G' \lhd G$.

The proof of the second part of the theorem goes as follows:

Let $H \lhd G$. Suppose, first, that G/H is abelian. Then, $aHbH = bHaH$ for all $a, b \in G$. But then also $abH = baH$, and therefore,

$$\begin{aligned} ab(ba)^{-1} &= aba^{-1}b^{-1} \\ &= [a, b] \in H \end{aligned}$$

In other words,

$$\begin{aligned} (aba^{-1}b^{-1})H &= (aH)(bH)(aH)^{-1}(bH)^{-1} \\ &= (aH)(aH)^{-1}(bH)(bH)^{-1} \\ &= H \end{aligned}$$

Thus, $G' \leq H$. The converse is proved similarly. ■

5.9 THE ISOMORPHISM THEOREMS

We conclude this chapter with three important theorems due to Emmy Noether.[5] As much as they are relevant for group theory itself, their analogues are also true for many other mathematical structures. They depict some additional relations between quotient groups, normal subgroups, and homomorphisms. You will recall that every quotient group G/N is a homomorphic image of G, so we ask whether every homomorphic image of G is isomorphic to a corresponding quotient group. Before addressing Isomorphism Theorems, as a warm-up, let's introduce some additional concepts.

[5] Amalie Emmy Noether (1882–1935), German mathematician.

Theorem 5.71 Let $\varphi : G \to H$ be a homomorphism with $\text{Ker}(\varphi) = K$. Then,

$$\varphi(a) = \varphi(b) \quad \text{iff} \quad aK = bK$$

Proof Suppose $\varphi(a) = \varphi(b)$. Then,

$$\begin{aligned} e &= \varphi(a)(\varphi(a))^{-1} = \varphi(a)(\varphi(b))^{-1} \\ &= \varphi(ab^{-1}) \end{aligned}$$

Hence, $ab^{-1} \in K$, that is, $aK = bK$.

The converse is obvious. ■

Let G and H be two groups, such that $M \lhd G$ and $N \lhd H$ are two respective normal subgroups. Furthermore, let

$$G^* = G/M \text{ and } H^* = H/N$$

be quotient groups. Then, the corresponding natural projections are

$$\pi_G : G \to G^* \text{ and } \pi_H : H \to H^*$$

Finally, let

$$\chi : (G, M) \to (H, N)$$

be some **homomorphism of "pairs,"** that is, a homomorphism $\chi : G \to H$, such that $\chi(M) \subseteq N$. We would like to see whether similar properties could be extended to the corresponding cosets. Thus,

Theorem 5.72 Let $g \in G$ be any element of G. Then,

$$\chi(gM) \subseteq \chi(g)N$$

Proof Let $m \in M$ be any element of M. Then,

$$\chi(gm) = \chi(g)\chi(m) \in \chi(g)\chi(M) \subseteq \chi(g)N$$

so we see that

$$\chi(gM) \subseteq \chi(g)N$$

Naturally, we then say that the map

$$gM \mapsto \chi(g)N$$

defines a function

$$\chi^* \ : \ G^* \ \rightarrow H^*$$

Note that χ^* depends on the definition of χ. We say that χ induces χ^*. ■

Next let's address whether χ^* is a homomorphism.

Theorem 5.73 Let G and H be two groups with $M \lhd G$ and $N \lhd H$. Let

$$G^* = G/M \ \text{ and } \ H^* = H/N$$

be two quotient groups. Then,

$$\chi^* \ : \ G^* \rightarrow H^*$$

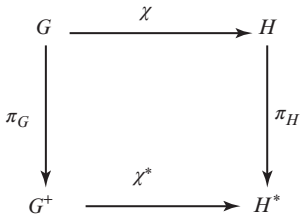
is a homomorphism of quotient groups.

Proof Let $x, y \in G$ be any two elements of G. Consider

$$\begin{aligned} \chi^*[(xM)(yM)] &= \chi^*[(xy)M] \\ &= \chi(xy)N \\ &= \chi(x)\chi(y)N \\ &= [\chi(x)N][\chi(y)N] \\ &= \chi^*(xM)\chi^*(yM) \end{aligned}$$

Thus, χ^* is a homomorphism indeed. We call it the **induced homomorphism** of χ. ■

Theorem 5.74 Let χ^* be the induced homomorphism of χ. Then, the following diagram commutes



that is, $\pi_H \circ \chi = \chi^* \circ \pi_G$.

Proof Take any $g \in G$ and consider

$$\begin{aligned}(\pi_H \circ \chi)(g) &= \pi_H[\chi(g)] \\ &= \chi(g)H \\ &= \chi^*(gM) \\ &= \chi^*[\pi_G(g)] \\ &= (\chi^* \circ \pi_G)(g)\end{aligned}$$ ■

Before we close this section, it might be instructive to prove one more proposition.

Theorem 5.75 Following the previous three theorems and the notation therein, we claim:

(i) $\mathrm{Im}\chi^* = \pi_H(\mathrm{Im}\chi)$
(ii) $\mathrm{Ker}\,\chi^* = \pi_G[\chi^{-1}(H)]$

Proof First note that π_G is an epimorphism. Now, we know from Theorem 5.74 that $\pi_H \circ \chi = \chi^* \circ \pi_G$, therefore,

(i)

$$\begin{aligned}\mathrm{Im}\chi^* &= \chi^*(G^*) \\ &= \chi^*[\pi_G(G)] \\ &= (\chi^* \circ \pi_G)(G) \\ &= (\pi_H \circ \chi)(G) \\ &= \pi_H[\chi(G)] \\ &= \pi_H(\mathrm{Im}\chi)\end{aligned}$$

(ii)

$$\begin{aligned}\mathrm{Ker}\,\chi^* &= \chi^{*-1}(eH^*) \\ &= \chi^{*-1}[\pi_H(e)] \\ &= \chi^{*-1}[\pi_H(\chi(e))] \\ &= \chi^{*-1}[(\pi_H \circ \chi)(e)] \\ &= \chi^{*-1}[(\chi^* \circ \pi_G)(e)] \\ &= (\chi^{*-1} \circ \chi^*)[\pi_G(e)] \\ &= \pi_G(\chi^{-1}(eH)) \\ &= \pi_G[\chi^{-1}(H)]\end{aligned}$$ ■

We have learned from Theorem 5.71 that if there exists a homomorphism $\varphi : G \to H$ with a kernel K, then all the elements of G with the same images in H live in the same coset of K (see Figure 5.14).

Consequently, we have three important theorems.

Theorem 5.76 (The first isomorphism theorem) Let $f : G \to H$ be a homomorphism of groups. Then,

$$G/\text{Ker}f \cong \text{Im}\, f$$

Proof Let's put $\text{Ker}f = K$, and define $\varphi : G/K \to H$ by

$$\varphi(xK) = f(x)$$

In other words, we are considering a map $xK \mapsto f(x)$.

To see that φ is well defined, first recall that $K = \text{Ker}\, f \lhd G$. Now, suppose

$$xK = yK, \quad x, y \in G$$

That means that $y^{-1}x \in K$, which furthermore that implies

$$\begin{aligned} e &= f(y^{-1}x) \\ &= f(y^{-1})f(x) \\ &= [f(y)]^{-1}f(x) \end{aligned}$$

Hence,

$$f(y) = f(x)$$

It follows that

$$\varphi(xK) = \varphi(yK)$$

So φ is well defined and is also injective.

Next, consider

$$\begin{aligned} \varphi((xK)(yK)) &= \varphi(xyK) \\ &= f(xy) \\ &= f(x)f(y) \\ &= \varphi(xK)\varphi(yK) \end{aligned}$$

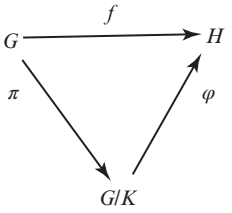
Hence, φ is a homomorphism. We immediately recognize that $\text{Im}\, \varphi = \text{Im}\, f$.

We conclude that φ is an isomorphism. ■

What we are saying with this theorem is that there is no significant difference between a quotient group and its homomorphic image $f(G)$. Furthermore, with the natural map

$$\pi \ : G \to G/K$$

the following diagram is commutative:



that is, $\varphi \circ \pi = f$. Indeed, observe that we can define

$$\varphi^{-1} \ : \text{Im}\, f \ \to G/K$$

by stating: $\forall\, x \in \text{Im}\, f, \exists g \in G$, such that $f(g) = x$ and $\varphi^{-1}(x) = gK$. It is easy to see that if $f(h) = x$, then $gK = hK$, and thus, φ^{-1} is well defined.

Theorem 5.77 Two cyclic groups are isomorphic iff they have the same order.

Proof Let $\langle x \rangle$ and $\langle y \rangle$ be two cyclic groups, such that $|\langle x \rangle| = |\langle y \rangle| = n$. Then, $\langle x \rangle \cong \langle y \rangle$. Consider a homomorphism $f : \mathbf{Z} \to \langle x \rangle$ defined by $f(k) = x^k \ \ k \in \mathbf{Z}$. Observe that φ is certainly surjective. Now, since $|\langle x \rangle| = n$, $\text{Ker}\, f = \langle n \rangle$. By Theorem 5.75, we have

$$\mathbf{Z}/\langle n \rangle \cong \langle x \rangle$$

Similarly,

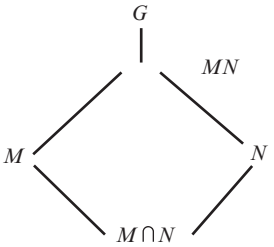
$$\mathbf{Z}/\langle n \rangle \cong \langle y \rangle$$

Thus, $\langle x \rangle \cong \langle y \rangle$. The converse is obvious. ■

Theorem 5.78 (The second isomorphism theorem) Let $M \leq G$, and $N \lhd G$. Then,

$$M/M \cap N \cong MN/N$$

Proof The following diagram is usually used as a mnemonic for "visualizing" the theorem[6]:



Observe that, since $N \lhd G$, $MN = NM$ is also a subgroup of G. Furthermore, $N \lhd MN$. Now consider the surjection

$$\upsilon : M \to MN/N$$

defined by

$$\upsilon(m) = mN, \quad \forall m \in M$$

Note that $v = \pi|_M$, that is, a restriction of the natural homomorphism $\pi : G \to G/N$ to M. So, $\operatorname{Ker} \upsilon = M \cap N$. By the first isomorphism theorem,

$$M/M \cap N \cong MN/N$$ ■

Theorem 5.79 (The third isomorphism theorem) Let $H \lhd G$, and $K \lhd G$ be two normal subgroups of G, such that $K \subseteq H$. Then,

$$(G/K)/(H/K) \cong (G/H)$$

Proof With the first isomorphism theorem in mind, we define

$$\varphi : G/K \to G/H$$

by $\varphi(xK) = xH$. Clearly, the mapping is a well-defined surjection since

$$\begin{aligned} \varphi((xK)(yK)) &= \varphi(xyK) \\ &= xyH \\ &= (xH)(yH) \end{aligned}$$

So, φ is a homomorphism.

[6]Because of this, the theorem is sometimes referred to as the "Diamond Isomorphism Theorem."

Next, suppose

$$xK = yK$$

That implies $x^{-1}y \in K$. Furthermore, since $K \subseteq H, x^{-1}y \in H$, therefore

$$xH = yH$$

that is, $x = yh$ for some $h \in H$.

Finally, consider

$$\begin{aligned} \text{Ker}\,\varphi &= \{xK | \varphi(xK) = eH = H\} \\ &= \{xK | x \in H\} \\ &= H/K \end{aligned}$$

The first isomorphism theorem immediately provides the desired result

$$(G/K)/(H/K) \cong (G/H)$$

■

Example/Exercise 5.135 As a little spin-off of the previous theorem, try to prove that

$$H/K \lhd G/H$$

Example 5.136 Let G be a group, such that for some $n \in \mathbf{Z}^+$, and for all $a, b \in G, (ab)^n = a^n b^n$. Also, define $G_n = \{a \in G | a^n = e\}$ and $G^n = \{a^n | a \in G\}$. Show that

(i) $G_n \lhd G$
(ii) $G^n \lhd G$
(iii) $G/G_n \cong G^n$

Solution

(i) Take $a, b \in G_n$ and $g \in G$. Then,

$$\begin{aligned} (ab^{-1})^n &= a^n (b^{-1})^n \\ &= e \end{aligned}$$

Therefore, $ab^{-1} \in G_n$, that is, $G_n \leq G$. Consider now,

$$\begin{aligned}(gag^{-1})^n &= g^n a^n (g^{-1})^n \\ &= g^n (g^{-1})^n \\ &= (gg^{-1})^n \\ &= e\end{aligned}$$

Thus, $G_n \lhd G$.

(ii) $G^n \lhd G$ can be proved similarly.

(iii) Consider a function $\varphi : G \to G_n$ defined by $\varphi(a) = a^n$. Then,

$$\begin{aligned}\varphi(ab) &= (ab)^n \\ &= a^n b^n \\ &= \varphi(a)\varphi(b), \quad \forall a, b \in G\end{aligned}$$

So, φ is a homomorphism. What about $\operatorname{Ker}\varphi$? Well,

$$\begin{aligned}\operatorname{Ker}\varphi &= \{a \in G | \varphi(a) = a^n = e\} \\ &= G_n\end{aligned}$$

By the first isomorphism theorem, we conclude

$$\begin{aligned}G/G_n &= G/\operatorname{Ker}\varphi \\ &\cong \operatorname{Im}\varphi = G^n\end{aligned}$$

■

Example 5.137 Show that the set $\operatorname{Aut}(G)$ of all automorphisms of a group G is a group under the composition of mappings as the group operation.

Solution Clearly, $\operatorname{Aut}(G) \neq \varnothing$. Let $\varphi, \psi \in \operatorname{Aut}(G)$ be two automorphisms. Then,

$$\begin{aligned}\varphi\psi(xy) &= \varphi(\psi(xy)) \\ &= \varphi(\psi(x)\psi(y)) \\ &= \varphi\psi(x)\varphi\psi(y)\end{aligned}$$

Thus, $\varphi\psi \in \operatorname{Aut}(G)$.

Now we consider

$$\begin{aligned}\varphi(\varphi^{-1}(x)\varphi^{-1}(y) &= \varphi\varphi^{-1}(x)\varphi\varphi^{-1}(y)\\ &= \varphi\varphi^{-1}(xy)\\ &= xy\end{aligned}$$

So,

$$\varphi^{-1}(x)\varphi^{-1}(y) = \varphi^{-1}(xy)$$

implying that $\varphi^{-1} \in \text{Aut}\,(G)$. This proves that Aut (G) is a group. (As a matter of fact, $\text{Aut}\,(G) \leq S_G$.) ■

Example 5.138 Show that the set of all inner automorphisms $\text{In}(G)$ is a normal subgroup of $\text{Aut}\,(G)$.

Solution Recall that by the inner automorphism of a group G, we mean a homomorphism $\iota_g : G \to G$, defined by $\iota_g(x) = gxg^{-1}$, $x, g \in G$. Now, define the map

$$\varphi : G \to \text{Aut}\,(G)$$

by $\varphi(a) = \iota_a$, $\forall a \in G$,
and consider

$$\begin{aligned}\iota_{ab}(x) &= abx(ab)^{-1}\\ &= a(bxb^{-1})a^{-1}\\ &= a\iota_b(x)a^{-1}\\ &= \iota_a(\iota_b(x))\\ &= \iota_a\iota_b(x) \quad \forall x \in G\end{aligned}$$

It follows that φ is a homomorphism, and therefore

$$\text{In}(G) = \text{Im}\,\varphi \leq \text{Aut}\,(G)$$

Is it a normal subgroup? Well, take any $\psi \in \text{Aut}\,(G)$ and consider

$$\begin{aligned}(\psi\ \iota_a\psi^{-1})(x) &= \psi(a\psi^{-1}(x)a^{-1})\\ &= \psi(a)\psi\psi^{-1}(x)\psi(a^{-1})\\ &= \psi(a)x\ \psi(a^{-1})\\ &= \iota_{\psi(a)}(x)\end{aligned}$$

So, $\psi\ \iota_a\psi^{-1} = \iota_{\psi(a)} \in \text{In}(G)$ from which we conclude $\text{In}(G) \lhd \text{Aut}\,(G)$. ■

Example 5.139 Show that

$$G/Z(G) \cong \text{In}(G)$$

Solution Let $\iota_a \in \text{In}(G)$ be an inner automorphism. Consider a homomorphism

$$\varphi : G \to \text{Aut}(G)$$

given by

$$\varphi(a) = \iota_a, \quad \forall a \in G$$

Since φ is a homomorphism, $\text{In}(G) = \text{Im}\,\varphi$ is a subgroup of $\text{Aut}(G)$. Being a group, it has an identity, and $\iota_a = \iota_e$ is the identity automorphism iff $axa^{-1} = x$ for all $x \in G$. Therefore,

$$\text{Ker}\,\varphi = Z(G)$$

But that is exactly what we need in order to apply the first isomorphism theorem, that is,

$$G/Z(G) = G/\text{Ker}\,\varphi \cong \text{Im}\,\varphi = \text{In}(G)$$ ■

5.10 DIRECT PRODUCT OF GROUPS

Suppose we are given two groups G and H. We would like to investigate the possibility of constructing another group as the "product" of G and H. Instinctively, one would consider a product

$$\Pi = G \times H = \{(g, h) | g \in G, h \in H\}$$

and ask how to organize Π to be a group. If $p_1, p_2 \in \Pi$ are any two elements of Π, the natural way to construct the product would be to define it as follows:

$$\begin{aligned} p_1 \cdot p_2 &= (g_1, h_1) \cdot (g_2, h_2) \\ &= (g_1 g_2, h_1 h_2) \end{aligned}$$

Similarly, instead of two groups, we can have a finite family of groups and define the product analogously. So we have

Definition 5.52 Let $G_1, G_2, \ldots, G_n$ be a finite family of groups. We say that

$$\Pi = G_1 \times G_2 \times \cdots \times G_n = \{(g_1, g_2, \ldots, g_n) | g_i \in G_i\}$$

is a **direct product of groups** $G_1, G_2, \ldots, G_n$.

We define multiplication in set Π as follows:

$$(g_1, g_2, \cdots, g_n)(g_1', g_2', \cdots, g_n') = (\, g_1 *_1 g_1', g_2 *_2 g_2', \cdots, g_n *_n g_n' \,)$$

where "$*_i$" indicates "multiplication" performed in the respective group G_i. Although the operations "$*_i$" can be different in each G_i in principle, we shall, as usual, simply write

$$(g_1, g_2, \cdots, g_n)(g_1', g_2', \cdots, g_n') = (g_1 g_1', g_2 g_2', \cdots, g_n g_n')$$

Hence, we have

Theorem 5.80 With the binary operation defined above, Π becomes a group.

Proof For simplicity, let's again consider the direct product of only two groups, G and H:

$$\Pi = G \times H = \{(g, h) | g \in G, h \in H\}$$

Take any $p_1, p_2, p_3 \in \Pi$, such that

$$p_1 = (g_1, h_1),\ p_2 = (g_2, h_2), \quad \text{and} \quad p_3 = (g_3, h_3)$$

Then,

$$\begin{aligned} p_1(p_2 p_3) &= (g_1, h_1)((g_2, h_2)(g_3, h_3)) \\ &= (g_1, h_1)(\, g_2 g_3, h_2 h_3) \\ &= (g_1 g_2 g_3, h_1 h_2 h_3) \\ &= ((g_1 g_2, h_1 h_2))(g_3, h_3) \\ &= (p_1 p_2) p_3 \end{aligned}$$

So, associativity holds. What about the neutral element?

Define $e = e_\Pi = (e_G, e_H)$. Then, for any $p = (g, h) \in \Pi$

$$\begin{aligned} ep &= (e_G, e_H)p \\ &= (e_G, e_H)(g, h) \\ &= (e_G g, e_H h) \\ &= (g, h) \end{aligned}$$

And, finally, defining the inverse as

$$p^{-1} = (g^{-1}, h^{-1})$$

we have

$$\begin{aligned} pp^{-1} &= (g,h)(g^{-1},h^{-1}) \\ &= (gg^{-1}, hh^{-1}) \\ &= (e_G, e_H) \\ &= e \end{aligned}$$

So, Π is a group indeed. An analogous proof can be given for the direct product of n groups. ■

Example 5.140 Consider $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$ with usual addition as the binary operation. It is immediately clear that $(\mathbf{R}^2; +)$ is a group. Similarly, one can consider $(\mathbf{R}^n; +)$. ■

Example 5.141 Consider

$$S^1 \times \mathbf{R} = \{(e^{2\pi \mathrm{i}}, r) | r \in \mathbf{R}\}$$

This group is called the **cylinder group**. ■

Example 5.142 Consider

$$S^1 \times S^1 = \{(e^{2\pi ir}, e^{2\pi is}) | r, s \in \mathbf{R}\}$$

This group is called the **torus group.** ■

Example 5.143

$$\mathbf{Z}_2 \times \mathbf{Z}_3 = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2)\}$$ ■

Example 5.144 Let G and H be two abelian groups. Prove that $G \times H$ is abelian too. ■

Proof Let $(g_1, h_1), (g_2, h_2) \in G \times H$ be any two elements of $G \times H$. Consider

$$\begin{aligned} (g_1, h_1) \cdot (g_2, h_2) &= (g_1 g_2, h_1 h_2) \\ &= (g_2 g_1, h_2 h_1) \\ &= (g_2, h_2) \cdot (g_1, h_1) \end{aligned}$$ ■

Theorem 5.81 Let $\Pi = G \times H$ be a direct product of groups. Then,

(i) G and H are normal subgroups of Π
(ii) $GH = \Pi$
(iii) $G \cap H = \{e_G, e_H\} = \{e\}$

Proof (i) Consider the following two maps:

$$\pi_1 : G \times H \to G, \text{ and } \pi_2 : G \times H \to H$$

defined by $\pi_1(g, h) = g$, and $\pi_2(g, h) = h$. If we could prove that $\operatorname{Ker} \pi_1 \cong H$ and $\operatorname{Ker} \pi_2 \cong G$, then the proof of the theorem will immediately follow.

Let's first prove that π_1 is a homomorphism. (A similar proof works for π_2.) For any $p_1, p_2 \in \Pi = G \times H$

$$\begin{aligned}
\pi_1(p_1 p_2) &= \pi_1((g_1 h_1)(g_2 h_2)) \\
&= \pi_1((g_1 g_2,\ h_1 h_2)) \\
&= g_1 g_2 \\
&= \pi_1(g_1 h_1)\pi_1(g_2 h_2) \\
&= \pi_1(p_1)\pi_1(p_2)
\end{aligned}$$

So, π_1 is a homomorphism. Next,

$$\begin{aligned}
\operatorname{Ker} \pi_1 &= \pi_1^{-1}(e_G) = \{(e_G, h) | h \in H\} \\
&= \{e_G\} \times H
\end{aligned}$$

But, clearly, $\{e_G\} \times H \cong H$, thus $\operatorname{Ker} \pi_1 \cong H$. Consequently, $\operatorname{Ker} \pi_1 \cong H$ and $\operatorname{Ker} \pi_2 \cong G$ imply that $G,\ H \lhd \Pi$.

Claims (ii) and (iii) are obvious. ■

Theorem 5.82 Let $A \lhd G$ and $B \lhd H$, then

(i) $A \times B \lhd G \times H$
(ii) $(G \times H)/(A \times B) \cong (G/A) \times (H/B)$

Proof Observe that the homomorphism

$$\varphi : G \times H \to (G/A) \times (H/B)$$

defined by $\varphi(g,h) = (gA, hB)$ is surjective, and

$$\begin{aligned}\operatorname{Ker}\varphi &= \varphi^{-1}(e_G A, e_H B)\\ &= A \times B\end{aligned}$$

so (i) immediately follows.

The first isomorphism theorem implies (ii). ■

Theorem 5.83 Let G, H, and K be three groups. Then,

(i) $G \times H \cong H \times G$
(ii) $G \times (H \times K) \cong (G \times H) \times K$

Theorem 5.84 Let $\varphi : G \to H$ and $\psi : G \to K$ be two homomorphisms of groups. Then the mapping

$$\chi : G \to H \times K$$

defined by

$$\chi(g) = (\varphi(g), \psi(g))$$

is also a homomorphism, called the **direct product of homomorphisms.**

5.11 SUPPLEMENTARY PROBLEMS

1. Determine whether the following are groups:
(i) $(\mathbf{R}; -)$
(ii) $(\mathbf{R}; +)$
(iii) $(\mathbf{Z}; +)$
(iv) $(\mathbf{Z}^+; +)$
(v) $(\mathbf{R} \setminus \{0\}; \cdot)$
(vi) $(\mathbf{Z} \setminus \{0\}; \cdot)$
(vii) $(\mathbf{R}^+ = \{x \in \mathbf{R} | x > 0\}; \cdot)$

2. Determine whether each set with the indicated operation forms a group.
(i) $(\mathrm{N} \cup \{0\}; a \ominus b = |a - b|)$
(ii) $(\mathbf{R}; \circledast =$ average of two numbers$)$
(iii) $(\mathbf{R}; a \oslash b = a/b)$

3. Let $S = \{a, b, c, d\}$ and the binary operation " $*$ " on S be partially given by the table

$*$	a	b	c	d
a	d	c		
b			a	
c				
d				

Complete the table and determine whether $(S; *)$ is a group.

4. Determine whether $(\mathbf{R};\ *)$ with the operation $*$ defined by

$$a * b = a + b + k, \quad k = \text{const}$$

is a group.

5. Let $G = \{a | a \in \mathbf{R}, -1 < a < 1\ \}$ and let the operation $*$ be defined by

$$a * b = \frac{a + b}{ab + 1}, \quad a, b \in G$$

Determine whether $(G; *)$ is a group or not.

6. Let $(G; *)$ be a group, and $a, b, c \in G$.

(i) Find the inverse of $a * (b * c)$.

(ii) Show that if $a * b * a * b = e$, then $b * a * b * a = e$.

7. Let G be a group, and let $a, b, c \in G$ be some elements of G. Show that if $(abc)(abc) = e$, then $(bca)(bca) = e$, and $(cab)(cab) = e$, that is, bca is its own inverse, and cab is its own inverse.

8. Prove: There exists no group containing exactly two elements that do not commute.

9. Given set $G = \{p_1, p_2, p_3, p_4\}$ where

$$p_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix},$$

$$p_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}, \text{ and } p_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}$$

show that G under composition is an abelian group.

10. Prove that the following permutations form a V-group:

$$(1); \quad \begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}3 & 4\end{pmatrix}; \quad \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 4\end{pmatrix}; \quad \begin{pmatrix}1 & 4\end{pmatrix}\begin{pmatrix}2 & 3\end{pmatrix}$$

11. Let $\mathbf{R}^n$ be the set of all ordered n-tuples, that is,

$$\mathbf{R}^n = \mathbf{R} \times \cdots \times \mathbf{R} = \{(x_1, \ldots, x_n) | x_i \in \mathbf{R}\}$$

Show that, with usual componentwise addition,

$$\begin{aligned} x + y &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) \\ &= (x_1 + y_1, \ldots, x_n + y_n) \end{aligned}$$

$(\mathbf{R}^n; +)$ is an abelian group.

12. Prove that a group G is abelian iff $(ab)^{-1} = a^{-1}b^{-1}, \ \forall a, b \in G$.

13. Let X be a given set. Show that $(\mathrm{P}(X); \Delta)$ is an abelian group. (Hint: $P(X)$ is a power set, and Δ indicates symmetric difference; see Chapter 1.)

14. Let $P_n = \{p \mid \deg p < n, n \in \mathbf{N}\} \cup \{0\}$be a set of all polynomials $p : \mathbf{R} \to \mathbf{R}$ defined in the usual way by

$$p(x) = \sum_{k=0}^{m} a_k x^k, \quad a_k \in \mathbf{R}$$

whose degree is less than $n \in \mathbf{N}$. Show that $(P_n; +)$ is a group.

15. Determine whether $P = \cup_{n=0}^{\infty} P_n$, with the usual addition of polynomials, is a group.

16. Prove that a set $T = \{3^x | x \in \mathbf{Z}\}$, with usual multiplication, is a group.

17. Consider a set $O = (-1, 1) = \{x \in \mathbf{R} | -1 < x < 1\}$. Show that under the binary operation "$*$" defined by

$$x * c = \frac{x + c}{1 + cx}$$

$(\mathrm{O}; *)$ is a group.

18. Let P be a polygon in the plane and let

$$\Pi(P) = \{\pi : P \to P | \ \pi \text{ is an isometry}\}$$

be a set of all isometric congruencies of a polygon P. Show that (P) is a group.

19. Let G be a group, such that $g \in G$. Show that $\{g, e\} \leq G$ iff $g = g^{-1}$.

20. Let H and K be two subgroups of G. Determine whether (i) $H \cup K$ and (ii) $H \cup K$ are subgroups of G.

21. Let H and K be two subgroups of G. Show that if $H \subseteq K$, then $H \leq K$.

22. Let $H, K \leq G$. Show that HK defined as follows:

$$HK = \{ab \mid a \in H, b \in K\}$$

is a subgroup of G.

23. Let $a, b \in \mathbf{Z}$, such that $a \neq 0$, $b \neq 0$, and let $= \{ma + nb | m, n \in \mathbf{Z}\}$ be the set of all linear combinations of a and b. Show that $L \leq (\mathbf{Z}; +)$.

24. Determine whether a set $H = \{0, 1, 2, 3\}$ is a subgroup of $\mathbf{Z}_8$.

25. Let $\mathbf{Z}$ be endowed with the binary operation $a \oplus b = a + b - 1$. Show that $(\mathbf{Z}; \oplus)$ is a group and determine whether $(\mathbf{Z}_{\text{odd}}; \oplus)$ is a subgroup of $(\mathbf{Z}; \oplus)$.

26. Consider groups $(\mathcal{F}(\mathbf{R}); +)$ and $(F; +)$ where

$$F = \{f \in \mathcal{F}(\mathbf{R}) | f(-x) = -f(x)\}$$

Show that $(F; +) \leq (\mathcal{F}(\mathbf{R}); +)$.

27. Consider the following sets

$$G = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a, b, c, d \in \mathbf{Z} \right\}$$

and

$$H = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G \mid a + b + c + d = 0 \right\}$$

Show that $(H; +) \leq (G; +)$.

28. Given a group $(\mathbf{R} \setminus \{0\}; \cdot) = (\mathbf{R}^*; \cdot)$, find all finite subgroups.

29. Show that a mapping $\varphi : (G; *) \to (G; *)$ defined by $\varphi(a) = a^2$ is a homomorphism iff G is abelian.

30. Let $\varphi : G \to H$ be a homomorphism of groups and let $A \leq G$, then the restriction $\varphi|_A : A \to H$ is also a homomorphism.

31. Show that a mapping $\varphi : (\mathbf{R}^+; \cdot) \to (\mathbf{R}^+; \cdot)$ defined by $\varphi(x) = \sqrt{x}$ is an automorphism.

32. Consider a mapping $\varphi : (\mathcal{F}(\mathbf{R}); +) \to (\mathbf{R}; +)$ defined by $\varphi(f) = f(0)$. Show that φ is a homomorphism.

33. If G is a group, show that a mapping $\varphi : G \to G$ defined by $\varphi(x) = x^{-1}$ is an automorphism iff G is abelian.

34. Prove that a homomorphism of groups $\varphi : G \to H$ is an injection iff $\operatorname{Ker}\varphi = e$.

35. Let α and β be two homomorphisms. Show that $\beta\circ\alpha$ is a trivial homomorphism iff $\operatorname{Im}\alpha \subseteq \operatorname{Ker}\beta$.

36. Show that a function $\varphi : \mathbf{R} \to \mathrm{S}^1$ defined by $\varphi(x) = e^{2\pi ix}$ is an epimorphism.

37. Let $\alpha : G \to H$ and $\beta : H \to K$ be two homomorphisms of groups. Show that if $\gamma = \beta \circ \alpha$ is an epimorphism, then β is an epimorphism too.

38. Show that $(\mathbf{R}; +) \cong (\mathbf{R}^+; \cdot)$, that is, show that a group of all real numbers with addition as the binary operation is isomorphic to the group of positive real numbers with multiplication as a binary operation.

39. Determine whether $(\mathbf{Z}; +)$ is isomorphic to $(\mathbf{Q}; +)$.

40. Determine whether $(\mathbf{Q}; +)$ is isomorphic to $(\mathbf{R}; \cdot)$.

41. Let $(G; *)$ be a group. Define another operation $\otimes$ on G, such that $a \otimes b = a * b, \quad \forall a, b \in G$. Show that $(G; \otimes)$ is a group isomorphic to $(G; *)$.

42. Consider a set $G = \{10^n | n \in \mathbf{Z}\}$ and a group $(G; \cdot)$. Show that $(G; \cdot) \cong (\mathbf{Z}; +)$.

43. Show that $\varphi : \mathbf{Z} \to m\mathbf{Z}$ is an isomorphism of groups.

44. Let G be a group. Show that the identity function $\imath : G \to G$, defined by $\imath(x) = x, \quad \forall x \in G$ is an isomorphism.

45. Consider a set $G = \{x + y\sqrt{2} \,|x, y \in \mathbf{Q}\}$ and a set

$$M = \left\{ \begin{bmatrix} x & 2y \\ y & x \end{bmatrix} \,|x, y \in \mathbf{Q} \right\}$$

Show that $(G; +) \cong (M; +)$.

46. Let $H \leq G$ be any subgroup of G. Show that $H \cong gHg^{-1}$ for any $g \in G$.

47. Show that the alternating group A_n is a normal subgroup of the symmetric group S_n.

48. Show that a commutator subgroup $[K] \leq G$ is a normal subgroup of G.

49. Consider the inner automorphism $\varphi_g : G \to G$ defined by $\varphi_g(x) = gxg^{-1}$. Show that φ_g is an isomorphism.

50. Show that if $\varphi : G \to H$ is an isomorphism of groups, and G is a cyclic group with generator a, then H is a cyclic group with generator $\varphi(a)$.

51. Show that $H \lhd G$ iff $\varphi_g(H) = H$ for every inner automorphism of a group G.

52. Let $A \leq Z(G)$ be any subgroup of the center of a group $Z(G)$. Show that $A \lhd G$.

53. Let $N \lhd G$ be a normal subgroup, and $H \leq G$ any subgroup of G. Show that NH is a subgroup of G.

54. Let $H, K \lhd G$. Show that $HK \lhd G$.

55. Let $H, K \lhd G$. Show that $HK = G$ and $H \cap K = e$ iff each $g \in G$ has a unique expression of the form $g = hk$, where $h \in H$ and $k \in K$.

56. Let $H, K \leq G$. Show that if $G = HK$, then $hk = kh$ for all $h \in H$ and $k \in K$.

57. Let $H, K \lhd G$. Prove that if $G = HK$ and $H \cap K = \{e\}$, then $G \cong H \times K$.

58. Let $\varphi : G \to H$ be a homomorphism of groups, and let $K \leq H$. Show that

$$\varphi^{-1}(K) = \{x \in G | \varphi(x) \in K\} \leq G$$

Furthermore, show that $\operatorname{Ker} \varphi \subseteq \varphi^{-1}(K)$.

59. Show that $\mathbf{Z}/m\mathbf{Z} \cong \mathbf{Z}m$.

60. Show that $\boldsymbol{V} \cong \mathbf{Z}_2 \times \mathbf{Z}_2$.

6

LINEAR ALGEBRA

6.1 INTRODUCTION

Linear Algebra is probably one of the most traditional abstract mathematical disciplines central to both pure and applied mathematics. Historically, it originates from the problem of finding solutions to a system of linear equations. Being "linear" suggests its intuitive simplicity, and soon enough people found a way to combine it with other mathematical disciplines and apply it to a number of other fields, starting from physics to the social sciences. Not surprisingly, Linear Algebra became a fascinating field of its own.

Before formally presenting the topic, let me append to this introduction something that may look rather familiar to many a reader and still hopefully provoke curiosity and interest for more insightful discussion.

First, by a **linear equation** with n unknowns $x_1, x_2, \ldots, x_n$, we mean an equation

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta \tag{6.1}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$, are called the coefficients (real, $\alpha_i \in \mathbf{R}$, or complex numbers, $\alpha_i \in \mathbf{C}$) of the x_i's and β (real or complex number) is called the constant term (sometimes, the free term). The solution of equation (6.1) is also a set of numbers from $\mathbf{R}$ (or $\mathbf{C}$)

$$x_1 = k_1,\ x_2 = k_2,\ \ldots,\ x_n = k_n$$

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

such that equation (6.1) is *satisfied*, that is the following statement is true:

$$\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_n k_n = \beta$$

A linear equation is said to be *degenerate* if all coefficients are zero, that is

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

So equation (6.1) becomes

$$0x_1 + 0x_2 + \cdots + 0x_n = \beta$$

Obviously, in that case any selection of $x_1, x_2, \ldots, x_n$ would satisfy equation (6.1) if $\beta = 0$, and we would say that our equation has no solution if $\beta \neq 0$.

Suppose now that you want to expand on your (usual) high school knowledge of solving a system of one, two, or three linear equations with a respective number of unknowns. Say you wonder how to solve a system of 4, 5, maybe 15 or, why not, 55 equations. In other words, you would like to find a method of solving a system of n linear equations with n variables (unknowns):

$$\left.\begin{array}{rc} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \cdots + \alpha_{1n}x_n = \beta_1 \\ \alpha_{21}x_1 + \alpha_{21}x_2 + \alpha_{23}x_3 + \cdots + \alpha_{2n}x_n = \beta_2 \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \alpha_{n3}x_3 + \cdots + \alpha_{nn}x_n = \beta_n \end{array}\right\} \tag{6.2}$$

In our expression (6.2), the α'_{ij}s are coefficients of x'_is; the x'_is are unknowns whose values are to be determined, and the β'_is are free terms. Suppose someone recognizing the severity of the problem comes with an ingenious idea and suggests to rewrite the system (6.2) by collecting all the coefficients in one "box," the unknowns in another, and finally the free terms in yet another "box," so that our system (6.2) looks as follows:

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \tag{6.3}$$

To make things even simpler, let's denote the "boxes" in (6.3) by A, X, and B, respectively, call them **matrices** and simply write everything as

$$AX = B \tag{6.4}$$

Equation (6.4) looks like a simple linear equation in one variable X, and it is rather tempting to say we know how to solve it. It's easy:

$$X = A^{-1}B$$

Well, that would be true if A, B, and X were "regular" objects (numbers), but they are not. Remember, they are "boxes" of numbers. So the question is this: How do we operate with those? Can we detect the general underlining structure those objects obey? Can we discover the rules that determine how to manipulate them? And, finally, in addition to solving the system (6.2), we started with, we ask: is it possible that some other mathematical objects fit into this scheme? So, here it goes.

6.2 VECTOR SPACE

Definition 6.1 Let $\boldsymbol{X} = \{ \boldsymbol{x}_1, \boldsymbol{x}_2, \dots \}$ be a set whose elements we call **vectors**,[1] such that $\boldsymbol{X} = (\boldsymbol{X}; +)$ is an additive abelian group, and let $\Phi = \{\alpha, \beta, \dots\}$ be a set whose elements we call **scalars**,[2] such that $(\Phi; +)$ is a field. Next, we define a mapping

$$f : \boldsymbol{X} \times \Phi \to \boldsymbol{X}$$

by $f(\alpha, \boldsymbol{x}) = \alpha\boldsymbol{x}$, for all $\alpha \in \Phi$, $\boldsymbol{x} \in X$. We want the following axioms to hold:

A.1 $\boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{x}_2 + \boldsymbol{x}_1, \ \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ ($\boldsymbol{X}$ is abelian group)

A.2 $(\boldsymbol{x}_1 + \boldsymbol{x}_2) + \boldsymbol{x}_3 = \boldsymbol{x}_1 + (\boldsymbol{x}_2 + \boldsymbol{x}_3), \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in \boldsymbol{X}$

A.3 There is unique $\mathbf{0} \in \boldsymbol{X}$, called the additive identity, such that $\mathbf{0} + \boldsymbol{x} = \mathbf{0} + \boldsymbol{x} = \boldsymbol{x}, \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}$

A.4. For each $\boldsymbol{x} \in \boldsymbol{X}$, $\exists(-\boldsymbol{x}) \in \boldsymbol{X}$, called the additive inverse, such that $\boldsymbol{x} + (-\boldsymbol{x}) = (-\boldsymbol{x}) + \boldsymbol{x} = \mathbf{0}$

A.5. $\alpha(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \alpha\boldsymbol{x}_1 + \alpha\boldsymbol{x}_2, \ \forall \alpha \in \Phi; \ \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$

A.6 $(\alpha + \beta)\boldsymbol{x} = \alpha\boldsymbol{x} + \beta\boldsymbol{x}, \ \forall \alpha, \beta \in \Phi; \ \forall \boldsymbol{x} \in \boldsymbol{X}$

A.7 $(\alpha\beta) = \alpha(\beta\boldsymbol{x}), \quad \forall\, \alpha, \beta \in \Phi; \ \forall \boldsymbol{x} \in \boldsymbol{X}$

A.8 $\exists 1 \in \Phi$ such that $1\boldsymbol{x} = \boldsymbol{x}1 = \boldsymbol{x}$

The triple $(X, \Phi, f) = \mathrm{V}$ satisfying A.1–A.8 we call a **vector** or a **linear space** over the field Φ.

[1] From now on we will use bold faced lower case Latin letters to designate vectors, and bold capitals to designate vector spaces.

[2] We will use the lower case Greek alphabet to designate scalars, and capitals for the corresponding fields.

NOTE:

1. From A.1–A.8, and the definition of the function f, it follows that if $\boldsymbol{x} \in \boldsymbol{X}$ is a vector and $\alpha \in \Phi$ is a scalar, then $\alpha\boldsymbol{x} \in$ V. Often, however, we simply say that the set X is a vector space if two algebraic operations – *vector addition* and *scalar multiplication* – are defined on it.
2. Observe that the conditions in the definition are not logically independent. For instance, the request for $\boldsymbol{X}$ to be an abelian group is redundant, in a sense, because it also follows from the axioms.
3. Once again, let's agree that from now on we will use bold-faced Latin letters for vectors and regular print Greek letters for scalars.

Example 6.1 R $= (\mathbf{R}, \mathbf{R}, +, \cdot)$, is a space of real numbers over real numbers. We take $x, y \in \mathbf{R}$, and the operation "+" is standard addition of real numbers. Also, take $\alpha \in \mathbf{R}$, and $\alpha\boldsymbol{x}$ is just multiplication of a real number. ■

Similarly,

Example 6.2 Consider a set $\mathbf{R}^{\boldsymbol{n}} = \{ (\alpha_1, \ldots, \alpha_n) | \alpha_i \in \mathbf{R} \}$
With standard component addition

$$(\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$$

and multiplication

$$\alpha(\alpha_1, \ldots, \alpha_n) = (\alpha\alpha_1, \ldots, \alpha\alpha_n)$$

we immediately recognize that axioms A.1–A.8 are satisfied, and we have a vector space $(\mathbf{R}^n, \mathbf{R}, +, \cdot)$. We call this space the **Cartesian *n*-space**. ■

To facilitate our intuition, let's consider $\mathbf{R}^2$, whose elements could be thought of as points or as arrows, as is illustrated in the following example.

Example 6.3 A typical element of $\mathbf{R}^2$ is a point $\boldsymbol{x} = (x_1, x_2)$ as in Figure 6.1.
Equally, we can think of an element of $\mathbf{R}^2$ as an arrow (a vector) (Figure 6.2).
It is easy to see that axioms A.1–A.8 are satisfied. With addition as defined in the previous example, the sum of two vectors is pictured in Figure 6.3.
Observe that we have moved vector $\boldsymbol{y}$ parallel to itself without changing its direction and length, so that its "tail" coincides with the "tip" of vector $\boldsymbol{x}$. (Doing this we have not changed the identity of vector $\boldsymbol{y}$.) Finally, we connected the "tail" of $\boldsymbol{x}$ with the "tip" of $\boldsymbol{y}$, thus obtaining the sum $\boldsymbol{x} + \boldsymbol{y}$. Examining the picture

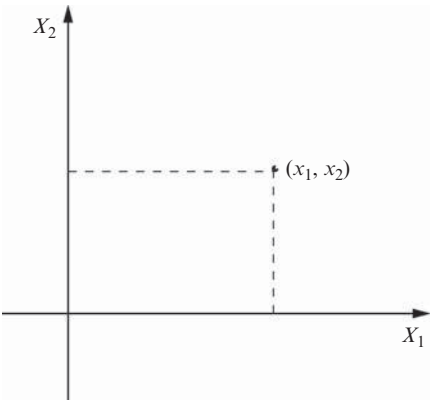


Figure 6.1

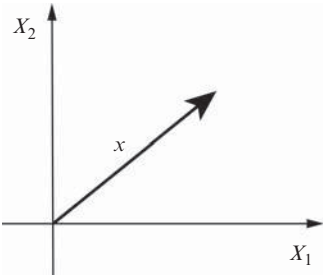


Figure 6.2

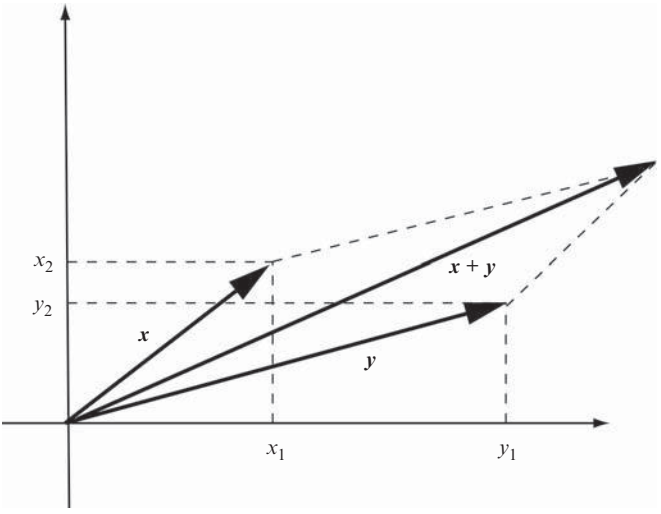


Figure 6.3

carefully, one can convince oneself that the sum of two vectors is exactly as desired, namely,

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= (x_1, x_2) + (y_1, y_2) \\ &= (x_1 + y_1,\ x_2 + y_2) \end{aligned}$$

We can dispense of the coordinate axes to make the picture less cluttered, and simply conceive the sum of two vectors as follows (Figure 6.4):

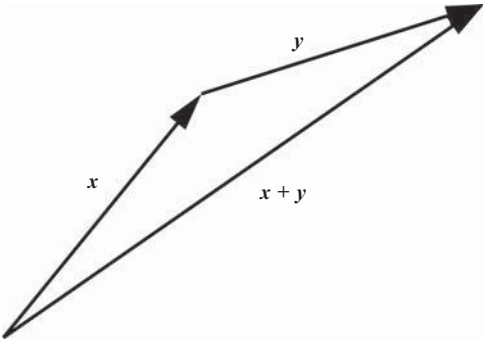


Figure 6.4

The multiplication of a vector $\boldsymbol{x}$ by a *scalar* looks something like this: λ:
Then, of course, with $\boldsymbol{x}$ from Figure 6.5, $-\boldsymbol{x} = (-1)\boldsymbol{x}$ is as in Figure 6.6

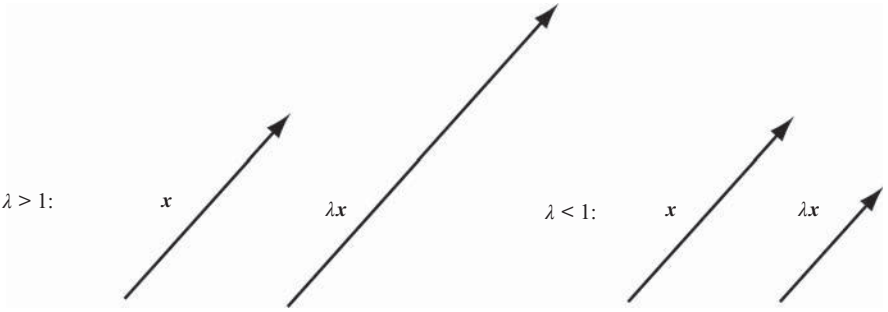


Figure 6.5

Figure 6.6

One should keep in mind that the pictures above are just aids to help us build our intuition for a better understanding of vectors and vector spaces. These pictures in no way represent more abstract vectors, in particular, vectors of n-dimensional spaces (Figure 6.6). ■

Example/Exercise 6.4 Convince yourself that $\mathbf{R}^3$, with the usual addition and multiplication, is a vector space.

Example 6.5 To define the n-dimensional analogues of $\mathbf{R}^3$, take Φ to be an arbitrary field. We can design the set of all n-tuples of elements in Φ, namely

$$\Phi^n = \{\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_n) | \ \phi_i \in \Phi\}$$

If we endow this set with addition defined by

$$\begin{aligned}\boldsymbol{\phi} + \boldsymbol{\phi'} &= (\phi_1, \phi_2, \ldots, \phi_n) + (\phi'_1, \phi'_2, \ldots, \phi'_n) \\ &= (\phi_1 + \phi'_1, \phi_2 + \phi'_2, \ldots, \phi_n + \phi'_n), \quad \forall \phi_i, \phi'_i \in \Phi\,,\end{aligned}$$

and scalar multiplication defined by

$$\begin{aligned}\lambda\boldsymbol{\phi} &= \lambda(\phi_1, \phi_2, \ldots, \phi_n) \\ &= (\lambda\phi_1, \lambda\phi_2, \ldots, \lambda\phi_n), \quad \forall\ \lambda, \phi_i \in \Phi\end{aligned}$$

and with an appropriately chosen "zero," say, $\mathbf{0} = (0, 0, \ldots, 0)$, and the "negative vector"

$$-\boldsymbol{\phi} = -(\phi_1, \phi_2, \ldots, \phi_n) = (-\phi_1, -\phi_2, \ldots, -\phi_n)$$

we obtain a vector space $\boldsymbol{\Phi}^n$. ■

Example 6.6 Let $P_n = \{p(x) | \ p$ is a polynomial, $\deg p \leq n \in \mathbf{N}\} \cup \{0\}$ be the set of all polynomials in one variable with real coefficients. Then, P_n with the usual addition of polynomials, and multiplication of polynomials by a real number, is a vector space. ■

Example 6.7 Let S be any set, and let $\mathbf{F}(S)$ be a set of all functions defined on S with values on some field Φ. With the usual addition of functions, and multiplication of functions by $\alpha \in \Phi$, we obtain a vector space $(\mathbf{F}, \Phi, +, \cdot)$. ■

Theorem 6.1 Let $\boldsymbol{X}$ be a vector space over a field Φ. Then,

(i) $0\boldsymbol{x} = \mathbf{0} \in \boldsymbol{X}, \quad \forall \boldsymbol{x} \in \boldsymbol{X},\ 0 \in \Phi$
(ii) $\alpha\mathbf{0} = \mathbf{0}, \quad \forall \alpha \in \Phi$
(iii) If $\alpha\boldsymbol{x} = \mathbf{0}$, then either $\alpha = 0$ or $\boldsymbol{x} = \mathbf{0}$

Proof

(i)
$$\begin{aligned} \boldsymbol{x} + \mathbf{0} &= \boldsymbol{x} \\ &= 1\boldsymbol{x} \\ &= (1+0)\boldsymbol{x} \\ &= 1\boldsymbol{x} + 0\boldsymbol{x} \\ &= \boldsymbol{x} + 0\boldsymbol{x} \end{aligned}$$

which implies $0\boldsymbol{x} = \mathbf{0}$.

(ii) Take any $\boldsymbol{x} \in \boldsymbol{X}$, and $\alpha \in \Phi$ and consider

$$\begin{aligned} \mathbf{0} + \alpha\boldsymbol{x} &= \alpha\boldsymbol{x} \\ &= \alpha(\mathbf{0} + \boldsymbol{x}) \\ &= \alpha\mathbf{0} + \alpha\boldsymbol{x} \end{aligned}$$

which implies $\alpha\mathbf{0} = \mathbf{0}$.

(iii) If $\alpha = 0$ there is nothing to prove. Suppose $\alpha \neq 0$. Since Φ is a field, there exists $\alpha^{-1} \in \Phi$, so

$$\begin{aligned} \alpha^{-1}(\alpha\boldsymbol{x}) &= \mathbf{0} \\ &= (\alpha^{-1}\alpha)\boldsymbol{x} \\ &= 1\boldsymbol{x} \\ &= \boldsymbol{x} \end{aligned}$$

Hence $\boldsymbol{x} = \mathbf{0}$, and that completes the proof. ■

Example/Exercise 6.8 Show that $(\alpha - \beta)\boldsymbol{x} = \alpha\boldsymbol{x} - \beta\boldsymbol{x}, \quad \forall \alpha, \beta \in \Phi, \ \boldsymbol{x} \in \boldsymbol{X}$.

Example/Exercise 6.9 Let $\boldsymbol{X}$ be a vector space over a field Φ, and let $\boldsymbol{x} \in \boldsymbol{X}$ and $\alpha \in \Phi$. Show that $\alpha\boldsymbol{x} = \boldsymbol{x}$ iff $\alpha = 1$ or $\boldsymbol{x} = 0$.

Example 6.10 Consider the set of all positive real numbers $\mathbf{R}^+ = X$, and define the "addition" of "vectors" as follows:

$$x + y = xy, \quad \forall x, y \in \mathbf{R}^+$$

where the product on the right side is just the usual product of numbers. Furthermore, define scalar multiplication by

$$\alpha x = x^\alpha$$

that is, a number x raised to the αth power. With these definitions, X becomes a vector space. ■

Example 6.11 Let $C[a, b] = \mathbf{R}_{[a,b]} = \{f | \ f : [a, b] \to \mathbf{R}\}$ be the set of all real and continuous functions on $[a, b]$. As with the addition of vectors, we define the usual addition of functions, that is, for all $x \in [a, b]$, and for all $f, g \in C[a, b]$

$$(f + g)(x) = f(x) + g(x)$$

For scalar multiplication we take

$$(\alpha f)(x) = \alpha f(x)$$

With these definitions, $C[a, b]$ becomes a vector space. ■

Theorem 6.2 A vector space has a *unique* additive identity.

Proof Let X be a vector space and suppose $\mathbf{0}, \mathbf{0}' \in X$ are two additive identities. Then,

$$\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}'$$

Thus, $\mathbf{0} = \mathbf{0}'$. ■

Theorem 6.3 Every element in a vector space has a *unique* additive inverse.

Proof Let X be a vector space and let $x \in X$ be any vector. Suppose $(-x)$ and $(-x')$ are two inverses of x. Then,

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

Thus, $(-x) = (-x')$, as desired. ■

Theorem 6.4 For every $\boldsymbol{x} \in \boldsymbol{X}$, $-\boldsymbol{x} = (-1)\boldsymbol{x}$, where $(-1) \in \Phi$.

Proof Let $\boldsymbol{x} \in \boldsymbol{X}$ be any vector. Then,

$$\begin{aligned} \boldsymbol{x} + (-1)\boldsymbol{x} &= 1\boldsymbol{x} + (-1)\boldsymbol{x} \\ &= (1 + (-1))\boldsymbol{x} \\ &= 0\boldsymbol{x} = \mathbf{0} \end{aligned}$$

Thus, $(-1)\boldsymbol{x}$ must be the additive inverse of $\boldsymbol{x}$. ■

Consequently, by $\boldsymbol{x} - \boldsymbol{y}$ we mean $\boldsymbol{x} + (-\boldsymbol{y})$ for every $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$.

Example 6.12 Let $S \subseteq \mathbf{R}^2$ be a set of all $(x, y) \in \mathbf{R}^2$ such that $x \geq 0$ and $y \geq 0$. Then, with ordinary operations in $\mathbf{R}^2$, S *is not* a vector (sub)space of $\mathbf{R}^2$.

Consider, for instance, $\boldsymbol{x} = (1, 1) \in \mathbf{R}^2$ and $\lambda = -1 \in \mathbf{R}$. Then, $\lambda\boldsymbol{x} = (-1, -1) \notin S$. Thus, since S is not closed under scalar multiplication, it is not a vector space. ■

6.3 LINEAR DEPENDENCE AND INDEPENDENCE

Definition 6.2 Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in X$ be a collection of vectors from a vector space X over a field Φ. We say that

$$\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_k\boldsymbol{x}_k, \qquad \alpha_i \in \Phi$$

is a **linear combination** of vectors $\boldsymbol{x}_i, \ldots, \boldsymbol{x}_k$. We say that the linear combination is **trivial** if all α_i are equal to zero.

Definition 6.3 Let $E = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ be a **finite** set of distinct vectors from X. We say that set E is **linearly independent** if

$$\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_k\boldsymbol{x}_k = \mathbf{0} \qquad (*)$$

with all $\alpha_i = 0$. In other words, set E is linearly independent iff equation (*) is obtainable only **trivially**. Consequently, we have

Definition 6.4 The set of vectors E that is not linearly independent is said to be **linearly dependent**. In other words, the set of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ is linearly dependent, if there is at least one α_i from equation (*) different from zero.

To illustrate this, let's suppose that in equation (*) $\alpha_j \neq 0$. Then, we can write

$$\alpha_j\boldsymbol{x}_j = -\alpha_1 x_1 - \cdots \alpha_{j-1}\boldsymbol{x}_{j-1} - \alpha_{j+1}\boldsymbol{x}_{j+1} - \cdots - \alpha_k\boldsymbol{x}_k$$

and therefore,

$$\boldsymbol{x}_j = \left(-\frac{\alpha_1}{\alpha_j}\right)\boldsymbol{x}_1 + \cdots + \left(-\frac{\alpha_k}{\alpha_j}\right)\boldsymbol{x}_k$$

that is, $\boldsymbol{x}_j$ is expressible as a linear combination of the remaining vectors from (*).

Example 6.13 Let $\boldsymbol{x}_1 = (1, 0)$ and $\boldsymbol{x}_2 = (0, 1)$ be two vectors from $\mathbf{R}^2$. We claim that the set $\{ \boldsymbol{x}_1, \boldsymbol{x}_2 \}$ is a set of linearly independent vectors. Let's see.

Consider

$$\begin{aligned}\alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 &= \mathbf{0}\\ &= \alpha_1(1, 0) + \alpha_2(0, 1)\\ &= (\alpha_1, 0) + (0, \alpha_2)\\ &= (\alpha_1, \alpha_2)\\ &= (0, 0)\end{aligned}$$

Since the zero vector $\mathbf{0}$ is expressible as a linear combination of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2$ only trivially, we conclude that $\boldsymbol{x}_1, \boldsymbol{x}_2$ are linearly independent. ■

Example/Exercise 6.14 Show that the set of vectors $\{(1, 1, 1), (0, 1, 1), (0, 0\ 1)\}$ are linearly independent.

Example 6.15 Determine whether the vectors $\boldsymbol{x}_1 = (2, 3)$, $\boldsymbol{x}_2 = (4, 6)$ from R^2 are linearly dependent or not.

Solution Consider

$$\alpha_1\,\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 = \mathbf{0}$$

that is,

$$\begin{aligned}\alpha_1(2, 3) + \alpha_2(4, 6) &= (0, 0)\\ &= (2\alpha_1, 3\alpha_1) + (4\alpha_2, 6\alpha_2)\\ &= (2\alpha_1 + 4\alpha_2, 3\alpha_1 + 6\alpha_2)\end{aligned}$$

Thus,

$$\begin{aligned}2\alpha_1 + 4\alpha_2 &= 0\\ 3\alpha_1 + 6\alpha_2 &= 0\end{aligned}$$

Observe that the above system has a trivial solution $\alpha_1 = \alpha_2 = 0$ but also a nontrivial one, say, $\alpha_1 = -2,\ \alpha_2 = 1$, hence vectors $\boldsymbol{x}_1 = (2,\ 3),\ \boldsymbol{x}_2 = (4,\ 6)$ are linearly dependent. ■

Example 6.16 Determine whether vectors $\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \boldsymbol{x}_3 \in \mathbf{R}^4$ defined below are linearly dependent or not.

$$\boldsymbol{x}_1 = (1,\ 2,\ 2,\ -1), \quad \boldsymbol{x}_2 = (4,\ 9,\ 9,\ -4), \quad \boldsymbol{x}_3 = (5,\ 8,\ 9,\ -5)$$

Solution Consider the linear combination

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \alpha_3 \boldsymbol{x}_3 = \mathbf{0}$$

that is,

$$\alpha_1(1,\ 2,\ 2,\ -1) + \alpha_2(4,\ 9,\ 9,\ -4) + \alpha_3(5,\ 8,\ 9,\ -5) = (0,\ 0,\ 0,\ 0)$$

We need to determine whether the corresponding system of linear equations has a trivial or nontrivial solution:

$$\begin{aligned} \alpha_1 + 4\alpha_2 + 5\alpha_3 &= 0 \\ 2\alpha_1 + 9\alpha_2 + 8\alpha_3 &= 0 \\ 2\alpha_1 + 9\alpha_2 + 9\alpha_3 &= 0 \\ -\alpha_1 - 4\alpha_2 - 5\alpha_3 &= 0 \end{aligned}$$

With some algebra, you can convince yourself that this system has only a trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = 0$, thus our $\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \boldsymbol{x}_3$ are linearly independent. ■

Example 6.17 Let the set $S = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ be the set of linearly independent vectors. Show that the set $\overline{S} = \{\boldsymbol{x}_1 + \boldsymbol{x}_2,\ \boldsymbol{x}_1 - \boldsymbol{x}_2\}$ is also linearly independent.

Solution Consider the linear combination

$$\alpha(\boldsymbol{x}_1 + \boldsymbol{x}_2) + \beta(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \mathbf{0}$$

from which follows

$$(\alpha + \beta)\,\boldsymbol{x}_1 + (\alpha - \beta)\,\boldsymbol{x}_2 = \mathbf{0}$$

Since $\{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ is linearly independent

$$\begin{aligned} \alpha + \beta &= 0 \\ \alpha - \beta &= 0 \end{aligned}$$

This system's only solution is $\alpha = \beta = 0$, thus $\overline{S}$ is linearly independent too. ■

Example 6.18 Determine whether vectors $\boldsymbol{x}_1 = (1+t),\ \boldsymbol{x}_2 = (1-t) \in \boldsymbol{P}_1$, where $\boldsymbol{P}_1 = \{p|\ p(t),\ t \in \mathbf{R}\}$ are linearly dependent or not.

Solution Let's examine the linear combination

$$\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 = \mathbf{0}$$

that is,

$$\begin{aligned} \alpha_1(1+t) + \alpha_2(1-t) &= (0,\ 0) \\ &= (\alpha_1 + \ \alpha_1 t) + (\ \alpha_2 - \alpha_2 t) \\ &= (\alpha_1 + \alpha_2,\ \alpha_1 t - \ \alpha_2 t) \\ &= (\alpha_1 + \alpha_2,\ (\alpha_1 - \alpha_2)t) \end{aligned}$$

from this we obtain the following system of equations

$$\begin{aligned} \alpha_1 + \alpha_2 &= 0 \\ \alpha_1 - \alpha_2 &= 0 \end{aligned}$$

Solving them, we get $\alpha_1 = \alpha_2 = 0$, and therefore our vectors $\boldsymbol{x}_1, \boldsymbol{x}_2$ are linearly independent. ■

Example/Exercise 6.19 Show that vectors

$$\boldsymbol{x}_1 = 1 - t, \quad \boldsymbol{x}_2 = 5 + 3t - 2t^2, \quad \boldsymbol{x}_3 = 1 + 3t - t^2 \in \boldsymbol{P}_2 = \{p|p(t),\ t \in \mathbf{R}\}$$

are linearly dependent.

Theorem 6.5 Let $\boldsymbol{P}_n(t)$ be a space of all polynomials of nth degree, where $t \in \mathbf{R}$. Then, the set of polynomials (i.e., vectors)

$$S = \{1,\ t,\ t^2,\ t^3,\ \ldots\ , t^n\}$$

is linearly independent in $\boldsymbol{P}_n(t)$.

Proof If we consider the elements of set S as vectors, that is

$$\boldsymbol{x}_0 = 1,\ \boldsymbol{x}_1 = t,\ \boldsymbol{x}_2 = t^2,\ \ldots\ ,\ \boldsymbol{x}_n = t^n$$

then, as before, to determine linear independence we need to show that the equation

$$\alpha_0 \boldsymbol{x}_0 + \alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \cdots + \alpha_n \boldsymbol{x}_n = \boldsymbol{0} \qquad (*)$$

has only the trivial solution $\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0$. Observe that equation (*) is equivalent to

$$\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_n t^n = 0, \quad \forall t \in \mathbf{R} \qquad (**)$$

The fundamental theorem of algebra tells us that a polynomial of nth degree has, at most, n roots, thus all coefficients in (**) must be zero. Otherwise, equation (**) would have infinitely many roots. Thus, since $\alpha_0 = \cdots = \alpha_n = 0$, equation (*) has only a trivial solution and therefore the vectors from set S are linearly independent in $\boldsymbol{P}_n(t)$. ■

Theorem 6.6 Any subset of a set of linearly independent vectors is linearly independent.

Proof Suppose $\boldsymbol{E} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subseteq \boldsymbol{X}$ is a set of linearly independent vectors from X, and let $\widetilde{E} = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_i\} \subseteq \boldsymbol{E}$ be any subset. Clearly, since all the elements of $\boldsymbol{E}$ are linearly independent, we have

$$\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_i \boldsymbol{x}_i + \alpha_{i+1} \boldsymbol{x}_{i+1} + \cdots + \alpha_k \boldsymbol{x}_k = \boldsymbol{0}$$

only with all $\alpha_1 = \alpha_2 = \cdots = \alpha_i = \alpha_{i+1} = \cdots = \alpha_k = \boldsymbol{0}$, thus

$$\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_i \boldsymbol{x}_i = \boldsymbol{0}$$

implies that $\widetilde{E} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i\}$ is a set of linearly independent vectors too. ■

Theorem 6.7 Set $\boldsymbol{E} = \{\boldsymbol{x}\} \subseteq \boldsymbol{X}$ containing only one vector from $\boldsymbol{X}$ is linearly *dependent* iff $\boldsymbol{x} = \boldsymbol{0}$.

Proof Suppose $\boldsymbol{E} = \{\boldsymbol{x}\}$ is a linearly dependent set. That means that there exists $\alpha \neq 0$ such that $\alpha \boldsymbol{x} = \boldsymbol{0}$. By Theorem 6.1, it implies that $\boldsymbol{x} = \boldsymbol{0}$, so the theorem holds. On the other hand, if $\boldsymbol{x} = \boldsymbol{0}$ then $\alpha \boldsymbol{x} = \boldsymbol{0}$, $\forall \alpha \in \Phi$. It follows, again, by Theorem 6.1i), that $\boldsymbol{E} = \{\boldsymbol{x}\}$ is linearly dependent. ■

Theorem 6.8 If a set of vectors $\boldsymbol{E}$ contains $\boldsymbol{0}$, then it is linearly dependent.

Proof With $\boldsymbol{x} = \boldsymbol{0}$ and $\alpha = 1$, the proof immediately follows from the previous two theorems. ■

6.4 BASIS AND DIMENSION OF A VECTOR SPACE

Definition 6.5 Let $\boldsymbol{X}$ be a linear space over a field Φ, and let $\boldsymbol{Y} \subseteq \boldsymbol{X}$ be some subset of $\boldsymbol{X}$. We say that $\boldsymbol{Y}$ **spans** $\boldsymbol{X}$, and we write $\boldsymbol{X} = L(\boldsymbol{Y})$, if every vector $\boldsymbol{x} \in \boldsymbol{X}$ can be expressed as a linear combination of vectors from $\boldsymbol{Y}$, that is

$$\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i, \quad \boldsymbol{x}_i \in \boldsymbol{Y},\ \alpha_i \in \Phi$$

Definition 6.6 A set of vectors $\boldsymbol{B} \subseteq \boldsymbol{X}$ is called a **basis** of a space $\boldsymbol{X}$, if $\boldsymbol{B}$ is linearly independent and if it spans $\boldsymbol{X}$, that is, $L(\boldsymbol{B}) = \boldsymbol{X}$.

Definition 6.6′ A set of vectors $\boldsymbol{B} \subseteq \boldsymbol{X}$ is called a **basis** of a space $\boldsymbol{X}$, if every vector $\boldsymbol{x} \in \boldsymbol{X}$ can be uniquely expressed as a linear combination of vectors from $\boldsymbol{B}$.

Example 6.20 One possible basis for $\mathbf{R}^2$ is $\boldsymbol{B} = \{ (1, 0), (0, 1) \}$. We have shown in Example 6.13 that set $\boldsymbol{B}$ is linearly independent. Let's see whether it spans $\boldsymbol{X}$. Consider any $(\alpha_1, \alpha_2) \in \mathbf{R}^2$. Then,

$$(\alpha_1, \alpha_2) = \alpha_1(1, 0) + \alpha_2(0, 1)$$

$\boldsymbol{B}$ is linearly independent and spans $\mathbf{R}^2$, thus it is a basis for $\mathbf{R}^2$. ■

We can easily generalize this to $\mathbf{R}^n$.

Example 6.21 Consider a set $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ where

$$\begin{aligned}
\boldsymbol{b}_1 &= (1, 0, \ldots, 0, 0) \\
\boldsymbol{b}_2 &= (0, 1, 0, \ldots, 0,) \\
&\vdots \quad \vdots \qquad \vdots \\
\boldsymbol{b}_n &= (0, 0, \ldots, 0, 1)
\end{aligned}$$

Following the logic of the previous example, it is easy to see that $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ is one possible basis for $\mathbf{R}^n$. ■

Example 6.22 Determine whether the set $\boldsymbol{B} = \{1, t - 1, (t - 2)(t - 1)\}$ is a basis for $\boldsymbol{P}_2(t)$.

Solution First, to check independence, let's suppose

$$\alpha_1 \cdot 1 + \alpha_2(t - 1) + \alpha_3(t - 2)(t - 1) = 0$$

Then,

$$\alpha_1 + \alpha_2 t - \alpha_2 + \alpha_3 t^2 - 3\alpha_3 t + 2\alpha_3 = 0$$
$$= (\alpha_1 - \alpha_2 + 2\alpha_3) + (\alpha_2 - 3\alpha_3)t + \alpha_3 t^2$$

From this, we get a system of three linear equations

$$\begin{aligned} \alpha_1 - \alpha_2 + 2\alpha_3 &= 0 \\ \alpha_2 - 3\alpha_3 &= 0 \\ \alpha_3 &= 0 \end{aligned}$$

Solving them, we obtain

$$\alpha_3 = 0, \quad \alpha_2 = 0, \quad \alpha_1 = 0$$

So far so good. Let's check whether set $\boldsymbol{B}$ spans $\boldsymbol{P}_2(t)$. Let $p(t) \in \boldsymbol{P}_2(t)$ such that $p(t) = a_2 t^2 + a_1 t + a_0$. Consider

$$a_2 t^2 + a_1 t + a_0 = \beta_2((t-2)(t-1)) + \beta_1(t-1) + \beta_0 \cdot 1$$

This gives us

$$\begin{aligned} a_0 &= \beta_0 - \beta_1 + 2\beta_2 \\ a_1 &= \beta_1 - 3\beta_2 \\ a_2 &= \beta_2 \end{aligned}$$

Solving these for betas we get

$$\begin{aligned} \beta_0 &= a_0 + a_1 + a_2 \\ \beta_1 &= a_1 + 3a_2 \\ \beta_2 &= a_2 \end{aligned}$$

Hence,

$$p(t) = a_2((t-2)(t-1)) + (a_1 + 3a_2)(t-1) + (a_0 + a_1 + a_2) \cdot 1$$

Thus, set $\boldsymbol{B} = \{1,\ t-1,\ (t-2)(t-1)\}$ is linearly independent and spans $\boldsymbol{P}_2(t)$. Therefore, it forms a basis for $\boldsymbol{P}_2(t)$. ■

At this point, natural questions that you might have asked yourself are as follows: What about the existence of the bases in general? How unique are they?

And, certainly, what do they tell us about the respective spaces? To answer these questions, we need two more definitions.

Definition 6.7 A vector space $\boldsymbol{X}$ is said to be **finite-dimensional**, if there exists at least one finite, linearly independent set of vectors that span $\boldsymbol{X}$, that is, if there exists one finite basis. Of course, a vector space that is not finite-dimensional is called **infinite-dimensional**.

Definition 6.8 We say that n is the **dimension** of a space $\boldsymbol{X}$, and we write $\dim\ \boldsymbol{X} = n$, if n is the cardinal number of a basis $\boldsymbol{B}$ of space $\boldsymbol{X}$. If the space is trivial, that is, $\boldsymbol{X} = \{\varnothing\}$ we say that its dimension is zero.

Example 6.23

(i) $\dim \mathbf{R}^3 = 3$
(ii) $\dim \mathbf{R}^n = n$
(iii) $\dim \boldsymbol{P}_n(\Phi) = n$ ■

Example 6.24 Unlike the space $\boldsymbol{P}_n(\Phi)$, the space $\boldsymbol{P}(\Phi)$ is infinite-dimensional. Indeed, consider any set S of the elements (polynomials) from $\boldsymbol{P}(\Phi)$. Let n denote the highest degree of any of the polynomials from set S. Thus, since every polynomial from the set is of a degree less than or equal to n, the set S cannot span $\boldsymbol{P}(\Phi)$ and therefore $\boldsymbol{P}(\Phi)$ is infinite-dimensional (see Theorem 6.5). ■

Theorem 6.9 Every nontrivial linear (finite-dimensional) space $\boldsymbol{X}$ has at least one basis.

Proof The theorem holds in general, but we will prove it only in the finite-dimensional case.

Let $\boldsymbol{X} \neq \{\varnothing\}$, and let $\boldsymbol{E} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \ldots, \boldsymbol{x}_n\}$ be a set that spans $\boldsymbol{X}$. Then, there also exists a subset of $\boldsymbol{E}$ consisting of linearly independent vectors. Indeed, suppose $\boldsymbol{x}_k \in \boldsymbol{E}$ is such that it can be represented as a linear combination of the remaining vectors from $\boldsymbol{E}$. We claim that $\boldsymbol{E} \setminus \{\boldsymbol{x}_k\}$ is still a set that spans the $\boldsymbol{X}$.

Consider

$$\boldsymbol{x}_k = \sum_{i \neq k=1}^{n} \alpha_i \boldsymbol{x}_i$$

Since $\boldsymbol{E} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \ldots, \boldsymbol{x}_n\}$ spans $\boldsymbol{X}$, any $\boldsymbol{x} \in \boldsymbol{X}$ can be expressed as

$$\begin{aligned} \boldsymbol{x} &= \sum_{i=1}^{n} \beta_i \boldsymbol{x}_i \\ &= \sum_{i \neq k=1}^{n} \beta_i \boldsymbol{x}_i + \beta_k \boldsymbol{x}_k \end{aligned}$$

$$= \sum_{i \neq k=1}^{n} \beta_i \boldsymbol{x}_i + \beta_k \sum_{i \neq k=1}^{n} \alpha_i \boldsymbol{x}_i$$

$$= \sum_{i \neq k=1}^{n} (\beta_i + \beta_k \alpha_i) \boldsymbol{x}_i$$

$$= \sum_{i \neq k=1}^{n} \gamma_i \boldsymbol{x}_i$$

We saw that we can remove all linearly dependent vectors from set $\boldsymbol{E}$ and still be able to span the space $\boldsymbol{X}$, that is, we obtained the set of linearly independent vectors that span the space $\boldsymbol{X}$ – that is, we got a basis of $\boldsymbol{X}$. ■

The proof of the next, very important, Theorem 6.10 depends on the following.

Lemma 6.1 Let $\boldsymbol{E} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ be an *ordered* set of linearly dependent vectors from $\boldsymbol{X}$, and let $\boldsymbol{x}_1 \neq 0$. Then at least one of the vectors from $\boldsymbol{E}$ can be expressed as a linear combination of the preceding vectors in $\boldsymbol{E}$ (see also Definition 6.4).

Proof Consider a nontrivial combination

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{x}_i = \boldsymbol{0} \qquad (*)$$

and let α_j be the last coefficient different from zero. j cannot be 1, for in that case we would have $\alpha_1 \boldsymbol{x}_1 = \boldsymbol{0}$ which, with $\alpha_1 \neq 0$, contradicts our assumption. Therefore, $j > 1$ and (*) reduces to

$$\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_j \boldsymbol{x}_j = \boldsymbol{0}$$

Thus, we can divide by α_j and an expression for $\boldsymbol{x}_j$ follows. ■

Theorem 6.10 Any two bases of a linear space $\boldsymbol{X}$ are equipotent (contain the same number of elements).

Proof We reason as follows: Let $\boldsymbol{X}$ be a finite-dimensional vector space, and let $\boldsymbol{B}^{(1)}$ and $\boldsymbol{B}^{(2)}$ be any two bases of $\boldsymbol{X}$. Then, the elements of $\boldsymbol{B}^{(1)}$ are linearly independent in $\boldsymbol{X}$, and the elements of $\boldsymbol{B}^{(2)}$ span $\boldsymbol{X}$. Consequently, $|\boldsymbol{B}^{(1)}| \leq |\boldsymbol{B}^{(2)}|$. But, we can also argue conversely, that is, just by interchanging the roles of $\boldsymbol{B}^{(1)}$ and $\boldsymbol{B}^{(2)}$ we can say that $|\boldsymbol{B}^{(2)}| \leq |\boldsymbol{B}^{(1)}|$. Thus,

$$|\boldsymbol{B}^{(1)}| = |\boldsymbol{B}^{(2)}|$$

More formally:

Since every vector space $\boldsymbol{X}$ has at least one basis, say,

$$\boldsymbol{B}^{(1)} = \{\boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}\}$$

we let

$$\boldsymbol{B}^{(2)} = \{\boldsymbol{b}_1^{(2)}, \ldots, \boldsymbol{b}_m^{(2)}\}$$

be another basis of $\boldsymbol{X}$. We would like to show that $|\boldsymbol{B}^{(1)}| = |\boldsymbol{B}^{(2)}|$, that is, $n = m$. First, without loss of generality, let's assume that $\boldsymbol{B}^{(1)} \cap \boldsymbol{B}^{(2)} = \emptyset$, and taking any

$$\boldsymbol{b}_k^{(2)} \in \boldsymbol{B}^{(2)}$$

let's consider the set

$$\{\boldsymbol{b}_k^{(2)}, \ \boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}\} \tag{6.5}$$

Since $\boldsymbol{b}_k^{(2)}$ can obviously be expressed as a linear combination of the remaining vectors of (6.5), set (1) is definitely a linearly dependent set of vectors that spans $\boldsymbol{X}$. According to the lemma above, there exists a $\boldsymbol{b}_i^{(1)}$, which can be expressed as a linear combination of its predecessors in (6.5). However, this vector can be removed (see Theorem 6.9) and we still have a set that spans $\boldsymbol{X}$:

$$\{\boldsymbol{b}_k^{(2)}, \ \boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_{i-1}^{(1)}, \boldsymbol{b}_{i+1}^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}\} \tag{6.6}$$

Let's choose another

$$\boldsymbol{b}_l^{(2)} \in \boldsymbol{B}^{(2)}$$

and consider the set

$$\{\boldsymbol{b}_l^{(2)}, \boldsymbol{b}_k^{(2)}, \boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_{i-1}^{(1)}, \boldsymbol{b}_{i+1}^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}\} \tag{6.7}$$

This is again a linearly dependent set that spans $\boldsymbol{X}$. Well, using Lemma 6.1 again, there is a vector in (6.7) that can be expressed as a linear combination of its preceding vectors. Note this can be neither $\boldsymbol{b}_l^{(2)}$ nor $\boldsymbol{b}_k^{(2)}$ since the set

$$\{\boldsymbol{b}_l^{(2)}, \boldsymbol{b}_k^{(2)}\}$$

is linearly independent (it is a part of the basis $\boldsymbol{B}^{(2)}$), so its elements have to be some of $\boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}$, say, $\boldsymbol{b}_j^{(1)}$. Thus, as before, it can be removed. Lemma 6.1 tells us that

$$\{\boldsymbol{b}_l^{(2)}, \boldsymbol{b}_k^{(2)}, \boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_{i-1}^{(1)}, \boldsymbol{b}_{i+1}^{(1)}, \ldots, \boldsymbol{b}_{j-1}^{(1)}, \boldsymbol{b}_{j+1}^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}\}$$

spans space $\boldsymbol{X}$ too. We continue this process to n steps. The number of vectors in $\boldsymbol{B}^{(2)}$ cannot exceed n, otherwise $\boldsymbol{b}_{n+1}^{(2)} \in \boldsymbol{B}^{(2)}$ would be expressible as a linear combination of

$$\{\boldsymbol{b}_n^{(2)},\ \boldsymbol{b}_{n-1}^{(2)}, \ldots, \boldsymbol{b}_2^{(2)},\ \boldsymbol{b}_1^{(2)}\}$$

so $\boldsymbol{B}^{(2)}$ wouldn't be a set of linearly independent vectors, which is contrary to our hypothesis. Hence, we proved that $|\boldsymbol{B}^{(2)}| \leq |\boldsymbol{B}^{(1)}|$. Similarly, we can prove that $|\boldsymbol{B}^{(1)}| \leq |\boldsymbol{B}^{(2)}|$ from which we conclude that $|\boldsymbol{B}^{(1)}| = |\boldsymbol{B}^{(2)}|$.

Considering the lengthiness of the proof, let's summarize everything slightly less precisely but nevertheless correctly:

Given two bases of the linear space $\boldsymbol{X}$,

$$\boldsymbol{B}^{(1)} = \{\boldsymbol{b}_1^{(1)}, \ldots, \boldsymbol{b}_n^{(1)}\}$$

and

$$\boldsymbol{B}^{(2)} = \{\boldsymbol{b}_1^{(2)}, \ldots, \boldsymbol{b}_m^{(2)}\}$$

implies that $\boldsymbol{B}^{(2)}$ is linearly independent set in $L(\boldsymbol{B}^{(1)}) = \boldsymbol{X}$, so $m \leq n$, but, at the same time, $\boldsymbol{B}^{(1)}$ is a linearly independent set in $L(\boldsymbol{B}^{(2)}) = \boldsymbol{X}$, so $n \leq m$, and therefore $n = m$. ■

Theorem 6.11 Let $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\} \subseteq \boldsymbol{X}$ be a set of linearly independent vectors from $\boldsymbol{X}$. Then, $\boldsymbol{A}$ is a subset of some basis of $\boldsymbol{X}$.

Proof In the case that $\boldsymbol{A}$ spans $\boldsymbol{X}$, there is nothing to prove. Suppose $\boldsymbol{A}$ doesn't span $\boldsymbol{X}$. Then, there exists a basis of $\boldsymbol{X}$, say, $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$, and we can form an ordered set

$$\boldsymbol{C} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\} \qquad (*)$$

Since $\boldsymbol{B}$ is a basis, it spans $\boldsymbol{X}$, thus (*) spans $\boldsymbol{X}$ too. Furthermore, (*) is clearly linearly dependent. According to Lemma 6.1, at least one vector from (*) can be expressed as a linear combination of its predecessors. It cannot be any of the $\boldsymbol{a}_i'$s, because, by assumption, they are linearly independent, thus it has to be one of $\boldsymbol{b}'$s, say, $\boldsymbol{b}_j$. By Theorem 6.9, $\boldsymbol{b}_j$ can be removed. We are left with the set

$$\overset{B}{{}_j}\boldsymbol{C} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_{j-1}, \boldsymbol{b}_{j+1}, \ldots, \boldsymbol{b}_n\} \qquad (**)$$

which again spans $\boldsymbol{X}$. If (**) is a linearly independent set, we are done. If not, we continue the process until after finitely many steps we arrive to the set

$$\boldsymbol{F} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_{i_1}, \ldots, \boldsymbol{b}_{i_l}\}$$

which is linearly independent and spans $\boldsymbol{X}$. Since $\boldsymbol{A} \subseteq \boldsymbol{F}$ our theorem is proved. ■

Theorem 6.12 A vector space $\boldsymbol{X}$ is finite-dimensional iff every linearly independent set of vectors in X is finite. A vector space is infinite-dimensional iff it contains an infinite linearly independent set of vectors.

Proof If $\boldsymbol{X}$ is finite–dimensional, then every linearly independent set of vectors in $\boldsymbol{X}$ is finite. Suppose that every linearly independent set in $\boldsymbol{X}$ is finite, but the space $\boldsymbol{X}$ is infinite-dimensional. That means that no finite set can span $\boldsymbol{X}$. Take $\boldsymbol{x}_1 \in \boldsymbol{X}$, $\boldsymbol{x}_1 \neq 0$, and note that $\boldsymbol{A}_1 = \{\boldsymbol{x}_1\}$ is obviously a linearly independent set. Since $\boldsymbol{X}$ is infinite-dimensional, $\boldsymbol{A}_1$ cannot span $\boldsymbol{X}$. Let's take another $\boldsymbol{x}_2 \in \boldsymbol{X}$, and consider $\boldsymbol{A}_2 = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$. $\boldsymbol{A}_2$ is clearly linearly independent but still doesn't span $\boldsymbol{X}$. We repeat this process and realize that in a finite number of steps, we cannot obtain a set $\boldsymbol{A}_n = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ that is linearly independent and spans $\boldsymbol{X}$. In other words, there is an infinite set $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots\}$ of linearly independent vectors in $\boldsymbol{X}$, which contradicts our assumption that every linearly independent set in $\boldsymbol{X}$ is finite. ■

Theorem 6.13 Let X be a finite-dimensional vector space, and let $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be some basis of $\boldsymbol{X}$. Then, any vector $\boldsymbol{x} \in X$ has a unique expression as a linear combination of vectors from $\boldsymbol{B}$ (see Definition 6.6′).

Proof Suppose there are two representations of $\boldsymbol{x}$ in base $\boldsymbol{B}$:

$$\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i$$

and

$$\boldsymbol{x} = \sum_{i=1}^{n} \beta_i \boldsymbol{b}_i$$

Then,

$$\begin{aligned} \boldsymbol{x} - \boldsymbol{x} &= \boldsymbol{0} \\ &= \sum_{i=1}^{n} (\alpha_i - \beta_i) \boldsymbol{b}_i \end{aligned}$$

Since $\boldsymbol{B}$ is a set of linearly independent vectors, it follows that $\alpha_i - \beta_i = 0$ for every $i = 1, \ldots, n$, and thus $\alpha_i = \beta_i$ for every $i = 1, \ldots, n$. Therefore, the expression for any $\boldsymbol{x} \in \boldsymbol{X}$ is unique. ■

Theorem 6.14 If X is an n-dimensional vector space, and if there is a set of linearly independent vectors $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ in $\boldsymbol{X}$, then they form a basis for $\boldsymbol{X}$.

Proof Suppose $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is not a basis of $\boldsymbol{X}$. Then, by Theorem 6.11, it is a subset of some basis $\boldsymbol{B}$ of $\boldsymbol{X}$, that is, $\boldsymbol{A} \subseteq \boldsymbol{B}$. In that case, we can supplement set $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ by another set of linearly independent vectors $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m\}$ to get a full basis of $\boldsymbol{X}$. But that would mean that $\dim(\boldsymbol{X}) = n + m \neq n$, which contradicts our assumption that $\boldsymbol{X}$ is n-dimensional. ■

Theorem 6.15 If a set of vectors $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ spans an n-dimensional vector space $\boldsymbol{X}$, then it forms a basis for $\boldsymbol{X}$.

Proof Since set $\boldsymbol{A}$ spans $\boldsymbol{X}$, some subset of it has to be a basis for $\boldsymbol{X}$. Thus, we can pick from $\boldsymbol{A}$ a subset $\boldsymbol{B} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$ with $m \leq n$, such that $\boldsymbol{B}$ is a basis for $\boldsymbol{X}$. But in that case $m = \dim(\boldsymbol{X}) = n$, and thus $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is a basis for $\boldsymbol{X}$. ■

Example 6.25 Determine whether the set of vectors $\boldsymbol{B} = \{\boldsymbol{b}_1 = (1, 0, 0), \boldsymbol{b}_2 = (1, 1, 0), \boldsymbol{b}_3 = (1, 1, 1)\}$ is a basis for $\mathbf{R}^3$.

Solution Since we have three vectors in set $\boldsymbol{B}$, in order to verify whether they form a basis for $\mathbf{R}^3$ (which is three-dimensional) we only need to check linear independence. As before, we write

$$\alpha_1(1,\ 0,\ 0) + \alpha_2(1,\ 1,\ 0) + \alpha_3(1,\ 1,\ 1) = (0,\ 0,\ 0)$$

that is,

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 0 \\ \alpha_2 + \alpha_3 &= 0 \\ \alpha_3 &= 0 \end{aligned}$$

Thus, $\alpha_1 = \alpha_2 = \alpha_3 = 0$, so the vectors of $\boldsymbol{B}$ are linearly independent and therefore form a basis for $\mathbf{R}^3$.

■

Example/Exercise 6.26 Consider a vector space $\mathbf{R}^3$ and a basis $\boldsymbol{B}$ from the previous example. Show that the vector $\boldsymbol{x} = (5, -6, 2) \in \mathbf{R}^3$ has the following expression in the basis $\boldsymbol{B}$:

$$\boldsymbol{x} = 11\boldsymbol{b}_1 - 8\boldsymbol{b}_2 + 2\boldsymbol{b}_1$$

Example/Exercise 6.27 Let $\boldsymbol{X}$ be an n-dimensional vector space. Show that any $n + 1$ or more vectors from $\boldsymbol{X}$ are linearly dependent.

6.5 SUBSPACES

Once we have a certain structure we want to see how to define, study, and judiciously use a substructure. We start with

Definition 6.9 Let $\boldsymbol{X}$ be a vector space over a field Φ, and let $\boldsymbol{L}$ be a subset of $\boldsymbol{X}$. We say that $\boldsymbol{L}$ is a **subspace** of $\boldsymbol{X}$ if $\boldsymbol{L}$ itself is a vector space over Φ with respect to the same operations of addition and scalar multiplication already defined on $\boldsymbol{X}$. We write $\boldsymbol{L} < \boldsymbol{X}$.

From this definition, it immediately follows that in order to show that some subset $\boldsymbol{L} \subseteq \boldsymbol{X}$ is a subspace, one needs to show that all the axioms of a vector space also hold on $\boldsymbol{L}$. Clearly, every vector space has at least two trivial subspaces, $\boldsymbol{L} = \boldsymbol{X}$ and $\boldsymbol{L} = \{0\}$. In the case that $\boldsymbol{L} < \boldsymbol{X}$ and $\boldsymbol{L} \neq \boldsymbol{X}$, we say that $\boldsymbol{L}$ is a proper subspace of $\boldsymbol{X}$. Thus, as a simple criterion for identifying subspaces, we have the following:

Theorem 6.16 A nonempty subset $\boldsymbol{L} \subseteq \boldsymbol{X}$ is a subspace of space $\boldsymbol{X}$ iff $\boldsymbol{L}$ is **closed** with respect to the operations defined on $\boldsymbol{X}$, that is, if the following conditions are satisfied:

(i) $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{L}, \quad \boldsymbol{x} + \boldsymbol{y} \in \boldsymbol{L}$

(ii) $\forall\, \boldsymbol{x} \in \boldsymbol{L}$, and $\alpha \in \Phi, \quad \alpha \boldsymbol{x} \in \boldsymbol{L}$

Clearly, conditions (i) and (ii) could have been compactly stated as

(iii) $\forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{L}, \quad \alpha, \beta \in \Phi, \quad \alpha \boldsymbol{x} + \beta \boldsymbol{y} \in \boldsymbol{L}$

Proof The necessary condition is obvious. Let's check the sufficient one:

Assuming that (i) and (ii) are satisfied, then in particular it follows from (ii) that $\forall\, \boldsymbol{x} \in \boldsymbol{L}\ (-1)\boldsymbol{x} = -\boldsymbol{x} \in \boldsymbol{L}$, thus $\boldsymbol{L} \subseteq \boldsymbol{X}$ is a subgroup of the additive group $\boldsymbol{X}$. Vector space axioms A1–A4 certainly hold for $\boldsymbol{X}$, so they have to be satisfied for $\boldsymbol{L} \subseteq \boldsymbol{X}$, and therefore $\boldsymbol{L}$ is a subspace as claimed. ■

Example 6.28 Consider a subset of $\mathbf{R}^n$ defined as follows:

$$\boldsymbol{L} = \{\boldsymbol{x} = (x_1, \ldots, x_n) \mid x_1 + \cdots + x_n = 0,\ x_i \in \mathbf{R}\}$$

Then with the usual component addition and scalar multiplication, $\boldsymbol{L}$ is a subspace of $\mathbf{R}^n$. Indeed,

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= (x_1, \ldots, x_n) + (y_1, \ldots, y_n) \\ &= (x_1 + y, \ldots, x_n + y_n) \end{aligned}$$

But

$$\begin{aligned} x_1 + y_1 + \cdots + x_n + y_n &= (x_1 + \cdots + x_n) + (y_1 + \cdots + y_n) \\ &= 0 + 0 \\ &= 0 \end{aligned}$$

Also,

$$\begin{aligned} \alpha \boldsymbol{x} &= \alpha(x_1, \ldots, x_n) \\ &= (\alpha x_1, \ldots, \alpha x_n) \end{aligned}$$

and

$$\begin{aligned} \alpha x_1 + \cdots + \alpha x_n &= \alpha(x_1 + \cdots + x_n) \\ &= \alpha \cdot 0 = 0 \end{aligned}$$

So, $\boldsymbol{L}$ is indeed a subspace of $\boldsymbol{X}$. ■

Example 6.29 Consider a vector space $\mathbf{R}^3$, and a set $\boldsymbol{L} = \{\boldsymbol{x} = (x_1, x_2, x_3) | \ x_1 = x_2 = x_3 \in \mathbf{R}\}$. Then, $\boldsymbol{L} < \mathbf{R}^3$. Indeed. Consider $\boldsymbol{x} = (x, x, x)$ and $\boldsymbol{y} = (y, y, y)$ from $\boldsymbol{L}$. Then,

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{y} &= (x, x, x) + (y, y, y) \\ &= (x + y, x + y, x + y) \in L \end{aligned}$$

Also,

$$\alpha \boldsymbol{x} = \alpha(x, x, x) = (\alpha x, \alpha x, \alpha x) \in \boldsymbol{L}, \quad \forall \alpha, x \in \mathbf{R}$$

Thus, $\boldsymbol{L} < \mathbf{R}^3$. ■

You can easily convince yourself that the claim in the following example is true.

Example 6.30 Let $0 < k \leq n$, and let $L \subseteq \mathbf{R}^n$ be defined by

$$\boldsymbol{L} = \{(x_1, \ldots, x_k, \ldots, x_n) | x_1 = x_2 = \cdots = x_k = 0\}$$

Then, L is a subspace of $\mathbf{R}^n$. ■

Example 6.31 Consider a vector space $\mathbf{R}^n$ and let's accept the convention

$$(x_1, \ldots, x_{n-1}) \equiv (x_1, \ldots, x_{n-1}, 0)$$

Then,

$$\mathbf{R}^{n-1} < \mathbf{R}^n$$

In other words, we have the following sequence of spaces/subspaces:

$$\mathbf{R} < \mathbf{R}^2 < \cdots < \mathbf{R}^n < \cdots < \mathbf{R}^\infty \quad \blacksquare$$

Example 6.32 Let $\boldsymbol{P}(t)$ be the space of all polynomials. Then, the space $\boldsymbol{P}_n(t)$of all polynomials of degree at most n is a subspace of $\boldsymbol{P}(t)$. ■

Example 6.33 Let $\boldsymbol{X}_1, \boldsymbol{X}_2 < X$ be two subspaces of $\boldsymbol{X}$. Show that the intersection of subspaces $\boldsymbol{X}_\cap = \boldsymbol{X}_1 \cap \boldsymbol{X}_2$ is again a subspace of $\boldsymbol{X}$.

Solution Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_\cap$ be any two vectors in $\boldsymbol{X}_\cap$, and let $\alpha, \beta \in \Phi$ be two scalars from Φ. Since $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_\cap$, it follows that $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_1$ and also $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}_2$. But $\boldsymbol{X}_1, \boldsymbol{X}_2$ are subspaces and therefore $\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in \boldsymbol{X}_1$ and $\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in \boldsymbol{X}_2$, which implies that $\alpha\boldsymbol{x} + \beta\boldsymbol{y} \in \boldsymbol{X}_\cap$. Thus, $\boldsymbol{X}_\cap$ is a subspace of X. ■

Similarly, we can prove that given a vector space $\boldsymbol{X}$ and any nonempty family of its subspaces $\boldsymbol{F} = \{\boldsymbol{X}_i\}$, the intersection $\boldsymbol{F}_\cap = \cap_i \boldsymbol{X}_i$ is again a subspace of $\boldsymbol{X}$.

Thus

Theorem 6.17 The intersection of any number of subspaces of a vector space $\boldsymbol{X}$ is a subspace of $\boldsymbol{X}$.

Example 6.34 Consider a space $\mathbf{R}^4$ and two of its subspaces

$$\boldsymbol{X}_1 = \{(x_1, \ldots, x_4) | \, x_1 - x_2 + x_3 - x_4 = 0\}$$

and

$$\boldsymbol{X}_2 = \{(x_1, \ldots, x_4) | \, x_1 + x_2 + x_3 + x_4 = 0\}$$

Find a basis for $\boldsymbol{X}_\cap = \boldsymbol{X}_1 \cap \boldsymbol{X}_2$.

Solution From the characterization of subspaces $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, we know that $\boldsymbol{x} \in \mathbf{R}^4$ belongs to $\boldsymbol{X}_\cap = \boldsymbol{X}_1 \cap \boldsymbol{X}_2$ iff

$$x_1 - x_2 + x_3 - x_4 = 0$$
$$x_1 + x_2 + x_3 + x_4 = 0$$

or

$$x_1 + x_3 = 0$$
$$x_2 + x_4 = 0$$

In other words,

$$x_1 = -x_3$$
$$x_2 = -x_4$$

Let's pick two vectors from $\boldsymbol{X}_\cap$ satisfying the conditions above, say,

$$\boldsymbol{y}_1 = (1,\ 0, -1,\ 0)$$
$$\boldsymbol{y}_2 = (0, -1,\ 0,\ 1)$$

We claim that $\boldsymbol{Y} = \{\boldsymbol{y}_1,\ \boldsymbol{y}_2\}$ is a basis for $\boldsymbol{X}_\cap$. To prove this, we have to show that $\boldsymbol{Y} = \{\boldsymbol{y}_1,\ \boldsymbol{y}_2\}$ is a set of linearly independent vectors and that they span $\boldsymbol{X}_\cap$, that is, $L\ (\boldsymbol{y}_1,\ \boldsymbol{y}_2) = \boldsymbol{X}_\cap$. Suppose not. Then there exist $\alpha_1, \alpha_2 \in \Phi$, both not zero, such that

$$\alpha_1 \boldsymbol{y}_1 + \alpha_2 \boldsymbol{y}_2 = 0$$

that is,

$$\begin{aligned} \alpha_1(1,\ 0, -1,\ 0) + \alpha_2(0, -1,\ 0,\ 1) &= (0,\ 0,\ 0,\ 0) \\ &= (\alpha_1, -\alpha_2, -\alpha_1,\ \alpha_2) \end{aligned}$$

which implies $\alpha_1 = \alpha_2 = 0$, contrary to our requirement that at least one of α-s be different from zero. Hence, $\boldsymbol{y}_1,\ \boldsymbol{y}_2$ must be linearly independent. Do they span $\boldsymbol{X}_\cap$? Well, if $\boldsymbol{x} \in \boldsymbol{X}_\cap$, then, as we have seen,

$$\begin{aligned} \boldsymbol{x} &= (\alpha_1,\ \alpha_2, -\alpha_1, -\alpha_2) \\ &= \alpha_1(1,\ 0, -1,\ 0) - \alpha_2(0, -1,\ 0,\ 1) \\ &= \alpha_1 \boldsymbol{y}_1 - \alpha_2 \boldsymbol{y}_2 \end{aligned}$$

So $\{\boldsymbol{y}_1,\ \boldsymbol{y}_2\}$ spans $\boldsymbol{X}_\cap$, that is, $\boldsymbol{X}_\cap \subseteq\ L(\boldsymbol{y}_1,\ \boldsymbol{y}_2)$ but at the same time $\{\boldsymbol{y}_1,\ \boldsymbol{y}_2\} \subseteq \boldsymbol{X}_\cap$, and we conclude that $L(\boldsymbol{y}_1,\ \boldsymbol{y}_2) =\ \boldsymbol{X}_\cap$. Thus, we proved that $\boldsymbol{y}_1,\ \boldsymbol{y}_2$ are linearly independent and they span $\boldsymbol{X}_\cap$, and therefore $\boldsymbol{Y} = \{\boldsymbol{y}_1,\ \boldsymbol{y}_2\}$ is a basis for $\boldsymbol{X}_\cap$ as claimed. ■

Theorem 6.18 If $\boldsymbol{X}$ is an n-dimensional vector space and $\boldsymbol{L} < \boldsymbol{X}$, then

$$\dim \boldsymbol{L} \leq n$$

Proof If $\boldsymbol{L} = \{0\}$, then $\dim \boldsymbol{L} = 0$. Suppose there exists an $\boldsymbol{x}_1 \in \boldsymbol{L},\ \boldsymbol{x} \neq \boldsymbol{0}$. If $\boldsymbol{x}_1$is the only vector in $\boldsymbol{L}$, that is, $\boldsymbol{L} = \{\boldsymbol{x}_1\}$, then $\boldsymbol{L}$ is obviously one-dimensional. In case there is another vector $\boldsymbol{x}_2 \in \boldsymbol{L}$, we have $\boldsymbol{L} = \{\boldsymbol{x}_1,\ \boldsymbol{x}_2\}$. We can continue

this way to obtain a set $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$. If the vectors $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ are linearly independent and do not span $\boldsymbol{L}$, then there exists $\boldsymbol{x}_{n+1} \in \boldsymbol{L}$ such that $\boldsymbol{x}_{n+1} \notin \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$. In that case, as we have seen before, the set

$$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n, \boldsymbol{x}_{n+1}\}$$

is linearly independent too. In other words, every linearly independent set of vectors in $\boldsymbol{L}$, that does not span $\boldsymbol{L}$, can be expanded to a larger independent set. This process cannot exceed n steps (cannot be infinite); otherwise, we would obtain more than n linearly independent vectors in $\boldsymbol{L}$, which would contradict Theorem 6.14. Thus, $\dim \boldsymbol{L} = m \leq n$. ■

Example 6.35 Let $\boldsymbol{L} < \mathbf{R}^3$ be some subspace. Then, $\dim \boldsymbol{L} \leq 3$. In other words, the dimension of $\boldsymbol{L}$ can be 0, 1, 2 or 3. Indeed, if $\dim \boldsymbol{L} = 0$ then $\boldsymbol{L} = \{0\}$, meaning we have a trivial case in which the whole space is a point. If $\dim \boldsymbol{L} = 1$, then $\boldsymbol{L}$ is a line through the origin. If $\dim \boldsymbol{L} = 2$, the subspace $\boldsymbol{L}$ is a plane through the origin. Finally, if $\dim \boldsymbol{L} = 3$, we have another trivial case $\boldsymbol{L} = \mathbf{R}^3$. ■

Definition 6.10 Let $\boldsymbol{L}_1, \boldsymbol{L}_2, \ldots, \boldsymbol{L}_n$ be subspaces of a space $\boldsymbol{X}$. We say that

$$\sum \boldsymbol{L}_i = \boldsymbol{L}_1 + \boldsymbol{L}_2 + \cdots + \boldsymbol{L}_n = \{\boldsymbol{l}_1 + \boldsymbol{l}_2 + \cdots + \boldsymbol{l}_n \mid \boldsymbol{l}_i \in \boldsymbol{L}_i, \ \ i = 1, 2, \ldots, m\}$$

is the sum of subspaces of $\boldsymbol{X}$.

Definition 6.11 A vector space $\boldsymbol{X}$ is said to be the **direct sum** of its subspaces $\boldsymbol{L}_1, \boldsymbol{L}_2, \ldots, \boldsymbol{L}_n$, denoted by $\boldsymbol{X} = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2 \oplus \cdots \oplus \boldsymbol{L}_n$ if each element of $\boldsymbol{X}$ can be *uniquely* expressed as a sum $\boldsymbol{x} = \boldsymbol{l}_1 + \boldsymbol{l}_2 + \cdots + \boldsymbol{l}_n$, where $\boldsymbol{l}_i \in \boldsymbol{L}_i$.

Example 6.36 Consider a space $\mathbf{R}^3$ and two of its subspaces $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ defined as follows:

$$\boldsymbol{L}_1 = \{(x, y, 0) \in \mathbf{R}^3 \mid x, y \in \mathbf{R}\}$$

and

$$\boldsymbol{L}_2 = \{(0, 0, z) \in \mathbf{R}^3 \mid z \in \mathbf{R}\}$$

Then, $\mathbf{R}^3 = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2$.

In other words, any $\boldsymbol{x} \in \mathbf{R}^3$ can be uniquely represented as $\boldsymbol{x} = \boldsymbol{l}_1 + \boldsymbol{l}_2$, where $\boldsymbol{l}_1 \in \boldsymbol{L}_1$ and $\boldsymbol{l}_2 \in \boldsymbol{L}_2$.

Indeed, let $\boldsymbol{x} = (x, y, z)$, then

$$\begin{aligned} \boldsymbol{x} = (x, y, z) &= (x, y, 0) + (0, 0, z) \\ &= \boldsymbol{l}_1 + \boldsymbol{l}_2 \end{aligned}$$

Suppose that there are $\boldsymbol{l}'_1 \in \boldsymbol{L}_1$ and $\boldsymbol{l}'_2 \in \boldsymbol{L}_2$ such that

$$\boldsymbol{x} = \boldsymbol{l}'_1 + \boldsymbol{l}'_2 = (x', y', 0) + (0, 0, z')$$

Then,

$$\begin{aligned}\mathbf{0} = (0,0,0) &= \boldsymbol{x} - \boldsymbol{x} \\ &= (x, y, 0) + (0, 0, z) - (x', y', 0) + (0, 0, z') \\ &= (x - x',\ y - y',\ z - z')\end{aligned}$$

Therefore,

$$x - x' = 0, \quad y - y' = 0, \quad z - z' = 0$$

implying $x = x'$, $y = y'$ and $z = z'$, that is, $\boldsymbol{l}_1 = \boldsymbol{l}'_1$ and $\boldsymbol{l}_2 = \boldsymbol{l}'_2$. Thus, $\mathbf{R}^3 = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2$, as claimed. ■

Thus,

Theorem 6.19 Suppose $\boldsymbol{L}_1, \boldsymbol{L}_2, \ldots, \boldsymbol{L}_n$ are subspaces of $\boldsymbol{X}$. Then, $\boldsymbol{X} = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2 \oplus \cdots \oplus \boldsymbol{L}_n$ if the following conditions are satisfied:

(i) $\boldsymbol{X} = \boldsymbol{L}_1 + \boldsymbol{L}_2 + \cdots + \boldsymbol{L}_n$ and
(ii) $0 = 0 + 0 + \cdots + 0$

Proof First, suppose $\boldsymbol{X} = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2 \oplus \cdots \oplus \boldsymbol{L}_n$, then (i) is certainly satisfied. To prove (ii) we need to prove that

$$0 = \boldsymbol{l}_1 + \boldsymbol{l}_2 + \cdots + \boldsymbol{l}_n, \quad \boldsymbol{l}_i \in \boldsymbol{L}_i$$

is true when

$$\boldsymbol{l}_1 = \boldsymbol{l}_2 = \cdots = \boldsymbol{l}_n = 0.$$

But this immediately follows from our assumption that $\boldsymbol{X} = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2 \oplus \cdots \oplus \boldsymbol{L}_n$, that is, the uniqueness of expression of every vector from $\boldsymbol{X} = \boldsymbol{L}_1 \oplus \boldsymbol{L}_2 \oplus \cdots \oplus \boldsymbol{L}_n$.

Next, assuming that (i) and (ii) are satisfied, we can express some $\boldsymbol{x} \in \boldsymbol{X}$ as

$$\boldsymbol{x} = \boldsymbol{l}_1 + \boldsymbol{l}_2 + \cdots + \boldsymbol{l}_n, \quad \boldsymbol{l}_i \in \boldsymbol{L}_i$$

We need to prove that this representation is unique. Well, suppose that this representation is not unique, that is, suppose that we can also write

$$\boldsymbol{x} = \boldsymbol{l}'_1 + \boldsymbol{l}'_2 + \cdots + \boldsymbol{l}'_n, \quad \boldsymbol{l}'_i \in \boldsymbol{L}_i$$

It follows that

$$\boldsymbol{x} - \boldsymbol{x} = 0 = (\boldsymbol{l}_1 - \boldsymbol{l}'_1) + (\boldsymbol{l}_2 - \boldsymbol{l}'_2) + \cdots + (\boldsymbol{l}_n - \boldsymbol{l}'_n)$$

where, of course, $(\boldsymbol{l}_i - \boldsymbol{l}'_i) \in \boldsymbol{L}_i$. From (ii), it immediately follows that $\boldsymbol{l}_i - \boldsymbol{l}'_i = 0$, that is, $\boldsymbol{l}_i = \boldsymbol{l}'_i$ and thus the representation of the vector $\boldsymbol{x}$ is unique. ■

Theorem 6.20 Let $\boldsymbol{L}, \boldsymbol{M} < \boldsymbol{X}$ be two subspaces of $\boldsymbol{X}$. Then,

$$\dim (\boldsymbol{L} + \boldsymbol{M}) = \dim \boldsymbol{L} + \dim \boldsymbol{M} - \dim(\boldsymbol{L} \cap \boldsymbol{M})$$

Proof Let $\dim \boldsymbol{L} = l$, $\dim \boldsymbol{M} = m$, and $\dim(\boldsymbol{L} \cap \boldsymbol{M}) = t$, and let

$$\boldsymbol{B}^{(0)} = \{\boldsymbol{b}_1^{(0)}, \boldsymbol{b}_2^{(0)}, \ldots, \boldsymbol{b}_t^{(0)}\}$$

be some basis of the space $\boldsymbol{L} \cap \boldsymbol{M}$. Therefore, the set of vectors $\{\boldsymbol{b}_k^{(0)}\}$ is linearly independent, both in $\boldsymbol{L}$ and $\boldsymbol{M}$. Thus, according to Theorem 6.11, we can always extend it to a basis of space $\boldsymbol{L}$ and a basis of space $\boldsymbol{M}$. So, there is a set of vectors $\{\boldsymbol{a}_i\} \subseteq \boldsymbol{L}$ such that

$$\boldsymbol{B}^{(1)} = \{\boldsymbol{b}_1^{(0)}, \boldsymbol{b}_2^{(0)}, \ldots, \boldsymbol{b}_t^{(0)}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r\} \tag{6.8}$$

is a basis of space $\boldsymbol{L}$.

Similarly, we can find a set of vectors $\{\boldsymbol{b}_j\} \subseteq \boldsymbol{M}$ such that

$$\boldsymbol{B}^{(2)} = \{\boldsymbol{b}_1^{(0)}, \boldsymbol{b}_2^{(0)}, \ldots, \boldsymbol{b}_t^{(0)}, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_s\} \tag{6.9}$$

is a basis of the space $\boldsymbol{M}$.

Thus,

$$\dim \boldsymbol{L} = l = t + r \tag{6.10}$$

and

$$\dim \boldsymbol{M} = m = t + s \tag{6.10'}$$

Consider a set

$$\boldsymbol{B} = \boldsymbol{B}^{(1)} \cup \boldsymbol{B}^{(2)} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_r, \boldsymbol{b}_1^{(0)}, \boldsymbol{b}_2^{(0)}, \ldots, \boldsymbol{b}_t^{(0)}, \boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_s\} \tag{6.11}$$

We would like to show that $\boldsymbol{B}$ is a basis of $\boldsymbol{L} + \boldsymbol{M}$. Well, let $\boldsymbol{x} \in \boldsymbol{L} + \boldsymbol{M}$ be such that $\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b}$, with $\boldsymbol{a} \in \boldsymbol{L}$ and $\boldsymbol{b} \in \boldsymbol{M}$. But, since we have already constructed the bases for $\boldsymbol{L}$ and $\boldsymbol{M}$, we can express $\boldsymbol{x}$ as a linear combination of vectors from $\boldsymbol{B}$.

Thus, $\boldsymbol{B}$ spans $\boldsymbol{L}+\boldsymbol{M}$. It remains to be seen whether $\boldsymbol{B}$ is a set of linearly independent vectors. To do that, we examine the linear combination

$$\alpha_1\boldsymbol{a}_1+\cdots+\alpha_r\boldsymbol{a}_r+\beta_1\boldsymbol{b}_1^{(0)}+\cdots+\beta_t\boldsymbol{b}_t^{(0)}+\gamma_1\,\boldsymbol{b}_1+\cdots+\,\gamma_s\boldsymbol{b}_s=\boldsymbol{0} \tag{6.12}$$

Take a vector $\boldsymbol{b}\in\boldsymbol{M}$. It can be written as

$$\boldsymbol{b}=\sum_{j=1}^{s}\gamma_j\boldsymbol{b}_j$$

but also as

$$\boldsymbol{b}=-\alpha_1\boldsymbol{a}_1-\cdots-\alpha_r\boldsymbol{a}_r-\beta_1\boldsymbol{b}_1^{(0)}-\cdots-\beta_t\boldsymbol{b}_t^{(0)} \tag{6.13}$$

Thus, $\boldsymbol{b}\in\boldsymbol{L}$, and therefore $\boldsymbol{b}\in\boldsymbol{L}\cap\boldsymbol{M}$. Consequently, in (6.13),

$$\alpha_1=\cdots=\alpha_r=0 \tag{6.14}$$

It follows from (6.12) that

$$\beta_1\boldsymbol{b}_1^{(0)}+\cdots+\beta_t\boldsymbol{b}_t^{(0)}+\gamma_1\,\boldsymbol{b}_1+\cdots\,+\gamma_s\boldsymbol{b}_s=\boldsymbol{0} \tag{6.15}$$

However, all the vectors in (6.15) are elements of the basis of $\boldsymbol{B}^{(2)}$; therefore, they are linearly independent. Hence,

$$\beta_1=\cdots=\beta_t=\gamma_1=\cdots=\gamma_s=0 \tag{6.16}$$

From (6.14) and (6.16), it immediately follows that expression (6.12) is necessarily trivial, hence $\boldsymbol{B}$ is a set of linearly independent vectors. This completes our proof. ■

Consequently, we also have

Theorem 6.21 Let $\boldsymbol{L},\boldsymbol{M}<X$ be two subspaces. Then, $\boldsymbol{L}+\boldsymbol{M}$ is a direct sum iff

$$\dim(\boldsymbol{L}+\boldsymbol{M})=\dim\boldsymbol{L}+\dim\boldsymbol{M}$$

Proof Suppose $\boldsymbol{L}+\boldsymbol{M}$ is a direct sum. Then, $\boldsymbol{L}\cap\boldsymbol{M}=\{\boldsymbol{0}\}$ so

$$\dim(\boldsymbol{L}+\boldsymbol{M})=\dim\{\boldsymbol{0}\}=0$$

It follows from the previous theorem that $\boldsymbol{L}+\boldsymbol{M}$ is a direct sum.

Conversely, if

$$\dim(\boldsymbol{L}+\boldsymbol{M})=\dim\boldsymbol{L}+\dim\boldsymbol{M}$$

Then, $\boldsymbol{L}\cap\boldsymbol{M}=\{\boldsymbol{0}\}$ and $\dim(\boldsymbol{L}+\boldsymbol{M})=0$, and the sum $\boldsymbol{L}+\boldsymbol{M}$ is a direct sum. ■

Theorem 6.22 Let $\boldsymbol{L} < X$ be any subspace of an n-dimensional space $\boldsymbol{X}$. Then there exists $\boldsymbol{M} < X$, called **direct complement of $\boldsymbol{L}$**, such that

$$\boldsymbol{X} = \boldsymbol{L} \oplus \boldsymbol{M}$$

Proof Certainly, the claim is true in trivial cases $\boldsymbol{L} = \{\boldsymbol{0}\}$ and $\boldsymbol{L} = \boldsymbol{X}$. So, let

$$\boldsymbol{B}_{\boldsymbol{L}} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_k\}$$

be some basis of $\boldsymbol{L} \neq \{\boldsymbol{0}\} \neq \boldsymbol{X}$. Then, again recalling Theorem 6.11, we supplement $\boldsymbol{B}_{\boldsymbol{L}}$ with additional vectors $\{\boldsymbol{b}_{k+1}, \ldots, \boldsymbol{b}_n\}$ to form a basis of $\boldsymbol{X}$. Then, any $\boldsymbol{x} \in \boldsymbol{X}$ can be written as

$$\begin{aligned}\boldsymbol{x} &= \sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i = \sum_{i=1}^{k} \alpha_i \boldsymbol{b}_i + \sum_{i=k+1}^{n} \alpha_i \boldsymbol{b}_i \\ &= \boldsymbol{l} + \boldsymbol{m}\end{aligned}$$

where $\boldsymbol{l} \in \boldsymbol{L}$ and $\boldsymbol{m} \in \boldsymbol{M}$. So, evidently, $\boldsymbol{x} \in \boldsymbol{L} + \boldsymbol{M}$. Suppose there exists a nonzero vector $\boldsymbol{y} \in \boldsymbol{L} \cap \boldsymbol{M}$. Then,

$$\boldsymbol{y} = \sum_{i=1}^{k} \beta_i \boldsymbol{b}_i = \sum_{i=k+1}^{n} \beta_i \boldsymbol{b}_i$$

from which follows

$$\sum_{i=1}^{k} \beta_i \boldsymbol{b}_i - \sum_{i=k+1}^{n} \beta_i \boldsymbol{b}_i = 0$$

However,

$$\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$$

is a basis for the space $\boldsymbol{X}$, and therefore $\beta_1 = \cdots = \beta_n = 0$. Thus, $\boldsymbol{y} = \boldsymbol{0}$ contradicts our assumption that $\boldsymbol{L} \cap \boldsymbol{M} \neq \{\boldsymbol{0}\}$. Hence, $\boldsymbol{X} = \boldsymbol{L} \oplus \boldsymbol{M}$ as claimed. ■

6.6 LINEAR TRANSFORMATIONS – LINEAR OPERATORS

Once we have a mathematical object with a particular structure, and after having examined its basic properties, we look at mappings between two or more of these objects. Similar to homomorphisms of groups, the most important mappings in the case of linear spaces are those that preserve their structure, that is, preserve vector addition and scalar multiplication. The study of Linear Transformations, or Linear Operators, is one of the most important parts of Linear Algebra.

Definition 6.12 Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two vector spaces over the same field Φ, and let

$$F \ : \boldsymbol{X} \rightarrow \boldsymbol{Y}$$

be some map from $\boldsymbol{X}$ to $\boldsymbol{Y}$, that is, a function defined on a vector space $\boldsymbol{X}$ with values in a vector space $\boldsymbol{Y}$. We say that F is a **linear transformation** or a **linear operator**[3] if it satisfies the following conditions:

(i) $F(\boldsymbol{x}+\boldsymbol{y}) = F(\boldsymbol{x}+\boldsymbol{y}), \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$
(ii) $F(\alpha \boldsymbol{x}) = \alpha F(\boldsymbol{x}), \qquad \forall \alpha \in \Phi, \ \forall \boldsymbol{x} \in \boldsymbol{X}$

I am sure that by now you recognize that conditions (i) and (ii) could be combined, and we could simply say that f is a linear operator iff

(iii) $F(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha F(\boldsymbol{x}) + \beta F(\boldsymbol{y}), \qquad \forall \alpha, \ \beta \in \Phi, \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$

Again, remembering groups and their homomorphisms, one notes that because of (i) every linear operator F is a homomorphism of additive groups of respective vector spaces. Clearly then, we have

Theorem 6.23 Let $F \ : \boldsymbol{X} \rightarrow \boldsymbol{Y}$ be a linear operator, then

$$F(\mathbf{0}_{\boldsymbol{X}}) = \mathbf{0}_{\boldsymbol{Y}}$$

Proof

$$\begin{aligned} F(\mathbf{0}_{\boldsymbol{X}}) &= F(0\mathbf{0}_{\boldsymbol{X}}) \\ &= 0F(\mathbf{0}_{\boldsymbol{X}}) \\ &= \mathbf{0}_{\boldsymbol{Y}} \end{aligned}$$

And we often simply write $F(\mathbf{0}) = \mathbf{0}$. ■

Similarly, one can easily prove that

$$F(-\boldsymbol{x}) = -F(\boldsymbol{x})$$

Example 6.37 Let $\boldsymbol{X}$ be a vector space over a field Φ. Define a map

$$H \ : \boldsymbol{X} \ \rightarrow \boldsymbol{X}$$

by $H(\boldsymbol{x}) = \lambda \boldsymbol{x}, \quad \forall \, \boldsymbol{x} \in \boldsymbol{X}, \ \text{and} \ \forall \lambda \in \Phi$. Show that H is a linear operator.

[3] Some authors distinguish between linear transformation and linear operator stating that the latter is a map $\boldsymbol{X} \rightarrow \boldsymbol{X}$.

Solution

$$\begin{aligned} H(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) &= \lambda(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) \\ &= (\lambda\alpha)\boldsymbol{x} + (\lambda\beta)\boldsymbol{y} \\ &= \alpha(\lambda \boldsymbol{x}) + \beta(\lambda\, \boldsymbol{y}) \\ &= \alpha H(\boldsymbol{x}) + \beta H(\boldsymbol{y}), \quad \alpha, \beta \in \Phi \end{aligned}$$

■

The following example is self-evident.

Example 6.38 The identity map $I : X \to X$ defined on some vector space by

$$I(\boldsymbol{x}) = \boldsymbol{x}, \quad \forall \in \boldsymbol{X}$$

is a linear operator. ■

Example 6.39 If you are familiar with ordinary differentiation, you can immediately recognize that the differentiation operator $\frac{d}{dx}$, acting on the set of all differentiable functions $\mathcal{F}$ and mapping it to the set $\mathcal{F}'$, is another well-known example of a linear operator. ■

Example 6.40 Consider function $F : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x, y, z) = (x, y)$. Show that F is a linear operator.

Solution Let $\boldsymbol{x} = (x_1,\ x_2,\ x_3)$, $\boldsymbol{y} = (y_1,\ y_2,\ y_3)$ be two vectors from $\mathbf{R}^3$, then

$$\begin{aligned} F(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) &= F(\alpha(x_1,\ x_2,\ x_3) + \beta(y_1,\ y_2,\ y_3)) \\ &= F(\alpha x_1 + \beta y_1,\ \alpha x_2 + \beta y_2,\ \alpha x_3 + \ \beta y_3) \\ &= \ \alpha x_1 + \beta y_1,\ \alpha x_2 + \beta y_2 \\ &= (\ \alpha x_1 + \ \alpha x_2,\ \beta y_1 + \ \beta y_2) \\ &= (\alpha(x_1 + x_2), \beta(y_1 + y_2)) \\ &= \alpha(x_1 + x_2) + \beta(y_1 + y_2) \\ &= \alpha F(\boldsymbol{x}) + \beta F(\boldsymbol{y}) \end{aligned}$$

Hence, F is a linear operator. ■

Example 6.41 Consider the map $\pi : \mathbf{R}^3 \to \mathbf{R}^3$ defined by

$$\pi(\boldsymbol{x}) = \pi(x_1, x_2, x_3) = (x_1, x_2)$$

It is easy to verify that π is a linear operator called (for obvious reasons) the projection. Take $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^3$ and consider

$$\begin{aligned}\pi(\boldsymbol{x}+\boldsymbol{y}) &= \pi[(x_1, x_2, x_3) + (y_1, y_2, y_3)] \\ &= \pi[x_1 + y_1,\ x_2 + y_2, x_3 + y_3] \\ &= (x_1 + y_1,\ x_2 + y_2, 0) \\ &= (x_1, x_2, 0) + (y_1, y_2, 0) \\ &= \pi(\boldsymbol{x}) + \pi(\boldsymbol{y})\end{aligned}$$

Also, for any scalar $k \in \mathbf{R}$

$$\begin{aligned}\pi(\alpha \boldsymbol{x}) &= \pi(\alpha x_1, \alpha x_2, \alpha x_3) \\ &= (\alpha x_1, \alpha x_2, 0) \\ &= \alpha(x_1, x_2) \\ &= \alpha \pi(\boldsymbol{x})\end{aligned}$$

Thus, π is indeed a linear operator. ■

Example/Exercise 6.42 Consider the map $\rho : \mathbf{R}^2 \to \mathbf{R}^2$ (called **reflection**) defined by

$$\rho(x_1, x_2) = (x_1, -x_2)$$

Show that ρ is a linear operator.

Example 6.43 Let $\boldsymbol{L} < \boldsymbol{X}$ be a subspace of $\boldsymbol{X}$, and let $i : \boldsymbol{L} \to \boldsymbol{X}$ be a map (called **inclusion**) defined by

$$i(\boldsymbol{l}) = \boldsymbol{l}, \quad \forall\, \boldsymbol{l} \in \boldsymbol{L}$$

We can easily see that i is a linear operator. ■

Example/Exercise 6.44 Show that the map $\rho : \mathbf{R}^2 \to \mathbf{R}^2$ (called **reflection**) defined by

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{y}, \boldsymbol{x})$$

is a linear operator.

Theorem 6.24 Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be a set if vectors from $\boldsymbol{X}$ and $\{\alpha_1, \ldots, \alpha_n\}$ a set of scalars from Φ. Then, a linear operator f satisfies the following:

$$F\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i\right) = \sum_{i=1}^{n} \alpha_i F(\boldsymbol{x}_i)$$

Proof We can easily prove this by induction. Clearly, equality is true for $n = 1$, so we assume it also holds for $n = k$. Let's check whether it is true for $n = k + 1$.

$$\begin{aligned} F\left(\sum_{i=1}^{k+1}\alpha_i \boldsymbol{x}_i\right) &= F\left(\sum_{i=1}^{k}\alpha_i \boldsymbol{x}_i + \alpha_{k+1}\boldsymbol{x}_{k+1}\right) \\ &= \sum_{i=1}^{k}\alpha_i F(\boldsymbol{x}_i) + F(\alpha_{k+1}\boldsymbol{x}_{k+1}) \\ &= F\left(\sum_{i=1}^{k}\alpha_i \boldsymbol{x}_i\right) + F(\alpha_{k+1}\boldsymbol{x}_{k+1}) \\ &= \sum_{i=1}^{k}\alpha_i F(\boldsymbol{x}_i) + \alpha_{k+1}F(\boldsymbol{x}_{k+1}) \\ &= \sum_{i=1}^{k+1}\alpha_i F(\boldsymbol{x}_i) \end{aligned}$$

Since the equality holds for $n = k + 1$ we conclude that it holds for any n. ■

Theorem 6.25 Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator, and suppose that $\boldsymbol{L} < \boldsymbol{X}$ is a subspace of $\boldsymbol{X}$. Furthermore, let

$$F(\boldsymbol{L}) = \{\, F(\boldsymbol{l}) |\, \boldsymbol{l} \in \boldsymbol{L}\} \subseteq \boldsymbol{Y}$$

be a set of all the images of $\boldsymbol{l} \in \boldsymbol{L}$. Then, $F(\boldsymbol{L})$ is a subspace of $\boldsymbol{Y}$.

Proof Suppose $\boldsymbol{l}'_1, \boldsymbol{l}'_2 \in F(\boldsymbol{L})$ are any two vectors such that $\boldsymbol{l}'_1 = F(\boldsymbol{l}_1)$ and $\boldsymbol{l}'_2 = F(\boldsymbol{l}_2)$. Then, since $\boldsymbol{L}$ is a subspace, $\alpha\boldsymbol{l}_1 + \beta\boldsymbol{l}_2 \in \boldsymbol{L}$, and therefore

$$\begin{aligned} F(\alpha\boldsymbol{l}_1 + \beta\boldsymbol{l}_2) &= \alpha F(\boldsymbol{l}_1) + \beta F(\boldsymbol{l}_2) \\ &= \alpha\boldsymbol{l}'_1 + \beta\boldsymbol{l}'_2 \in \boldsymbol{Y} \end{aligned}$$

So, $F(\boldsymbol{L})$ is a subspace of $\boldsymbol{Y}$, as claimed. ■

Theorem 6.26 The composition of linear operators is a linear operator.

Proof Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ and $G : \boldsymbol{Y} \to \boldsymbol{W}$ be two linear operators, and let $H = G \circ H$ be their composition, that is

$$H : \boldsymbol{X} \to \boldsymbol{W}$$

We have

$$\begin{aligned} H(\alpha \boldsymbol{x}) &= (G \circ F)(\alpha \boldsymbol{x}) \\ &= G[F(\alpha \boldsymbol{x})] \\ &= G[\alpha F(\boldsymbol{x})] \\ &= \alpha[G(F(\boldsymbol{x})] \\ &= \alpha[(G \circ F)(\boldsymbol{x})] \\ &= \alpha H(\boldsymbol{x}) \end{aligned}$$

Thus, H is a linear operator. ■

Some comments on notation and terminology may now be in order. We will be studying only linear operators, since they are the most important in algebra, so the term "linear" will be occasionally omitted as is often the case in literature. Also, for the sake of simplicity and when there is no confusion, very often one writes $A\boldsymbol{x}$ instead of $A(\boldsymbol{x})$ for an operator A acting on a vector $\boldsymbol{x}$.

As you might have expected, we are now in a position to construct even more interesting vector spaces, illustrated by the following

Example 6.45 Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two vector spaces over the same field Φ. Let[4]

$$\text{Hom}(\boldsymbol{X}, \boldsymbol{Y}) = \{ F \ : \boldsymbol{X} \to \boldsymbol{Y} \mid F \text{ is a linear operator} \} = \{ \boldsymbol{X} \to \boldsymbol{Y} \}$$

be the set of all linear operators from X to Y. Then, Hom $(\boldsymbol{X}, \boldsymbol{Y})$ is a vector space. Indeed, if we consider all operators $F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$ to be vectors, we can naturally say:

(i) $(F + G)(\boldsymbol{x}) = F\boldsymbol{x} + G\boldsymbol{x} = G\boldsymbol{x} + F\boldsymbol{x} = (G + F)(\boldsymbol{x}), \ \ \forall\, F, G \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$

Thus, $F + G = G + F$.

(ii) $(F + (G + H))(\boldsymbol{x}) = (\,(F + G) + H\,)(\boldsymbol{x}), \ \ \forall\, F, G, H \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$

Thus, $(G + H) = (F + G) + H$.

(iii) $(\alpha(F + G))(\boldsymbol{x}) = \alpha F\boldsymbol{x} + \alpha G\boldsymbol{x}, \quad \forall F, G \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y}), \quad \forall \alpha \in \Phi$

Thus, $\alpha(F + H) = \alpha F + \alpha G$.

(iv) $((\alpha + \beta)F)(\boldsymbol{x}) = (\alpha F + \beta F)(\boldsymbol{x}) = \alpha F\boldsymbol{x} + \beta F\boldsymbol{x}, \ \ \forall \alpha, \beta \in \Phi,$
$F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$

Thus, $(\alpha + \beta)F = \alpha F + \beta F$.

[4]Sometimes this set is denoted by $\mathcal{L}(X,\ Y)$ or Lin$(\boldsymbol{X}, \boldsymbol{Y})$.

(v) For each $F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$, $\exists(-F) \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$, such that $(F + (-F))(\boldsymbol{x}) = F\boldsymbol{x} - F\boldsymbol{x} = \boldsymbol{O}\boldsymbol{x} = \boldsymbol{0}$.

(vi) $(O + F)(\boldsymbol{x}) = F\boldsymbol{x}$, where $O(\boldsymbol{x}) = O\boldsymbol{x} = \boldsymbol{0}$, $\forall \boldsymbol{x} \in \boldsymbol{X}$ and $O \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$ and $\boldsymbol{0} \in \boldsymbol{Y}$

Thus, $O + F = F$, $\forall F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$.

(vii) $((\alpha\beta)F)(\boldsymbol{x}) = (\alpha\beta)F(\boldsymbol{x}) = \alpha(\beta F)\boldsymbol{x}$, $\forall \alpha, \beta \in \Phi$, $F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$

Thus,

$$(\alpha\beta)F = \alpha(\beta F)$$

And also

$$O + F = F, \quad \forall F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$$

(viii) $\exists\, 1 \in \Phi$, such that $1F = F1 = F$, $\forall F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$. ■

Formally, we have

Definition 6.13 Endowed with the usual addition of operators and multiplication by a scalar, $\text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$ is a vector space.

In other words, we claim

Theorem 6.27 Let $\boldsymbol{X}$, $\boldsymbol{Y}$ be two vector spaces over a field Φ. Then, the set of all linear operators defined on $\boldsymbol{X}$ and with values in $\boldsymbol{Y}$ becomes a vector space with the usual addition of operators and multiplication by a scalar.

Proof The proof immediately follows from the two lemmas below (see also Example 6.45). ■

Lemma 6.2 Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ and $G : \boldsymbol{X} \to \boldsymbol{Y}$ be two linear operators. Then,

$$(F + G) : \boldsymbol{X} \to \boldsymbol{Y}$$

is a linear operator too.

Proof Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ be any two vectors, and $\alpha, \beta \in \Phi$ any two scalars. Then,

$$\begin{aligned}(F + G)(\alpha\,\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) &= F(\alpha\,\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) + G(\alpha\,\boldsymbol{x}_1 + \beta\boldsymbol{x}_2)\\ &= \alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2) + \alpha G(\boldsymbol{x}_1) + \beta G(\boldsymbol{x}_2)\\ &= \alpha(F(\boldsymbol{x}_1) + G(\boldsymbol{x}_1)) + \beta(F(\boldsymbol{x}_2) + G(\boldsymbol{x}_2))\\ &= \alpha(F + G)(\boldsymbol{x}_1) + \beta(F + G)(\boldsymbol{x}_2)\end{aligned}$$

Hence, $(F + G)$ is a linear operator. ■

Lemma 6.3 The set Hom($\boldsymbol{X}$, $\boldsymbol{Y}$) with the usual addition of operators is an abelian group.

Proof First, observe that if $F : \boldsymbol{X} \to \boldsymbol{Y}$ is a linear operator, then $\forall \boldsymbol{x} \in \boldsymbol{X}$ and $\lambda \in \Phi$

$$\lambda F : \boldsymbol{X} \to \boldsymbol{Y}$$

is well defined by

$$(\lambda F)(\boldsymbol{x}) = \lambda F(\boldsymbol{x})$$

Thus, for every operator F, there is an "inverse" operator $G = -F$. A neutral element immediately follows as a nil-operator

$$O : \boldsymbol{X} \to \boldsymbol{Y}$$

defined by $O(\boldsymbol{x}) = \boldsymbol{0}$. So, Hom($\boldsymbol{X}$, $\boldsymbol{Y}$) is a group. It is easy to verify that it is an abelian. ■

Without proof we state the following important

Theorem 6.28 If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are *finite-dimensional* linear spaces over a field Φ, then

$$\dim \text{Hom}(\boldsymbol{X}, \boldsymbol{Y}) = \dim \boldsymbol{X} \cdot \dim \boldsymbol{Y}$$

Definition 6.14 Let $\boldsymbol{X}$ be a vector space over a field Φ. The set $\mathcal{L}(\boldsymbol{X})$ of all linear mappings from a vector space $\boldsymbol{X}$ into itself is called an **algebra** if, in addition to axioms A1–A8, $\mathcal{L}(\boldsymbol{X})$ is endowed with "multiplication" defined as a composition of mappings. In other words, an algebra $\mathcal{L}(\boldsymbol{X})$ over a field Φ is a vector space over Φ, if for all $F, G, H \in \mathcal{L}(\boldsymbol{X})$and for all $\alpha \in \Phi$ the following is satisfied:

(i) $F(G + H) = FG + FH$
(ii) $(G + H)F = GF + HF$
(iii) $\alpha(FG) = (\alpha F)G = F(\alpha G)$

The algebra is said to be **associative** if, in addition to (i)–(iii), the following is also satisfied:

(iv) $F(GH) = (FG)H$

Theorem 6.29 Let $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be vector spaces over a field Φ, and suppose $F : \boldsymbol{X} \to \boldsymbol{Y}$ and $G : \boldsymbol{Y} \to \boldsymbol{Z}$ are linear operators. Then, $G \circ F = GF$ is a linear operator too.

Proof Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ be any two vectors from X, and $\alpha, \beta \in \Phi$. Then,

$$\begin{aligned}(G \circ F)(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) &= G(\, F(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2)) = G(\alpha F \boldsymbol{x}_1 + \beta F \boldsymbol{x}_2) \\ &= \alpha G(F \boldsymbol{x}_1) + \beta G(F \boldsymbol{x}_2) \\ &= \alpha (GF) \boldsymbol{x}_1 + \beta (GF) \boldsymbol{x}_2 \\ &= \alpha (G \circ F) \boldsymbol{x}_1 + \beta (G \circ F)\, \boldsymbol{x}_2\end{aligned}$$

Thus, $(G \circ F)$ is a linear operator. ■

Example 6.46 Let $\mathcal{L}(\boldsymbol{X})$ be an algebra, and let $F \in \mathcal{L}(\boldsymbol{X})$. Then, with $F^0 = I$, $F^2 = F \circ F$, $F^3 = F^2 \circ F = F \circ F \circ F$, … we can construct another linear operator

$$p(F) = \alpha_0 I + \alpha_1 F + \alpha_2 F^2 + \cdots + \alpha_n F^n$$ ■

Definition 6.15 Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator. We say that the **kernel** of F is a set

$$\operatorname{Ker} F = F^{-1}(\boldsymbol{0}_{\boldsymbol{Y}}) = \{\, \boldsymbol{x} \in \boldsymbol{X} |\; F(\boldsymbol{x}) = \boldsymbol{0}_{\boldsymbol{Y}} \}$$

As noted before, unless it is really necessary, we will drop the index on the zero vector and instead of $\boldsymbol{0}_{\boldsymbol{Y}}$simply write $\boldsymbol{0}$.

Definition 6.16 Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator. We say that the **image of** F is a set

$$\operatorname{Im} F = \{\, \boldsymbol{y} \in \boldsymbol{Y} |\; \boldsymbol{y} = F(\boldsymbol{x}), \quad \boldsymbol{x} \in \boldsymbol{X} \}$$

Definition 6.17 An operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ is said to be **singular** if there exists a nonzero $\boldsymbol{x} \in \boldsymbol{X}$, such that $F(\boldsymbol{x}) = 0$, that is, if $\operatorname{Ker} F \neq \{0\}$. Conversely, if $\operatorname{Ker} F = \{0\}$ we say that the operator F is **nonsingular** (**regular**).

Theorem 6.30 If $F : \boldsymbol{X} \to \boldsymbol{Y}$ is a linear operator, then $\operatorname{Ker} F$ is a subspace of $\boldsymbol{X}$.

Proof Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ be any two vectors in $\boldsymbol{X}$, and let $\alpha, \beta \in \Phi$ be two scalars from Φ such that

$$F(\boldsymbol{x}_1) = F(\boldsymbol{x}_2) = \boldsymbol{0}$$

that is, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \operatorname{Ker} \boldsymbol{X}$

Then,

$$\begin{aligned}F(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) &= \alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2) \\ &= \alpha \boldsymbol{0} + \beta \boldsymbol{0} \\ &= \boldsymbol{0}\end{aligned}$$

Thus, $\operatorname{Ker} \boldsymbol{X}$ is a subspace of X. ■

Definition 6.18 If $F : \boldsymbol{X} \to \boldsymbol{X}$ is a linear operator, such that for each $\boldsymbol{x} \in \boldsymbol{L} < X$, $F(\boldsymbol{x}) \in \boldsymbol{L}$, we say that $\boldsymbol{L}$ is **invariant** under F.

Example 6.47 Three trivial examples are as follows: (i) $\text{Ker}\, F$, (ii) $\text{Im}\, F$, and (iii) $\boldsymbol{X}$ as a whole is invariant under F. This is easy to see:

(i) Since F is a linear operator, for each $\boldsymbol{x} \in \text{Ker}\, F$, $F(\boldsymbol{x}) = \boldsymbol{0}$. And, $\boldsymbol{0} \in \text{Ker}\, F$ since $F(\boldsymbol{0}) = \boldsymbol{0}$. But, $\text{Ker}\, F < \boldsymbol{X}$, thus $\text{Ker}\, F$ is invariant under F.

Similarly, you can convince yourself about (ii) and (iii). ■

Theorem 6.31 If $F : \boldsymbol{X} \to \boldsymbol{Y}$ is a linear operator, then $\text{Im}\, F$ is a subspace of $\boldsymbol{Y}$.

Proof Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator from $\boldsymbol{X}$ to $\boldsymbol{Y}$. Then, by Theorem 6.30 $F(\boldsymbol{0}) = \boldsymbol{0}$, which implies that $\boldsymbol{0} \in \text{Im}\, F$.

Suppose $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \text{Im}\, F$, then there exists $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ such that $F(\boldsymbol{x}_1) = \boldsymbol{y}_1$ and $F(\boldsymbol{x}_2) = \boldsymbol{y}_2$.

Thus,

$$F(\boldsymbol{x}_1 + \boldsymbol{x}_2) = F(\boldsymbol{x}_1) + F(\boldsymbol{x}_2) = \boldsymbol{y}_1 + \boldsymbol{y}_2$$

and therefore $(\boldsymbol{y}_1 + \boldsymbol{y}_2) \in \text{Im}\, F$.

Similarly, if $\boldsymbol{y} \in \text{Im}\, F$ and $\alpha \in \Phi$, then there exists $\boldsymbol{x} \in \boldsymbol{X}$ such that $F(\boldsymbol{x}) = \boldsymbol{y}$. It follows that

$$F(\alpha \boldsymbol{x}) = \alpha F(\boldsymbol{x}) = \alpha \boldsymbol{y}$$

So, $\alpha \boldsymbol{y} \in \text{Im}\, F$. Hence, $\text{Im}\, F$ contains $\boldsymbol{0}$. It is closed under addition and scalar multiplication, and thus $\text{Im}\, F < \boldsymbol{Y}$. ■

Theorem 6.32 Let $F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$. Then, F is injective iff $\text{Ker}\, F = \{\boldsymbol{0}\}$.

Proof First, suppose F is injective. We need to prove that $\text{Ker}\, F = \{\boldsymbol{0}\}$. Well, since $\text{Ker}\, F < \boldsymbol{X}$, $\{\boldsymbol{0}\} \subseteq \text{Ker}\, F$. But, F being injective implies

$$F(\boldsymbol{x}) = \boldsymbol{0}_{\boldsymbol{Y}} = F(\boldsymbol{0}_{\boldsymbol{X}})$$

Thus, $\text{Ker}\, F = \{\boldsymbol{0}\}$.

Conversely, suppose $\text{Ker}\, F = \{\boldsymbol{0}\}$, and suppose $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ such that

$$F(\boldsymbol{x}_1) = F(\boldsymbol{x}_2)$$

Then,

$$\boldsymbol{0}_{\boldsymbol{Y}} = F(\boldsymbol{x}_1) - F(\boldsymbol{x}_2) = F(\boldsymbol{x}_1 - \boldsymbol{x}_2) = F(\boldsymbol{0}_{\boldsymbol{X}})$$

Thus, F is injective, as desired. ■

Definition 6.19 A linear operator $F \in \mathrm{Hom}(\boldsymbol{X}, \boldsymbol{Y})$is said to be **surjective** if $\mathrm{Im}\, F = \boldsymbol{Y}$.

Theorem 6.33 Let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \subseteq \boldsymbol{X}$ be a set of linearly independent vectors in a space $\boldsymbol{X}$, and let $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\} \subseteq \boldsymbol{Y}$ be any set of vectors in $\boldsymbol{Y}$. Then, there exists at least one linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ such that

$$F(\boldsymbol{x}_i) = \boldsymbol{y}_i, \quad \forall i = 1, \ldots, k$$

Proof Suppose the set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ is not a basis for X. Then, according to Theorem 6.11, we can always augment it to the basis of X as, say,

$$\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_n\}$$

Let's extend the set $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k\}$ to the set $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{y}_{k+1}, \ldots, \boldsymbol{y}_n\} \subseteq \boldsymbol{Y}$. As we know, any $\boldsymbol{x} \in \boldsymbol{X}$ can be uniquely represented as a linear combination

$$\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i$$

So, let's introduce a function $F : \boldsymbol{X} \to \boldsymbol{Y}$ defined by

$$\begin{aligned} F(\boldsymbol{x}) &= F\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i\right) \\ &= \sum_{i=1}^{n} \alpha_i \boldsymbol{y}_i \in \boldsymbol{Y} \end{aligned} \qquad (*)$$

We need to show that F is a linear operator. Suppose we put in (*) $\alpha_i = 1$ and $\alpha_{j\neq i} = 0$, then

$$F(\boldsymbol{x}_i) = \boldsymbol{y}_i, \quad \forall i = 1, \ldots, n$$

exactly as needed. Next, we consider

$$\begin{aligned} F(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) &= F\left(\alpha \sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i + \beta \sum_{i=1}^{n} \beta_i \boldsymbol{x}_i\right) \\ &= F\left(\sum_{i=1}^{n} \left(\alpha\alpha_i + \beta\beta_i\right) \boldsymbol{x}_i\right) \\ &= \sum_{i=1}^{n} (\alpha\alpha_i + \beta\beta_i) \boldsymbol{y}_i \end{aligned}$$

$$
\begin{aligned}
&= \alpha \sum_{i=1}^{n} \alpha_i \boldsymbol{y}_i + \beta \sum_{i=1}^{n} \beta_i \boldsymbol{y}_i \\
&= \alpha F\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{x}_i \right) + \beta \left(\sum_{i=1}^{n} \beta_i \boldsymbol{x}_i \right) \\
&= \alpha F(\boldsymbol{x}) + \beta F(\boldsymbol{y})
\end{aligned}
$$

Thus, F is a linear operator as claimed. ■

It is worth keeping in mind that the operator F described above depends on a particular extension of the set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ to the basis of $\boldsymbol{X}$. Some other extension would produce a different operator. Let's articulate this more precisely with

Theorem 6.34 Let $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be some basis of a vector space $\boldsymbol{X}$, and let $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n\} \subseteq \boldsymbol{Y}$ be any ordered set in $\boldsymbol{Y}$. Then there exists a unique linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ such that

$$F(\boldsymbol{b}_i) = \boldsymbol{y}_i, \quad \forall\, i = 1, \ldots, n$$

Proof In the previous theorem, we showed the existence of the operator F; thus, we only have to prove uniqueness. Suppose that there exists another linear operator $G : \boldsymbol{X} \to \boldsymbol{Y}$, such that $G(\boldsymbol{b}_i) = \boldsymbol{y}_i, \ \forall\, i = 1, \ldots, n$. Then, for any $\boldsymbol{x} \in \boldsymbol{X}$

$$
\begin{aligned}
G(\boldsymbol{x}) &= G\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i \right) \\
&= \sum_{i=1}^{n} \alpha_i G(\boldsymbol{b}_i) \\
&= \sum_{i=1}^{n} \alpha_i \boldsymbol{y}_i \\
&= \sum_{i=1}^{n} \alpha_i F(\boldsymbol{b}_i) \\
&= F\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i \right) \\
&= F(\boldsymbol{x})
\end{aligned}
$$

■

Theorem 6.35 Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two vector spaces, and let $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a basis for a space $\boldsymbol{X}$. Furthermore, let $F : \boldsymbol{B} \to \boldsymbol{Y}$ be any map, then there exists a unique linear operator $G : \boldsymbol{X} \to \boldsymbol{Y}$ which extends F, that is, $G/\boldsymbol{B} = F$.

Proof The proof follows immediately from Theorem 6.31 if we put $F(\boldsymbol{b}_i) = \boldsymbol{y}_i \in \boldsymbol{Y}$. ■

Theorem 6.36 Let $\boldsymbol{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ be a standard basis of $\mathbf{R}^n$, and let $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subseteq \boldsymbol{X}$ be some set of vectors from $\boldsymbol{X}$. If $F : \mathbf{R}^n \to \boldsymbol{X}$ is a linear extension such that $F(\boldsymbol{b}_i) = \boldsymbol{x}_i$, then $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent iff $\operatorname{Ker} F = \{\mathbf{0}\}$.

Proof Suppose first that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent, then for some $\boldsymbol{r} = (r_1, \ldots, r_n) \in \mathbf{R}^n$

$$\begin{aligned} F(\boldsymbol{r}) &= F(r_1\boldsymbol{b}_1 + \cdots + r_n\boldsymbol{b}_n) \\ &= r_1F(\boldsymbol{b}_1) + \cdots + r_nF(\boldsymbol{b}_n) \\ &= r_1\boldsymbol{x}_1 + \cdots + r_n\boldsymbol{x}_n \end{aligned}$$

If $F(\boldsymbol{r}) = \mathbf{0}$ then $r_1\boldsymbol{x}_1 + \cdots + r_n\boldsymbol{x}_n = 0$, and, since we assumed that $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent, $r_1 = r_2 = \cdots = r_n = 0$ and therefore $\boldsymbol{r} = \mathbf{0}$, that is, $\operatorname{Ker} F = \{\mathbf{0}\}$.

On the other hand, if we assume that $\operatorname{Ker} F = \{\mathbf{0}\}$ and we take

$$\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_n\boldsymbol{x}_n = \mathbf{0}$$

which means that

$$\begin{aligned} F(\alpha_1\boldsymbol{b}_1 + \cdots + \alpha_n\boldsymbol{b}_n) &= \alpha_1F(\boldsymbol{b}_1) + \cdots + \alpha_nF(\boldsymbol{b}_n) \\ &= \alpha_1\boldsymbol{x}_1 + \cdots + \alpha_n\boldsymbol{x}_n = \mathbf{0} \end{aligned}$$

that is, $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ and therefore $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent. ■

6.7 ISOMORPHISM OF LINEAR SPACES

A natural question provoked by the previous discussion arises: Is there a way to classify linear spaces and how could one do that? Similar to ideas in group theory, one would like to know in what sense could two spaces be considered the same, and in what sense are they irreconcilably different. The following definition comes with no surprise.

Definition 6.20 Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be linear spaces over the same field Φ and let

$$F : \boldsymbol{X} \to \boldsymbol{Y}$$

be a linear operator. We say that F is an **isomorphism of linear spaces** if

(i) F is a linear operator and
(ii) F is a bijection.

We say that the isomorphism $F : \boldsymbol{X} \to \boldsymbol{X}$ is an **automorphism of $\boldsymbol{X}$**; we say that the operator F is a **regular operator (or nonsingular)**. If the linear operator is not regular, we say that it is **singular**. Equivalently, we have

Definition 6.20′ If $F : \boldsymbol{X} \to \boldsymbol{Y}$ is a linear transformation of vector spaces, we say that F is an isomorphism of spaces if there exists a linear operator

$$G : \boldsymbol{Y} \to \boldsymbol{X}$$

such that

(i) $G(F(\boldsymbol{x})) = \boldsymbol{x}, \quad \forall \boldsymbol{x} \in \boldsymbol{X}$
(ii) $F(G(\boldsymbol{y})) = \boldsymbol{y}, \quad \forall \boldsymbol{y} \in \boldsymbol{Y}$

We say that G is an inverse of F, and we write $G = F^{-1}$.

If there is an isomorphism $\boldsymbol{X} \to \boldsymbol{Y}$, we say that spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$ are isomorphic and we write $\boldsymbol{X} \cong \boldsymbol{Y}$.

Theorem 6.37

(i) Identity operator is an isomorphism.
(ii) The inverse of an isomorphism operator is an isomorphism too.
(iii) The composition of isomorphisms is an isomorphism too.

Proof Claims (i) and (iii) are trivial – we have already proved similar statements in group theory. So let's prove (ii).

Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be an isomorphism of spaces. Remember, vector spaces are also additive groups, thus F is an isomorphism of the respective groups X and Y. Therefore, as we know,

$$F^{-1} : Y \to X$$

is an isomorphism of groups too. Take any vector $\boldsymbol{y} \in \boldsymbol{Y}$, such that $F^{-1}(\boldsymbol{y}) = \boldsymbol{x} \in \boldsymbol{X}$ and $\alpha \in \Phi$, and consider

$$\begin{aligned} F^{-1}(\alpha\, \boldsymbol{y}) &= F^{-1}(\alpha F(\boldsymbol{x})) \\ &= F^{-1}(F(\alpha \boldsymbol{x})) \\ &= (F^{-1} \circ F)(\alpha \boldsymbol{x}) \\ &= \alpha \boldsymbol{x} \\ &= \alpha F^{-1}(\boldsymbol{y}) \end{aligned}$$

So, F^{-1} is a linear operator and an isomorphism too. ■

As you might have expected, the relation "$\cong$," that is, the relation "to be isomorphic," is a relation of equivalence on a class of all linear spaces over the same field Φ. That is

(i) $\boldsymbol{X} \cong \boldsymbol{X}$.
(ii) If $\boldsymbol{X} \cong \boldsymbol{Y}$, then $\boldsymbol{Y} \cong \boldsymbol{X}$.
(iii) If $\boldsymbol{X} \cong \boldsymbol{Y}$ and $\boldsymbol{Y} \cong \boldsymbol{Z}$, then $\boldsymbol{X} \cong \boldsymbol{Z}$.

Let's prove this claim.

(i) is trivial.
(ii) $\boldsymbol{X} \cong \boldsymbol{Y}$ means that there exists a function (an operator) $F : \boldsymbol{X} \to \boldsymbol{Y}$ such that

$$F(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) = \alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2), \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}, \ \ \forall \alpha, \beta \in \Phi$$

But since F is an isomorphism there also exists a function $F^{-1} : \boldsymbol{Y} \to \boldsymbol{X}$. Now, take two vectors $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \boldsymbol{Y}$ such that

$$\boldsymbol{y}_1 = F(\boldsymbol{x}_1) \quad \text{and} \quad \boldsymbol{y}_2 = F(\boldsymbol{x}_2)$$

Then, with $\gamma, \delta \in \Phi$, we have

$$\begin{aligned} F^{-1}(\gamma\, \boldsymbol{y}_1 + \delta \boldsymbol{y}_2) &= F^{-1}(\gamma F(\boldsymbol{x}_1) + \delta F(\boldsymbol{x}_2)) \\ &= F^{-1}(F(\gamma \boldsymbol{x}_1 + \delta \boldsymbol{x}_2)) \\ &= \gamma \boldsymbol{x}_1 + \delta \boldsymbol{x}_2 \\ &= \gamma F^{-1}(\,\boldsymbol{y}_1) + \delta F^{-1}(\boldsymbol{y}_2) \qquad (*) \end{aligned}$$

We can always rename F^{-1} by putting $F^{-1} = G$, and interpret (*) as

$$G(\gamma \boldsymbol{y}_1 + \delta \boldsymbol{y}_2) = \gamma G(\boldsymbol{y}_1) + \delta G(\boldsymbol{y}_2)$$

Hence, there is an isomorphism $G : \boldsymbol{Y} \to \boldsymbol{X}$, and reflexivity is proved. Finally, to prove (iii), let's assume that $G : \boldsymbol{X} \to \boldsymbol{Y}$ and $G : \boldsymbol{Y} \to \boldsymbol{Z}$ are isomorphisms. We already know that in that case $H = G \circ F$ is an isomorphism too. Indeed, for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$, and $\alpha, \beta \in \Phi$

$$\begin{aligned} H(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) &= (G \circ F)(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) \\ &= G[F(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2)] \\ &= G[\alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2)] \\ &= \alpha G[F(\boldsymbol{x}_1)] + \beta G[F(\boldsymbol{x}_2)] \\ &= \alpha (G \circ F)(\boldsymbol{x}_1) + \beta (G \circ F)(\boldsymbol{x}_2) \\ &= \alpha H(\boldsymbol{x}_1) + \beta H(\boldsymbol{x}_2) \end{aligned}$$

Since H^{-1} exists, and is a linear operator too, transitivity is established.

From Definition 6.20, the following theorem immediately follows:

Theorem 6.38 If $F : \boldsymbol{X} \to \boldsymbol{Y}$ is an isomorphic operator, and $\boldsymbol{K} \subseteq \boldsymbol{Y}$ is a subspace of $\boldsymbol{Y}$, then, $F^{-1}(\boldsymbol{K}) \subseteq \boldsymbol{X}$ is a subspace of $\boldsymbol{X}$.

We have previously established the fact that $\operatorname{Im} F$ and $\operatorname{Ker} F$ are subspaces of their respective spaces, that is, they are linear spaces on their own. They also have dimensions. So we have

Definition 6.21

(i) We say that

$$r = r(F) = \dim(\operatorname{Im} F)$$

is the **rank** of a linear operator F.

(ii) We say that

$$d = d(F) = \dim(\operatorname{Ker} F)$$

is the **defect** of a linear operator F.

Consequently, the following theorem is noteworthy, but first we need to recall two definitions (see Definitions 6.10 and 6.11).

Definition 6.22 If $\boldsymbol{L},\ \boldsymbol{K} < \boldsymbol{X}$ are two subspaces of $\boldsymbol{X}$, we say that

$$\sum(\boldsymbol{L},\ \boldsymbol{K}) = \{\, \boldsymbol{l} + \boldsymbol{k} \,|\boldsymbol{l} \in \boldsymbol{L},\ \ \boldsymbol{k} \in \boldsymbol{K}\,\}$$

is a **sum** of subspaces $\boldsymbol{L}$ and $\boldsymbol{K}$.

Definition 6.23 Let $\boldsymbol{X}$ be a linear space, and let $\boldsymbol{L},\ \boldsymbol{K} < \boldsymbol{X}$ be its two subspaces. We say that

$$\boldsymbol{L} \oplus \boldsymbol{K} = \boldsymbol{X}$$

is a **direct sum** of subspaces $\boldsymbol{L}$ and $\boldsymbol{K}$ iff $\boldsymbol{L} + \boldsymbol{K} < \boldsymbol{X}$ and $\boldsymbol{L} \cap \boldsymbol{K} = \{\boldsymbol{0}\}$.

Theorem 6.39 If $F : \boldsymbol{X} \to \boldsymbol{Y}$ is a linear operator, then the sum of the rank and defect of F is equal to the dimension of the space $\boldsymbol{X}$, that is,

$$r(F) + d(F) = \dim \boldsymbol{X}$$

Proof Let $\operatorname{Im} F < \boldsymbol{Y}$ be a subspace of $\boldsymbol{Y}$ such that $\dim(\operatorname{Im} F) = r$, and let $\operatorname{Ker} F < \boldsymbol{X}$ be a subspace of $\boldsymbol{X}$ such that $\dim(\operatorname{Ker} F) = d$. Furthermore, let

$$\boldsymbol{B}_{\operatorname{Im} F} = \{\, \boldsymbol{b}_1,\ \ldots\ , \boldsymbol{b}_r \,\}$$

be some basis of $\operatorname{Im} F$. We can always choose a set of vectors $\boldsymbol{k}_1,\ \ldots\ ,\ \boldsymbol{k}_r \in \boldsymbol{X}$ such that

$$F(\boldsymbol{k}_i) = \boldsymbol{b}_i, \quad \forall\ i = 1,\ \ldots\ ,\ r$$

We would like to prove that the vectors $\boldsymbol{k}_1,\ \ldots\ ,\ \boldsymbol{k}_r \in \boldsymbol{X}$ are linearly independent. Well, consider

$$\sum_{i=1}^{r} \alpha_i \boldsymbol{k}_i = \boldsymbol{0}_{\boldsymbol{X}}$$

which implies

$$\begin{aligned} F\left(\sum_{i=1}^{r} \alpha_i \boldsymbol{k}_i\right) &= F(\boldsymbol{0}_{\boldsymbol{X}}) \\ &= \sum_{i=1}^{r} \alpha_i F(\boldsymbol{k}_i) \\ &= \sum_{i=1}^{r} \alpha_i \boldsymbol{b}_i \\ &= \boldsymbol{0}_{\boldsymbol{Y}} \end{aligned}$$

But $\boldsymbol{B}_{\mathrm{Im}\,F}$ is a basis for $\mathrm{Im}\,F$, hence $\alpha_1 = \cdots = \alpha_r = 0$. Therefore, vectors, let's call them $\boldsymbol{B}_{\boldsymbol{K}} = \{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_r\}$, are linearly independent too. Now, let $\boldsymbol{K} < \boldsymbol{X}$ be a subset of $\boldsymbol{X}$ spanned by vectors from $\boldsymbol{B}_{\boldsymbol{K}}$. But $\boldsymbol{k}_1, \ldots, \boldsymbol{k}_r$ are linearly independent, and therefore they form some basis for $\boldsymbol{K}$. Clearly, then

$$\dim \boldsymbol{K} = r$$

We claim that

$$\boldsymbol{X} = \boldsymbol{K} \oplus \mathrm{Ker}\,F$$

Let's see! First, recall that the sum is direct iff

$$\dim \boldsymbol{X} = \dim \boldsymbol{K} + \dim(\mathrm{Ker}\,F)$$

which also means that

(i) $\boldsymbol{K} \cap \mathrm{Ker}\,F = \{\boldsymbol{0}\}$
(ii) $\boldsymbol{K} + \mathrm{Ker}\,F = \boldsymbol{X}$

So, to prove (i) let $\boldsymbol{k} \in \boldsymbol{K} \cap \mathrm{Ker}\,F$ be any vector, which we write

$$\boldsymbol{k} = \sum_{i=1}^{r} \alpha_i \boldsymbol{k}_i$$

for it is an element of $\boldsymbol{K}$, but also

$$F(\boldsymbol{k}) = \boldsymbol{0}$$

since $\boldsymbol{k}$ belongs to $\mathrm{Ker}\,F$ too. Therefore,

$$\begin{aligned} F(\boldsymbol{k}) &= F\left(\sum_{i=1}^{r} \alpha_i \boldsymbol{k}_i\right) \\ &= \sum_{i=1}^{r} \alpha_i F(\boldsymbol{k}_i) \\ &= \sum_{i=1}^{r} \alpha_i \boldsymbol{b}_i \\ &= \boldsymbol{0} \end{aligned}$$

Given that the set $\boldsymbol{B}_{\boldsymbol{K}}$ is linearly independent, it follows that $\alpha_1 = \cdots = \alpha_r = 0$ and therefore $\boldsymbol{k} = \boldsymbol{0}$, that is, $\boldsymbol{K} \cap \mathrm{Ker}\,F = \boldsymbol{0}$.

To prove (ii) choose any vector $\boldsymbol{x} \in \boldsymbol{X}$. Then,

$$F(\boldsymbol{x}) \in F(\boldsymbol{X}) = \operatorname{Im} F$$

And as usual, we can represent $F(\boldsymbol{x})$ as a linear combination

$$F(\boldsymbol{x}) = \sum_{i=1}^{r} \beta_i \boldsymbol{b}_i$$

Let's also pick a vector $\boldsymbol{k} \in \boldsymbol{K}$ defined as

$$\boldsymbol{k} = \sum_{i=1}^{r} \beta_i \boldsymbol{k}_i$$

So we have

$$\begin{aligned} F(\boldsymbol{k}) &= F\left(\sum_{i=1}^{r} \beta_i \boldsymbol{k}_i \right) \\ &= \sum_{i=1}^{r} \beta_i F(\boldsymbol{k}_i) \\ &= \sum_{i=1}^{r} \beta_i \boldsymbol{b}_i \\ &= F(\boldsymbol{x}) \end{aligned}$$

Finally, suppose there is a vector $\boldsymbol{b} \in \boldsymbol{X}$ defined by

$$\boldsymbol{b} = \boldsymbol{x} - \boldsymbol{k}$$

Then,

$$\begin{aligned} F(\boldsymbol{b}) &= F(\boldsymbol{x} - \boldsymbol{k}) \\ &= F(\boldsymbol{x}) - F(\boldsymbol{k}) \\ &= \boldsymbol{0} \end{aligned}$$

so $\boldsymbol{b} \in \operatorname{Ker} F$, which implies that $\boldsymbol{x} = \boldsymbol{k} + \boldsymbol{b},\ \boldsymbol{k} \in \boldsymbol{K},\ \boldsymbol{b} \in \operatorname{Ker} F$, and therefore

$$\boldsymbol{X} = \boldsymbol{K} + \operatorname{Ker} F$$ ■

Two corollaries immediately follow.

Corollary 6.1 If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are finite-dimensional vector spaces such that $\dim \boldsymbol{X} > \dim \boldsymbol{Y}$, then no linear operator from $\boldsymbol{X}$ to $\boldsymbol{Y}$ is injective.

Proof Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be finite-dimensional vector spaces such that $\dim \boldsymbol{X} > \dim \boldsymbol{Y}$, and let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator. Then,

$$\begin{aligned}\dim(\operatorname{Ker} F) = \dim \boldsymbol{X} - \dim(\operatorname{Im} F) &\geq \dim \boldsymbol{X} - \dim \boldsymbol{Y} \\ &> 0\end{aligned}$$

But if $\dim(\operatorname{Ker} F) > 0$, then $\operatorname{Ker} F$ must contain vectors other $\mathbf{0}$, and thus F is not injective. ■

Corollary 6.2 If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are finite-dimensional vector spaces such that $\dim \boldsymbol{X} < \dim \boldsymbol{Y}$, then no linear operator from $\boldsymbol{X}$ to $\boldsymbol{Y}$ is surjective.

Proof Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be finite-dimensional vector spaces such that $\dim \boldsymbol{X} < \dim \boldsymbol{Y}$ and let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator. Then,

$$\dim(\operatorname{Im} F) = \dim \boldsymbol{X} - \dim(\operatorname{Ker} F) \leq \dim \boldsymbol{X} < \dim \boldsymbol{Y}$$

Thus, $\operatorname{Im} F \neq \boldsymbol{Y}$, that is, F is not surjective. ■

Theorem 6.40 A linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ is an isomorphism iff

(i) $\operatorname{Ker} F = \{\mathbf{0}\}$ and
(ii) $\operatorname{Im} F = \boldsymbol{Y}$

Proof Let $\operatorname{Ker} F = \{\mathbf{0}\}$ and $\operatorname{Im} F = \boldsymbol{Y}$, then $\forall \boldsymbol{y} \in \boldsymbol{Y}$ there exists a unique vector $\boldsymbol{x} \in \boldsymbol{X}$ such that

$$F(\boldsymbol{x}) = \boldsymbol{y}$$

Now, consider a function $G : \boldsymbol{Y} \to \boldsymbol{X}$ defined by $G(\boldsymbol{y}) = \boldsymbol{x} \in \boldsymbol{X}$. Since $\boldsymbol{x}$ is a unique vector satisfying this equation, G is well defined. We need to prove that G is a linear operator. So, let's take two vectors $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \boldsymbol{Y}$ such that for the unique vectors $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$

$$F(\boldsymbol{x}_1) = \boldsymbol{y}_1 \quad \text{and} \quad F(\boldsymbol{x}_2) = \boldsymbol{y}_2$$

Then, we have

$$\begin{aligned}F(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) &= \alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2) \\ &= \alpha \boldsymbol{y}_1 + \beta \boldsymbol{y}_2\end{aligned}$$

On the other hand, by definition of G,

$$\begin{aligned} G(\alpha \boldsymbol{y}_1 + \beta \boldsymbol{y}_2) &= \alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2 \\ &= \alpha G(\boldsymbol{y}_1) + \beta G(\boldsymbol{y}_2) \end{aligned}$$

Hence, G is a linear operator. Is it an isomorphism? Well, observe that

$$G(F(\boldsymbol{x})) = \boldsymbol{x}, \quad \forall \boldsymbol{x} \in \boldsymbol{X}$$

and also

$$F(G(\boldsymbol{y})) = \boldsymbol{y}, \quad \forall \boldsymbol{y} \in \boldsymbol{Y}$$

So we see that F is an isomorphism and G is its inverse. But if $\operatorname{Ker} F = \{\mathbf{0}\}$ we can assume that $F(\boldsymbol{x}) = \mathbf{0}$. Then, with G being an inverse of F, we have

$$\begin{aligned} \boldsymbol{x} &= G(F(\boldsymbol{x})) \\ &= G(\mathbf{0}) \\ &= \mathbf{0} \end{aligned}$$

which completes our proof. ■

Theorem 6.41 Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator such that $\operatorname{Ker} F = \{\mathbf{0}\}$. Then, $\forall \boldsymbol{y} \in \operatorname{Im} F$ there exists a unique $\boldsymbol{x} \in \boldsymbol{X}$ such that $F(\boldsymbol{x}) = \boldsymbol{y}$.

Proof Since we have assumed that $\operatorname{Im} F \neq \varnothing$, there is at least one vector, say, $\boldsymbol{y} \in \boldsymbol{Y}$ such that $\boldsymbol{y} = F(\boldsymbol{x}_1)$ with $\boldsymbol{x}_1 \in \boldsymbol{X}$. Suppose there exists another vector $\boldsymbol{x}_2 \in \boldsymbol{X}$ such that $F(\boldsymbol{x}_2) = \boldsymbol{y}$. Then, we can write

$$\begin{aligned} F(\boldsymbol{x}_1 - \boldsymbol{x}_2) &= F(\boldsymbol{x}_1) - F(\boldsymbol{x}_2) \\ &= \boldsymbol{y} - \boldsymbol{y} \\ &= \mathbf{0} \end{aligned}$$

This implies that $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in \operatorname{Ker} F = \{\mathbf{0}\}$ and therefore $\boldsymbol{x}_1 - \boldsymbol{x}_2 = 0$. Hence, $\boldsymbol{x}_1 = \boldsymbol{x}_2$ as claimed. ■

Example 6.48 Let $F : \mathbf{R}^2 \to \mathbf{R}^2$ be a linear operator defined by

$$F(\boldsymbol{x}) = F(x, y) = (x - y, x - 2y)$$

Determine whether or not F is nonsingular.

Solution We need to find whether $\text{Ker}\, F = \{\mathbf{0}\}$ or not.
Let

$$\begin{aligned} F(\boldsymbol{x}) = F(x, y) &= \mathbf{0} \\ &= (x - y, x - 2y) = (0, 0) \end{aligned}$$

Hence,

$$\begin{aligned} x - y &= 0 \\ x - 2y &= 0 \end{aligned}$$

The only solution of these equations is $x = y = 0$. So, $\mathbf{0} \in \text{Ker}\, F$. It is easy to see that $\mathbf{0}$ is the only element of $\text{Ker}\, F$. Thus, $\text{Ker}\, F = \{\mathbf{0}\}$ and therefore F is nonsingular. ■

Example/Exercise 6.49 Let $F : \mathbf{R}^2 \to \mathbf{R}^2$ be a linear operator defined by

$$F(\boldsymbol{x}) = F(x, y) = (2x - 4y, 3x - 6y)$$

Show that F is a singular operator.

Theorem 6.42 Let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be an isomorphism of spaces, and let $\boldsymbol{A} \subseteq \boldsymbol{X}$ be some set of vectors from $\boldsymbol{X}$ such that $F(\boldsymbol{A}) = \boldsymbol{B} \subseteq \boldsymbol{Y}$. Then vectors in $\boldsymbol{A}$ are linearly independent iff vectors in $\boldsymbol{B}$ are linearly independent too. Furthermore, set $\boldsymbol{A}$ spans $\boldsymbol{X}$ iff set $\boldsymbol{B}$ spans $\boldsymbol{Y}$.

Proof Suppose $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s\}$ is a set of linearly independent vectors such that

$$F(\boldsymbol{a}_i) = \boldsymbol{b}_i \in \boldsymbol{B}.$$

Consider

$$\begin{aligned} \sum_{i=1}^{s} \beta_i \boldsymbol{b}_i &= \mathbf{0}_{\boldsymbol{Y}} \\ &= \sum_{i=1}^{s} \beta_i F(\boldsymbol{a}_i) \\ &= F\left(\sum_{i=1}^{s} \beta_i \boldsymbol{a}_i \right) \end{aligned}$$

But since F is a bijection by definition, that is, $F(\mathbf{0}_X) = \mathbf{0}_Y$, we have

$$\sum_{i=1}^{s} \beta_i \boldsymbol{a}_i = \mathbf{0}_X.$$

We have assumed that $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s\}$ is a set of linearly independent vectors, thus $\beta_i = 0$ for all $i = 1, 2, \ldots, s$ and therefore $\boldsymbol{B}$ is a set of linearly independent vectors too. Since F^{-1} is also an isomorphism, we can analogously prove the converse claim.

Finally, let's assume that $\boldsymbol{A} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s\}$ spans $\boldsymbol{X}$, then for any $\boldsymbol{b} \in \boldsymbol{Y}$

$$F^{-1}(\boldsymbol{b}) = \sum_{i=1}^{s} \alpha_i \boldsymbol{a}_i$$

Thus

$$\begin{aligned} \boldsymbol{b} &= F(F^{-1}(\boldsymbol{b})) \\ &= F\left(\sum_{i=1}^{s} \alpha_i \boldsymbol{a}_i\right) \\ &= \sum_{i=1}^{s} \alpha_i F(\boldsymbol{a}_i) \\ &= \sum_{i=1}^{s} \alpha_i \boldsymbol{b}_i \end{aligned}$$

So set $\boldsymbol{B}$ spans $\boldsymbol{Y}$. The converse can be proved similarly. ■

This theorem leads straightforwardly to the proof of the following two corollaries.

Corollary 6.3 If $F : \boldsymbol{X} \to \boldsymbol{Y}$ is an isomorphism of spaces, then every basis of $\boldsymbol{X}$ is mapped to some basis of $\boldsymbol{Y}$, or conversely, every basis of $\boldsymbol{Y}$ is an image of a basis of $\boldsymbol{X}$.

Corollary 6.4 If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are isomorphic vector spaces, then $\dim \boldsymbol{X} = \dim \boldsymbol{Y}$.

Another important theorem is

Theorem 6.43 Two linear spaces are isomorphic iff they have the same dimension.

Proof Considering the previous corollaries, it suffices to show that if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are two vector spaces such that $\dim \boldsymbol{X} = \dim \boldsymbol{Y}$, then $\boldsymbol{X}$ and $\boldsymbol{Y}$ are isomorphic.

Let $\dim \boldsymbol{X} = \dim \boldsymbol{Y} = n$ then there is a basis $\boldsymbol{B_X} = \{\boldsymbol{b}_{X1}, \ldots, \boldsymbol{b}_{Xn}\}$ for $\boldsymbol{X}$, and a basis $\boldsymbol{B_Y} = \{\boldsymbol{b}_{Y1}, \ldots, \boldsymbol{b}_{Yn}\}$ for $\boldsymbol{Y}$. Define a linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ by

$$F(\boldsymbol{b}_{Xi}) = \boldsymbol{b}_{Yi}, \quad i = 1, \ldots, n$$

Observe that F is a surjection. We need to prove that it is a bijection. Take two vectors from $\boldsymbol{X}$

$$\boldsymbol{x}_1 = \sum_{i=1}^{n} \alpha_i \boldsymbol{b}_{Xi}$$

$$\boldsymbol{x}_2 = \sum_{i=1}^{n} \beta_i \boldsymbol{b}_{Xi}$$

such that

$$\begin{aligned} F(\boldsymbol{x}_1) &= F\left(\sum_{i=1}^{n} \alpha_i \boldsymbol{b}_{Xi}\right) \\ &= F\left(\sum_{i=1}^{n} \beta_i \boldsymbol{b}_{Xi}\right) \\ &= F(\boldsymbol{x}_2) \end{aligned}$$

However, F is a linear operator so we can write

$$\sum_{i=1}^{n} \alpha_i F(\boldsymbol{b}_{Xi}) = \sum_{i=1}^{n} \beta_i F(\boldsymbol{b}_{Xi})$$

and therefore,

$$\sum_{i=1}^{n} (\alpha_i - \beta_i)\, \boldsymbol{b}_{Xi} = \boldsymbol{0}$$

Considering that the vectors from $\boldsymbol{B_X}$ are linearly independent, we have that $\alpha_i - \beta_i = 0$ for all $i = 1, \ldots, n$. Thus, $\boldsymbol{x}_1 = \boldsymbol{x}_2$, which tells us that F, besides being a surjection, is also an injection, that is, F is a bijection. We conclude that F is an isomorphism, that is, $\boldsymbol{X} \cong \boldsymbol{Y}$. ■

This is a significant result, telling us that all spaces with the same dimension have the same abstract structure. Thus, we are free to choose and do our calculations on the one which is the most suitable for the problem at hand. Consequently, the following corollary is very useful.

Corollary 6.5 Every n-dimensional linear space over a field Φ is isomorphic to the coordinate space $\mathbf{R}^n$.

6.8 LINEAR TRANSFORMATIONS AND MATRICES

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be vector spaces over a field Φ, and let $F : \boldsymbol{X} \to \boldsymbol{Y}$ be an operator from $\boldsymbol{X}$ into $\boldsymbol{Y}$. Furthermore, let $\boldsymbol{B_X} = \{\boldsymbol{b}_{X1}, \ldots, \boldsymbol{b}_{Xn}\}$ and $\boldsymbol{B_Y} = \{\boldsymbol{b}_{Y1}, \ldots, \boldsymbol{b}_{Ym}\}$ be the two respective bases. Recall that we know an operator if we know how it acts on the basis vectors $\boldsymbol{B_X}$, so we can write

$$F(\boldsymbol{b}_{Xk}) = \sum_{i=1}^{m} \alpha_{ik} \boldsymbol{b}_{Yi}, \quad k = 1, \ldots, n, \quad \alpha_{ik} \in \Phi$$

Observe that the operator F is uniquely determined by the set of scalars $\alpha_{ik} \in \Phi$. Let's write them in the rectangular scheme

$$F = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

called the **matrix** of the operator F with respect to two bases $\boldsymbol{B_X}$ and $\boldsymbol{B_Y}$. By close inspection, you note that the kth column contains coefficients of the vector $F(\boldsymbol{b}_{Xk})$ in the $\boldsymbol{B_Y}$ basis. More formally we say:

Definition 6.24 A rectangular array of numbers (elements of a given field Φ) composed of m rows and n columns

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

is called an $(m \times n)$ or (m, n) **matrix**, where elements $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in}$ form the ith row, and elements

$$\begin{matrix} \alpha_{1k} \\ \alpha_{2k} \\ \vdots \\ \alpha_{mk} \end{matrix}$$

form the kth column of A.

Formally, and more precisely:

Definition 6.24′ Let Φ be any field whose elements we call scalars, and let $m, n \in \mathbf{N}$ be natural numbers. If by

$$D_{mn} = \{1,\ 2,\ \dots\ ,\ m\} \times \{1,\ 2,\ \dots\ ,\ n\}$$

we indicate a Cartesian product, then we call the map

$$A : D_{mn} \to \Phi$$

a **matrix of (m, n) type**, where the functional value

$$A(i, k) = \alpha_{ik} \in \Phi.$$

Example 6.50 Let $F : \mathbf{R}^3 \to \mathbf{R}^3$ be a linear operator, and let

$$\boldsymbol{B} = \{\boldsymbol{b}_1 = (1,\ 0,\ 0),\ \ \boldsymbol{b}_2 = (0,\ 1,\ 0),\ \ \boldsymbol{b}_3 = (0,\ 0,\ 1)\}$$

be, as usual, our standard basis for $\mathbf{R}^3$. We define F as follows:

$$\begin{aligned} F(\boldsymbol{b}_1) &= (\alpha_{11},\ \alpha_{21},\ \alpha_{31}) \\ F(\boldsymbol{b}_2) &= (\alpha_{12},\ \alpha_{22},\ \alpha_{32}) \\ F(\boldsymbol{b}_3) &= (\alpha_{13},\ \alpha_{23},\ \alpha_{33}) \end{aligned}$$

Now we rewrite this by expressing the right-hand side through the same basis vectors

$$\begin{aligned} F(\boldsymbol{b}_1) &= \alpha_{11}\boldsymbol{b}_1 + \alpha_{21}\boldsymbol{b}_2 + \alpha_{31}\boldsymbol{b}_3 \\ F(\boldsymbol{b}_2) &= \ \alpha_{12}\boldsymbol{b}_1 + \alpha_{22}\boldsymbol{b}_2 + \alpha_{32}\boldsymbol{b}_3 \\ F(\boldsymbol{b}_3) &= \alpha_{13}\boldsymbol{b}_1 + \alpha_{23}\boldsymbol{b}_2 + \alpha_{33}\boldsymbol{b}_3 \end{aligned}$$

Observe that this is just an explicit expression of

$$F(\boldsymbol{b}_{Xk}) = \sum_{i=1}^{m} \alpha_{ik}\boldsymbol{b}_{Xi}, \quad k = 1,\ \dots\ , n, \quad \alpha_{ik} \in \Phi$$

from the previous page. Certainly, we could have chosen any vector $\boldsymbol{x} = (x_1, x_2, x_3) \in \mathbf{R}^3$ and expressed it as a linear combination of basis vectors

$$\boldsymbol{x} = x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + x_3\boldsymbol{b}_3$$

Thus,

$$\begin{aligned} F(\boldsymbol{x}) &= F(x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + x_3\boldsymbol{b}_3) \\ &= x_1F(\boldsymbol{b}_1) + x_2F(\boldsymbol{b}_2) + x_3F(\boldsymbol{b}_3) \\ &= x_1(\alpha_{11}\boldsymbol{b}_1 + \alpha_{21}\boldsymbol{b}_2 + \alpha_{31}\boldsymbol{b}_3) \\ &\quad + x_2(\alpha_{12}\boldsymbol{b}_1 + \alpha_{22}\boldsymbol{b}_2 + \alpha_{32}\boldsymbol{b}_3) \\ &\quad + x_3(\alpha_{13}\boldsymbol{b}_1 + \alpha_{23}\boldsymbol{b}_2 + \alpha_{33}\boldsymbol{b}_3) \end{aligned}$$

and, finally, collecting like terms, we have

$$\begin{aligned} F(\boldsymbol{x}) &= (\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3)\boldsymbol{b}_1 + (\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3)\boldsymbol{b}_2 \\ &\quad + (\alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3)\boldsymbol{b}_3 \\ &= \sum_{i=1}^{3}\sum_{k=1}^{3}\alpha_{ik}x_k\boldsymbol{b}_i \end{aligned}$$

As you can see, we can calculate the "value" of a linear operator F acting on any vector $\boldsymbol{x}$ by considering an array of numbers α_{ik} corresponding to a given basis. This array of numbers is exactly the matrix $A = (\alpha_{ik})$ described again in the following definition. ■

Definition 6.25 We say that matrix A is of **type** $(\boldsymbol{m}, \boldsymbol{n})$ if it has m rows and n columns. Very often we simply write $A = (\alpha_{ik})$ where we call α_{ik} a **representative** of the elements of A.

Definition 6.26 If a matrix A is of (n, n) type, that is, if it has the same number of rows as columns, we say that A is a **square** matrix.

Example 6.51 The matrix $A = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 5 & 1 \\ 1 & 1 & 7 \end{bmatrix}$ is a $(3, 3)$-square matrix. ■

Definition 6.27 If a matrix A is of type $(1,\ n)$, that is, if it has just one row and n columns, we say that A is a **row matrix**. Similarly, if a matrix A is of type $(m, 1)$, that is, if it has m rows and just one column, we say that A is a **column matrix**.

Definition 6.28 We say that the (m, n) matrix

$$O = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

is a **zero matrix**.

Example 6.52 The matrix $O = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is a $(2, 4)$-zero matrix. ■

Definition 6.29 Given a square matrix A, we call the n-tuple $(\alpha_{11},\ \alpha_{22},\ \ldots\ ,\alpha_{nn})$ the **main diagonal** of A.

Definition 6.30 Given a square matrix A we call the sum of all main diagonal elements the **trace of A**, that is,

$$\operatorname{tr} A = \sum_{i=1}^{n} \alpha_{ii}$$

Example 6.53

$$\operatorname{tr} A = \operatorname{tr} \begin{bmatrix} 2 & 3 & 0 \\ 1 & 5 & 1 \\ 1 & 1 & 7 \end{bmatrix} = 2 + 5 + 7 = 14$$

■

Definition 6.31 We say that a square matrix with elements $\alpha_{ik} = \delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$

$$I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

is the **identity matrix** if it has 1s on the main diagonal and 0s elsewhere.

At this point, it may be a good idea to formally introduce Kronecker's[5] delta (or the Kronecker delta) function.

Definition 6.32 A function δ_{ik} defined by

$$\delta : \mathbf{N} \times \mathbf{N} \to \{0,\ 1\}$$

such that

$$\delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases}$$

is called the **Kronecker delta function**.

[5]Leopold Kronecker (1823–1891), German mathematician.

Example 6.54 The matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = (\delta_{ik})$$

is a $(3, 3)$ identity matrix. ■

Definition 6.33 A square matrix $S = (\alpha_{ij})$ is called a **scalar matrix** if $A = \lambda I$, $\lambda \in \Phi$.

Example 6.55 The matrix

$$S = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = 3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 3(\delta_{ik})$$

is a scalar matrix. ■

Definition 6.34 A square matrix is called a **diagonal matrix** if its only nonzero elements are on the main diagonal.

Example 6.56 The matrices

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{and} \quad E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are both diagonal. ■

Definition 6.35 We say that two matrices $A = (\alpha_{ij})$ and $B = (\beta_{ij})$, over the same field Φ, are equal iff they are of the same type and if they have the same elements at the same "locations," that is,

$$A = B \text{ iff } \text{type } A = \text{type } B, \quad \text{and} \quad \alpha_{ij} = \beta_{ij}$$

Example 6.57 Matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ are not equal even though they are of the same type, since they do not have the same elements at the same locations, that is, $\alpha_{ij} \neq \beta_{ij}$.

Similarly, the matrix

$$C = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \neq D = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \text{ since type } C \neq \text{type } D$$

■

Definition 6.36 A matrix $A = (\alpha_{ij})$ of type (m, n) uniquely determines a matrix B of type (n, m) defined by

$$\alpha_{ij} = \beta_{ji}, \quad \forall\, i = 1, \ldots, m;\ j = 1, \ldots, n$$

In other words, we obtain matrix B by interchanging the rows and columns of matrix A.

Matrix B is called the **transpose matrix of matrix** A and is denoted by A^τ.

Example 6.58 $A = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 0 & 5 \end{bmatrix}$ and $A^\tau = \begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 5 \end{bmatrix}$.

Obviously, $(A^\tau)^\tau = A$. ■

Definition 6.37 A square matrix A is said to be **symmetric** if $A^\tau = A$, that is, $(\alpha_{ki}) = (\alpha_{ik})$. A matrix is an **antisymmetric** or **skew-symmetric** if $A^\tau = -A$, that is, $(\alpha_{ki}) = (-\alpha_{ik})$.

Example 6.59 $A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{bmatrix}$ and $A^\tau = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{bmatrix}$.

Thus, $A = A^\tau$, that is, matrix A is symmetric. ■

Example 6.60 Matrix A

$$A = \begin{bmatrix} 3 & 2 & 7 \\ 2 & 1 & 0 \\ 7 & 0 & 4 \end{bmatrix}$$

is symmetric, and matrix B

$$B = \begin{bmatrix} 0 & 1 & -3 \\ -1 & 0 & 5 \\ 3 & -5 & 0 \end{bmatrix}$$

is antisymmetric. ■

Obviously, if $A^\tau = A$ and $A^\tau = -A$, then $A = O$.

6.9 LINEAR SPACE $\boldsymbol{M_{mn}}$

The discussion in the previous section inspires the next natural idea: Why not endow a set of all (m, n)-matrices with the necessary structure and create a vector space $\boldsymbol{M_{mn}}$ of all (m, n)-type matrices.

Definition 6.38 Let $\boldsymbol{M_{mn}}$ be the set of all (m, n) matrices over the field Φ, and let $A = (\alpha_{ik})$, $B = (\beta_{ik}) \in \boldsymbol{M_{mn}}$ be any two matrices. We define the **sum** $C = (\gamma_{ik})$ of matrices A and B, $C = A + B$ by

$$\gamma_{ik} = \alpha_{ik} + \beta_{ik}, \qquad \forall i = 1, \ldots, m; \;\; k = 1, \ldots, n$$

We see that addition of matrices is well defined iff they are of the same type.

Example 6.61 Let $A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -2 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & 0 \end{bmatrix}$ then,

$$A + B = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -2 & 2 \end{bmatrix} + \begin{bmatrix} 4 & 1 & -1 \\ 2 & 5 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 \\ 2 & 3 & 2 \end{bmatrix}$$ ■

Definition 6.39 Let $A = (\alpha_{ik}) \in \boldsymbol{M_{mn}}$ be any (m, n)-matrix, and $\lambda \in \Phi$. We define the product of a scalar λ and a matrix A by multiplying each entry in the matrix A by the scalar:

$$\lambda \cdot A = \lambda(\alpha_{ik}) = (\lambda \alpha_{ik})$$

Example 6.62

$$3 \cdot A = 3 \cdot \begin{bmatrix} 2 & 1 & 3 \\ 0 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 9 \\ 0 & -6 & 6 \end{bmatrix}$$ ■

Theorem 6.44 The set $\boldsymbol{M_{mn}}$ of all (m, n)-matrices over a field Φ, with addition of matrices as the binary operation, is an abelian group.

Proof Closure and associativity are obviously satisfied. The neutral element is a nil-matrix O, and the opposite matrix of the matrix $A = (\alpha_{ik})$ is the matrix $-A = (-\alpha_{ik})$, thus we have a group. ■

Consequently, we expect the following theorem to hold too.

Theorem 6.45 The set $\boldsymbol{M_{mn}}$ of all (m, n)-matrices over a field Φ, with addition of matrices and multiplication of matrices by a scalar, is a vector space. The dimension of this space is

$$\dim \boldsymbol{M_{mn}} = mn.$$

Proof We have already established the fact that for any two (m, n)-matrices $A,\ B\ \in \boldsymbol{M_{mn}}, A + B \in \boldsymbol{M_{mn}}$. From Definitions 6.38 and 6.39, it immediately follows that $\forall\ \lambda, \mu \in \Phi$, and $A,\ B \in \boldsymbol{M_{mn}}$:

(i) $\lambda(\mu A) = (\lambda\mu)A$
(ii) $1 \cdot A = A$
(iii) $(\lambda + \mu)A = \lambda A + \mu A$
(iv) $\lambda(A + B) = \lambda A + \lambda B$

Therefore, $\boldsymbol{M_{mn}}$ is a vector space. What about the dimension of this space? Well, consider a set of matrices

$$\mathbf{B} = \{ B_{ik} \mid i = 1, \ldots, m;\ \ k = 1, \ldots, n \} \subseteq \boldsymbol{M_{mn}},\ B_{ik} = (\beta_{rs})$$

such that $\beta_{rs} = \begin{cases} 1, & \text{if } r = i,\ s = k \\ 0, & \text{otherwise} \end{cases}$.

It is easy to see that matrices B_{ik}, constructed this way, represent the basis for $\boldsymbol{M_{mn}}$. Now, consider the space Φ^{mn} of mn-tuples (think of $\mathbf{R}^{mn}$, for instance) and remember that an (m, n) matrix is a particular recording of ordered mn-tuples from the field Φ. Therefore, there is always a map

$$\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix} \mapsto (\alpha_{11}, \ldots, \alpha_{1n}, \ldots, \alpha_{m1}, \ldots, \alpha_{mn}) \in \Phi^{mn}$$

that is, an isomorphism of spaces $\boldsymbol{M_{mn}}$ and Φ^{mn}. The space Φ^{mn} is clearly mn-dimensional, hence, by Theorem 6.43, the space $\boldsymbol{M_{mn}}$ is mn-dimensional too. ■

Example 6.63 Show that the standard (canonical) basis for $\boldsymbol{M}_{22}$ is

$$\mathbf{B} = \left\{ \boldsymbol{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \boldsymbol{B}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \boldsymbol{B}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \boldsymbol{B}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Solution As usual, we have to prove that $\boldsymbol{B}$ is a linearly independent set and that matrices $\boldsymbol{B}_i$ span $\boldsymbol{M}_{22}$.

Consider first

$$\alpha_1\boldsymbol{B}_1 + \alpha_2\boldsymbol{B}_2 + \alpha_3\boldsymbol{B}_3 + \alpha_4\boldsymbol{B}_4 = 0, \quad \alpha_i \in \Phi,$$

that is,

$$\alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Thus, our equation has only a trivial solution $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ and therefore vectors $\boldsymbol{B}_i$ are linearly independent. Next, observe that any $M \in \boldsymbol{M}_{22}$, say,

$$M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

can be expressed as a linear combination of basis vectors $\boldsymbol{B}_i$, that is,

$$M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \beta \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \delta \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

We see that $\boldsymbol{B}_i$ are linearly independent and they span the space $\boldsymbol{M}_{22}$, hence, they form a basis for $\boldsymbol{M}_{22}$. ■

Example/Exercise 6.64 Show that vectors

$$\left\{ \boldsymbol{A}_1 = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}, \ \boldsymbol{A}_2 = \begin{bmatrix} 3 & 3 \\ 2 & 0 \end{bmatrix}, \ \boldsymbol{A}_3 = \begin{bmatrix} 0 & 3 \\ 2 & 2 \end{bmatrix} \right\} \in \boldsymbol{M}_{22}$$

are linearly independent.

6.10 MATRIX MULTIPLICATION

Unlike addition and multiplication by a scalar, matrix multiplication is somewhat complicated. Although the procedure of this operation may look rather strange at first, it does have a very logical rationale, as you will see shortly.

Before introducing an example and a general rule of matrix multiplication, let's state a fundamental principle:

Two matrices A and B can be multiplied, that is, the product $A \cdot B$ is defined, iff type $A = (m, n)$ and type $B = (n, p)$.

Let's start with a simple example.

Example 6.65 Let $A = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix}$ be a $(1, n)$ matrix and $B = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$ a $(n, 1)$ matrix, with $\alpha_i, \ \beta_i \in \Phi$. We define the product $A \cdot B$ as follows:

$$A \cdot B = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = \sum_{i=1}^{n} \alpha_i \beta_i \in \Phi$$

Observe that the result is a $(1, 1)$ matrix, that is, a scalar from Φ. ■

Example 6.66 Find $A \cdot B$ if $A = \begin{bmatrix} 2 & 1 & 3 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 3 \\ 3 \\ 0 \\ 1 \end{bmatrix}$.

Solution $A \cdot B = \begin{bmatrix} 2 & 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \\ 0 \\ 1 \end{bmatrix} = 6 + 3 + 0 + 0 = 9$ ■

Observe that the products are well defined in both examples, that is, the number of columns in A is equal to the number of rows in B. Also note that with multiplication defined as above, the product $B \cdot A$ is not obtainable. Let's generalize this.

Definition 6.40 Let $A = (\alpha_{ik})$ be an (m, n) matrix and $B = (\beta_{ik})$ an (n, p)-matrix. We say that a matrix $C = (\gamma_{ik})$ is a product of A and B, that is, $C = A \cdot B$ if

$$\gamma_{ik} = \sum_{j=1}^{n} \alpha_{ij} \beta_{jk}, \quad \forall i = 1, \ldots, m; \;\; k = 1, \ldots, p$$

Explicitly,

$$A \cdot B = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots \\ \alpha_{i1} & \cdots & \alpha_{in} \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix} \cdot \begin{bmatrix} \beta_{11} & \cdots & \beta_{1k} & \cdots & \beta_{1p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{n1} & \cdots & \beta_{nk} & \cdots & \beta_{np} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_j \alpha_{1j} \beta_{j1} & \cdots & \sum_j \alpha_{1j} \beta_{jp} \\ \vdots & \sum_j \alpha_{ij} \beta_{jk} & \vdots \\ \sum_j \alpha_{mj} \beta_{j1} & \cdots & \sum_j \alpha_{mj} \beta_{jp} \end{bmatrix}$$

Example 6.67 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 1 & 0 \end{bmatrix}$. Calculate $A \cdot B$.

Solution

$$A \cdot B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 14 & 11 & 3 \\ 17 & 32 & 29 & 12 \end{bmatrix}$$

■

Example 6.68 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. Calculate $A \cdot B$ and $B \cdot A$.

Solution

$$A \cdot B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

$$B \cdot A = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$$

Note that although type A = type B, $A \cdot B \neq B \cdot A$. ■

We conclude: **Matrix multiplication (in general) is noncommutative**. Occasionally, it might happen that $A \cdot B = B \cdot A$. In that case, we say that matrices A and B **commute**.

Example 6.69

$$A \cdot B = \begin{bmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{bmatrix} \begin{bmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 5 & 7 & -15 \\ -12 & 0 & 20 \\ 17 & 7 & -35 \end{bmatrix}$$

$$B \cdot A = \begin{bmatrix} 2 & -5 & 1 \\ 0 & 3 & -2 \\ 1 & 2 & -4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ -4 & -4 & -4 \\ 5 & 6 & 7 \end{bmatrix} = \begin{bmatrix} 27 & 30 & 33 \\ -22 & -24 & -26 \\ -27 & -30 & -33 \end{bmatrix}$$

■

Example 6.70 Here are two matrices that commute:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

Indeed,

$$A \cdot B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = A \cdot B = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$$

■

However, for every (n, n) matrix A and (n, n)-identity matrix I

$$I \cdot A = A \cdot I = A$$

Example 6.71

$$I \cdot A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 2 & 4 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= A \cdot I = A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 3 & 2 & 4 \end{bmatrix}$$ ■

Theorem 6.46 Let A, B, and C be matrices. Then, whenever respective multiplication is defined, matrix multiplication is associative, that is, $A(BC) = (AB)C$.

Proof Let $A = (\alpha_{ij})$ be a (m, n) matrix, $B = (\beta_{jk})$ a (n, p) matrix, and $C = (\gamma_{kl})$ a (p, q) matrix. First note that both products $A(BC)$ and $(AB)C$ are well defined, and the result in both cases is a (m, q) matrix.

Let the product $BC = (\sigma_{jl})$, then

$$\sigma_{jl} = \sum_{k=1}^{p} \beta_{jk}\gamma_{kl}$$

Also, let

$$A(BC) = (\xi_{il})$$

So, we have

$$\xi_{il} = \sum_{j=1}^{n} \alpha_{ij}\sigma_{jl}$$
$$= \sum_{j=1}^{n} \alpha_{ij} \sum_{k=1}^{p} \beta_{jk}\gamma_{kl}$$
$$= \sum_{j=1}^{n} \sum_{k=1}^{p} \alpha_{ij}\beta_{jk}\gamma_{kl} \qquad (*)$$

Now, if we denote $AB = (\tau_{ik})$ then,

$$\tau_{ik} = \sum_{j=1}^{n} \alpha_{ij}\beta_{jk}$$

and, with $(AB)C = (\eta_{il})$, we have

$$\begin{aligned}\eta_{il} &= \sum_{k=1}^{p} \tau_{ik}\gamma_{kl} \\ &= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} \alpha_{ij}\beta_{jk} \right) \gamma_{kl} \\ &= \sum_{j=1}^{n} \sum_{k=1}^{p} \alpha_{ij}\beta_{jk}\gamma_{kl} \qquad (**)\end{aligned}$$

From (*) and (**), we see that $\xi_{il} = \eta_{il}$, $\forall\, i = 1, \ldots, m$; $l = 1, \ldots, q$. Hence, we have proved that $A(BC) = (AB)C$. ■

Theorem 6.47 Let A, B, and C be matrices, and let $\lambda \in \Phi$. Then, whenever products and sums are defined,

(i) $(\lambda A)B = \lambda(AB) = A(\lambda B)$
(ii) $(A + B)C = AC + BC$
(iii) $A(B + C) = AB + AC$

Proof The proof immediately follows from the definitions of matrix addition and multiplication. ■

Once matrix multiplication is defined, one immediately accepts powers of matrices. We have

Definition 6.41 For any square matrix A over a field Φ,

$$A^2 = A \cdot A, \quad A^3 = A^2 \cdot A, \ \ldots, \ A^n = A^{n-1} \cdot A, \text{ and } A^0 = I$$

Thus, polynomials in matrix A are also defined.

Definition 6.42 Let

$$p_n(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$$

be any polynomial where a_i are scalars in Φ, then $p_n(A)$ is the following matrix:

$$p_n(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I$$

Example 6.72 Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, and let $p_2(t) = 2t^2 - 3t + 5$, then

$$\begin{aligned} p_2(A) &= 2A^2 - 3A + 5I \\ &= 2\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 3\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + 5\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= 2\begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 16 & 14 \\ 21 & 37 \end{bmatrix} \end{aligned}$$

■

Definition 6.43 A matrix A is said to be a **root** or a **zero** of a polynomial $p_n(t)$, if $p_n(A)$ is a zero matrix, that is, $p_n(A) = O$.

Example 6.73 Let $p_n(t) = t^2 + 3t - 10$, and let $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$, then

$$\begin{aligned} p_n(A) &= A^2 + 3A - 10 \\ &= \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} + 3\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} - 10\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 7 & -6 \\ -9 & 22 \end{bmatrix} + \begin{bmatrix} 3 & 6 \\ 9 & -12 \end{bmatrix} - \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Thus, A is a zero of the polynomial $p_n(t) = t^2 + 3t - 10$. ■

6.11 SOME MORE SPECIAL MATRICES. GENERAL LINEAR GROUP

Definition 6.44 A square matrix A is said to be **idempotent** if $A^2 = A$.

Example 6.74 Matrices $\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and, of course, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ are three among many idempotent matrices. ■

Example/Exercise 6.75 Show that if a matrix A is idempotent, then A^τ is also idempotent.

Definition 6.45 A square matrix A is said to be **nilpotent** if there is a natural number $n \in \mathbf{N}$, such that $A^n = O$. The smallest such natural number n is called the **index of nilpotence** of A.

Example 6.76 If $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, then

$$A^2 = AA = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ and}$$

$$A^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, A is a nilpotent matrix with index of nilpotence 3. ■

Example/Exercise 6.77 Show that, if matrix A is nilpotent, then A^τ is nilpotent too.

Definition 6.46 A square matrix A is said to be **invertible** or **nonsingular**, or **regular**, if there exists a matrix X such that

$$AX = XA = I$$

We call matrix X the inverse of A and denote it by A^{-1}.

A matrix that is not invertible is called **singular**.

Example 6.78 Suppose we are given a matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. We would like to show that it is invertible, that is, we would like to find A^{-1}.

According to the definition above if $A^{-1} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ exists, then

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Multiplying the matrices on the left-hand side we get

$$\begin{bmatrix} \alpha + 2\gamma & \beta + 2\delta \\ 3\alpha + 4\gamma & 3\beta + 4\delta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus,

$$\begin{aligned} \alpha + 2\gamma &= 1 \\ \beta + 2\delta &= 0 \\ 3\alpha + 4\gamma &= 0 \\ 3\beta + 4\delta &= 1 \end{aligned}$$

Solving this system, we get

$$\alpha = -2, \quad \beta = 1, \quad \gamma = 3/2, \quad \delta = -1/2$$

So, matrix $A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}$. Indeed,

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

You can easily convince yourself that

$$A^{-1}A = I$$ ■

You are absolutely right in thinking that the above example of finding the inverse matrix is rather clumsy, and definitely not elegant. We will discuss a much better way shortly. For now, let's continue exploring some additional features of matrices.

Theorem 6.48 The inverse of a matrix is unique.

Proof Let A be a matrix whose inverse is X. Suppose there is another inverse Y, that is, suppose that

$$AX = XA = I \quad \text{and} \quad AY = YA = I$$

Then,

$$X = XI = X(AY) = (XA)Y = IY = Y$$ ■

Example/Exercise 6.79 Show that if A is a symmetric and invertible matrix, then A^{-1} is also symmetric.

Definition 6.47 A square matrix A is said to be **involutory** if $A^2 = I$ or equivalently, if $A^{-1} = A$.

Example 6.80 Consider a matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and calculate A^2

$$A^2 = AA = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

■

Example/Exercise 6.81 Convince yourself that the following matrices are involutory:

(i) $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$

(ii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

Example 6.82 If a matrix A is of the form $\begin{bmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{bmatrix}$, then it is involutory. Indeed,

$$\begin{bmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{bmatrix} \begin{bmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{bmatrix} = \begin{bmatrix} a^2 + b\frac{1-a^2}{b} & ab - ab \\ a\frac{1-a^2}{b} - a\frac{1-a^2}{b} & b\frac{1-a^2}{b} + a^2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

■

Theorem 6.49 If A is a regular matrix, then $(A^{-1})^{-1} = A$.

Proof By definition of inverse, for any regular matrix A

$$A^{-1}A = I$$

On the other hand, if X is an inverse of $(A^{-1})^{-1}$, then

$$X\,(A^{-1})^{-1} = I$$

Thus,

$$X = (A^{-1})^{-1} = A.$$ ■

Theorem 6.50 A product AB of regular matrices A and B is a regular matrix, that is, there exists $(AB)^{-1}$ and it is equal to $B^{-1}A^{-1}$.

Proof Let A and Bbe two regular matrices, that is, A^{-1} and B^{-1} exist. If AB is a regular matrix, then

$$(AB)(AB)^{-1} = (AB)^{-1}(AB) = I$$

Consider now

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1}\\ &= AIA^{-1}\\ &= I\end{aligned}$$

Similarly,

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B\\ &= B^{-1}IB\\ &= I\end{aligned}$$

Thus, AB is regular and $(AB)^{-1} = B^{-1}A^{-1}$. ■

Theorem 6.51 A nilpotent matrix is not invertible.

Proof Suppose there exists a matrix A that is nilpotent and also invertible. Let B be an inverse of A. Since A is nilpotent there exists $n \in \mathbf{N}$, such that $A^n = O$. Then,

$$\begin{aligned}O = A^nB &= A^{n-1}AB\\ &= A^{n-1}I\\ &= A^{n-1}\end{aligned}$$

$$\begin{aligned} &= A^{n-2}AB = A^{n-2}I \\ &= A^{n-2} \\ &\vdots \\ O = AB &= I \end{aligned}$$

And this, of course, is impossible. We conclude that the nilpotent matrix is not invertible. ■

Theorem 6.52 If a square matrix A is idempotent and invertible, then $A = I$.

Proof Let A be an idempotent matrix, that is, $A^2 = A$, and let B be its inverse. Then,

$$\begin{aligned} A &= IA = BAA \\ &= BA^2 = BA \\ &= I \end{aligned}$$

So, $A = I$ as claimed. ■

Definition 6.48 We say that a matrix A is **orthogonal**, if

$$AA^\tau = A^\tau A = I$$

It is clear from this definition that every orthogonal matrix is regular and that

$$A^\tau = A^{-1}$$

Example 6.83 Consider a matrix

$$A = \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix}$$

and let's calculate

$$\begin{aligned} AA^\tau &= \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix} \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

Similarly,

$$AA^{\tau} = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix} \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Hence, matrix A is orthogonal. ■

Example 6.84 Matrix A

$$A = \begin{bmatrix} 1/9 & 8/9 & -4/9 \\ 4/9 & -4/9 & -7/9 \\ 8/9 & 1/9 & 4/9 \end{bmatrix}$$

is another example of an orthogonal matrix. Indeed,

$$AA^{\tau} = \begin{bmatrix} 1/9 & 8/9 & -4/9 \\ 4/9 & -4/9 & -7/9 \\ 8/9 & 1/9 & 4/9 \end{bmatrix} \begin{bmatrix} 1/9 & 4/9 & 8/9 \\ 8/9 & -4/9 & 1/9 \\ -4/9 & -7/9 & 4/9 \end{bmatrix}$$

$$= A^{\tau}A = \begin{bmatrix} 1/9 & 4/9 & 8/9 \\ 8/9 & -4/9 & 1/9 \\ -4/9 & -7/9 & 4/9 \end{bmatrix} \begin{bmatrix} 1/9 & 8/9 & -4/9 \\ 4/9 & -4/9 & -7/9 \\ 8/9 & 1/9 & 4/9 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

■

Example 6.85 Matrix A $A = \begin{bmatrix} \sin\alpha & \cos\alpha \\ \cos\alpha & -\sin\alpha \end{bmatrix}$ is an orthogonal matrix for any $\alpha \in \mathbf{R}$. Indeed,

$$AA^{\tau} = A^{\tau}A$$

$$= \begin{bmatrix} \sin\alpha & \cos\alpha \\ \cos\alpha & -\sin\alpha \end{bmatrix} \begin{bmatrix} \sin\alpha & \cos\alpha \\ \cos\alpha & -\sin\alpha \end{bmatrix}$$

$$= \begin{bmatrix} \sin^2\alpha + \cos^2\alpha & 0 \\ 0 & \cos^2\alpha + \sin^2\alpha \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

■

Theorem 6.53 Let A and B be two matrices over a field Φ, such that the additions and multiplications that follow are well defined, then

(i) $(A+B)^\tau = A^\tau + B^\tau$
(ii) $(A^\tau)^\tau = A$
(iii) $(\lambda A)^\tau = \lambda A^\tau, \ \forall\ \lambda \in \Phi$
(iv) $(AB)^\tau = B^\tau A^\tau$

Proof We will prove (iv). The rest is left as an easy exercise.

Let $A = (\alpha_{ik})$ and $B = (\beta_{kj})$ be two (n, n) matrices, and let $AB = (\gamma_{ij})$. Then, the ij-entry of AB is

$$\gamma_{ij} = \alpha_{i1}\beta_{1j} + \alpha_{i2}\beta_{2j} + \ \cdots\ + \alpha_{in}\beta_{nj} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}.$$

Observe that this is exactly the ji-entry of $(AB)^\tau$.

Now, transposing A and B, we get $A^\tau = (\alpha_{ki})$ and $B^\tau = (\beta_{jk})$. Thus, the γ_{ij}-entry of $B^\tau\, A^\tau$ is

$$\begin{aligned}(\beta_{1j},\ \beta_{2j},\ \ldots\ ,\ \beta_{nj})(\alpha_{i1}, \alpha_{i2},\ \ldots\ , \alpha_{in})^\tau &= \beta_{1j}\alpha_{i1} + \beta_{2j}\alpha_{i2} + \ \cdots + \beta_{nj}\alpha_{in} \\ &= \sum_{k=1}^{n} \beta_{kj}\alpha_{ik} = \gamma_{ij}\end{aligned}$$

So, $(AB)^\tau = B^\tau\, A^\tau$ as claimed. ■

Example/Exercise 6.86 Let A and B be two orthogonal matrices. Show that $(A + B)$ is not an orthogonal matrix.

Definition 6.49 A square matrix A over a real field is said to be **normal**[6] if it commutes with its transpose A^τ, that is, if $AA^\tau = A^\tau A$.

Example 6.87 Matrix $A = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix}$ is normal since

$$AA^\tau = \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 6 & 3 \\ -3 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix} = A^\tau A = \begin{bmatrix} 45 & 0 \\ 0 & 45 \end{bmatrix}$$ ■

Example/Exercise 6.88 Show that all symmetric matrices are normal.

The following theorem comes as no surprise considering everything that has been said so far.

[6] In the case of complex matrices, i.e. if $\Phi = C$, by "normal" we mean $A^*A = AA^*$, where A^* is the complex conjugate transpose of A.

Theorem 6.54 A set of all regular (n, n) matrices over a field Φ forms a group with respect to matrix multiplication. This group is clearly non-abelian.

Proof The proof follows from Theorems 6.49 and 6.50. ■

Hence,

Definition 6.50 A group of all regular (n, n) matrices over a field Φ, with matrix multiplication as a binary operation, is called a **general linear group** and is denoted $\mathrm{GL}(n, \Phi)$.

Theorem 6.55 The set of all orthogonal (n, n) matrices over the field Φ is a subgroup of $\mathrm{GL}(n, \Phi)$. We denote this group by $O(n, \Phi)$.

Proof Suppose A and B are orthogonal matrices, that is, $A^{-1} = A^{\tau}$, and $B^{-1} = B^{\tau}$. We need to show that AB^{-1} is also an orthogonal matrix. Well,

$$\begin{aligned}(AB^{-1})^{-1} &= (B^{-1})^{-1}A^{-1}\\ &= BA^{\tau}\\ &= (B^{\tau})^{\tau}A^{\tau}\\ &= (AB^{\tau})^{\tau}\\ &= (AB^{-1})^{\tau}\end{aligned}$$

We see that $(AB^{-1})^{-1} = (AB^{-1})^{\tau}$, thus, AB^{-1} is indeed an orthogonal matrix. ■

Hence,

Definition 6.51 The set of all orthogonal (n, n)-matrices over a field Φ, with matrix multiplication as a binary operation, is said to be an **orthogonal group**, denoted $O(n, \Phi)$.

Theorem 6.56 If $A = (\alpha_{ik})$ is an orthogonal matrix, then

(i) $\displaystyle\sum_{j=1}^{n} \alpha_{ij}\alpha_{kj} = \delta_{ik}, \quad \forall\, i, k = 1, \ldots, n$

and

(ii) $\displaystyle\sum_{j=1}^{n} \alpha_{ji}\alpha_{jk} = \delta_{ik}, \quad \forall\, i, k = 1, \ldots, n.$

Proof Since A is orthogonal, $A^\tau = (\beta_{ik}) = (\alpha_{ki})$. Thus, considering that $I = (\delta_{ik})$ and $AA^\tau = I$, we have

$$\begin{aligned}\delta_{ik} &= \sum_{j=1}^{n} \alpha_{ij}\beta_{jk} \\ &= \sum_{j=1}^{n} \alpha_{ij}\alpha_{kj}\end{aligned}$$

Similarly, you can prove (ii). You can see from this that orthogonal matrices have another important feature. Namely, if $i = k$, the equations above yield

$$1 = \sum_{j=1}^{n} \alpha_{ij}^2$$

that is, the sum of squares of any row of an orthogonal matrix is equal to 1.

And if $i \neq k$

$$0 = \sum_{j=1}^{n} \alpha_{ij}\alpha_{kj}$$

that is, the sum of products of corresponding elements of two different rows is equal to zero. We can make analogous statements for the columns of an orthogonal matrix. ■

Example 6.89 Suppose

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$$

is an orthogonal matrix. Then, by definition,

$$\begin{aligned} AA^\tau &= \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} \\ \alpha_{13} & \alpha_{23} & \alpha_{33} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

So, we get

$$\alpha_{11}^2 + \alpha_{22}^2 + \alpha_{33}^2 = 1$$
$$\alpha_{21}^2 + \alpha_{22}^2 + \alpha_{23}^2 = 1$$
$$a_{31}^2 + \alpha_{32}^2 + \alpha_{33}^2 = 1$$

and also

$$\alpha_{11}\alpha_{21} + \alpha_{12}\alpha_{22} + \alpha_{13}\alpha_{23} = 0$$
$$\alpha_{21}\alpha_{11} + \alpha_{22}\alpha_{12} + \alpha_{23}\alpha_{13} = 0$$

etc.

Suppose we identify the rows of our matrix A with respective vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$, that is, we call $(\alpha_{11}, \alpha_{12}, \alpha_{13}) = \boldsymbol{a}_1$, $(\alpha_{21}, \alpha_{22}, \alpha_{23}) = \boldsymbol{a}_2$, and $(\alpha_{31}, \alpha_{32}, \alpha_{33}) = \boldsymbol{a}_3$. Then, we can write the "scalar" products of vectors as

$$\boldsymbol{a}_1 \cdot \boldsymbol{a}_1 = 1, \quad \boldsymbol{a}_2 \cdot \boldsymbol{a}_2 = 1, \quad \boldsymbol{a}_3 \cdot \boldsymbol{a}_3 = 1$$

and also

$$\boldsymbol{a}_1 \cdot \boldsymbol{a}_2 = \boldsymbol{a}_1 \cdot \boldsymbol{a}_3 = \boldsymbol{a}_2 \cdot \boldsymbol{a}_3 = 0$$

In other words, we have

$$\boldsymbol{a}_i \cdot \boldsymbol{a}_j = \delta_{ij}$$

which is exactly what is claimed by Theorem 6.56. ■

The analogy with the space $\mathbf{R}^3$ is evident, so we generalize by

Definition 6.52 We say that a set of vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbf{R}^n$ is an **orthonormal set** if $\boldsymbol{a}_i \cdot \boldsymbol{a}_j = \delta_{ij}$, $\forall\, i, j = 1, \ldots, n$.

Theorem 6.57 If $A = (\alpha_{ij})$ is a matrix with $\alpha_{ij} \in \mathbf{R}$, the following statements are equivalent:

(i) A is orthogonal.
(ii) The rows of A form an orthonormal set.
(iii) The columns of A form an orthonormal set.

Proof That (i) implies (ii) is evident from Definition 6.48, and the previous theorem (see also Example 6.83). We will show that (ii) and (iii) imply (i). Let $A = (\alpha_{ik})$ be a square matrix such that its rows and columns form an orthonormal set. Furthermore, let $A^\tau = (\beta_{ik})$ and $AA^\tau = (\gamma_{ik})$. Then,

$$\begin{aligned}\gamma_{ik} &= \sum_{j=1}^{n} \alpha_{ik}\beta_{jk} \\ &= \sum_{j=1}^{n} \alpha_{ij}\alpha_{kj} \\ &= \delta_{ik}, \quad \forall i,\ k = 1, \ldots, n\end{aligned}$$

Therefore,

$$AA^{\tau} = (\gamma_{ik}) = (\delta_{ik}) = I$$

Similarly, we can show that $A^{\tau}A = I$, and we conclude that A is an orthogonal matrix. ■

Definition 6.53 A square matrix $A = (\alpha_{ik})$ is said to be **upper triangular** if $\alpha_{ik} = 0$ for all $i > k$, that is,

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{nn} \end{bmatrix}$$

Similarly, we define a **lower triangular** matrix. The proof of the following theorem is obvious.

Theorem 6.58 If A and B are two upper triangular matrices then, $A + B$, λA and AB are also upper triangular matrices.

6.12 RANK OF A MATRIX

Suppose we are given a matrix

$$A = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix} \qquad (*)$$

With our knowledge of linear spaces and their relation to matrices, we are now wondering whether there is a way to associate with every matrix a specific number revealing some additional aspects of matrices and their respective linear operators. Let's approach this matter considering a particular linear space, namely

$$\boldsymbol{M}_{\boldsymbol{m}1} = \left\{ \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{bmatrix} \mid \alpha_i \in \Phi \right\}$$

which is the space of all column matrices with m rows. As you remember, $\boldsymbol{M_{m1}} \cong \Phi^m$ (think for instance of Φ^m as $\mathbf{R}^m$). Consider now a matrix $A \in \boldsymbol{M_{mn}}$ as an element of the linear space $\boldsymbol{M_{mn}}(\Phi)$, and let's label every column of A as a particular matrix

$$C_k = \begin{bmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{mk} \end{bmatrix}, \quad k = 1, \ldots, n$$

This way we establish the correspondence $A \mapsto \{C_1, \ldots, C_n\} \in \boldsymbol{M_{m1}}$, that is, we can associate with every matrix A an n-tuple of vectors from $\boldsymbol{M_{m1}}$. We got a **column-representation** of A. Now we are ready for

Definition 6.54 We say that $r(A)$ is a **rank** (by columns) of a matrix A if $r(A)$ is equal to the maximal number of linearly independent columns from A, that is, if $r(A)$ is equal to the maximal number of linearly independent vectors in $\{C_1, \ldots, C_n\}$.

We can also consider $\{C_1, \ldots, C_n\}$, as a subspace of $\boldsymbol{M_{m1}}$ spanned by the vectors $C_1, \ldots, C_n$, and define the rank of A as follows.

Definition 6.54′ We say that $r(A)$ is the **rank** (by columns) of matrix A if

$$r(A) = \dim\{C_1, \ldots, C_n\}$$

Similarly, we can consider a space $\boldsymbol{M_{1n}} = \left\{ \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix} | \alpha_i \in \Phi \right\}$ of all row matrices with n elements, and then select vectors

$$R_i = [\alpha_{i1}, \ldots, \alpha_{in}], \quad i = 1, \ldots, m$$

as rows of our matrix A in $(*)$. This way, as earlier, we can establish the correspondence $A \mapsto \{R_1, \ldots, R_m\}$ as a **row-representation** of matrix A. Clearly, the set $\{R_1, \ldots, R_m\}$ is a subspace of $\boldsymbol{M_{1n}}$, and we are tempted to say

Definition 6.55 We say that $r'(A)$ is a **rank** (by rows) of matrix A if

$$r'(\mathrm{A}) = \dim\{R_1, \ldots, R_m\}$$

It turns out that $r(A) = r'(A)$. Without proof, we state

Theorem 6.59 The maximal number of linearly independent rows of any matrix A is equal to the maximal number of linearly independent columns. That is, for any matrix A the rank by rows is equal to the rank by columns, that is, $r'(A) = r(A)$.

If you accept this theorem then the following corollary seems obvious.

Corollary 6.6 Matrix A and its transpose A^τ have the same rank, that is, $r(A) = r(A^\tau)$. Observe that $r(O) = 0$ and $r(I) = n$.

Before discussing the actual calculation of the rank, let's introduce

Elementary Operations on Matrices

(E_1) : Interchanging any two rows (columns) of the matrix:

$$[R_i \leftrightharpoons R_j \, (C_i \leftrightharpoons C_j)]$$

(E_2) : Multiplying any row (column) of the matrix by a nonzero number k:

$$[kR_i \leftrightharpoons R_i \;\; (kC_i \leftrightharpoons C_i)]$$

(E_3) : Multiplying any row by any number k and adding the result to any other row:

$$[kR_i + R_j \leftrightharpoons R_j \;\; (kC_i + C_j \leftrightharpoons C_j)]$$

Definition 6.56 An (m, n) matrix A is said to be equivalent to an (m, n) matrix B, written $A \sim B$ if B can be obtained from A by a sequence of elementary operations. In other words, $A \sim B$ if there exists a finite sequence of matrices of the same type $A_1, A_2, \ldots, A_{k-1}, A_k = B$ such that A_i is obtained from A_{i-1} by some of the operations $(E_1) - (E_3)$.

I hope you have recognized, and can easily prove, that "$\sim$" as defined above is an equivalence relation. Indeed,

(i) $A \sim A$ for any matrix A.
(ii) If $A \sim B$ then, $B \sim A$.
(iii) If $A \sim B$ and $B \sim C$ then $A \sim C$.

That $A \sim A$ is trivial. For (ii) observe that any operation that led from A to B can be reversed. For instance, the interchange of rows R_i and R_j can be "reversed," that is, the interchange is its own inverse. Similarly, replacing R_i by kR_i can be "reversed" by replacing R_i by $1/k\, R_i$, that is, "replace R_i by kR_i" and "replace R_i by $1/k\, R_i$," are inverse operations. Finally, "replace R_j by $kR_i + R_j$" and "replace R_j by $-kR_i + R_j$" are also inverse operations. Thus, $B \sim A$. With (i) and (ii), (iii) immediately follows.

Theorem 6.60 Let $A \sim B$, then $r(A) = r(B)$.

Proof We need to prove that elementary transformations do not change the rank. It is clear that transformations (E_1) and (E_2) have no effect on the rank of a matrix. What about (E_3)? Let's see. Suppose matrix B is obtained from matrix A by adding a second column to its first column, that is, $C_1 + C_2 \leftrightharpoons C_1$. Thus, the column representations of A and B are as follows:

$$A \mapsto (C_1,\ C_2,\ \dots\ ,\ C_n) \tag{*}$$

$$B \mapsto (C_1 + C_2,\ C_2,\ \dots\ ,\ C_n) \tag{**}$$

Consider now the two subspaces of $\boldsymbol{M}_{\boldsymbol{m}1}$ spanned by $(*)$ and $(**)$:

$$\boldsymbol{L} = [(C_1,\ C_2,\ \dots\ ,\ C_n)] < \boldsymbol{M}_{\boldsymbol{m}1}$$

and

$$\boldsymbol{K} = [(C_1 + C_2,\ C_2,\ \dots\ ,\ C_n)] < \boldsymbol{M}_{\boldsymbol{m}1}$$

We would like to prove that $\boldsymbol{L} = \boldsymbol{K}$, which would furthermore imply that

$$\begin{aligned} \dim \boldsymbol{L} &= \dim \boldsymbol{K} \\ &= r(A) = r(B) \end{aligned}$$

If we take some $\boldsymbol{x} \in \boldsymbol{L}$, then

$$\boldsymbol{x} = \alpha_1 C_1 + \alpha_2 C_2 + \ \cdots + \alpha_n C_n$$

but also, with a little trick,

$$\begin{aligned} \boldsymbol{x} &= \alpha_1 C_1 + \alpha_1 C_2 - \alpha_1 C_2 + \alpha_2 C_2 + \ \cdots + \alpha_n C_n \\ &= \alpha_1 (C_1 + C_2) + (\alpha_2 - \alpha_1) C_2 + \alpha_3 C_3 + \cdots + \alpha_n C_n \end{aligned}$$

Hence, $\boldsymbol{x} \in \boldsymbol{K}$ and therefore $\boldsymbol{L} \subseteq \boldsymbol{K}$. On the other hand, if we take some $\boldsymbol{y} \in \boldsymbol{K}$ we can express it as

$$\begin{aligned} \boldsymbol{y} &= \beta_1 (C_1 + C_2) + \beta_2 C_2 + \cdots + \beta_n C_n \\ &= \beta_1 C_1 + (\beta_1 + \beta_2) C_2 + \cdots + \beta_n C_n \end{aligned}$$

This obviously implies that $\boldsymbol{y} \in \boldsymbol{L}$, and therefore $\boldsymbol{K} \subseteq \boldsymbol{L}$. We have proved that $\boldsymbol{L} = \boldsymbol{K}$, which means $\dim \boldsymbol{L} = \dim \boldsymbol{K}$ and thus $r(A) = r(B)$ as claimed. ■

Now let's address the problem of actually determining the rank of a matrix. We need two more definitions.

Definition 6.57 A matrix $A = (\alpha_{ij})$ is said to be in **echelon form** or to be an **echelon matrix** if the following conditions are satisfied:

(i) All zero rows, if any, are at the bottom of the matrix.
(ii) Each leading nonzero entry in a row is to the right of the leading nonzero entry in the preceding row.

Definition 6.58 A matrix $A = (\alpha_{ij})$ is said to be in **row canonical form** or **reduced echelon form** if it is an echelon matrix and if it satisfies two additional conditions:

(iii) Each leading **nonzero entry (pivot)** is equal to 1.
(iv) Each pivot is the only nonzero entry in its column.

Example 6.90 The following three matrices are in echelon form:

$$A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 5 & 7 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad C = \begin{bmatrix} 3 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 5 \end{bmatrix}$$

■

Example 6.91 The following matrices are in row canonical form:

$$A = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad B = \begin{bmatrix} 0 & 1 & 3 & 2 \\ 0 & 0 & 1 & 4 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

■

Example 6.92 Consider a matrix

$$A = \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix}$$

and reduce it to row canonical form.

Solution Using the elementary operation we reduce matrix A first to echelon form as follows:

$$A \xrightarrow{-2R_1+R_2\to R_2} \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix} \xrightarrow{-3R_1+R_3\to R_3} \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 & 7 \end{bmatrix}$$

$$\xrightarrow{\frac{-3}{2}R_2+R_3\to R_3} \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

The matrix is now in echelon form. What remains is to reduce it to row canonical form. Again, applying the elementary operations, we get

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \xrightarrow{-\frac{1}{2}R_3\ \to R_3} \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{-6R_3+R_2\ \to R_2} \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{-2R_3+R_1\ \to R_1} \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}R_2\ \to R_2} \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{3R_2+R_1\to R_1} \begin{bmatrix} 1 & 2 & 0 & 7 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

And this is the desired result. ■

We anticipate that the problem of finding the rank of an (m, n) matrix A reduces to the problem of transforming a matrix A to its equivalent, say, a matrix D whose rank is self-evident. Here is the idea more explicitly. Suppose we succeed in obtaining an (m, n) matrix $D = (\alpha_{ik})$ such that

$$\alpha_{ik} = \begin{cases} 1 & i = k \le r \\ 0 & \text{otherwise} \end{cases} \qquad r \le m, n$$

that is, we have

$$D_r = \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

which is a canonical matrix whose rank is evidently r. The following theorem assures us that we can do this for any matrix.

Theorem 6.61 Let $A = (\alpha_{ik})$ be any (m, n) matrix. Then there exists a canonical (m, n) matrix D_r, such that $A \sim D_r$.

Proof Let $A = (\alpha_{ik})$ be any matrix, such that $A \neq O$, that is, at least one entry $\alpha_{ik} \neq 0$. Otherwise, A would be equal to D_0 and there would be nothing to prove. Permuting the rows and columns we move this element to the upper left corner and then, dividing the first row by this element, we get matrix B, which has 1 at the upper left corner, that is,

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix} \sim \begin{bmatrix} 1 & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mn} \end{bmatrix} = B$$

We repeat this procedure to get

$$B \sim \cdots \sim \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & \gamma_{m2} & \cdots & \gamma_{mn} \end{bmatrix} = C$$

In the case that $\gamma_{ik} = 0$, for all i and k we are done, $A \sim D_1$. If not, that is, at least one $\alpha_{ik} \neq 0$, we continue the same way until after finitely many steps we reach the matrix

$$D_r = \begin{bmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix}$$

whereby $\boldsymbol{I}_r$ we mean (r, r)-identity matrix, and by $\boldsymbol{O}$ we mean the remaining zeros in the matrix D_r. Matrix D_r has the same rank as $\boldsymbol{I}_r$, that is r.

So we see that $A \sim B \sim C \sim \cdots \sim D_r$, and the rank $r(A) = r(D_r) = r$. ■

Consequently, we have another useful definition

Definition 6.59 The rank of a matrix A is equal to the number of pivots of A in echelon form.

Example 6.93 Find the rank of the following matrix:

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 1 & 0 \\ -2 & -1 & 3 \\ -1 & 4 & -2 \end{bmatrix}$$

Solution In this case it might be practical to consider A^τ, and then proceed with the transformations:

$$A \sim A^\tau = \begin{bmatrix} 1 & 2 & -2 & -1 \\ 2 & 1 & -1 & 4 \\ -3 & 0 & 3 & -2 \end{bmatrix} \xrightarrow{-2R_1+R_2\to R_2} \begin{bmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 6 & -3 & -5 \end{bmatrix}$$

$$\xrightarrow{2R_2+R_3\to R_3} \begin{bmatrix} 1 & 2 & -2 & -1 \\ 0 & -3 & 3 & 6 \\ 0 & 0 & 3 & 7 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2\to R_2} \begin{bmatrix} 1 & 2 & -2 & -1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 3 & 7 \end{bmatrix}$$

We obtained the echelon matrix with three nonzero rows, thus, $r(A) = 3$. However, let's continue to reach reduced echelon form:

$$\xrightarrow{\frac{1}{3}R_3\to R_3} \begin{bmatrix} 1 & 2 & -2 & -1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & \frac{7}{3} \end{bmatrix} \xrightarrow{-2R_2\to R_1} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & \frac{7}{3} \end{bmatrix}$$

$$\xrightarrow{R_3+R_2\to R_2} \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & -0 & \frac{1}{3} \\ 0 & 0 & 1 & \frac{7}{3} \end{bmatrix}$$

Thus, $r(A) = 3$ indeed. ■

Example 6.94 Find the rank of the following matrix:

$$A = \begin{bmatrix} 0 & -2 & 1 & 1 & -5 \\ 2 & 1 & 3 & 2 & 1 \\ 2 & 1 & -1 & 6 & 5 \\ 2 & -1 & 12 & -5 & -12 \end{bmatrix}$$

Solution

$$A = \begin{bmatrix} 0 & -2 & 1 & 1 & -5 \\ 2 & 1 & 3 & 2 & 1 \\ 2 & 1 & -1 & 6 & 5 \\ 2 & -1 & 12 & -5 & -12 \end{bmatrix} \xrightarrow{2^{-1}C_1 \leftrightharpoons C_3} \begin{bmatrix} 1 & -2 & 0 & 1 & -5 \\ 3 & 1 & 1 & 2 & 1 \\ -1 & 1 & 1 & 6 & 5 \\ 12 & -1 & 1 & -5 & -12 \end{bmatrix}$$

$$\xrightarrow{2C_1+C_2\to C_2} \begin{bmatrix} 1 & 0 & 0 & 1 & -5 \\ 3 & 7 & 1 & 2 & 1 \\ -1 & -1 & 1 & 6 & 5 \\ 12 & 23 & 1 & -5 & -12 \end{bmatrix} \to$$

$$\xrightarrow{C_4-C_1\to C_4}\begin{bmatrix}1&0&0&0&-5\\3&7&1&-1&1\\-1&-1&1&7&5\\12&23&1&-17&-12\end{bmatrix}$$

$$\xrightarrow{5C_1+C_5\to C_5}\begin{bmatrix}1&0&0&0&-5\\3&7&1&-1&16\\-1&-1&1&7&0\\12&23&1&-17&48\end{bmatrix}\xrightarrow{-3R_1+R_2\to R_2;\ R_1+R_3\to R_3;-12R_1+R_4\to R_4}$$

$$\to\begin{bmatrix}1&0&0&0&0\\0&7&1&-1&16\\0&-1&1&7&0\\0&23&1&-17&48\end{bmatrix}\xrightarrow{16^{-1}C_5}\begin{bmatrix}1&0&0&0&0\\0&7&1&-1&1\\0&-1&1&7&0\\0&23&1&-17&3\end{bmatrix}$$

$$\xrightarrow{C_2\leftrightharpoons C_3}\begin{bmatrix}1&0&0&0&0\\0&1&7&-1&1\\0&1&-1&7&0\\0&1&23&-17&3\end{bmatrix}\xrightarrow{R_3-R_2\to R_3;\ \ R_4-R_2\to R_4}$$

$$\to\begin{bmatrix}1&0&0&0&0\\0&1&7&-1&1\\0&0&-8&8&-1\\0&0&16&-16&2\end{bmatrix}\to$$

$$\xrightarrow{-7C_2+C_3\to C_3;\ \ C_2+C_4\to C_4;\ C_5-C_2\to C_5}\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&-8&8&-1\\0&0&16&-16&2\end{bmatrix}\xrightarrow{C_3+C_4\to C_4}$$

$$\to\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&-8&0&-1\\0&0&16&0&2\end{bmatrix}\xrightarrow{2R_3+R_4\to R_4}\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&-8&0&-1\\0&0&0&0&0\end{bmatrix}$$

$$\xrightarrow{(-8)^{-1}C_3;\ \ (-8)^{-1}C_3+C_5\to C_5}\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0\end{bmatrix}$$

So, $r(A) = 3$. ■

Example 6.95 Find the rank of the following matrix:

$$A=\begin{bmatrix}0&2&3\\2&2&4\end{bmatrix}$$

Solution

$$A = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 2 & 4 \end{bmatrix} \xrightarrow{R_1 \curvearrowright R_2} \begin{bmatrix} 2 & 2 & 4 \\ 0 & 2 & 3 \end{bmatrix} \xrightarrow{2^{-1}R_1} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 3 \end{bmatrix} \xrightarrow{C_2 - C_1 \to C_2;\ C_3 - 2C_1 \to C_3}$$

$$\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \end{bmatrix} \xrightarrow{2^{-1}R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3/2 \end{bmatrix} \xrightarrow{-3/2C_2 + C_3 \to C_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

So $r(A) = 2$. ■

Example 6.96 Find the rank of the following matrix:

$$A = \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{bmatrix}$$

Solution

$$A = \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 1 & 4 & 3 & -1 & -4 \\ 2 & 3 & -4 & -7 & -3 \\ 3 & 8 & 1 & -7 & -8 \end{bmatrix} \xrightarrow{-R_1 + R_2 \to R_2;\ -2R_1 + R_3 \to R_3;\ -3R_1 + R_3 \to R_3}$$

$$\to \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & -3 & -6 & -3 & 3 \\ 0 & -1 & -2 & -1 & 1 \end{bmatrix} \xrightarrow{3R_2 + R_3 \to R_3;\ R_2 + R_4 \to R_4} \begin{bmatrix} 1 & 3 & 1 & -2 & -3 \\ 0 & 1 & 2 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence, $r(A) = 2$. ■

We end this section by stating four useful theorems.

Theorem 6.62 The rank of an (m, n) matrix cannot be greater than m or n.

Theorem 6.63 For any matrix A, the rank by rows is equal to the rank by columns.

Theorem 6.64 $r(A) = r(A^\tau)$.

Theorem 6.65 Two (m, n) matrices are equivalent iff they have the same rank.

6.13 DETERMINANTS

We now introduce in our study of linear operators and matrices an extremely important, in fact indispensable, tool – the **determinant**. The idea is this: Is it

possible to assign to every (n, n) matrix over a field Φ, a number (a scalar) from the same field? And, if it is, what is the significance of that number in our study of vector spaces?

In other words, we are looking for an object defined as follows.

Definition 6.60 Let $\boldsymbol{M_{nn}}$ be a space of all (n, n) matrices over the field Φ. We define a map called a **functional**

$$\det : \boldsymbol{M_{nn}} \to \Phi$$

which associates with every matrix $A = (\alpha_{ij}) \in \boldsymbol{M_{nn}}$ a "number"

$$\det A = |A| \in \Phi$$

and we say that $\det A = |A|$ is the **determinant** of A.

However, in order to proceed in our search for the determinants, we also need

Definition 6.61 Given a (n, n) matrix $A = (\alpha_{ij})$ we say that $M_{i,j}(A)$, or simply $M_{i,j}$, is $\boldsymbol{i,j}$**-minor** of A if $M_{i,j}(A)$ is a matrix obtained from matrix A by deleting the ith row and jth column of A.

Example 6.97 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 2 \\ 1 & 5 & 7 \end{bmatrix}$ then $M_{1,1} = \begin{bmatrix} 0 & 2 \\ 5 & 7 \end{bmatrix}, M_{1,2} = \begin{bmatrix} 3 & 2 \\ 1 & 7 \end{bmatrix}, M_{1,3} = \begin{bmatrix} 3 & 0 \\ 1 & 5 \end{bmatrix}$, and so on. ■

Definition 6.62 Let $A = (\alpha_{ij})$ be a (n, n) matrix. We (inductively) define the determinant of A as follows:

$\det(A) = \alpha$ if A is $(1, 1)$ matrix. If $n > 1$ then

$$\begin{aligned} \det(A) &= \alpha_{11} \det(M_{1,1}) - \alpha_{12} \det(M_{1,2}) + \cdots + (-1)^{n+1} \alpha_{1n} \det(M_{1,n}) \\ &= \sum_{j=1}^{n} (-1)^{1+j} \alpha_{1j} \det(M_{1,j}) \end{aligned}$$

Fortunately, as Laplace has shown, we can similarly evaluate $\det(A)$ by any row or column. First, here is another definition of a minor.

Definition 6.61′ Given a (n, n) matrix $A = (\alpha_{ij})$ we say that $M_{i,j}(A)$, or simply $M_{i,j}$, is a **minor of the element** $\boldsymbol{\alpha_{ij}}$ **of** $\boldsymbol{A}$, if $M_{i,j}$ is a **determinant** of sub-matrix of A obtained by deleting its ith row and jth column.

Definition 6.62 Let[7] $A = (\alpha_{ik})$ be a (n, n) matrix, then

$$A_{ik} = (-1)^{i+k} \det M_{i,k}$$

is called a **cofactor** of the matrix element α_{ik}.

Example 6.98 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, then:

$$M_{1,1} = \begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} \quad \text{and} \quad A_{11} = + \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} = 45 - 48 = -3$$

$$M_{1,2} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} \quad \text{and} \quad A_{12} = - \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -(36 - 42) = 6$$

$$\vdots \qquad\qquad \vdots$$

$$M_{3,1} = \begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix} \quad \text{and} \quad A_{31} = + \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 12 - 15 = -3$$

and so on. ■

Example 6.99 Let $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$, then some of the cofactors are:

$$A_{11} = + \begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} = -18, \quad A_{23} = - \begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix} = 5, \quad A_{33} = + \begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix} = -8$$

and so on. You can easily calculate the rest. ■

Theorem 6.66 (Laplace) Let $A = (\alpha_{ik})$ be any (n, n) matrix

$$\det A = \sum_{j=1}^{n} \alpha_{ij} A_{ij}, \qquad \forall\, i = 1, \ldots, n$$

or

$$\det A = \sum_{j=1}^{n} \alpha_{jk} A_{jk} \quad \forall\, k = 1, \ldots n$$

Example 6.100 Calculate the determinant of the following matrix

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$$

[7] In the case that a minor is defined as in Definition 6.61′ then, of course, $A_{ik} = (-1)^{i+k} M_{i,k}$.

Solution

$$\det A = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix}$$

$$= \alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12} \begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13} \begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix}$$

$$= \alpha_{11}(\alpha_{22}\alpha_{33} - \alpha_{23}\alpha_{32}) - \alpha_{12}(\alpha_{21}\alpha_{33} - \alpha_{23}\alpha_{31}) + \alpha_{13}(\alpha_{21}\alpha_{32} - \alpha_{22}\alpha_{31})$$

■

Example 6.101 Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 2 \\ 3 & 1 & 0 \end{bmatrix}$, calculate $\det A$.

Solution

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 3 & 2 \\ 3 & 1 & 0 \end{vmatrix}$$

$$= 1 \cdot (3 \cdot 0 - 2 \cdot 1) - 2 \cdot (0 \cdot 0 - 2 \cdot 3) + 3 \cdot (0 \cdot 1 - 3 \cdot 3)$$

$$= -17$$

■

Example 6.102 Calculate

$$\begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 4 & 2 \\ 0 & 1 & 0 & 4 \\ 1 & 0 & 2 & 1 \end{vmatrix}$$

Solution

$$\begin{vmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 4 & 2 \\ 0 & 1 & 0 & 4 \\ 1 & 0 & 2 & 1 \end{vmatrix} = 1 \begin{vmatrix} 1 & 4 & 2 \\ 1 & 0 & 4 \\ 0 & 2 & 1 \end{vmatrix} - 2 \begin{vmatrix} 0 & 4 & 2 \\ 0 & 0 & 4 \\ 1 & 2 & 1 \end{vmatrix} + (-1) \begin{vmatrix} 0 & 1 & 2 \\ 0 & 1 & 4 \\ 1 & 0 & 1 \end{vmatrix} - 3 \begin{vmatrix} 0 & 1 & 4 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{vmatrix}$$

$$= 1(1(0 - 8) - 4(1 - 0) + 2(2 - 0))$$
$$- 2(0(0 - 8) - 4(0 - 4) + 2(0 - 0))$$
$$- 1(0(1 - 0) - 1(0 - 4) + 2(0 - 1))$$
$$- 3(0(2 - 0) - 1(0 - 0) + 4(0 - 1))$$

$$= -30$$

■

We are ready to define the determinant of any order precisely. However, we need to recall certain things from the previous chapters.

First, recall that we write an arbitrary permutation $p \in S_n$ either as

$$p = \begin{pmatrix} 1 & 2 & \cdots & n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}$$

or simply as

$$p = j_1 j_2 \cdots j_n$$

where $j_i = p(i)$. Also, recall that the *parity* of a permutation is called *even* or *odd*, depending on the number of inversions $I(p)$ of p. So, the parity of a permutation is

$$\begin{aligned} \operatorname{sgn} p &= \begin{cases} 1 & \text{if } p \text{ is even} \\ -1 & \text{if } p \text{ is odd} \end{cases} \\ &= (-1)^{I(p)} \end{aligned}$$

With all that said, we proceed to precisely define the determinant.

Take a matrix $A = (\alpha_{ij}) \in \boldsymbol{M}_{nn}$ over a field Φ. Next, consider a product $\prod_{k=1}^{n} \alpha_{1p(k)}$ of n elements of A, making sure that *one and only one* element comes from each row and *one and only one* element from each column. Explicitly written, this product is

$$\alpha_{1p(1)} \alpha_{2p(2)} \cdots \alpha_{np(n)}$$

Since there are n terms (elements) in this product, there are $n!$ of such products one can construct from matrix A. We want to collect them all. Thus, the object we are looking for is defined as follows:

Definition 6.63 The determinant of a matrix $A = (\alpha_{ij}) \in \boldsymbol{M_{nn}}$ over a field Φ is a functional

$$\det : \boldsymbol{M_{nn}} \to \Phi$$

defined by

$$\det A = |A| = \sum_{p \in S_n} (-1)^{I(p)} \alpha_{1p(1)} \alpha_{2p(2)} \ \cdots \ \alpha_{np(n)}$$

Example 6.103 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The set of all permutations $\prod_2$ of S_2 has two elements

$$p_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad p_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

So $\operatorname{sgn} p_1 = 1$ and $\operatorname{sgn} p_2 = -1$. Thus,

$$\det A = ad - bc \qquad \blacksquare$$

Example 6.104 Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Now the set $\prod_3$ has the following elements:

$$p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

$$p_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad p_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad p_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Observe that p_1, p_2, p_3 are even and p_4, p_5, p_6 are odd permutations. Therefore,

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{31}a_{22}a_{13}$$

$\blacksquare$

The calculations of determinants look quite laborious and boring indeed. Fortunately, matrices/determinants have properties that very often expedite those calculations. Here they are:

Theorem 6.67 (Properties of determinants)

(i) If a matrix A has a row of zeros then $\det A = 0$.
(ii) If a matrix B is obtained from A by interchanging any two rows (columns) then $\det B = -\det A$.
(iii) If a matrix A has two rows equal (two columns equal), then $\det A = 0$.
(iv) $\det A^\tau = \det A$.
(v) If a matrix A is orthogonal, then $\det A = \pm 1$.
(vi) If a matrix B is obtained from matrix A by multiplying *one* of its rows (columns) by a scalar $\lambda \in \Phi$, then $\det B = \lambda \det A$.
(vii) If a matrix B is obtained from a matrix A by adding a multiple of one row (column) to a different row (column), then $\det B = \det A$.
(viii) If a matrix A has two proportional rows (columns), then $\det A = 0$.
(ix) If a matrix A is triangular, then $\det A$ is the product of the diagonal entries.
(x) **Binet–Cauchy Theorem:**
$\det(AB) = \det A \det B = \det B \det A = \det(BA)$
(xi) If A is an (n, n) matrix, and $\lambda \in \Phi$, then $\det(\lambda A) = \lambda^n \det A$.

Proof The proofs directly follow from the definition of the determinant. For instance, statement (ii) follows from the fact that interchanging of any two neighboring rows (columns) changes the parity. Thus, "any two" interchanges can be obtained by an odd number of successive interchanges of the neighboring rows (columns). Similarly, (iii) follows from (ii) and the observation that $\det A = -\det A$, thus $\det A = 0$. (v) also directly follows from the definition of the determinant. Then, (vii) follows from (iii) and (v). (x) follows from (v), and so on. ■

Now we are ready to prove another useful theorem.

Theorem 6.68 Let A be a regular matrix, then

$$\det A^{-1} = \frac{1}{\det A}$$

Proof Since $AA^{-1} = \mathrm{I}$, by the Binet–Cauchy formula

$$\begin{aligned} \det(AA^{-1}) &= \det(I) \\ &= \det(A)\det(A^{-1}) = 1 \end{aligned}$$

and thus

$$\det(A^{-1}) = \frac{1}{\det A}$$

■

Example 6.105 Show that

$$\det(A^{-1}A^{\tau}A) = \det A$$

Solution

$$\begin{aligned} \det(A^{-1}A^{\tau}A) &= \det A^{-1} \det A^{\tau} \det A \\ &= \frac{1}{\det A} \det A \det A \\ &= \det A \end{aligned}$$

■

Example 6.106 Show that if A and B are two (n, n) matrices, with A being singular, then AB is also a singular matrix.

Solution If matrix A is singular, then $\det A = 0$ (See Theorem 6.72). Therefore,

$$\begin{aligned}\det(AB) &= \det A \ \det B \\ &= 0 \cdot \det B \\ &= 0\end{aligned}$$

Thus, AB is a singular matrix, as claimed. Note, though, that the converse is not true. Indeed,

$$\begin{aligned}\det(AB) &= 0 \\ &= \det A \ \det B\end{aligned}$$

implies that either $\det\ A = 0$ or $\det\ B = 0$, or both are equal to zero. ■

6.14 THE INVERSE AND THE RANK OF A MATRIX REVISITED

Recall (see Definition 6.46) that an (n, n) matrix A is said to be regular if there exists a matrix A^{-1} such that $A\,A^{-1} = A^{-1}A = I$. Now we are in possession of a very handy criterion for deciding whether a matrix is regular or not.

Definition 6.64 Let $A = (\alpha_{ij})$ be an (n, n) matrix over a field Φ, and let A_{ij} be a cofactor of α_{ij}. The **classical adjoint of** A, denoted $\text{adj}A$ or $\widetilde{A}$, is a transpose of the matrix of cofactors of A, namely

$$\text{adj}A = \widetilde{A} = (A_{ij})^{\tau} = (A_{ji})$$

Example 6.107 Let $A = \begin{bmatrix} 2 & 3 & -4 \\ 0 & -4 & 2 \\ 1 & -1 & 5 \end{bmatrix}$, then the cofactors are

$$A_{11} = -18, \quad A_{12} = 2, \quad A_{13} = 4, \quad A_{21} = -11, \quad A_{22} = 14, \quad A_{23} = 5$$
$$A_{31} = -10, \quad A_{32} = -4, \quad A_{33} = -8$$

So, the transpose matrix of cofactors is

$$\text{adj}\ A = \widetilde{A} = \begin{bmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{bmatrix}$$

■

Example 6.108 Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 0 & 4 & 6 \end{bmatrix}$, then the cofactors are

$$A_{11} = -2, \quad A_{12} = -12, \quad A_{13} = 8, \quad A_{21} = 4, \quad A_{22} = 6, \quad A_{23} = -4$$
$$A_{31} = -1, \quad A_{32} = 0, \quad A_{33} = 1$$

So, the transpose matrix of cofactors is

$$\text{adj } A = \widetilde{A} = \begin{bmatrix} -2 & 4 & -1 \\ -12 & 6 & 0 \\ 8 & -4 & 1 \end{bmatrix}$$ ■

Example 6.109 Let $A = \begin{bmatrix} 3 & 2 & -1 \\ 1 & 6 & 3 \\ 2 & -4 & 0 \end{bmatrix}$, then the cofactors are

$$A_{11} = 12, \quad A_{12} = 6, \quad A_{13} = -16, \quad A_{21} = 4, \quad A_{22} = 2, \quad A_{23} = 16, \quad A_{31} = 12$$
$$A_{32} = -10, \quad A_{33} = 16$$

So,

$$\text{adj}A = \widetilde{A} = \begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix}$$ ■

Example/Exercise 6.110 Show that

$$(\widetilde{A}^{\tau}) = (\widetilde{A})^{\tau}$$

Theorem 6.69 Let $A = (\alpha_{ik})$ be any (n, n) matrix, then the sum of the products of elements of any row (column) and cofactors of the corresponding elements of any other row (column) is equal to zero, that is,

$$\sum_{k=1}^{n} \alpha_{ik} A_{jk} = 0$$

Proof Consider a matrix $B = (\beta_{rk})$ defined by

$$\beta_{rk} = \begin{cases} \alpha_{rk}, & r \neq k \\ \alpha_{ik}, & r = j \end{cases} \qquad \forall k = 1, \ldots, n$$

Calculate det B expanding on jth row

$$\det B = \sum_{k=1}^{n} \beta_{jk} B_{jk} = \sum_{k=1}^{n} \alpha_{ik} A_{jk}$$

Since ith and jth rows of the matrix B are identical, $\det B = 0$, and we have the proof of our theorem. ■

As a consequence we have

Theorem 6.70 If $A = (\alpha_{ik})$ is any (n, n)-matrix, then

(i) $\displaystyle\sum_{k=1}^{n} \alpha_{ik} A_{jk} = \delta_{ij} \det A$

(ii) $\displaystyle\sum_{i=1}^{n} \alpha_{ik} A_{il} = \delta_{kl} \det A$

Theorem 6.71 If $A = (\alpha_{ik})$ is any (n, n)-matrix, then

$$A\widetilde{A} = \widetilde{A}A = \det A \cdot I.$$

Proof Let $A\widetilde{A} = C = (\gamma_{ik})$. Then, because of the previous theorem,

$$\gamma_{ik} = \sum_{j=1}^{n} \alpha_{ij} A_{kj} = \delta_{ik} \det A$$

Thus,

$$\begin{aligned} C = (\gamma_{ik}) &= (\delta_{ik} \det A) \\ &= \det A \cdot (\delta_{ik}) = \det A \cdot I \end{aligned}$$

■

Theorem 6.72 A (n, n) matrix A is regular iff $\det(A) \neq 0$.

Proof Let A be a regular matrix and A^{-1} its inverse, and suppose that $\det A = 0$. Then, by the Binet–Cauchy formula, we would have

$$\begin{aligned} \det(AA^{-1}) &= \det I = 1 \\ &= \det A \ \det A^{-1} \\ &= 0 \ \det A^{-1} = 0 \end{aligned}$$

which, of course, is a contradiction. Conversely, let $\det A \neq 0$. Then, since

$$A\widetilde{A} = \widetilde{A}A = (\det A) \cdot I$$

Multiplying both sides of this equation by $(\det A)^{-1}$, we get

$$\begin{aligned}(\det A)^{-1}A\widetilde{A} &= A[(\det A)^{-1}\widetilde{A}] \\ &= [(\det A)^{-1}\widetilde{A}\,]A \\ &= (\det A)^{-1}(\det A) \\ &= I\end{aligned}$$

We see that the equation $AX = XA = I$ has a solution. Thus, A is a regular matrix. ■

Example/Exercise 6.111 Show that if A is a regular matrix, then $\det A^{-1} = (\det A)^{-1}$.

Example 6.112 Let $A = \begin{bmatrix} 2 & -1 & 1 \\ 4 & 1 & -3 \\ 2 & -1 & 3 \end{bmatrix}$. Show that

$$\det A^{-1} = \frac{1}{\det A}.$$

Solution Cofactors are as follows:

$$A_{11} = 0, \quad A_{12} = -18, \quad A_{13} = -6, \quad A_{21} = 2, \quad A_{22} = 4, \quad A_{23} = 0, \quad A_{31} = 2,$$
$$A_{32} = 10, \quad A_{33} = 6$$

So,

$$\widetilde{A} = \begin{bmatrix} 0 & 2 & 2 \\ -18 & 4 & 10 \\ -6 & 0 & 6 \end{bmatrix} \quad \text{and} \quad \det A = \det \begin{bmatrix} 2 & -1 & 1 \\ 4 & 1 & -3 \\ 2 & -1 & 3 \end{bmatrix} = 12$$

Thus,

$$A^{-1} = \frac{1}{\det A}\widetilde{A} = \frac{1}{12}\begin{bmatrix} 0 & 2 & 2 \\ -18 & 4 & 10 \\ -6 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 1/6 & 1/6 \\ -3/2 & 1/3 & 5/6 \\ -1/2 & 0 & 1/2 \end{bmatrix}$$

Therefore,

$$\det A^{-1} = \det \begin{bmatrix} 0 & 1/6 & 1/6 \\ -3/2 & 1/3 & 5/6 \\ -1/2 & 0 & 1/2 \end{bmatrix} = \frac{1}{12}$$
$$= \frac{1}{\det A}$$

■

Theorem 6.73 If A is a regular matrix, then

$$A^{-1} = (\det A)^{-1} \cdot \widetilde{A}$$

Proof The proof follows immediately from the previous theorem. ■

Example 6.113 Let $A = \begin{bmatrix} 5 & 7 & -2 \\ 3 & 2 & 1 \\ -2 & 4 & 6 \end{bmatrix}$. Find A^{-1} using Theorem 6.73.

Solution First, we calculate the cofactors:

$$A_{11} = 8, \quad A_{12} = -20, \quad A_{13} = 16, \quad A_{21} = -50, \quad A_{22} = 26, \quad A_{23} = -34,$$
$$A_{31} = 11, \quad A_{32} = -11, \quad A_{33} = -11$$

So,

$$\widetilde{A} = \begin{bmatrix} 8 & -50 & 11 \\ -20 & 26 & -11 \\ 16 & -34 & -11 \end{bmatrix}$$

We also need

$$\det A = \det \begin{bmatrix} 5 & 7 & -2 \\ 3 & 2 & 1 \\ -2 & 4 & 6 \end{bmatrix} = -132$$

Hence,

$$A^{-1} = (\det A)^{-1} \cdot \widetilde{A}$$
$$= -\frac{1}{132} \begin{bmatrix} 8 & -50 & 11 \\ -20 & 26 & -11 \\ 16 & -34 & -11 \end{bmatrix}$$

■

Example 6.114 Let $A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$. Show that $\widetilde{A} \cdot A = \det A \cdot I$.

Solution The cofactors of A are as follows:

$$A_{11} = -1, \quad A_{12} = -1, \quad A_{13} = 1, \quad A_{21} = 2, \quad A_{22} = 2$$
$$A_{23} = -4, \quad A_{31} = -1, \quad A_{32} = -3, \quad A_{33} = 3$$

Consequently,

$$\widetilde{A} = \begin{bmatrix} -1 & 1 & -1 \\ -1 & 2 & -3 \\ 1 & -4 & 3 \end{bmatrix}$$

So,

$$\widetilde{A} \cdot A = \begin{bmatrix} -1 & 1 & -1 \\ -1 & 2 & -3 \\ 1 & -4 & 3 \end{bmatrix} \begin{bmatrix} 3 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = -2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= -2 \cdot I$$

On the other hand,

$$\det A = \det \begin{bmatrix} 3 & 1 & 2 \\ 0 & 1 & 1 \\ -1 & 1 & 0 \end{bmatrix} = -2$$

Thus,

$$\widetilde{A} \cdot A = \det A \cdot I$$

as desired. ■

Example/Exercise 6.115 Let $A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}$, where $a_{ii} \neq 0$.

Show that $A^{-1} = \begin{bmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & \cdots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_{nn}^{-1} \end{bmatrix}$.

Theorem 6.74 If A is any (n, n)-matrix, then

$$\det A \cdot \det \widetilde{A} = (\det A)^n$$

Proof Apply the Binet–Cauchy formula and Theorem 6.71. ■

Theorem 6.75 If A is a (n, n)-matrix, then $\det A \neq 0$, iff its rows (columns) are linearly independent.

We close this subsection with a few additional remarks.

First, observe that Theorem 6.72 allows the alternative definition of the $\mathrm{GL}(n, \Phi)$ group, namely,

$$\mathrm{GL}(n, \Phi) = \{A \in \boldsymbol{M_{nn}} | \ \det A \neq 0\}$$

Also, we define the **unimodular group** $UM(n, \Phi)$as

$$\mathrm{UM}(n, \Phi) = \{A \in \boldsymbol{M_{nn}} | \ \det A = \pm 1\}$$

and the **special linear group**

$$\mathrm{SL}(n, \Phi) = \{A \in \boldsymbol{M_{nn}} | \ \det A = 1\}$$

6.15 MORE ON LINEAR OPERATORS

We are now in position to study linear operators in even more detail. First, recall that, given two vector spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$ over a field Φ, we call a mapping

$$F : \boldsymbol{X} \rightarrow \boldsymbol{Y}$$

a linear transformation or a linear operator if for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ and all $\alpha, \beta \in \Phi$

$$F(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha F(\boldsymbol{x}) + \beta F(\boldsymbol{y}) \in \boldsymbol{Y}$$

Example 6.116 Suppose we are given a transformation $F : \mathbf{R}^n \rightarrow \mathbf{R}^m$ as a rule that assigns to every vector $\boldsymbol{x} \in \mathbf{R}^n$ a unique vector $F(\boldsymbol{x}) \in \mathbf{R}^m$, such that for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$ and all $\alpha, \beta \in \mathbf{R}$

$$F(\alpha \boldsymbol{x} + \beta \boldsymbol{y}) = \alpha F(\boldsymbol{x}) + \beta F(\boldsymbol{y}) \in \boldsymbol{R}^m$$

We say that F is a linear operator from space $\mathbf{R}^n$ to $\boldsymbol{R}^m$, and we often write this as $\mathbf{R}^n \xrightarrow{F} \boldsymbol{R}^m$. ■

Keeping in mind what we have learned about matrices so far, we ask: why not consider a mapping

$$A : \Phi^{\boldsymbol{n}} \rightarrow \Phi^{\boldsymbol{m}}$$

defined by $A\,\boldsymbol{x} = \boldsymbol{y}$, where $A \in \boldsymbol{M_{mn}}$ is an (m, n)-matrix, Φ is a given field, and $\Phi^{\boldsymbol{n}}$, $\Phi^{\boldsymbol{m}}$ are vector spaces with elements $\boldsymbol{x} = (\alpha_1, \ \ldots \ , \ \alpha_n)$ and $\boldsymbol{y} = (\beta_1, \ \ldots \ , \ \beta_m)$,

respectively. An important thing to note: although $\boldsymbol{x}$ and $\boldsymbol{y}$ are often written as n-tuples (m-tuples) we think of them as column vectors

$$\boldsymbol{x} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \in \Phi^n \quad \text{and} \quad \boldsymbol{y} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \in \Phi^m$$

Example 6.117 Let $A = \begin{bmatrix} 1 & -4 & 5 \\ 2 & 3 & -6 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix}$, then

$$A\boldsymbol{x} = \begin{bmatrix} 1 & -4 & 5 \\ 2 & 3 & -6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix}$$

$$= \begin{bmatrix} -36 \\ 41 \end{bmatrix} = \boldsymbol{y}$$

It is not difficult to convince yourself that the aforementioned mapping is a linear operator. Indeed, if A is any (m, n)-matrix, and $A : \Phi^n \to \Phi^m$ a mapping defined by $A\boldsymbol{x} = \boldsymbol{y}$, $\boldsymbol{x} \in \Phi^n$, $\boldsymbol{y} \in \Phi^m$, then

$$\begin{aligned} A(\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2) &= \lambda_1 A(\boldsymbol{x}_1) + \lambda_2 A(\boldsymbol{x}_2) \\ &= \lambda_1 \boldsymbol{y}_1 + \lambda_2 \boldsymbol{y}_2 \in \Phi^m \end{aligned}$$

■

We can also reverse the question and ask the following: Assuming that we know the result of an action of matrix A on some known vector $\boldsymbol{x}$, how can we find matrix A? The following examples illustrate this.

Example 6.118 Suppose we know that matrix A acts on any vector $\boldsymbol{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \in \mathbf{R}^2$ as follows:

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 2\alpha_1 + \alpha_2 \\ 3\alpha_2 \end{bmatrix}$$

How can we find matrix A?

Solution The simplest way to determine matrix A is to see its effect on the vectors of the standard basis of $\mathbf{R}^2$. In this case, we have

$$A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad \text{and} \quad A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

and thus,

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$$

It is easy to verify this:

$$A\boldsymbol{x} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 2\alpha_1 + \alpha_2 \\ 3\alpha_2 \end{bmatrix}$$

■

Example 6.119 Suppose we are interested in the matrix formulation of the transformation $D : \mathbf{R}^2 \to \mathbf{R}^2$ that maps every point in $\mathbf{R}^2$ into a point d times from the origin, that is,

$$D\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = d \begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{x}'$$

The matrix formulation of this equation is obviously

$$\begin{bmatrix} d & 0 \\ 0 & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = d \begin{bmatrix} x \\ y \end{bmatrix}$$

If $d > 1$, we call the transformation a *dilatation*. For $0 < d < 1$, we have a *contraction* (Figure 6.7). ■

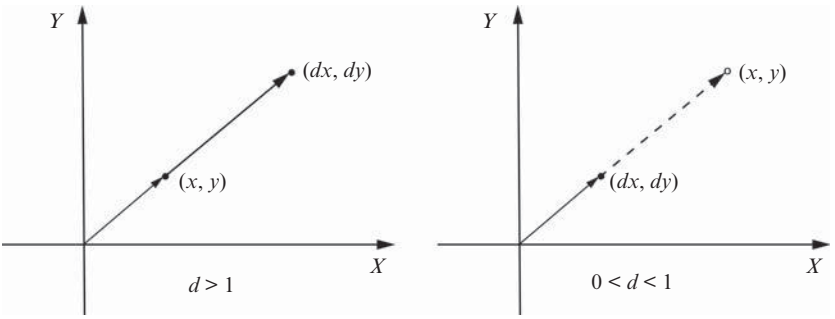


Figure 6.7

Similarly, a dilatation in $\mathbf{R}^3$ is described in

Example 6.120 If $D : \mathbf{R}^3 \to \mathbf{R}^3$ is defined by

$$D = \begin{bmatrix} d & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix},$$

then for every $\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{R}^3$,

$$D\boldsymbol{x} = \begin{bmatrix} d & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \boldsymbol{x}'$$

which looks something like this (Figure 6.8). ■

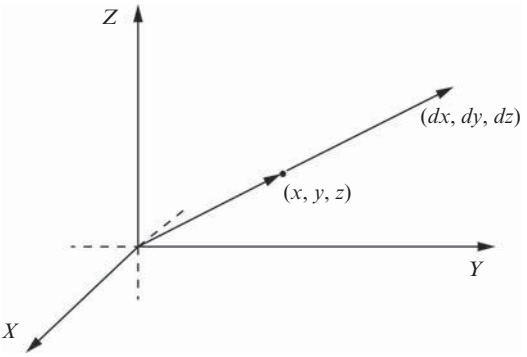


Figure 6.8

Example 6.121 Let $E : \mathbf{R}^2 \to \mathbf{R}^2$ be a transformation whose matrix is

$$E = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$$

so that its action on a vector $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2$ is defined by

$$E \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ex \\ y \end{bmatrix}$$

That is, every point $(x, y) \in \mathbf{R}^2$ is mapped to a point $(ex, y) \in \mathbf{R}^2$, where, if $e > 1$, we say that we have an *expansion* in the x-direction, and if $0 < e < 1$, we have a *contraction* (Figure 6.9). ■

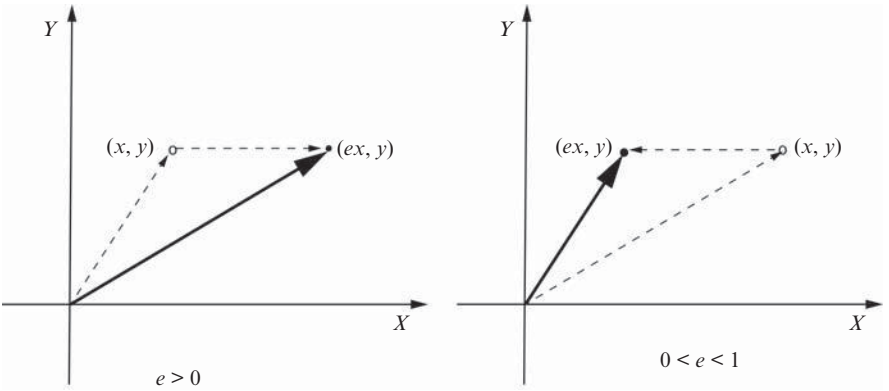


Figure 6.9

Similarly, we could have an expansion and a contraction only in the y-direction.

Example/Exercise 6.122 Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be a transformation whose matrix is $T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. Describe the action of matrix T on a vector $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbf{R}^2$.

Example 6.123 Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be a transformation defined by

$$T(\boldsymbol{x}) = T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix} = \boldsymbol{x}', \quad \forall \boldsymbol{x} \in \mathbf{R}^2$$

where its matrix representation is $T = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. We call this transformation a **reflection** about the X-axis (Figure 6.10). ■

The examples that follow address an important linear transformation: "**rotation**" in $\mathbf{R}^2$ and $\mathbf{R}^3$. For the sake of completeness, and in order to make use of matrices involving trigonometric functions less abrupt, we introduce the issue with[8]

Example 6.124 Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be a linear operator representing a counterclockwise rotation of a vector in $\mathbf{R}^2$ through an angle θ. We would like to find a matrix of T with respect to the usual basis

$$\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

[8]This example can be skipped in the first reading.

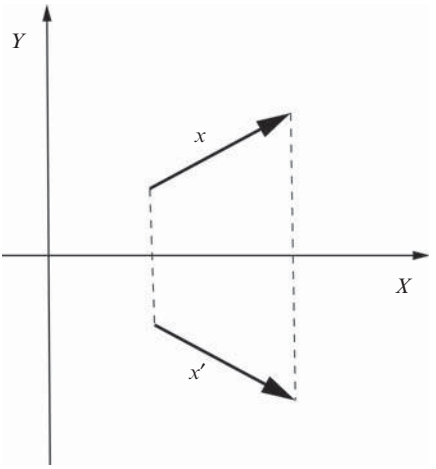


Figure 6.10

First, to make things more transparent, let's say that our vectors $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $T(\boldsymbol{x}) = \boldsymbol{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ look something like this (Figure 6.11):

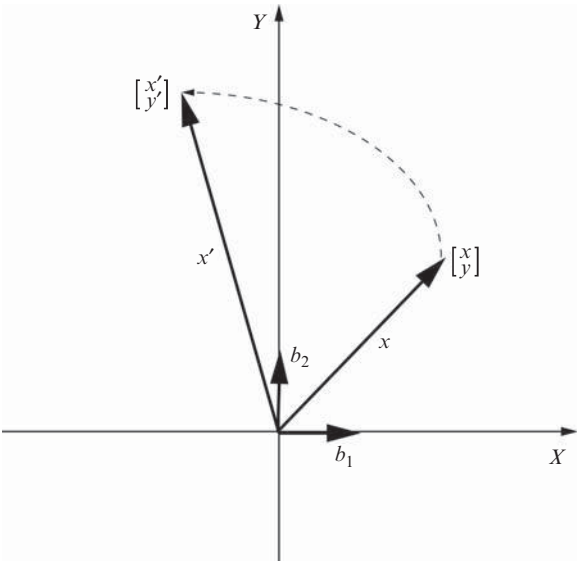


Figure 6.11

Since a linear transformation is specified by its effects on the basis vectors $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ whose length is equal to one, we can, without loss of

generality, simplify the picture and consider a rotation on the unit circle, namely (Figure 6.12).

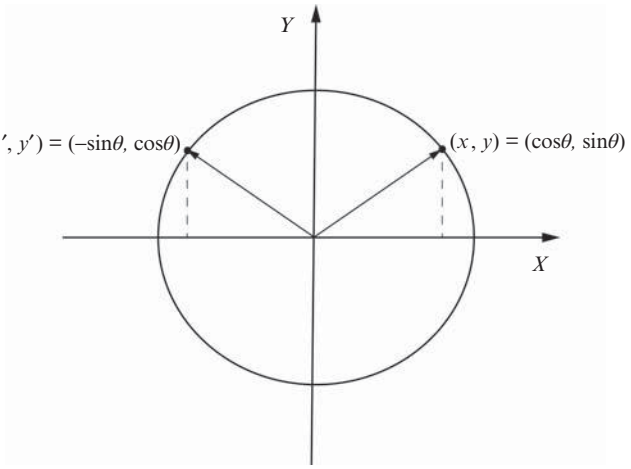


Figure 6.12

So,

$$T(\boldsymbol{b}_1) = T\begin{bmatrix}1\\0\end{bmatrix} = T(1,0) = (\cos\theta, \sin\theta)$$

$$T(\boldsymbol{b}_2) = T\begin{bmatrix}0\\1\end{bmatrix} = T(0,1) = \left(\cos\left(\theta + \frac{\pi}{2}\right), \sin\left(\theta + \frac{\pi}{2}\right)\right) = (-\sin\theta, \cos\theta)$$

Therefore, the matrix representation of T is

$$T = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$ ■

In general, a counterclockwise rotation of any vector from $\mathbf{R}^2$ is described in the following

Example 6.125 Let (x, y) be a point in $\mathbf{R}^2$, and let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be a transformation whose matrix is

$$T = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

Let's examine what happens to a point (x, y) (i.e., a vector $\boldsymbol{x} = \begin{bmatrix}x\\y\end{bmatrix}$) under the action of a transformation T (Figure 6.13).

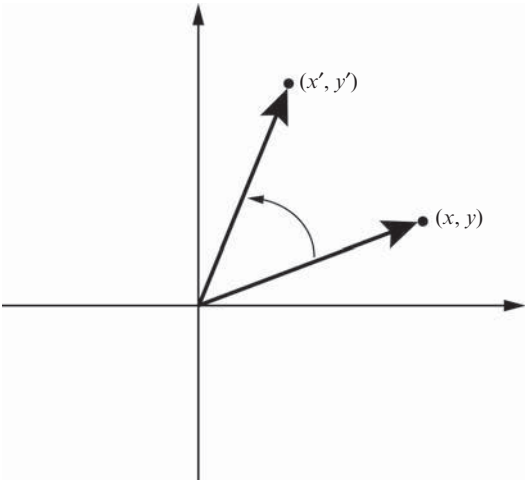


Figure 6.13

We have

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$
$$= \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} x' \\ y' \end{bmatrix} = \boldsymbol{x}'$$

So, a vector $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ is rotated through an angle θ to a vector $\boldsymbol{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, that is, the point (x, y) is mapped to the point (x', y'). ■

Example/Exercise 6.126 Find the image $T(\boldsymbol{x})$ under a rotation of $\frac{\pi}{3}$ about the origin if $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Of course, we can also "rotate" a vector in $\mathbf{R}^3$.

Example 6.127 Let $T_Z : \mathbf{R}^3 \to \mathbf{R}^3$ be a transformation defined by the matrix

$$T_Z = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and let $\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ be some vector in $\mathbf{R}^3$.

Consider

$$T_Z\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}$$
$$=\begin{bmatrix}x\cos\theta - y\sin\theta\\ y\sin\theta + y\cos\theta\\ z\end{bmatrix}=\begin{bmatrix}x'\\y'\\z'\end{bmatrix}=\boldsymbol{x}'$$

That looks something like this (Figure 6.14).

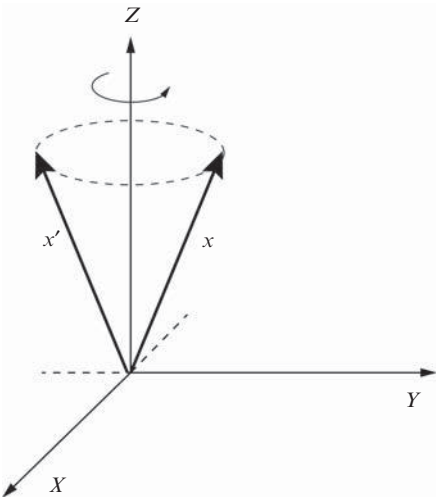


Figure 6.14

That is, a counterclockwise rotation about the positive Z-axis through an angle θ.Analogously, we can "rotate" about the positive X-axis or positive Y-axis with matrices$T_X=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 1 & \sin\theta & \cos\theta\end{bmatrix}$ and $T_Y=\begin{bmatrix}\cos & 0 & \sin\theta\\ 0 & 1 & 0\\ \sin & 0 & \cos\theta\end{bmatrix}$ respectively. ■

Of course, the transformations don't have to be $\mathbf{R}^n \to \mathbf{R}^n$. We can as well have a transformation $\mathbf{R}^n \to \mathbf{R}^m$.

Example 6.128 Let $F = [1 \quad -2 \quad 1]$ be a matrix representing some linear transformation $\mathbf{R}^3 \to \mathbf{R}$, and let

$$\boldsymbol{x}=\begin{bmatrix}6\\-4\\9\end{bmatrix}\in\mathbf{R}^3$$

Then,

$$F(\boldsymbol{x}) = [1 \; -2 \; \; 1] \begin{bmatrix} 6 \\ -4 \\ 9 \end{bmatrix} = 23$$

that is, matrix F maps a vector from $\mathbf{R}^3$ to a point in $\mathbf{R}$. ■

Example 6.129 Let

$$F = \begin{bmatrix} 5 & 3 & -2 \\ 0 & 4 & -1 \end{bmatrix}$$

be a matrix representing a transformation $F : \mathbf{R}^3 \to \mathbf{R}^2$, so that for every $\boldsymbol{x} = \begin{bmatrix} x \\ y \\ x \end{bmatrix} \in \mathbf{R}^3$,

$$F(\boldsymbol{x}) = \begin{bmatrix} 5 & 3 & -2 \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ x \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} = \boldsymbol{x}' \in \mathbf{R}^2$$

For instance, if $\boldsymbol{x} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \in \mathbf{R}^3$, then,

$$F(\boldsymbol{x}) = \begin{bmatrix} 5 & 3 & -2 \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \boldsymbol{x}' \in \mathbf{R}^2$$

■

As you might have anticipated by now, every matrix in fact defines a certain transformation. However, first, one wants to make sure that the transformations are well defined. Also, there is a finesse to be noted: Although a matrix *defines* (*represents*) a transformation (linear operator), *it is not* the same as an operator. So, to be more precise, one should distinguish between an **operator** $\widehat{T}$ (say, T with a hat) and its **corresponding matrix** T. However, in order to avoid clutter, we are often lazy and use the same symbols for operators and their corresponding matrices.

Now we can move one step further and consider more general situations.

Let X be any n-dimensional vector space with one of its ordered basis

$$\boldsymbol{B} = \{\boldsymbol{b}_1, \ \ldots \ , \ \boldsymbol{b}_n\}$$

We call $\boldsymbol{B}$ the **coordinate basis** or **coordinate system** for the space $\boldsymbol{X}$. Then, as we know, any vector $\boldsymbol{x} \in \boldsymbol{X}$ can be uniquely expressed as a linear combination of

vectors from $\boldsymbol{B}$, that is,

$$\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{b}_i$$

We say that $(\alpha_1, \ldots, \alpha_n)$, or simply (α_i), are the **coordinates** of vector $\boldsymbol{x}$ in base $\boldsymbol{B}$. Similarly, we say that the matrix

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = [\alpha_i]$$
$$= (\alpha_1, \ldots, \alpha_n)^\tau$$

is a **coordinate matrix** of a vector $\boldsymbol{x}$ in basis $\boldsymbol{B}$.[9] Furthermore, let

$$F : X \to \Phi^n$$

be an operator defined by $F(\boldsymbol{x}) = (\alpha_i)$, and

$$G : \boldsymbol{X} \to \boldsymbol{M}_{n1}$$

be an operator defined by $g(\boldsymbol{x}) = [\alpha_i]$. It is evident that F and G are isomorphisms of the respective linear spaces, and we establish the identification

$$\boldsymbol{x} = F(\boldsymbol{x}) = G(\boldsymbol{x})$$

by writing

$$\boldsymbol{x} = (\alpha_i) \quad \text{or} \quad \boldsymbol{x} = [\alpha_i]$$

that is,

$$\boldsymbol{x} = (\alpha_1, \alpha_2, \ldots, \alpha_n) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

Hence,

Theorem 6.76 Let $\boldsymbol{X}$ be an n-dimensional vector space over a field Φ. Then, $\boldsymbol{X}$ and Φ^n are isomorphic.

Theorem 6.77 Let $A = (\alpha_{ij})$ and $B = (\beta_{ij})$ be two (m, n)-matrices, and let $X \in \boldsymbol{M}_{n1}$ be any column matrix. If

$$AX = BX$$

[9]One should be careful not to be confused by, admittedly, somewhat inconsistent notation for 1-column (row) matrices.

Then

$$A = B$$

Proof We show this by checking the following: Take a special column, say, C_k, such that for every $k = 1, \ldots, n \quad C_k = [\delta_{ik}]$. Then,

$$AC_k = BC_k$$

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{k1} & \alpha_{k2} & \cdots & \alpha_{kn} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{kn} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{kk} \\ \vdots \\ \alpha_{mn} \end{bmatrix} = \begin{bmatrix} \beta_{1k} \\ \vdots \\ \beta_{kk} \\ \vdots \\ \beta_{mn} \end{bmatrix}$$

that is, $\alpha_{ik} = \beta_{ik}, \ \forall k = 1, \ldots, n$, and thus $A = B$. ■

Example 6.130 Consider a set of vectors $V = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\} \subseteq \boldsymbol{X}$ from a given space $\boldsymbol{X}$. From our earlier discussion, we know that we can associate with every vector $\boldsymbol{x}_i, \ i = 1, \ldots, r$, a coordinate matrix, a column,

$$\boldsymbol{x}_i = \begin{bmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{ni} \end{bmatrix}$$

representing a vector $\boldsymbol{x}_i$ in some basis $\boldsymbol{B}$. Obviously, we can do that for any $\boldsymbol{x}_i \in V$, and thus we associate with set V a matrix

$$M(V) = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nr} \end{bmatrix}$$

whose columns are evidently the coordinates of vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r$. ■

This example reveals another good feature of coordinate matrices. Namely, we naturally expect that the maximal number of linearly independent vectors from set V is equal to the rank of matrix $M(V)$. Also, if $r = \dim \boldsymbol{X}$ and $\det M(V) \neq 0$, then V is a basis of $\boldsymbol{X}$.

We reason as follows: Since every vector in space $\boldsymbol{X}$ is uniquely expressible as a linear combination of basis vectors, we expect that the action of F on basis vectors $\boldsymbol{B}$ would provide adequate information on the transformation of any vector in general with respect to that basis. Hence, we have

Theorem 6.78 Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be vector spaces over a field Φ, and let $\boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n\}$ be some basis of $\boldsymbol{X}$. If $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n \in \boldsymbol{Y}$, then there exists a unique linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ such that $F(\boldsymbol{b}_i) = \boldsymbol{y}_i, \ i = 1, 2, \ldots, n$.

Proof Since $\boldsymbol{B}$ is a basis of $\boldsymbol{X}$, any vector $\boldsymbol{x} \in \boldsymbol{X}$ can be uniquely expressed as

$$\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2 + \cdots + \alpha_n \boldsymbol{b}_n, \qquad \alpha_i \in \Phi$$

So, we define $F : \boldsymbol{X} \to \boldsymbol{Y}$ by

$$F(\boldsymbol{x}) = \alpha_1 \boldsymbol{y}_1 + \alpha_2 \boldsymbol{y}_2 + \cdots + \alpha_n \boldsymbol{y}_n$$

Now, consider

$$\begin{aligned} F(\boldsymbol{b}_i) &= F(0 \cdot \boldsymbol{b}_1 + 0 \cdot \boldsymbol{b}_2 + \cdots + 1 \cdot \boldsymbol{b}_i + 0 \cdot \boldsymbol{b}_{i+1} + \cdots + 0 \cdot \boldsymbol{b}_{n1}) \\ &= \boldsymbol{y}_i \end{aligned}$$

Thus, F is well defined.

Next, take any two vectors from $\boldsymbol{X}$:

$$\boldsymbol{x}_1 = \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2 + \cdots + \alpha_n \boldsymbol{b}_n, \quad \boldsymbol{x}_2 = \beta_1 \boldsymbol{b}_1 + \beta_2 \boldsymbol{b}_2 + \cdots + \beta_n \boldsymbol{b}_n, \quad \alpha_i, \beta_i \in \Phi$$

Since X is a vector space we certainly have

$$\begin{aligned} \boldsymbol{x}_1 + \boldsymbol{x}_2 &= (\alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2 + \cdots + \alpha_n \boldsymbol{b}_n) + (\beta_1 \boldsymbol{b}_1 + \beta_2 \boldsymbol{b}_2 + \cdots + \beta_n \boldsymbol{b}_n) \\ &= (\alpha_1 + \beta_1)\boldsymbol{b}_1 + (\alpha_2 + \beta_2)\boldsymbol{b}_2 + \cdots + (\alpha_n + \beta_n)\boldsymbol{b}_n \\ &= \gamma_1 \boldsymbol{b}_1 + \gamma_2 \boldsymbol{b}_2 + \cdots + \gamma_n \boldsymbol{b}_n \end{aligned}$$

where $\gamma_i = \alpha_i + \beta_i$.

Now, with $\lambda_1, \lambda_2 \in \Phi$, consider

$$\begin{aligned} &F(\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2) \\ &\quad = F((\lambda_1(\alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2 + \cdots + \alpha_n \boldsymbol{b}_n) + \lambda_2(\beta_1 \boldsymbol{b}_1 + \beta_2 \boldsymbol{b}_2 + \cdots + \beta_n \boldsymbol{b}_n)) \\ &\quad = F((\lambda_1 \alpha_1)\boldsymbol{b}_1 + \cdots + (\lambda_1 \alpha_n)\boldsymbol{b}_n + (\lambda_2 \beta_1)\boldsymbol{b}_1 + \cdots + (\lambda_2 \beta_n)\boldsymbol{b}_n) \\ &\quad = F((\lambda_1 \alpha_1 + \lambda_2 \beta_1)\boldsymbol{b}_1 + (\lambda_1 \alpha_2 + \lambda_2 \beta_2)\boldsymbol{b}_2 + \cdots + (\lambda_1 \alpha_n + \lambda_2 \beta_n)\boldsymbol{b}_n) \end{aligned}$$

$$
\begin{aligned}
&= (\lambda_1\alpha_1 + \lambda_2\beta_1)\boldsymbol{y}_1 + (\lambda_1\alpha_2 + \lambda_2\beta_2)\boldsymbol{y}_2 + \cdots + (\lambda_1\alpha_n + \lambda_2\beta_n)\boldsymbol{y}_2 \\
&= \lambda_1 F(\alpha_1\boldsymbol{b}_1) + \cdots + \lambda_1 F(\alpha_n\boldsymbol{b}_n) + \lambda_2 F(\beta_1\boldsymbol{b}_1) + \cdots + \lambda_2 F(\beta_n\boldsymbol{b}_n) \\
&= \lambda_1 F(\alpha_1\boldsymbol{b}_1 + \cdots + \alpha_n\boldsymbol{b}_n) + \lambda_2 F(\beta_1\boldsymbol{b}_1 + \cdots + \beta_n\boldsymbol{b}_n) \\
&= \lambda_1 F(\boldsymbol{x}_1) + \lambda_2 F(\boldsymbol{x}_2)
\end{aligned}
$$

Thus, F is a linear operator.

Finally, to show uniqueness, suppose there exists another linear operator $G : \boldsymbol{X} \to \boldsymbol{Y}$ such that $G(\boldsymbol{b}_i) = \boldsymbol{y}_i$. Then,

$$
\begin{aligned}
G(\boldsymbol{x}) &= G(\alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2 + \cdots + \alpha_n\boldsymbol{b}_n) \\
&= \alpha_1 G(\boldsymbol{b}_1) + \alpha_2 G(\boldsymbol{b}_2) + \cdots + \alpha_n G(\boldsymbol{b}_n) \\
&= \alpha_1\boldsymbol{y}_1 + \alpha_2\boldsymbol{y}_2 + \cdots + \alpha_n\boldsymbol{y}_n \\
&= \alpha_1 F(\boldsymbol{b}_1) + \alpha_2 F(\boldsymbol{b}_2) + \cdots + \alpha_n F(\boldsymbol{b}_n) \\
&= F(\alpha_1\boldsymbol{b}_1 + \alpha_2\boldsymbol{b}_2 + \cdots + \alpha_n\boldsymbol{b}_n) \\
&= F(\boldsymbol{x})
\end{aligned}
$$

Thus, $G = F$ indeed. ■

Let's illustrate all of this with a few examples.

Example 6.131 Consider a vector space $\mathbf{R}^n$ and its **standard** (**canonical**) basis $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, where

$$
\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \boldsymbol{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}
$$

Furthermore, let $\boldsymbol{x} \in \mathbf{R}^n$ be any vector. Then,

$$
\begin{aligned}
\boldsymbol{x} &= \alpha_1\boldsymbol{e}_1 + \cdots + \alpha_n\boldsymbol{e}_n \\
&= \begin{bmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \alpha_n \end{bmatrix} \\
&= \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}
\end{aligned}
$$

We obtained a coordinate representation (coordinate matrix) of our vector $\boldsymbol{x}$.

Now, suppose we are interested in the action of a linear operator

$$F : \mathbf{R}^n \to \mathbf{R}^n$$

on the vector $\boldsymbol{x}$, that is,

$$\begin{aligned} F(\boldsymbol{x}) &= F(\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n) \\ &= \alpha_1 F(\boldsymbol{e}_1) + \cdots + \alpha_n F(\boldsymbol{e}_n) \end{aligned}$$

which we can write

$$= [F(\boldsymbol{e}_1) \cdots F(\boldsymbol{e}_n)] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

or

$$= \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}^T \begin{bmatrix} F\left(\boldsymbol{e}_1\right) \\ \vdots \\ F(\boldsymbol{e}_n) \end{bmatrix} = [\alpha_1 \ \ldots \ \alpha_n] \begin{bmatrix} F\left(\boldsymbol{e}_1\right) \\ \vdots \\ F(\boldsymbol{e}_n) \end{bmatrix}$$

On the other hand, every $F(\boldsymbol{e}_i)$ is a vector in $\mathbf{R}^n$ again, so it should have its representation in the same basis, say,

$$\begin{aligned} F(\boldsymbol{e}_1) &= \alpha_{11} \boldsymbol{e}_1 + \cdots + \alpha_{1n} \boldsymbol{e}_n \\ F(\boldsymbol{e}_2) &= \alpha_{21} \boldsymbol{e}_1 + \cdots + \alpha_{2n} \boldsymbol{e}_n \\ &\vdots \\ F(\boldsymbol{e}_n) &= \alpha_{n1} \boldsymbol{e}_1 + \cdots + \alpha_{nn} \boldsymbol{e}_n \end{aligned}$$

We can obviously write this as

$$\begin{bmatrix} F\left(\boldsymbol{e}_1\right) \\ \vdots \\ F(\boldsymbol{e}_n) \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_n \end{bmatrix}$$

So, we associated with the operator F a matrix $\begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}$

And we write

$$[F]_E = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}$$

as a matrix representation of the operator F with respect to the E basis. ■

Example 6.132 Consider a vector space $\mathbf{R}^2$ and one of its bases

$$\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \ \boldsymbol{b}_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right\}$$

Furthermore, let $F : \mathbf{R}^2 \to \mathbf{R}^2$ be a linear operator whose matrix representation with respect to some other basis is

$$F = \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix}$$

Now, let's see how it acts on $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ from the $\boldsymbol{B}$ basis:

$$\begin{aligned} F(\boldsymbol{b}_1) &= \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ -6 \end{bmatrix} \\ &= \beta_{11}\boldsymbol{b}_1 + \beta_{12}\boldsymbol{b}_2 \\ &= \beta_{11} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta_{12} \begin{bmatrix} 2 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} \beta_{11} + 2\beta_{12} \\ 2\beta_{11} + 5\beta_{12} \end{bmatrix} \end{aligned}$$

We have

$$\begin{aligned} -1 &= \beta_{11} + 2\beta_{12} \\ -6 &= 2\beta_{11} + 5\beta_{12} \end{aligned}$$

Hence, $\beta_{11} = 7$ and $\beta_{12} = -4$. Therefore,

$$F(\boldsymbol{b}_1) = 7\boldsymbol{b}_1 - 4\boldsymbol{b}_2$$

Similarly,

$$\begin{aligned} F(\boldsymbol{b}_2) &= \begin{bmatrix} 3 & -2 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} -4 \\ -7 \end{bmatrix} \end{aligned}$$

$$= \beta_{21}\boldsymbol{b}_1 + \beta_{22}\boldsymbol{b}_2$$
$$= \beta_{21}\begin{bmatrix}1\\2\end{bmatrix} + \beta_{22}\begin{bmatrix}2\\5\end{bmatrix}$$
$$= \begin{bmatrix}\beta_{21} + 2\beta_{22}\\2\beta_{21} + 5\beta_{22}\end{bmatrix}$$

and so

$$-4 = \beta_{21} + 2\beta_{22}$$

$$-7 = 2\beta_{21} + 5\beta_{22}$$

from which we obtain $\beta_{21} = -6$ and $\beta_{22} = 1$. Thus,

$$F(\boldsymbol{b}_2) = -6\boldsymbol{b}_1 + \boldsymbol{b}_2$$

Consequently, the matrix representation of our operator with respect to the $\boldsymbol{B}$ basis is

$$F_{\boldsymbol{B}} = [\beta_{ij}]^\tau = \begin{bmatrix}7 & -4\\-6 & 1\end{bmatrix}^\tau = \begin{bmatrix}7 & -6\\-4 & 1\end{bmatrix} \quad \blacksquare$$

Example 6.133 Suppose we ask the question differently: Let $F : \mathbf{R}^3 \to \mathbf{R}^3$ be a linear operator defined by

$$F(\boldsymbol{x}) = F\begin{bmatrix}x\\y\\z\end{bmatrix} = \boldsymbol{x}' = \begin{bmatrix}2y + z\\x - 4y\\3x\end{bmatrix}, \quad \forall \boldsymbol{x} \in \mathbf{R}^3$$

What would be a matrix representation of such an operator with respect to the following basis:

$$\boldsymbol{B} = \left\{\boldsymbol{b}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \boldsymbol{b}_2 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \quad \boldsymbol{b}_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}\right\} \ ?$$

Well, since for any $\boldsymbol{x} \in \mathbf{R}^3$

$$F\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}2y + z\\x - 4y\\3x\end{bmatrix}$$

then, in particular,

$$
\begin{aligned}
F(\boldsymbol{b}_1) &= F\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}0\\1\\3\end{bmatrix} \\
&= -1\begin{bmatrix}1\\0\\0\end{bmatrix} - 2\begin{bmatrix}1\\1\\0\end{bmatrix} + 3\begin{bmatrix}1\\1\\1\end{bmatrix} = -1\boldsymbol{b}_1 - 2\boldsymbol{b}_2 + 3\boldsymbol{b}_3 \\
F(\boldsymbol{b}_2) &= F\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}2\\-3\\3\end{bmatrix} \\
&= 5\begin{bmatrix}1\\0\\0\end{bmatrix} - 6\begin{bmatrix}1\\1\\0\end{bmatrix} + 3\begin{bmatrix}1\\1\\1\end{bmatrix} = 5\boldsymbol{b}_1 - 6\boldsymbol{b}_2 + 3\boldsymbol{b}_3 \\
F(\boldsymbol{b}_3) &= F\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}3\\-3\\3\end{bmatrix} \\
&= 6\begin{bmatrix}1\\0\\0\end{bmatrix} - 6\begin{bmatrix}1\\1\\0\end{bmatrix} + 3\begin{bmatrix}1\\1\\1\end{bmatrix} = 6\boldsymbol{b}_1 - 6\boldsymbol{b}_2 + 3\boldsymbol{b}_3
\end{aligned}
$$

Thus, our matrix representing the linear operator F is

$$
F = \begin{bmatrix}-1 & 5 & 6\\ -2 & -6 & -6\\ 3 & 3 & 3\end{bmatrix}
$$

■

We will do another example, with some extra twists and turns, to make the procedures more convincing.

Example 6.134 Consider again a linear operator $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by

$$
F(\boldsymbol{x}) = F\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}2x+3y\\4x-5y\end{bmatrix}, \quad \forall \boldsymbol{x} = \begin{bmatrix}x\\y\end{bmatrix} \in \mathbf{R}^2 \tag{6.17}
$$

We would like to find a matrix representation of an operator F with respect to the following basis:

$$
\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix}1\\-2\end{bmatrix},\ \boldsymbol{b}_2 = \begin{bmatrix}2\\-5\end{bmatrix} \right\}
$$

First, let's see what the F operator does to the basis vectors $\boldsymbol{b}_1,\ \boldsymbol{b}_2$. According to (6.17)

$$F(\boldsymbol{b}_1) = F\begin{bmatrix}1\\-2\end{bmatrix} = \begin{bmatrix}-4\\14\end{bmatrix} = \boldsymbol{b}'_1 \tag{6.18}$$

$$F(\boldsymbol{b}_2) = F\begin{bmatrix}2\\-5\end{bmatrix} = \begin{bmatrix}-11\\33\end{bmatrix} = \boldsymbol{b}'_2 \tag{6.19}$$

Vectors $\boldsymbol{b}'_1 = F(\boldsymbol{b}_1)$ and $\boldsymbol{b}'_2 = F(\boldsymbol{b}_2)$, as any vector from $\mathbf{R}^2$, should be expressible via basis vectors of $\boldsymbol{B}$. Thus,

$$\begin{aligned}\boldsymbol{b}'_1 &= \begin{bmatrix}-4\\14\end{bmatrix} = \alpha_{11}\,\boldsymbol{b}_1 + \alpha_{12}\boldsymbol{b}_2\\ &= 8\begin{bmatrix}1\\-2\end{bmatrix} - 6\begin{bmatrix}2\\-5\end{bmatrix}\end{aligned} \tag{6.20}$$

$$\begin{aligned}\boldsymbol{b}'_2 &= \begin{bmatrix}-11\\33\end{bmatrix} = \alpha_{21}\,\boldsymbol{b}_1 + \alpha_{22}\boldsymbol{b}_2\\ &= 11\begin{bmatrix}1\\-2\end{bmatrix} - 11\begin{bmatrix}2\\-5\end{bmatrix}\end{aligned} \tag{6.21}$$

But, remember, $\alpha_{11},\ \alpha_{12}$ are the coordinates of $\boldsymbol{b}'_1$, and α_{21}, α_{22} of $\boldsymbol{b}'_2$.

Writing them as columns, we obtain the desired matrix of the operator F, that is,

$$F = \begin{bmatrix}8 & 11\\-6 & -11\end{bmatrix}$$

Let's look at all of this from yet another angle. Suppose we take any vector $\boldsymbol{x} = \begin{bmatrix}x\\y\end{bmatrix}$ from our space and write it as a linear combination of basis vectors $\boldsymbol{B}$,

$$\begin{aligned}\boldsymbol{x} = \begin{bmatrix}x\\y\end{bmatrix} &= \alpha\,\boldsymbol{b}_1 + \beta\boldsymbol{b}_2\\ &= \alpha\begin{bmatrix}1\\-2\end{bmatrix} + \beta\begin{bmatrix}2\\-5\end{bmatrix}\\ &= \begin{bmatrix}\alpha + 2\beta\\-2\alpha - 5\beta\end{bmatrix}\end{aligned} \tag{6.22}$$

We have

$$\begin{aligned}x &= \alpha + 2\beta\\ y &= -2\alpha - 5\beta\end{aligned}$$

Solving this for α and β, we get

$$\alpha = 5x + 2y \quad \text{and} \quad \beta = -2x - y \tag{6.23}$$

Thus (substituting (6.23) into (6.22)), *any vector* $\boldsymbol{x}$ can be written in the $\boldsymbol{B}$ basis as

$$\begin{aligned} \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} &= (5x + 2y)\,\boldsymbol{b}_1 + (-2x - y)\boldsymbol{b}_2 \\ &= (5x + 2y) \begin{bmatrix} 1 \\ -2 \end{bmatrix} + (-2x - y) \begin{bmatrix} 2 \\ -5 \end{bmatrix} \end{aligned} \tag{6.24}$$

Now, let's take vector $\boldsymbol{b}'_1 = \begin{bmatrix} -4 \\ 14 \end{bmatrix}$, since we already know its linear expansion (6.20), as our arbitrary vector, and see if we are going to get the same result

$$\begin{aligned} \begin{bmatrix} -4 \\ 14 \end{bmatrix} &\stackrel{?}{=} (\,5 \cdot (-4) + 2 \cdot 14\,) \begin{bmatrix} 1 \\ -2 \end{bmatrix} + (-2 \cdot (-4) - 14\,) \begin{bmatrix} 2 \\ -5 \end{bmatrix} \\ &= 8 \begin{bmatrix} 1 \\ -2 \end{bmatrix} - 6 \begin{bmatrix} 2 \\ -5 \end{bmatrix} \end{aligned}$$

As expected, this is indeed expression (6.20). ■

With these examples, we touched upon a very important issue of the transition from one (old) basis, say,

$$\boldsymbol{B} = \{\boldsymbol{b}_1, \ \ldots \ , \boldsymbol{b}_n\}$$

to another (new) basis

$$\boldsymbol{B}' = \{\boldsymbol{b}'_1, \ \ldots \ , \boldsymbol{b}'_n\}$$

Let's put it differently. If $\boldsymbol{B}$ and $\boldsymbol{B}'$ are two bases of a vector space $\boldsymbol{X}$, then any vector $\boldsymbol{x} \in \boldsymbol{X}$ has its coordinate matrix in the basis $\boldsymbol{B}$

$$\boldsymbol{x}_{\boldsymbol{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = [\alpha_i]$$

The same vector has its coordinate matrix in the $\boldsymbol{B}'$ basis

$$\boldsymbol{x}_{\boldsymbol{B}'} = \begin{bmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{bmatrix} = [\alpha'_i]$$

We would like to know how the coordinates of vector $\boldsymbol{x}$ change in transition from one basis (coordinate system) to another. Evidently, every $\boldsymbol{b}'_k$ from $\boldsymbol{B}'$ has its depiction in $\boldsymbol{B}$ (and vice versa), say,

$$\boldsymbol{b}'_k = \begin{bmatrix} \beta_{1k} \\ \vdots \\ \beta_{nk} \end{bmatrix}, \quad k = 1, \ldots, n$$

so our **transition matrix (change-of-basis matrix)** describing $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$ has these vectors as columns, that is,

$$T = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{bmatrix}$$

In other words,

$$\begin{aligned} \boldsymbol{b}'_1 &= \beta_{11}\boldsymbol{b}_1 + \beta_{12}\boldsymbol{b}_2 + \cdots + \beta_{1n}\boldsymbol{b}_n \\ \boldsymbol{b}'_2 &= \beta_{21}\boldsymbol{b}_1 + \beta_{22}\boldsymbol{b}_2 + \cdots + \beta_{2n}\boldsymbol{b}_n \\ &\vdots \\ \boldsymbol{b}'_n &= \beta_{n1}\boldsymbol{b}_1 + \beta_{n2}\boldsymbol{b}_2 + \cdots + \beta_{nn}\boldsymbol{b}_n \end{aligned}$$

Here are a few simple examples.

Example 6.135 Let $\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and $\boldsymbol{B}' = \left\{ \boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \boldsymbol{b}'_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$ be two bases of $\mathbf{R}^2$. Find the transition matrix $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$.

Solution First, let's express the vectors of $\boldsymbol{B}$ in terms of the vectors of $\boldsymbol{B}'$.

$$\begin{aligned} \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} &= \alpha_{11}\, \boldsymbol{b}'_1 + \alpha_{12}\, \boldsymbol{b}'_2 \\ &= \alpha_{11} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha_{12} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \\ &= -1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} \end{aligned}$$

$$\begin{aligned}\boldsymbol{b}_2 = \begin{bmatrix}0\\1\end{bmatrix} &= \alpha_{21}\,\boldsymbol{b}'_1 + \alpha_{22}\,\boldsymbol{b}'_2\\ &= \alpha_{21}\begin{bmatrix}1\\1\end{bmatrix} + \alpha_{22}\begin{bmatrix}2\\1\end{bmatrix}\\ &= 2\cdot\begin{bmatrix}1\\1\end{bmatrix} - 1\cdot\begin{bmatrix}2\\1\end{bmatrix}\end{aligned}$$

Thus, the transition matrix is

$$T = \begin{bmatrix}-1 & 2\\ 1 & -1\end{bmatrix}$$

It is important to observe that the question of what is the "old" and what is the "new" basis is arbitrary. We could, as equally, have asked how to express vectors from $\boldsymbol{B}'$ in terms of vectors from $\boldsymbol{B}$. Let's do that:

$$\begin{aligned}\boldsymbol{b}'_1 = \begin{bmatrix}1\\1\end{bmatrix} &= \beta_{11}\,\boldsymbol{b}_1 + \beta_{12}\,\boldsymbol{b}_2\\ &= \beta_{11}\begin{bmatrix}1\\0\end{bmatrix} + \beta_{12}\begin{bmatrix}0\\1\end{bmatrix}\\ &= 1\cdot\begin{bmatrix}1\\0\end{bmatrix} + 1\cdot\begin{bmatrix}0\\1\end{bmatrix}\\ \boldsymbol{b}'_2 = \begin{bmatrix}2\\1\end{bmatrix} &= \beta_{21}\,\boldsymbol{b}_1 + \beta_{22}\,\boldsymbol{b}_2\\ &= \beta_{21}\begin{bmatrix}1\\0\end{bmatrix} + \beta_{22}\begin{bmatrix}0\\1\end{bmatrix}\\ &= 2\cdot\begin{bmatrix}1\\0\end{bmatrix} + 1\cdot\begin{bmatrix}0\\1\end{bmatrix}\end{aligned}$$

Hence, in this case, the transition matrix is $\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}$. You can easily convince yourself that

$$\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix} = \begin{bmatrix}-1 & 2\\ 1 & -1\end{bmatrix}^{-1} = T^{-1}$$

■

Once again, the reader should be well advised not to be misled by this, somewhat inconsistent, terminology regarding the "old" $\boldsymbol{B}$, and the "new" $\boldsymbol{B}'$ bases. Namely, the aforementioned example points to the fact that T^{-1} transforms $\boldsymbol{B}$ to $\boldsymbol{B}'$ and T transforms $\boldsymbol{B}'$ to $\boldsymbol{B}$.

Example 6.136 Let $\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \boldsymbol{b}_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \right\}$ and $\boldsymbol{B}' = \left\{ \boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},\ \boldsymbol{b}'_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$ be two bases of $\mathbf{R}^2$. Find the transition matrix $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$.

Solution

$$\boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \beta_{11}\,\boldsymbol{b}_1 + \beta_{12}\,\boldsymbol{b}_2 = \beta_{11} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta_{12} \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

$$\boldsymbol{b}'_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \beta_{21}\,\boldsymbol{b}_1 + \beta_{22}\,\boldsymbol{b}_2 = \beta_{21} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \beta_{22} \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$

Solving for β_{ij} we get: $\beta_{11} = -8,\ \ \beta_{12} = 3,\ \ \beta_{21} = -11,\ \ \beta_{22} = 4$.

Thus,

$$\boldsymbol{b}'_1 = -8\,\boldsymbol{b}_1 + 3\,\boldsymbol{b}_2$$
$$\boldsymbol{b}'_2 = -11\,\boldsymbol{b}_1 + 4\,\boldsymbol{b}_2$$

So, the transition matrix is

$$T = \begin{bmatrix} -8 & -11 \\ 3 & 4 \end{bmatrix}$$

.Once again, observe that we could have done the following:

$$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \alpha_{11}\,\boldsymbol{b}'_1 + \alpha_{12}\,\boldsymbol{b}'_2 = \alpha_{11} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \alpha_{12} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\boldsymbol{b}_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \alpha_{21}\,\boldsymbol{b}'_1 + \alpha_{22}\,\boldsymbol{b}'_2 = \alpha_{21} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \alpha_{22} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

Solving for α_{ij} we get: $\alpha_{11} = 4,\ \ \alpha_{12} = -3,\ \ \alpha_{21} = 11,\ \ \alpha_{22} = -8$. Hence,

$$\boldsymbol{b}_1 = 4\boldsymbol{b}'_1 - 3\boldsymbol{b}'_2$$
$$\boldsymbol{b}_2 = 11\boldsymbol{b}'_1 - 8\boldsymbol{b}'_2$$

So, the transition matrix is $\begin{bmatrix} 4 & 11 \\ -3 & -8 \end{bmatrix}$. Again, note that

$$\begin{bmatrix} 4 & 11 \\ -3 & -8 \end{bmatrix} = \begin{bmatrix} -8 & -11 \\ 3 & 4 \end{bmatrix}^{-1} = T^{-1}$$

■

Succinctly put, the previous examples illustrate the following "rule": *The columns of the transitional matrix from one ("old") basis to another ("new") basis are the coordinate vectors of the "old" basis relative to the "new" basis.*

The previous examples entice us to ask the next natural question: how does an arbitrary vector $\boldsymbol{x}_{\boldsymbol{B}}$ expressed in one basis look when expressed in terms of the other? Consider the following

Example 6.137 Let $\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\}$ and $\boldsymbol{B}' = \left\{ \boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{b}'_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ be two bases of $\mathbf{R}^2$. Suppose that vector $\boldsymbol{x}$ in the $\boldsymbol{B}$ basis is given by $\boldsymbol{x}_{\boldsymbol{B}} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. We would like to find $\boldsymbol{x}_{\boldsymbol{B}'}$.

To get the transition matrix, we first express the vectors of the (old) basis $\boldsymbol{B}$ in terms of (new) basis vectors $\boldsymbol{b}'_1$ and $\boldsymbol{b}'_2$:

$$\begin{aligned} b_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} &= \alpha_{11}\boldsymbol{b}'_1 + \alpha_{12}\boldsymbol{b}'_2 \\ &= \alpha_{11}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_{12}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

Thus, $\alpha_{11} = 1$, and $\alpha_{12} = 2$.

Similarly,

$$\begin{aligned} \boldsymbol{b}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix} &= \alpha_{21}\boldsymbol{b}'_1 + \alpha_{22}\boldsymbol{b}'_2 \\ &= \alpha_{21}\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_{22}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

Thus, $\alpha_{21} = 3$ and $\alpha_{22} = -1$. Therefore, the transition matrix is

$$T = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$$

Hence,

$$\boldsymbol{x}_{\boldsymbol{B}'} = T(\boldsymbol{x}_{\boldsymbol{B}}) = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 2 \end{bmatrix}$$ ■

In anticipation of a general formalism addressing all the aforementioned questions, let's examine yet another example.

Example 6.138 Consider the following two bases[10] of $\mathbf{R}^3$:

$$\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

and

$$\boldsymbol{B}' = \left\{ \boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{b}'_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \boldsymbol{b}'_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right\}$$

Find the matrices of transition $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$ and $\boldsymbol{B}' \xrightarrow{S} \boldsymbol{B}$.

Solution

$$\begin{aligned}
\boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} &= \alpha_{11}\boldsymbol{b}_1 + \alpha_{12}\boldsymbol{b}_2 + \alpha_{13}\boldsymbol{b}_3 \\
&= \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \end{bmatrix} \Rightarrow \alpha_{11} = 1, \quad \alpha_{12} = 0, \quad \alpha_{13} = 1 \\
\boldsymbol{b}'_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} &= \alpha_{21}\boldsymbol{b}_1 + \alpha_{22}\boldsymbol{b}_2 + \alpha_{23}\boldsymbol{b}_3 \\
&= \begin{bmatrix} \alpha_{21} \\ \alpha_{22} \\ \alpha_{23} \end{bmatrix} \Rightarrow \alpha_{21} = 2, +\alpha_{22} = 1, \quad \alpha_{23} = 2 \\
\boldsymbol{b}'_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} &= \alpha_{31}\boldsymbol{b}_1 + \ \alpha_{32}\boldsymbol{b}_2 + \alpha_{33}\boldsymbol{b}_3 \\
&= \begin{bmatrix} \alpha_{31} \\ \alpha_{32} \\ \alpha_{33} \end{bmatrix} \Rightarrow \alpha_{31} = 1, +\alpha_{32} = 2, \quad \alpha_{33} = 2
\end{aligned}$$

[10]Remember that we often call such basis the "usual" or "canonic" basis.

So, the transition matrix is

$$T = [\alpha_{ij}]^{\tau} = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}^{\tau} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

We look for matrix S in the same way.

$$\begin{aligned} \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} &= \beta_{11}\boldsymbol{b}'_1 + \beta_{12}\boldsymbol{b}'_2 + \beta_{13}\boldsymbol{b}'_3 \\ &= \beta_{11}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta_{12}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + \beta_{13}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} \beta_{11} + 2\beta_{12} + \beta_{13} \\ \beta_{12} + 2\beta_{13} \\ \beta_{11} + 2\beta_{12} + 2\beta_{13} \end{bmatrix} \end{aligned}$$

Thus,

$$\begin{aligned} 1 &= \beta_{11} + 2\beta_{12} + \beta_{13} \\ 0 &= \beta_{12} + 2\beta_{13} \\ 0 &= \beta_{11} + 2\beta_{12} + 2\beta_{13} \end{aligned}$$

So we have $\beta_{11} = -2, \ \ \beta_{12} = 2, \ \ \beta_{13} = -1$

$$\begin{aligned} \boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} &= \beta_{21}\boldsymbol{b}'_1 + \beta_{22}\boldsymbol{b}'_2 + \beta_{23}\boldsymbol{b}'_3 \\ &= \beta_{21}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta_{22}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + \beta_{23}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} \beta_{21} + 2\beta_{22} + \beta_{23} \\ \beta_{22} + 2\beta_{23} \\ \beta_{21} + 2\beta_{22} + 2\beta_{23} \end{bmatrix} \end{aligned}$$

and thus,

$$\begin{aligned} 0 &= \beta_{21} + 2\beta_{22} + \beta_{23} \\ 1 &= \beta_{22} + 2\beta_{23} \\ 0 &= \beta_{21} + 2\beta_{22} + 2\beta_{23} \end{aligned}$$

Solving this system, we get

$$\beta_{21} = -2, \quad \beta_{22} = 1, \quad \beta_{23} = 0$$

Finally,

$$\begin{aligned} \boldsymbol{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} &= \beta_{31}\boldsymbol{b}'_1 + \beta_{32}\boldsymbol{b}'_2 + \beta_{33}\boldsymbol{b}'_3 \\ &= \beta_{31}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \beta_{32}\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + \beta_{33}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} \beta_{31} + 2\beta_{32} + \beta_{33} \\ \beta_{32} + 2\beta_{33} \\ \beta_{31} + 2\beta_{32} + 2\beta_{33} \end{bmatrix} \end{aligned}$$

Thus,

$$\begin{aligned} 0 &= \beta_{31} + 2\beta_{32} + \beta_{33} \\ 0 &= \beta_{32} + 2\beta_{33} \\ 1 &= \beta_{31} + 2\beta_{32} + 2\beta_{33} \end{aligned}$$

Solving this system, we obtain $\beta_{31} = 3, \ \beta_{32} = -2, \ \beta_{33} = 1$. Hence, the transition matrix

$$S = [\beta_{ij}]^\tau = \begin{bmatrix} -2 & 2 & -1 \\ -2 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix}^\tau = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}$$

Now, since T transforms $\boldsymbol{B}$ to $\boldsymbol{B}'$ and S transforms $\boldsymbol{B}'$ to $\boldsymbol{B}$, we would like S to be T^{-1}. Is that the case? Let's see.

$$T \cdot S = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix}\begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

$$S \cdot T = \begin{bmatrix} -2 & -2 & 3 \\ 2 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

Indeed, $S = T^{-1}$. ■

Is this just an accident or does it hold in general? Well, here is what we have hoped for:

Theorem 6.79 Let $A_{\boldsymbol{B}}$ and $A'_{\boldsymbol{B}'}$ be the two coordinate matrices of a vector $\boldsymbol{x} \in \boldsymbol{X}$ in the $\boldsymbol{B}$ and $\boldsymbol{B}'$ bases, respectively, and let T be a transformation matrix from $\boldsymbol{B}$ to $\boldsymbol{B}'$, that is, $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$. Then,

$$A_{\boldsymbol{B}} = T\mathrm{A}'_{\boldsymbol{B}'}$$

Or, equivalently,

$$A'_{\boldsymbol{B}'} = T^{-1}A_{\boldsymbol{B}}$$

Proof Suppose that the expression for a vector $\boldsymbol{x}$ in the $\boldsymbol{B}$ basis is

$$\boldsymbol{x} = \sum_{i=1}^{n} \alpha_i \boldsymbol{x_i} \tag{*}$$

then in the $\boldsymbol{B}'$ basis we have

$$\begin{aligned} \boldsymbol{x} &= \sum_{k=1}^{n} \alpha'_k \boldsymbol{x}'_k \\ &= \sum_{k=1}^{n} \alpha'_k \left(\sum_{i=1}^{n} \beta_{ik} \boldsymbol{x}_i \right) \\ &= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \beta_{ik} \alpha'_k \right) \boldsymbol{x}_i \end{aligned} \tag{**}$$

From (*) and (**) we see that

$$\alpha_i = \sum_{k=1}^{n} \beta_{ik} \alpha'_k, \qquad \forall\, i = 1, \ldots, n$$

So, we conclude

$$A_{\boldsymbol{B}} = TA'_{\boldsymbol{B}'}$$

Similarly, we could get

$$\alpha'_i = \sum_{k=1}^{n} \gamma_{ik} \alpha_k$$

where, of course, $[\gamma_{ik}] = T^{-1}$ and

$$A'_{\boldsymbol{B}'} = T^{-1}A_{\boldsymbol{B}}$$

■

We have seen that if we multiply the coordinates of a vector $\boldsymbol{x} \in \boldsymbol{X}$ in the (old) $\boldsymbol{B}$ basis by T^{-1} we get the coordinates of the same vector in the (new) $\boldsymbol{B}'$ basis. In other words, knowing the transition "in one direction," say, $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$, we can right away find the transition $\boldsymbol{B}' \xrightarrow{S=T^{-1}} \boldsymbol{B}$ by determining T^{-1}.

Example 6.139 Let's take the matrix of T of the $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$ transition from Example 6.138 and apply a little trick to get the matrix of T^{-1} of $\boldsymbol{B}' \xrightarrow{T^{-1}} B$. We "augment" matrix T with the identity matrix to get matrix $M = [T|I]$, that is,

$$M = [T|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{array}\right]$$

We perform elementary row operations on matrix M to get

$$M = \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -2 & 3 \\ 0 & 1 & 0 & 2 & 1 & -2 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{array}\right]$$
$$= [I|T^{-1}]$$

As you can see, the right-hand part of matrix M turned out to be exactly the T^{-1} we were looking for. ■

Example 6.140 Consider the two bases of $\mathbf{R}^3$:

$$\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}, \quad \boldsymbol{b}_2 = \begin{bmatrix} 25 \\ -7 \\ 12 \end{bmatrix}, \quad \boldsymbol{b}_3 = \begin{bmatrix} 34 \\ -16 \\ 17 \end{bmatrix} \right\}$$

$$\boldsymbol{B}' = \left\{ \boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \quad \boldsymbol{b}'_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{b}'_1 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Find the matrix $\boldsymbol{B} \xrightarrow{T} \boldsymbol{B}'$.

Solution We recall that the matrix we are looking for has to have the following property:

For every $\boldsymbol{x} \in \mathbf{R}^3$, $\boldsymbol{x}_{B'} = T\boldsymbol{x}_{B}$. Thus, in particular, $\boldsymbol{b}_1$ written in the new basis $\boldsymbol{B}'$ is

$$\boldsymbol{b}_1 = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} = \beta_{11}\,\boldsymbol{b}'_1 + \beta_{12}\,\boldsymbol{b}'_2 + \beta_{13}\,\boldsymbol{b}'_3$$

We express $\boldsymbol{b}_2$ and $\boldsymbol{b}_3$ similarly. Therefore, columns of T should be $(\boldsymbol{b}_1)_{B'}$, $(\boldsymbol{b}_2)_{B'}$, $(\boldsymbol{b}_3)_{B'}$. As you might have anticipated, calculating all of the β_{ij}'s could be very tedious, so we try another approach. Consider the following "augmented" matrix

$$[\boldsymbol{b}'_1\,\boldsymbol{b}'_2\,\boldsymbol{b}'_3 \mid \boldsymbol{b}_1\,\boldsymbol{b}_2\,\boldsymbol{b}_3] = \left[\begin{array}{ccc|ccc} 1 & -2 & 3 & 5 & 25 & 34 \\ -3 & 0 & 1 & -1 & -7 & -16 \\ 2 & 1 & 0 & 3 & 12 & 17 \end{array}\right]$$

which we now reduce to

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 5 & 7 \\ 0 & 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 & 8 & 11 \end{array}\right] = [I \mid T]$$

Let's check whether our transition matrix T indeed works as we desire. We can apply the transformation T on *any* vector $\boldsymbol{x} \in \mathbf{R}^3$, so why not try it on, say, vector $\boldsymbol{x}$, whose representation in the $\boldsymbol{B}$ basis is $\boldsymbol{x} = 1 \cdot \boldsymbol{b}_1 + 0 \cdot \boldsymbol{b}_2 + 0 \cdot \boldsymbol{b}_3$ (this is the vector we already know).

$$\begin{aligned} T(\boldsymbol{x}) &= T(1 \cdot \boldsymbol{b}_1 + 0 \cdot \boldsymbol{b}_2 + 0 \cdot \boldsymbol{b}_3) \\ &= \begin{bmatrix} 1 & 5 & 7 \\ 1 & 2 & 3 \\ 2 & 8 & 11 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \\ &= 1 \cdot \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix} + 1 \cdot \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \\ &= 1 \cdot \boldsymbol{b}'_1 + 1 \cdot \boldsymbol{b}'_2 + 2 \cdot \boldsymbol{b}'_3 \\ &= \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} = \boldsymbol{b}_1 \end{aligned}$$

■

At this point, a little digression, an observation, may be in order.

Recall that if two spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$ are isomorphic, that is, there exists at least one isomorphism $F : \boldsymbol{X} \to \boldsymbol{Y}$, then any basis of $\boldsymbol{X}$ is mapped to some basis of Y. We say that $\boldsymbol{B_X} \xrightarrow{F} \boldsymbol{B_Y}$, or that any basis of $\boldsymbol{Y}$ is an image of some basis of $\boldsymbol{X}$. This fact establishes another neat feature in the case that the transition is from *canonical* basis $\boldsymbol{E}$ of Φ^n to some other basis of $\mathbf{R}^n$. Namely, columns of the transition matrix $\boldsymbol{E} \xrightarrow{T} \boldsymbol{B'}$ are simply the basis vectors of $\boldsymbol{B'}$.

We illustrate this with the following example.

Example 6.141 Consider the space $\mathbf{R}^3$ and its basis

$$\boldsymbol{B'} = \left\{ \boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},\ \boldsymbol{b}'_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix},\ \boldsymbol{b}'_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right\}.$$

Find the change-of-basis matrix T from the canonical basis $\boldsymbol{E}$ to the $\boldsymbol{B'}$ basis.

Solution The vectors from $\boldsymbol{B'}$ are easily expressible via the vectors of the canonical basis

$$\boldsymbol{E} = \left\{ \boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\quad \boldsymbol{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

$$\boldsymbol{b}'_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \boldsymbol{e}_1 + \boldsymbol{e}_3$$

$$\boldsymbol{b}'_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = 2\,\boldsymbol{e}_1 + \boldsymbol{e}_2 + 2\boldsymbol{e}_3$$

$$\boldsymbol{b}'_3 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \boldsymbol{e}_1 + 2\boldsymbol{e}_2 + 2\boldsymbol{e}_3$$

We immediately obtain the transition matrix

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

Note that the main point of this exercise is to recognize that the columns of T are exactly the coordinates of $\boldsymbol{b}'_1,\ \boldsymbol{b}'_2,\ \boldsymbol{b}'_3$. ■

Theorem 6.80 The change-of-basis matrix T, from the canonical $\boldsymbol{E}$ basis of Φ^n to any basis $\boldsymbol{B}'$ of $\mathbf{R}^n$ (any basis of Φ^n), is a matrix whose columns are the basis vectors of $\boldsymbol{B}'$.

Going back to Examples 6.139 and 6.140, note that the "trick" we used there can also be used in general when searching for the inverse matrix. Here is an example.

Example 6.142 Let's find the inverse of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$$

Solution

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

Thus,

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

One can easily convince oneself that this is indeed so by checking the following:

$$A^{-1}A = AA^{-1} = I$$

■

Example/Exercise 6.143 Find the inverse of

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 5 & 5 & 1 \end{bmatrix}$$

Solution

$$A^{-1} = \begin{bmatrix} 13/8 & -1/2 & -1/8 \\ -15/8 & 1/2 & 3/8 \\ 5/4 & 0 & -1/4 \end{bmatrix}$$

The same method could be used to determine whether or not a given matrix is singular.

Example 6.144 Show that the following matrix A is singular:

$$A = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}$$

Solution

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 2 & 4 & -1 & 0 & 1 & 0 \\ -1 & 2 & 5 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 8 & 9 & 1 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right]$$

We obtained a row with all zeros on the left side, telling us that A is not invertible. ■

Continuing our discussion, we ask the following: Let's say a linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ has a matrix representation $[F]_{(\boldsymbol{B_X}, \boldsymbol{B_Y})} = (\alpha_{ik})$ with respect to a pair of bases $(\boldsymbol{B_X}, \boldsymbol{B_Y})$. Suppose furthermore that $\boldsymbol{x} \in \boldsymbol{X}$ is any vector in $\boldsymbol{X}$, such that its coordinate matrix in the $\boldsymbol{B_X}$ basis is $[\boldsymbol{x}]_{\boldsymbol{B_X}} = [\alpha_i]$. Finally, let $[F(\boldsymbol{x})]_{\boldsymbol{B_Y}} = [\beta_i]$ be the coordinate matrix of the image of $\boldsymbol{x}$ in the $\boldsymbol{B_Y}$ basis. Is there a way to determine $[\beta_i]$ by the matrices (α_{ik}) and $[\alpha_i]$? Here is the answer.

Theorem 6.81 If $[F]_{(\boldsymbol{B_X}, \boldsymbol{B_Y})}$ is a matrix representation of a linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ with respect to basis $\boldsymbol{B_X}$ and $\boldsymbol{B_Y}$, then the coordinate matrix of an image of any $\boldsymbol{x} \in \boldsymbol{X}$ is given by

$$[F(\boldsymbol{x})]_{\boldsymbol{B_Y}} = [F]_{(\boldsymbol{B_X}, \boldsymbol{B_Y})} [\boldsymbol{x}]_{\boldsymbol{B_X}}$$

where $[\boldsymbol{x}]_{\boldsymbol{B_X}} = [\alpha_i]$ is the coordinate matrix of $\boldsymbol{x}$ in the $\boldsymbol{B_X}$ basis.

Proof Since

$$\begin{aligned} \boldsymbol{x} &= \sum_{k=1}^{n} \alpha_k \boldsymbol{b}_{Xk} \\ F(\boldsymbol{x}) &= F\left(\sum_{k=1}^{n} \alpha_k \boldsymbol{b}_{Xk} \right) \\ &= \sum_{k=1}^{n} \alpha_k F(\boldsymbol{b}_{Xk}) \\ &= \sum_{k=1}^{n} \alpha_k \left(\sum_{i=1}^{m} \alpha_{ik} \boldsymbol{b}_{Yi} \right) \\ &= \sum_{i=1}^{m} \left(\sum_{k=1}^{n} \alpha_{ik} \alpha_k \right) \boldsymbol{b}_{Yi} \end{aligned}$$

On the other hand, the image of $\boldsymbol{x}$,

$$F(\boldsymbol{x}) = \sum_{i=1}^{m} \beta_i \boldsymbol{b}_{Yi}$$

Comparing the last two expressions we see that

$$\beta_i = \sum_{k=1}^{n} \alpha_{ik} \alpha_k, \quad \forall\, i = 1, \ldots, m$$

Thus,

$$[F(\boldsymbol{x})]_{\boldsymbol{B}_Y} = [F]_{(\boldsymbol{B}_X, \boldsymbol{B}_Y)} [\boldsymbol{x}]_{\boldsymbol{B}_X}$$

as claimed. ■

Suppose now that the matrix representation of the linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$, with respect to the pair of bases $(\boldsymbol{B}_X, \boldsymbol{B}_Y)$, is known to be $[F]_{(\boldsymbol{B}_X, \boldsymbol{B}_Y)}$. We would like to know the matrix representation of the same operator in a different set of bases $(\boldsymbol{B}'_X, \boldsymbol{B}'_Y)$, that is, we are looking for $[F]_{(\boldsymbol{B}'_X, \boldsymbol{B}'_Y)}$. The answer is given by the following

Theorem 6.82 For brevity, let's denote $[F]_{(\boldsymbol{B}_X, \boldsymbol{B}_Y)}$ simply by $[F]$, and similarly, $[F]_{(\boldsymbol{B}'_X, \boldsymbol{B}'_Y)}$ we will call $[F]'$. Then,

$$[F]' = T^{-1} [F] S$$

where S and T are matrices of the transitions $\boldsymbol{B}_X \xrightarrow{S} \boldsymbol{B}'_X$ and $\boldsymbol{B}_Y \xrightarrow{T} \boldsymbol{B}'_Y$.

Proof Let $\boldsymbol{x} \in$ be any vector in $\boldsymbol{X}$, and let $[\boldsymbol{x}]_{\boldsymbol{B}_X}$ and $[\boldsymbol{x}]_{\boldsymbol{B}'_X}$ be its coordinate matrices in $\boldsymbol{B}_X$ and $\boldsymbol{B}'_X$. Similarly, let $[F(\boldsymbol{x})]_{\boldsymbol{B}_Y}$ and $[F(\boldsymbol{x})]_{\boldsymbol{B}'_Y}$ be coordinate matrices of the image of $\boldsymbol{x}$ in $\boldsymbol{B}_Y$and $\boldsymbol{B}'_Y$. Then,

$$[F(\boldsymbol{x})]_{\boldsymbol{B}_Y} = [F][\boldsymbol{x}]_{\boldsymbol{B}_X}$$

and

$$[F(\boldsymbol{x})]_{\boldsymbol{B}'_Y} = [F]'[\boldsymbol{x}]_{\boldsymbol{B}'_X}$$

Since S is a transformation $\boldsymbol{B}_X \to \boldsymbol{B}'_X$, following the previous theorem, we write

$$[\boldsymbol{x}]_{\boldsymbol{B}_X} = S[\boldsymbol{x}]_{\boldsymbol{B}'_X}$$

Similarly, since T transforms $\boldsymbol{B}_Y \to \boldsymbol{B}'_Y$ we write

$$[F(\boldsymbol{x})]_{\boldsymbol{B}_Y} = T[F(\boldsymbol{x})]_{\boldsymbol{B}'_Y}$$

Collecting it all together we have

$$\begin{aligned}
[F]'[\boldsymbol{x}]_{\boldsymbol{B}'_X} &= [F(\boldsymbol{x})]_{\boldsymbol{B}'_Y} \\
&= T^{-1}[F(\boldsymbol{x})]_{\boldsymbol{B}_Y} \\
&= T^{-1}[F][\boldsymbol{x}]_{\boldsymbol{B}_X} \\
&= T^{-1}[F]S[\boldsymbol{x}]_{\boldsymbol{B}'_X} \\
&= (T^{-1}[F]S)\,[\boldsymbol{x}]_{\boldsymbol{B}'_X}
\end{aligned}$$

and we see that $[F]' = T^{-1}[F]S$, as claimed. ■

Let's pause for a while and sum up what we have learned so far.

Let X and Y be two vector spaces over a field Φ with respective dimensions n and m. If $\boldsymbol{B}_X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ is a basis of X, and $\boldsymbol{B}_Y = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$ is a basis of Y, then the linear operator $F : \boldsymbol{X} \to \boldsymbol{Y}$ is defined by

$$F(\boldsymbol{x}_k) = \sum_{i=1}^{m} \alpha_{ik} \boldsymbol{y}_i$$

As we said before, since the linear operator is uniquely determined by its action on the basis vectors, it is also completely characterized by its matrix representation

$$[F]_{(\boldsymbol{B}_X,\boldsymbol{B}_Y)} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix}$$

where the columns of this (m, n) matrix represent the coordinates of $F(\boldsymbol{x}_k)$, that is, the images of basis vectors of $\boldsymbol{B}_X$ in the $\boldsymbol{B}_Y$ basis. So far so good. One is now inspired to ask: Are we able to combine linear operators as we combine functions? In other words, let

$$F : \boldsymbol{X} \to \boldsymbol{Y} \quad \text{and} \quad G : \boldsymbol{X} \to \boldsymbol{Y}$$

be two linear operators over the same field Φ. We claim that

$$(F + G) : \boldsymbol{X} \to \boldsymbol{Y}$$

and

$$\lambda F : \boldsymbol{X} \to \boldsymbol{Y}, \quad \lambda \in \Phi$$

are linear operators too. Indeed, let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ be any two vectors, and let $\alpha, \beta \in \Phi$ be any two scalars, then

$$\begin{aligned} (F + G)(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) &= F(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) + G(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) \\ &= \alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2) + \alpha G(\boldsymbol{x}_1) + \beta G(\boldsymbol{x}_2) \\ &= \alpha[F(\boldsymbol{x}_1) + G(\boldsymbol{x}_1)] + \beta[F(\boldsymbol{x}_2) + G(\boldsymbol{x}_2)] \\ &= \alpha(F + G)(\boldsymbol{x}_1) + \beta(F + G)(\boldsymbol{x}_2) \end{aligned}$$

Similarly,

$$\begin{aligned} \lambda F(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) &= \lambda[\alpha F(\boldsymbol{x}_1) + \beta F(\boldsymbol{x}_2) \\ &= \alpha[\lambda F(\boldsymbol{x}_1)] + \beta[\lambda\, F(\boldsymbol{x}_2)] \end{aligned}$$

So, $(F + G)$ and λF are linear operators, as claimed. Thus, we have (almost) proved.

Theorem 6.83 Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two vector spaces over a field Φ. Then, the set of all linear operators from $\boldsymbol{X}$ to $\boldsymbol{Y}$, with standard addition and multiplication, as described earlier, forms a vector space over the same field Φ. As you remember, we call this space $\text{Hom}(\boldsymbol{X}, \boldsymbol{Y})$ (see Definition 6.13).

Before continuing with the next example you may want to revisit Examples 6.119–6.128.

Example 6.145 Suppose we want to find the matrix representation of an operator $\mathcal{T} : \mathbf{R}^2 \to \mathbf{R}^2$ consisting of an *R-reflection* about the X-axis, followed by a *P-rotation* through $\pi/2$, followed by a *D-dilatation* by a factor 3. In other words, we are looking for an operator $\mathcal{T}$, which is a composition

$$\mathcal{T} = D \circ P \circ R$$

Since the matrix representations of D, P, and R are

$$D = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \quad P = \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

the matrix representation of $\mathcal{T}$ is

$$\begin{aligned} \mathcal{T} = D \circ P \circ R &= \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} \end{aligned}$$

■

Definition 6.65 A linear operator $F : \boldsymbol{X} \to \boldsymbol{X}$ is said to be invertible, or that it has an inverse, if there exists F^{-1} such that

$$FF^{-1} = F^{-1}F = I$$

It is good to remember that (a mapping) F is invertible iff F is a bijection, in which case F^{-1} is also a linear operator, as we expected. We illustrate this with the following

Example 6.146 Suppose $F : \boldsymbol{X} \to \boldsymbol{X}$ is a bijective linear operator. Let's show that F^{-1} is also a linear operator.

Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ be any two vectors, then, since F is a bijection, there exist $\boldsymbol{x}', \boldsymbol{y}' \in \boldsymbol{X}$ such that $F(\boldsymbol{x}) = \boldsymbol{x}'$ and $F(\boldsymbol{y}) = \boldsymbol{y}'$. Consider now,

$$\begin{aligned} F(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) &= \alpha F(\boldsymbol{x}) + \beta F(\boldsymbol{y}) \\ &= \alpha\boldsymbol{x}' + \beta\boldsymbol{y}' \end{aligned}$$

On the other hand, by definition of the inverse map,

$$F^{-1}(\alpha\boldsymbol{x}') = \alpha\boldsymbol{x} \quad \text{and} \quad F^{-1}(\beta\boldsymbol{y}') = \alpha\boldsymbol{x}$$

as well as

$$F^{-1}(\boldsymbol{x}' + \boldsymbol{y}') = \boldsymbol{x} + \boldsymbol{y}.$$

Therefore,

$$\begin{aligned} F^{-1}(\alpha \boldsymbol{x}' + \beta \boldsymbol{y}') &= \alpha \boldsymbol{x} + \alpha \boldsymbol{y} \\ &= F^{-1}(\alpha \boldsymbol{x}') + F^{-1}(\beta \boldsymbol{y}') \\ &= \alpha F^{-1}(\boldsymbol{x}') + \beta F^{-1}(\boldsymbol{y}'). \end{aligned}$$

So, F^{-1} is a linear operator. ■

Example 6.147 Let $F : \mathbf{R}^2 \to \mathbf{R}^2$ be a linear operator defined by

$$F(\boldsymbol{x}) = F \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ 3x + 2y \end{bmatrix} = \boldsymbol{x}'$$

Find F^{-1}.

Solution First, if F is invertible, then it has to be regular. Indeed, consider

$$F \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{0}' = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then,

$$\begin{aligned} 2x + y &= 0 \\ 3x + 2y &= 0 \end{aligned}$$

Solving this system, we get $x = 0$ and $y = 0$. Thus, F is regular and therefore has an inverse, that is, if $F(\boldsymbol{x}) = \boldsymbol{x}'$, then $F^{-1}(\boldsymbol{x}') = \boldsymbol{x}$.

So, let

$$\begin{bmatrix} 2x + y \\ 3x + 2y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

That is

$$\begin{aligned} 2x + y &= x' \\ 3x + 2y &= y' \end{aligned}$$

Solving this system for x and y we get

$$\begin{aligned} x &= 2x' - y' \\ y &= -3x' + 2y' \end{aligned}$$

Hence, the inverse operator we are looking for is defined by

$$F^{-1}(\boldsymbol{x}') = F\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 2x' - y' \\ -3x' + 2y' \end{bmatrix}$$

Of course, considering that x' and y' are variables (dummies) from $\mathbf{R}$, it is more convenient to simply write

$$F^{-1}(\boldsymbol{x}) = F\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - y \\ -3x + 2y \end{bmatrix}$$ ■

6.16 SYSTEMS OF LINEAR EQUATIONS I

We started this chapter by briefly mentioning the problem of solving a system of n linear equations with n variables, as one of the motivating factors for studying Linear Algebra. Now we address this issue in more detail, using the tools and techniques we acquired in the previous sections.

First, let's recall some basic concepts.

Definition 6.66 A linear equation in n unknowns $x_1, x_2, \ldots, x_n$ over a field Φ is an equation that in **standard form** reads as follows:

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta \qquad (*)$$

where $\alpha_i, \beta \in \Phi$ are any scalars from the given field. We say that the scalars α_i are **coefficients** of equation (*) and β is a **free term** or a **constant term**.

Definition 6.67 A solution of equation (*) is an n-tuple $C = (x_1 = \gamma_1, x_2 = \gamma_2, \ldots, x_n = \gamma_n)$ of scalars $\gamma_i \in \Phi$ such that

$$\alpha_1 \gamma_1 + \alpha_2 \gamma_2 + \cdots + \alpha_n \gamma_n = \beta$$

We are interested in the **finite** system of equations of the form (*), so we have

Definition 6.68 A **system of linear equations** $\mathcal{L}$ over a field Φ, is a collection of equations

$$\begin{aligned}
\alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n &= \beta_1 \\
\alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n &= \beta_2 \\
\vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots & \\
\alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n &= \beta_i \\
\vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots & \\
\alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n &= \beta_m
\end{aligned} \qquad (**)$$

where $\alpha_{ik} \in \Phi$ are **coefficients**, and $\beta_k \in \Phi$ are **free terms**. We will often write the system (**) in a compact form

$$\sum_{k=1}^{n} \alpha_{ik} x_k = \beta_i, \quad i = 1, \ldots, m \qquad (**)'$$

Definition 6.69 A system of linear equations is called **homogeneous** if all β_i's in (**) are equal to zero.

A system of linear equations is called **nonhomogeneous** if at least one of the β_i's from (**) is different from zero.

Definition 6.70 An ordered n-tuple $C = (\gamma_1, \ldots, \gamma_n)$ of scalars $\gamma_i \in \Phi$ is said to be a solution of the system $\mathcal{L}$ iff it satisfies every equation of (**).

Example 6.148 Here is an example of a simple 3×4 (nonhomogeneous) system

$$\begin{aligned} x_1 + x_2 + 4x_3 + 3x_4 &= 5 \\ 2x_1 + 3x_2 + x_3 - 2x_4 &= 1 \\ x_1 + 2x_2 - 5x_3 + 4x_4 &= 3 \end{aligned}$$

The system obviously consists of three equations with four unknowns and its solution, as can be easily verified, is an ordered quadruple $C = (-8, 6, 1, 1)$, which satisfies all three equations. ■

We have indicated before that one can associate corresponding matrices with any system of linear equations $\mathcal{L}$, hoping to facilitate the process of finding a solution for the system. First, we collect all the coefficients of $\mathcal{L}$ in matrix form

$$A = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix}$$

making sure that the scalars α_i are in the same position as in (**). We call this matrix the **matrix of coefficients** of our system. Next, we construct the **augmented matrix**,

$$\overline{A} = \left[\begin{array}{ccc|c} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & \ddots & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} & \beta_m \end{array}\right]$$

by supplying the matrix A with an additional column $B = [\beta_i]$ of free terms of the system (**). Finally, we write all the unknowns in matrix form

$$X = [x_k] = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Thus the system, $\mathcal{L}$, compactly expressed in matrix form, reads

$$AX = B$$

Consequently, saying that $C = (\gamma_1, \ldots, \gamma_n) = [\gamma_k]$ is the solution of the system means the following is also true:

$$AC = B$$

Evoking our experience with linear operators, we can look at the problem of solving the system of linear equations, $\mathcal{L}$, by considering the transformation

$$F : \mathbf{R}^n \to \mathbf{R}^m$$

where the matrix representation of F in a standard basis is the matrix of coefficients A, matrix B is a vector in $\mathbf{R}^m$, and the solution C is a vector in $\mathbf{R}^n$. Let's elaborate on this in more general terms. Let

$$\boldsymbol{M}_{\boldsymbol{n}1}(\Phi) = \left\{ \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \mid \alpha_i \in \Phi \right\}$$

be a linear space of all $(n, 1)$-matrices over a field Φ. Consider a standard basis of this n-dimensional space

$$\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$$

where, as usual, $\boldsymbol{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, $i = 1, \ldots, n$

Let's pick any vector $\boldsymbol{x} \in \boldsymbol{M_{n1}}$. Its coordinate matrix in the basis $\boldsymbol{E}$ is

$$\begin{aligned}\boldsymbol{x} &= \begin{bmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{bmatrix} \\ &= \alpha_1 \begin{bmatrix}1 \\ 0 \\ \vdots \\ 0\end{bmatrix} + \cdots + \alpha_n \begin{bmatrix}0 \\ \vdots \\ 0 \\ \vdots \\ 1\end{bmatrix} \\ &= \alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n \\ &= [\boldsymbol{x}]_{\boldsymbol{E}}\end{aligned}$$

So, as expected, *any vector from* $\boldsymbol{M_{n1}}$ *is uniquely represented by its coordinate matrix with respect to a given basis*, that is, $\boldsymbol{x} = \begin{bmatrix}x_1 \\ \vdots \\ x_n\end{bmatrix} = X \in \boldsymbol{M_{n1}}$. Similarly, we can consider an m-dimensional space $\boldsymbol{M_{m1}}$ over the same field Φ, and analogously choose a corresponding standard basis, say,

$$\boldsymbol{E'} = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_m\}$$

In the same way, we consider B the matrix of free terms as some vector $\boldsymbol{b}$ in space $\boldsymbol{M_{m1}}$, that is, $[\boldsymbol{b}]_{\boldsymbol{E'}} = B$.

Let's say we are interested in a linear operator

$$F : \boldsymbol{M_{n1}} \to \boldsymbol{M_{m1}}$$

whose matrix representation is a matrix $A = [\alpha_{ik}]$. As we have learned

$$\begin{aligned}F(\boldsymbol{e}_k) &= \sum_{i=1}^{m} \alpha_{ik} \boldsymbol{e}'_i \\ &= \sum_{i=1}^{m} \alpha_{ik} \begin{bmatrix}0 \\ \vdots \\ 1 \\ \vdots \\ 0\end{bmatrix} = \begin{bmatrix}\alpha_{1k} \\ \vdots \\ \alpha_{mk}\end{bmatrix} \in \boldsymbol{M_{m1}}\end{aligned}$$

So, the matrix equation

$$AX = B$$

can be understood as an "operator equation"

$$F(\boldsymbol{x}) = \boldsymbol{b}$$

or, taking into consideration the standard basis,

$$F(X) = B$$

Finally, instead of the spaces $\boldsymbol{M}_{\boldsymbol{n}1}$ and $\boldsymbol{M}_{\boldsymbol{m}1}$, we could have used any n-, m-dimensional vector spaces and obtain the same "geometrical" interpretation of our system $\mathcal{L}$.

Of course, now we would like to know if there are some conditions that have to be satisfied in order to know in advance whether a system of linear equations $\mathcal{L}$ has a solution or not, and if it does, how many. First,

Definition 6.71 A system of linear equations $\mathcal{L}$ is said to be **consistent** if it has one or more solutions, and it is said to be **inconsistent if** it has no solution.

Example 6.149 The following systems:

$$\begin{aligned} x_1 + 2x_2 &= 9 \\ 2x_1 - 3x_2 &= 4 \end{aligned}$$

and

$$\begin{aligned} x_1 + x_2 + 4x_3 + 3x_4 &= 5 \\ 2x_1 + 3x_2 + x_3 - 2x_4 &= 1 \\ x_1 + 2x_2 - 5x_3 + 4x_4 &= 3 \end{aligned}$$

are examples of two consistent systems. On the other hand, the next two systems,

$$\begin{aligned} x_1 + 2x_2 &= 9 \\ x_1 + 2x_2 &= 15 \end{aligned}$$

and

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + \quad + x_3 &= 1 \\ 2x_1 + x_2 + 2x_3 &= 0 \end{aligned}$$

are inconsistent. ■

Faced with the question of the existence of a solution of a linear system over an infinite field Φ, we expect to have a unique solution, no solution at all, or an infinite number of solutions. Recall our matrix equation

$$AX = B \tag{*}$$

and its operator analog

$$F(X) = B \tag{**}$$

Clearly, a column matrix $X = [x_k]$ is a solution of (*) iff a vector $\boldsymbol{x} = X \in \boldsymbol{M_{n1}}$ satisfies equation (**). Let's denote the set of all solutions of (*)/(**) by

$$\mathrm{S} = F^{-1}(B)$$

We expect that our system $\mathcal{L}$ has a solution (i.e., equation (*) has a solution), iff $S = F^{-1}(B) \neq \varnothing$. In other words, iff $B \in F(\boldsymbol{M_{n1}}) = \mathrm{Im}(F)$. Thus,

Theorem 6.84 A linear system $\mathcal{L}$ has a solution iff the vector $B = [\boldsymbol{b}_i] \in F(\boldsymbol{M}_{n1}) = \mathrm{Im}(F)$ where $F : \mathbf{R}^n \to \mathbf{R}^m$.

The aforementioned theorem succinctly explicates the necessities for the existence of a solution of a system of linear equations. However, it is not very convenient for practical calculations. In other words, a critical question still remains: How can we tell whether $B \in \mathrm{Im}(F)$? Let's elaborate on this a little more. First, if we choose $\boldsymbol{E} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ as a standard basis for a space $\boldsymbol{M}_{n1}$, and consider the columns of the coefficient matrix A as vectors in $\mathbf{R}^m$, that is,

$$\boldsymbol{a}_1 = \begin{bmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} \alpha_{12} \\ \vdots \\ \alpha_{m2} \end{bmatrix}, \ \ldots, \quad \boldsymbol{a}_n = \begin{bmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{bmatrix}$$

then $\mathrm{Im}(F)$ is spanned by

$$F(\boldsymbol{e}_1) = \boldsymbol{a}_1, \quad F(\boldsymbol{e}_2) = \boldsymbol{a}_2, \ \ldots, \ F(\boldsymbol{e}_n) = \boldsymbol{a}_n$$

Thus, vector B is an element of $\mathrm{Im}(F)$ iff it can be expressed as

$$\begin{aligned} B &= \sum_{k=1}^{n} \gamma_k F(\boldsymbol{e}_k) \\ &= \sum_{k=1}^{n} \gamma_k \begin{bmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{mk} \end{bmatrix} \end{aligned}$$

Put differently, and in more of a "user-friendly" way, we say that $B \in \text{Ker}(F)$ iff the last column of the augmented matrix of a linear system $\mathcal{L}$

$$\overline{A} = \left[\begin{array}{ccc|c} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & \ddots & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} & \beta_m \end{array}\right]$$

is a linear combination of the previous columns, that is, the column vectors of matrix A. So we have

Theorem 6.85 Let $\mathcal{L}$ be a system of linear equations whose matrix representation is

$$AX = B$$

Then, a system $\mathcal{L}$ has a solution iff vector $B = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$ can be expressed as a linear combination of the column vectors of A.

Example 6.150 Consider a linear system

$$\begin{aligned} x_1 + x_2 + \ x_3 &= 1 \\ x_1 + \quad + \ x_3 &= 1 \\ 2x_1 + x_2 + 2x_3 &= 0 \end{aligned}$$

Its augmented matrix

$$\overline{A} = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 2 & 0 \end{array}\right]$$

We anticipate that the column space of A is spanned by the vectors

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\alpha}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

but since $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_3$, $\boldsymbol{\alpha}_1$, and $\boldsymbol{\alpha}_2$ are sufficient for the span.

On the other hand, since vector $\boldsymbol{\beta} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, we need to determine whether $\boldsymbol{\beta}$ belongs to the space spanned by $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, that is, whether we can express $\boldsymbol{\beta}$ as

a linear combination of $\boldsymbol{a}_i'$s. Let's see:

$$\begin{bmatrix}1\\1\\0\end{bmatrix} = \lambda_1\begin{bmatrix}1\\1\\2\end{bmatrix} + \lambda_2\begin{bmatrix}1\\0\\1\end{bmatrix}, \quad \lambda_1,\ \lambda_2 \in \mathbf{R}$$

gives us

$$\begin{aligned} 1 &= \lambda_1 + \lambda_2 \\ 1 &= \lambda_1 \\ 0 &= 2\lambda_1 + \lambda_2 \end{aligned}$$

which implies $\lambda_1 = 1$, and $\lambda_2 = 0 = -2$, and that is impossible. Thus, $\boldsymbol{\beta}$ does not belong to the space spanned by the vectors $\boldsymbol{a}_i$, and our system has no solution. ■

Example/Exercise 6.151 Show that the following system:

$$\begin{aligned} 2x_1 + 6x_2 - x_3 + x_4 &= 2 \\ x_1 + 3x_2 + x_3 &= 5 \\ -x_1 - 3x_2 - x_3 &= 0 \end{aligned}$$

has no solution.

We conclude this discussion with the following

Theorem 6.86 (Kronecker–Capelli) A system of linear equations is solvable iff the matrix of the system A and the augmented matrix $\widetilde{A}$ have the same rank.

Proof In Theorem 6.84, we had the expression

$$\begin{aligned} B &= \sum_{k=1}^{n} \gamma_k F(\boldsymbol{e}_k) \\ &= \sum_{k=1}^{n} \gamma_k \begin{bmatrix}\alpha_{1k}\\ \vdots \\ \alpha_{mk}\end{bmatrix} \end{aligned} \tag{*}$$

which equivalently can be written as

$$B = \begin{bmatrix}\beta_1\\ \vdots \\ \beta_i \\ \vdots \\ \beta_m\end{bmatrix} = \sum_{k=1}^{n} \begin{bmatrix}\gamma_k\alpha_{1k}\\ \vdots \\ \gamma_k\alpha_{ik} \\ \vdots \\ \gamma_k\alpha_{mk}\end{bmatrix}$$

and thus,

$$\beta_i = \sum_{k=1}^{n} \alpha_{ik}\gamma_k, \quad i = 1, \ldots, m$$

We conclude that the coefficients $(\gamma_1, \ldots, \gamma_n)$ from (*) represent one of the solutions of the corresponding linear system. ■

The next definition, and the two theorems that follow, will prove to be very useful.

Definition 6.72 A system of linear equations whose matrix form is $AX = B$, is said to be a **Cramer system** if

(i) the number of equations is equal to the number of unknowns, that is, $m = n$.
(ii) $r(A) = n$, that is, the rank of A is maximal, that is, $\det(A) \neq 0$.

Theorem 6.87 Let $\mathcal{L}$ be a system of m linear equations whose matrix form is $AX = B$. If

$$r(A) = m$$

then the system $\mathcal{L}$ has a solution.

Proof If $r(A) = m$, then $r(\overline{A}) = r(A)$, since this rank is already maximal for both matrices. The claim of the theorem follows from Theorem 6.86 ■

Theorem 6.88 Let $\mathcal{L}$ be a solvable system of linear equations, that is, a solution set

$$S = F^{-1}(B) \neq 0$$

where $F : \boldsymbol{M_{n1}} \to \boldsymbol{M_{m1}}$. Then the solution of the system is unique iff F is an injection.

Proof We deduced from the previous theorem that if $r(\mathrm{A}) = r(F) = \dim(\mathrm{Im}(F)) = m$, then $\mathrm{Im}(F) = \boldsymbol{M_{m1}}$, that is, F is surjective and therefore $F^{-1}(B) \neq 0$, as expected. Now we need to prove that F is injective. Let

$$C = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} \in \boldsymbol{M_{n1}}$$

be a *unique* solution of our system

$$AX = B$$

Or, equivalently,

$$F(X) = B$$

Clearly then,

$$F(C) = B$$

Suppose, now, that F is not injective. Then there exists a vector

$$D = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_n \end{bmatrix} \in \boldsymbol{M}_{n1}, \quad D \neq O$$

such that

$$F(D) = O$$

Well, consider

$$\begin{aligned} F(C + D) &= F(C) + F(D) \\ &= B + O \end{aligned}$$

This would imply that $C + D$ is also a solution of our system, contradicting our assumption that the solution is unique. It follows that F is injective. Conversely, let's assume that F injective. Then, for every $B \in \boldsymbol{M}_{m1}$, our operator equation

$$F(X) = B$$

has only one solution. That completes the proof. ■

Simply put, a bijective (injective + surjective) and thus regular, operator allows us to state

Theorem 6.89 The solvable system of n linear equations with n unknowns has a unique solution iff the operator F is regular.

Consequently, we also have

Theorem 6.90 A Cramer system always has a unique solution.

Proof If a system $\mathcal{L}$ is a Cramer system, then

$$\begin{aligned} r(A) &= n = m \\ &= (\#\text{ of unknowns} = \#\text{ of equations}) \end{aligned}$$

and thus the system has at least one solution. Theorems 6.87 and 6.88 show that the solution is unique. ■

To get a practical procedure for solving a system of linear equations let's sum up everything we have learned so far:

Let $\mathcal{L}$ be a Cramer system of linear equations

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n &= \beta_2 \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots & \\ \alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{nn}x_n &= \beta_n \end{aligned} \tag{*}$$

In matrix form this reads

$$AX = B \tag{**}$$

where

$$A = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Since we have assumed that our system $\mathcal{L}$ is a Cramer system, $\det A \neq 0$, and the solution of $\mathcal{L}$ is

$$X = A^{-1}B$$

A very neat procedure for obtaining the solution of such a system is given by

Theorem 6.91 (Cramer's rule) Let $\mathcal{L}$ be a Cramer system of linear equations (*) whose matrix expression is (as above)

$$AX = B \tag{**}$$

Furthermore, let $D = \det A$, and let D_k be a determinant obtained from $\det A$ by replacing the kth column of A with the column of free terms $[\beta_i]$, that is

$$D_k = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1k-1} & \beta_1 & \alpha_{1k+1} & \cdots & \alpha_{1n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nk-1} & \beta_n & \alpha_{nk+1} & \cdots & \alpha_{nn} \end{bmatrix}.$$

Then, the unique solution of the system $\mathcal{L}$ is

$$\gamma_k = D_k D^{-1}, \quad k = 1, \ldots, n$$

Proof In order for $C = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}$ to be a solution of $\mathcal{L}$ we need to check whether γ_i satisfies the ith equation of (*) for every i. Let's see

$$\begin{aligned}
\sum_{k=1}^{n} \alpha_{ik}\gamma_k &= \sum_{k=1}^{n} \alpha_{ik} D_k D^{-1} \\
&= D^{-1} \sum_{k=1}^{n} \alpha_{ik} D_k \\
&= D^{-1} \sum_{k=1}^{n} \alpha_{ik} \left(\sum_{j=1}^{n} \beta_j A_{jk} \right) \\
&= D^{-1} \sum_{j=1}^{n} \beta_j \left(\sum_{k=1}^{n} \alpha_{ik} A_{jk} \right) \\
&= D^{-1} \left(\sum_{j=1}^{n} \beta_j \delta_{ij} D \right) \\
&= D^{-1} \beta_i D = \beta_i
\end{aligned}$$

where A_{jk} is a cofactor of the element α_{ik}. Since this holds for every i, that is, an n-tuple $(\gamma_1, \ldots, \gamma_n)$ satisfies all the equations

$$\sum_{k=1}^{n} \alpha_{ik} x_k = \beta_i, \quad i = 1, \ldots, n$$

Thus, we have proved the theorem. ■

This is an important result, so let's explicate on it a little more.

First, recall Theorem 6.73: If A is a regular matrix, then

$$A^{-1} = \frac{1}{\det A} \widetilde{A}$$

where $\widetilde{A} = \text{adj}(A)$.

To make sure that this is clear, some things are worth repeating. By Theorem 6.71

$$A\widetilde{A} = \det A \cdot I$$

The left-hand side, written explicitly is

$$A\widetilde{A} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & \ddots & \vdots \\ A_{1n} & \cdots & A_{nn} \end{bmatrix}$$

The product of the ith row and jth column of $A\widetilde{A}$ is clearly

$$\alpha_{i1}A_{j1} + \alpha_{i2}A_{j2} + \cdots + \alpha_{in}A_{jn} \qquad (*)$$

Now observe, and this is crucial, that in the case where $i = j$, (*) gives us a cofactor expansion of $\det A$ along the ith row. If, on the other hand, if $i \neq j$, then the αs and cofactors come from different rows of our matrix A, so the value of (*) is zero. Hence,

$$A\widetilde{A} = \begin{bmatrix} \det A & 0 & \cdots & 0 & 0 \\ 0 & \det A & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ \vdots & \vdots & \cdots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \det A \end{bmatrix}$$

$$= \det A \cdot I$$

So, our equation

$$A\widetilde{A} = \det A \cdot I$$

can be rewritten as

$$A\frac{1}{\det A}\widetilde{A} = I$$

Finally, multiplying both sides by A^{-1} yields our result

$$A^{-1} = \frac{1}{\det A}\widetilde{A}$$

Now we go back to our system of linear equations, $\mathcal{L}$, and Cramer's solutions. We have seen that if $\mathcal{L}$, written in matrix form, is

$$AX = B$$

and $\det(A) \neq 0$, then the system has a unique solution

$$X = A^{-1}B$$

$$= \frac{1}{\det A}\widetilde{A}B$$

Now, by Theorem 6.91

$$x_1 = \gamma_1 = \frac{D_1}{D}, \quad x_2 = \gamma_2 = \frac{D_2}{D}, \ \ldots\ , x_n = \gamma_n = \frac{D_n}{D}$$

where, as we recall, $D = \det(A)$ and

$$D_k = \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{1k-1} & \beta_1 & \alpha_{1k+1} & \cdots & \alpha_{1n} \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nk-1} & \beta_n & \alpha_{nk+1} & \cdots & \alpha_{nn} \end{vmatrix}$$

Here are some examples.
First, something rather simple.

Example 6.152 Suppose we have the following linear system:

$$\begin{bmatrix} 1 & 4 & 2 \\ 3 & -3 & 6 \\ 2 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ -4 \end{bmatrix}$$

and we want to find the value of, say, x_2. Well, by the previous theorem we have

$$x_2 = \frac{D_2}{D} = \frac{\begin{vmatrix} 1 & 3 & 2 \\ 3 & 5 & 6 \\ 2 & -4 & 5 \end{vmatrix}}{\begin{vmatrix} 1 & 4 & 2 \\ 3 & -3 & 6 \\ 2 & 0 & 5 \end{vmatrix}} = \frac{4}{15}$$

■

Example 6.153 Let's solve the following system of linear equations $\mathcal{L}$ using Cramer's method:

$$\begin{aligned} x_1 + \quad & +2x_3 = 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Matrix A is evidently $A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix}$ and $D = \det(A) = 44$ and therefore,

$$D_1 = \begin{vmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{vmatrix}, \quad D_2 = \begin{vmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{vmatrix}, \quad D_3 = \begin{vmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{vmatrix}$$

$$= -40 \qquad = 72 \qquad = 18$$

Thus,

$$x_1 = \frac{D_1}{D} = \frac{-40}{44}; \quad x_2 = \frac{D_2}{D} = \frac{72}{44}; \quad x_3 = \frac{152}{44}.$$ ■

Example 6.154 Solve by Cramer's method.

$$\begin{aligned} 2x_1 + 3x_2 - \quad x_3 &= 1 \\ 3x_1 + 5x_2 + 2x_3 &= 8 \\ x_1 - 2x_2 - 3x_3 &= -1 \end{aligned}$$

Solution

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 3 & 5 & 2 \\ 1 & -2 & -3 \end{bmatrix}, \quad D = \det(A) = 22$$

$$D_1 = \begin{vmatrix} 1 & 3 & -1 \\ 8 & 5 & 2 \\ -1 & -2 & -1 \end{vmatrix}, \quad D_2 = \begin{vmatrix} 2 & 1 & -1 \\ 3 & 8 & 2 \\ 1 & -1 & -3 \end{vmatrix}, \quad D_3 = \begin{vmatrix} 1 & 3 & 1 \\ 3 & 5 & 8 \\ 1 & -2 & -1 \end{vmatrix}$$

$$= 66 \qquad = -22 \qquad = 44$$

Thus,

$$x_1 = \frac{D_1}{D} = \frac{66}{22} = 3; \quad x_2 = \frac{D_2}{D} = \frac{-22}{22} = -1; \quad x_3 = \frac{D_3}{D} = \frac{44}{22} = 2$$ ■

Example 6.155 Solve by Cramer's method.

$$\begin{aligned} x_1 + 3x_2 + x_3 &= -2 \\ 2x_1 + 5x_2 + x_3 &= -5 \\ x_1 + 2x_2 + 3x_3 &= 6 \end{aligned}$$

Solution

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 1 \\ 1 & 2 & 3 \end{bmatrix}; \quad D = -3; \quad D_1 = -3, \quad D_2 = 6, \quad D_3 = -9$$

Thus,

$$x_1 = 1, \quad x_2 = -2, \quad x_3 = 3$$ ■

Example/Exercise 6.156 Show that the following system of linear equations is a Cramer system and find its solution:

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 1 \\ x_1 + 2x_2 - 2x_3 &= 2 \\ 3x_1 - x_2 - x_3 &= 3 \end{aligned}$$

6.17 SYSTEMS OF LINEAR EQUATIONS II

In the previous section, we have defined (Definition 6.69 a homogeneous system of linear equations, $\mathcal{L}$, as a system whose all free terms β_i ($i = 1, \ldots, m$) are equal to zero, that is, a system

$$\sum_{k=1}^{n} \alpha_{ik} x_k = 0, \quad i = 1, \ldots, m \qquad (*)$$

which, written in matrix form, reads

$$AX = O \qquad (**)$$

Put differently and in more general terms, if

$$F : \mathbf{R}^n \to \mathbf{R}^m$$

is a linear operator whose matrix in a standard basis is A, then a solution of a homogeneous equation (**) is a linear subspace of $\boldsymbol{R}^n$, that is,

$$S = \mathrm{Ker}(F)$$

Needless to say, (**) always has a trivial solution, namely

$$x_1 = x_2 = \cdots = x_n = 0$$

But what about nontrivial solutions? Let's see. First, recall that in "operator language," what equations (*) and (**) are actually saying is that the operator

$$F : \boldsymbol{M}_{\boldsymbol{n}1} \to \boldsymbol{M}_{\boldsymbol{m}1}$$

is such that the set of solutions is equal to $\mathrm{Ker}(F)$, that is,

$$S = F^{-1}(O) = \mathrm{Ker}(F)$$

where $O = B = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \boldsymbol{M_{m1}}$. In other words, $C = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} \in \boldsymbol{M_{n1}}$ is a solution of (**) iff $C \in \text{Ker}(F)$. But remember, $\text{Ker}(F) < \boldsymbol{M_{n1}}$, and $\dim(\text{Ker}(F)) = d(F)$, thus by Theorem 6.39

$$r(F) + d(F) = \dim(\boldsymbol{M_{n1}}) = n$$
$$d(F) = n - r(F)$$

So we have

Theorem 6.92 A *homogeneous system* of m linear equations with n unknowns and a matrix of coefficients A, such that $r(F) = r(A)$ has

(i) only a trivial solution iff $r(A) = n$,
(ii) also nontrivial solutions iff $r(A) < n$,
(iii) also nontrivial solutions if $m < n$.

Theorem 6.93 (Rouche) A *homogeneous system* of n linear equations, with n unknowns and a matrix of coefficients A, has a nontrivial solution iff $\det A = 0$.

Proof Suppose $\det(A) \neq 0$. Then we have $r(A) = m = n$ and, by the previous theorem, our system would have only a trivial solution. ■

We illustrate this with the following example.

Example 6.157 Let's show that a system $AX = O$, where $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 1 & 3 \end{bmatrix}$ has only a trivial solution.

Written explicitly, our system reads

$$AX = O$$
$$= \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Thus,

$$X = A^{-1}O$$
$$= \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 1 & 3 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

That means that a trivial solution exists if A^{-1} exists, that is, A is regular. But an (n, n) matrix is regular iff $r(A) = n$, that is, the rank is maximal. In our case, the rank should be 3. Well, let's see:

By applying elementary row operations, we get

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 1 & 3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

The number of pivots is 3 (see Definition 6.59), so $r(A) = 3$. Hence, A is regular, and our system has only a trivial solution.

Of course, we could have invoked Rouche's theorem and simply shown that

$$\det(A) \neq 0$$

concluding that our *homogeneous system* has only a trivial solution. ■

Now we turn again to a nonhomogeneous system, that is, to a system

$$\sum_{k=1}^{n} \alpha_{ik} x_k = \beta_i, \quad i = 1, \ldots, m \qquad (*)'$$

where at least one of free terms is different from zero.

In other words, we are considering

$$AX = B, \qquad B \neq O \qquad (**)'$$

Evidently, we can associate with every nonhomogeneous system $(*)'$ a **corresponding homogeneous** system (*), or equivalently, to each system $(**)'$ we can associate a system (**). A natural question to ask next is : what is the relationship between corresponding homogeneous and nonhomogeneous systems? Well, let $F : \mathbf{R}^n \to \mathbf{R}^m$ be a linear operator corresponding to our system $(*)'/(**)'$. More transparently, we write this as

$$F : \boldsymbol{M}_{n1} \to \boldsymbol{M}_{m1}$$

Remember, if we assume that our system $(*)'$ is solvable, then a set of solutions is

$$S = F^{-1}(B) \neq 0$$

and also

$$r(\overline{A}) = r(A)$$

If $C_0 \in S$ is any vector from S. such that

$$F(C_0) = B$$

we say that C_0 is a **particular solution** of the nonhomogeneous system $(*)'$. If $C \in S$ is another solution of $(*)'$, that is, $F(C) = B$, then

$$\begin{aligned} F(C - C_0) &= F(C) - F(C_0) \\ &= B - B \\ &= 0 \end{aligned}$$

Thus,

$$C - C_0 \in \text{Ker}(F)$$

and we see that

$$C \in C_0 + \text{Ker}(F), \quad \forall\, C \in S$$

Now we are ready to proceed with another practical algorithm for solving systems of linear equations.

Gauss–Jordan Method

Let's remind ourselves of the techniques and procedures that we have encountered before in a different context.

Elementary Transformations on a System of Linear Equations

(i) Interchange two equations.
(ii) Multiply an equation through by a nonzero constant.
(iii) Add a multiple of one equation to another equation.

Equivalently,

Elementary Row Operations on Matrices

(E_1) Interchange two rows of a matrix
(E_2) Multiply the elements of a row by a nonzero constant
(E_3) Add a multiple of the elements of one row to the corresponding elements of another row

The aforementioned procedures are formally justified by

Theorem 6.94 After a finite number of elementary transformations (i–iii) is performed on a linear system $\mathcal{L}$, the newly obtained system, $\mathcal{L}'$, is equivalent to $\mathcal{L}$.

Proof Equivalence under (i) and (ii) is obvious. We only need to show equivalence under transformation (iii).

Let our system $\mathcal{L}$ be

$$\sum_{k=1}^{n} \alpha_{ik} x_k = \beta_i, \qquad i = 1, \ldots, m \tag{*}$$

Suppose that we add the jth equation to the ith equation ($i \neq j$), to get a new system

$$\sum_{k=1}^{n} \alpha'_{ik} x_k = \beta'_i, \quad i = 1, \ldots, m \tag{**}$$

Note that (*) and (**) differ only in the ith equation, which now reads

$$\sum_{k=1}^{n} (\alpha_{ik} + \alpha_{jk}) x_k = \beta_i + \beta_j$$

If $(\gamma_1, \ldots, \gamma_n)$ is any solution of the system labeled (*), then

$$\sum_{k=1}^{n} \alpha_{ik} \gamma_k = \beta_i \quad \text{and} \quad \sum_{k=1}^{n} \alpha_{jk} \gamma_k = \beta_j$$

and thus,

$$\begin{aligned} \sum_{k=1}^{n} \alpha'_{ik} \gamma_k &= \sum_{k=1}^{n} (\alpha_{ik} + \alpha_{jk}) \gamma_k \\ &= \sum_{k=1}^{n} \alpha_{ik} \gamma_k + \sum_{k=1}^{n} \alpha_{jk} \gamma_k \\ &= \beta_i + \beta_j \\ &= \beta'_i \end{aligned}$$

We see that $(\gamma_1, \ldots, \gamma_n)$ is also a solution of (**). Conversely, if $(\gamma'_1, \ldots, \gamma'_n)$ is a solution of (**), we can write

$$\sum_{k=1}^{n} (\alpha_{ik} + \alpha_{jk}) \gamma'_k = \beta_i + \beta_j$$

and

$$\sum_{k=1}^{n} \alpha_{jk} \gamma'_k = \beta_j$$

Then,

$$\sum_{k=1}^{n}(\alpha_{ik}+\alpha_{jk})\gamma'_k - \sum_{k=1}^{n}\alpha_{jk}\gamma'_k = \beta_i + \beta_j - \beta_j$$
$$= \beta_i$$

and we see that $(\gamma'_1, \ldots, \gamma'_n)$ satisfies the ith equation of (*). Thus, $(\gamma'_1, \ldots, \gamma'_n)$ is a solution of (*). ■

Definition 6.73 Two systems of linear equations are said to be **equivalent systems** if one can be obtained from the other by a *finite* number of elementary transformations.

Similarly,

Definition 6.74 Two matrices are said to be **row equivalent** if one can be obtained from the other by a finite number of elementary row operations.

Consequently, the following theorem holds.

Theorem 6.95 Two systems of linear equations are equivalent iff their augmented matrices are row equivalent.

Finally, let's recall the concept of a matrix in (reduced) echelon form (Definition 6.57, 6.58): A matrix is in echelon form if

(i) All rows consisting entirely of zeros are at the bottom of the matrix.
(ii) The leading nonzero entry of each row is 1, and it is to the right of the leading entry in the preceding row.
(iii) All other elements in a column that contains a leading 1 are zero.

Equipped with the aforementioned technique, we are ready to move toward solving a linear system $AX = B$ by the **Gauss–Jordan method**.

Suppose a system of linear equations

$$\sum_{k=1}^{n}\alpha_{ik}x_k = \beta_i, \qquad i = 1, \ldots, m \tag{6.25}$$

is such that $r = r(A)$ is the rank of matrix A. We would like to replace (6.25) by an *equivalent one*, say,

$$\sum_{k=1}^{n}\alpha'_{ik}x_k = \beta'_i \qquad , i = 1, \ldots, r \tag{6.26}$$

which is simpler to solve. So, we take an augmented matrix of our system $\overline{A}$, whose rank is r, and transform it to its equivalent matrix $\overline{A'}$ in echelon form, from which we can easily extract the information we are looking for:

$$\overline{A} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & \cdots & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} & \beta_m \end{bmatrix} \sim \cdots$$

$$\sim \begin{bmatrix} 1 & 0 & \cdots & 0 & \alpha'_{1r+1} & \cdots & \alpha'_{1n} & \beta'_1 \\ 0 & 1 & \cdots & 0 & \alpha'_{2r+1} & \cdots & \alpha'_{2n} & \beta'_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \alpha'_{rr+1} & \cdots & \alpha'_{rn} & \beta'_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta'_{r+1} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \beta'_m \end{bmatrix} = \overline{A'} \tag{6.27}$$

$\overline{A'}$ is more advantageous for us since it corresponds to a rather simple system:

$$\begin{aligned} x_1 + \quad & \alpha'_{1r+1}x_{r+1} + \cdots + \alpha'_{1n}x_n = \beta'_1 \\ x_2 + \quad & \alpha'_{2r+1}x_{r+1} + \cdots + \alpha'_{2n}x_n = \beta'_2 \\ \cdots \quad & \cdots \quad \cdots \quad \cdots \\ x_r + \quad & \alpha'_{rr+1}x_{r+1} + \cdots + \alpha'_r x_n = \beta'_r \\ 0 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_r + & 0 \cdot x_{r+1} + \cdots + 0 \cdot x_n = \beta'_{r+1} \\ \cdots \quad \cdots \quad \cdots \quad & \cdots \quad \cdots \\ 0 \cdot x_1 + 0 \cdot x_2 + \cdots + \cdots + & \cdots + \cdots + 0 \cdot x_n = \beta'_m \end{aligned} \tag{6.28}$$

Keeping in mind that (6.28) is still equivalent to (6.25), we note:

(i) if any of $\beta'_{r+1}, \ldots, \beta'_m$ is different from zero then (6.28), and therefore (6.25), is unsolvable;

(ii) if, on the other hand, $\beta'_{r+1} = \beta'_{r+2} = \cdots = \beta'_m = 0$, (6.28) is solvable and is reducible to a simpler system with only r equations (while still equivalent to the original system (6.25)):

$$\begin{aligned} x_1 + \quad & \alpha'_{1r+1}x_{r+1} + \cdots + \alpha'_{1n}x_n = \beta'_1 \\ x_2 + \quad & \alpha'_{2r+1}x_{r+1} + \cdots + \alpha'_{2n}x_n = \beta'_2 \\ \cdots \quad & \cdots \quad \cdots \quad \cdots \quad \cdots \\ x_r + \quad & \alpha'_{rr+1}x_{r+1} + \cdots + \alpha'_r x_n = \beta'_r \end{aligned} \tag{6.29}$$

From (6.29), we can immediately read the following solutions:

$$\begin{aligned}
x_1 &= \beta'_1 - \alpha'_{1r+1}t_1 - \cdots - \alpha'_{1n}t_d \\
x_2 &= \beta'_2 - \alpha'_{2r+1}t_1 - \cdots - \alpha'_{2n}t_d \\
&\vdots \quad \vdots \qquad \vdots \qquad\qquad \vdots \\
x_r &= \beta'_r - \alpha'_{rr+1}t_1 - \cdots - \alpha'_r t_d \\
x_{r+1} &= t_1 \\
x_{r+2} &= t_2 \\
&\vdots \\
x_n &= t_d
\end{aligned} \tag{6.30}$$

where $d = n - r$, and $t_1, \ldots, t_d \in \Phi$ are free parameters.

Let's pause here for a while and work on a few examples.

Example 6.158 Solve the following system of linear equations:

$$\begin{aligned}
3x_1 - 3x_2 + 3x_3 &= 9 \\
2x_1 - x_2 + 4x_3 &= 7 \\
3x_1 - 5x_2 - x_3 &= 7
\end{aligned}$$

Solution

$$\overline{A} = \left[\begin{array}{ccc|c} 3 & -3 & 3 & 9 \\ 2 & -1 & 4 & 7 \\ 3 & -5 & -1 & 7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 1 & 3 \\ 2 & -1 & 4 & 7 \\ 3 & -5 & -1 & 7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 1 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & -2 & -4 & -2 \end{array}\right]$$
$$\sim \left[\begin{array}{ccc|c} 1 & 0 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

So we have

$$\begin{aligned}
x_1 + \quad 3x_3 &= 4 \\
x_2 + 2x_3 &= 1
\end{aligned}$$

that is,

$$\begin{aligned}
x_1 &= 4 - 3x_3 \\
x_2 &= 1 - 2x_3
\end{aligned}$$

Now we are free to assign an arbitrary value to x_3, say, $t \in \mathbf{R}$, and we have a **general solution** of our system

$$\begin{aligned} x_1 &= 4 - 3t \\ x_2 &= 1 - 2t \end{aligned}$$

■

Example 6.159 Solve the following system of linear equations:

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + \quad 2x_5 \quad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - \ 2x_4 + 4x_5 - 3x_6 &= -1 \\ 5x_3 + 10x_4 + \quad 15x_6 &= -1 \\ 2x_1 + 6x_2 + \quad 8x_4 + 4x_5 + 18x_6 &= 6 \end{aligned}$$

Solution

$$\overline{A} = \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{array}\right] \sim \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{array}\right]$$

$$\sim \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \end{array}\right] \sim \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{cccccc|c} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The corresponding equations are

$$\begin{aligned} x_1 + 3x_2 + \quad 4x_4 + 2x_5 \quad &= 0 \\ x_3 + 2x_4 \quad &= 0 \\ x_6 &= \frac{1}{3} \end{aligned}$$

and thus,

$$\begin{aligned} x_1 &= -3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= \frac{1}{3} \end{aligned}$$

Assigning parameters r, s, and t to free variables x_2, x_4, and x_5, respectively, we get the general solution

$$\begin{aligned} x_1 &= -3r - 4s - 2t \\ x_2 &= r \\ x_3 &= -2s \\ x_4 &= s \\ x_5 &= t \\ x_6 &= \frac{1}{3} \end{aligned}$$

which, in vector notation, is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix} + r \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

■

Example 6.160 Solve the following system of linear equations:

$$\begin{aligned} 8x_1 - 2x_2 + 4x_3 + 3x_4 + x_5 &= 2 \\ x_2 - 4x_3 + x_4 - 2x_5 &= -10 \\ 2x_1 + x_2 - 4x_4 &= 1 \end{aligned}$$

Solution

$$\overline{A} = \left[\begin{array}{ccccc|c} 8 & -2 & 4 & 3 & 1 & 2 \\ 0 & 1 & -4 & 1 & -2 & -10 \\ 2 & 1 & 0 & -4 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} 2 & 1 & 0 & -4 & 0 & 1 \\ 0 & 1 & -4 & 1 & -2 & -10 \\ 8 & -2 & 4 & 3 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 2 & 1 & 0 & -4 & 0 & 1 \\ 0 & 1 & -4 & 1 & -2 & -10 \\ 0 & -6 & 4 & 19 & 1 & -2 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} 2 & 0 & 4 & -5 & 2 & 11 \\ 0 & 1 & -4 & 1 & -2 & -10 \\ 0 & 0 & -20 & 25 & -11 & -62 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 2 & 0 & 0 & 0 & -1/5 & -7/5 \\ 0 & 1 & 0 & -4 & 1/5 & 12/5 \\ 0 & 0 & 1 & -25/20 & 11/20 & 62/20 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -1/10 & -7/10 \\ 0 & 1 & 0 & -4 & 1/5 & 12/5 \\ 0 & 0 & 1 & -5/4 & 11/20 & 62/20 \end{array}\right]$$

The corresponding equations are

$$\begin{aligned} x_1 \qquad\qquad\qquad - \frac{1}{10}x_5 &= -\frac{7}{10} \\ x_2 \qquad - 4x_4 + \frac{1}{5}x_5 &= \frac{12}{5} \\ x_3 - \frac{5}{4}x_4 + \frac{11}{20}x_5 &= \frac{62}{20} \end{aligned}$$

Thus, we have

$$\begin{aligned} x_1 &= -\frac{7}{10} + \frac{1}{10}x_5 \\ x_2 &= \frac{12}{5} + 4x_4 - \frac{1}{5}x_5 \\ x_3 &= \frac{31}{10} + \frac{5}{4}x_4 - \frac{11}{20}x_5 \end{aligned}$$

For the free variables x_4 and x_5, we choose s and t, respectively, and the general solution of our system is

$$\begin{aligned} x_1 &= -\frac{7}{10} + \frac{1}{10}t \\ x_2 &= \frac{12}{5} + 4s - \frac{1}{5}t \\ x_3 &= \frac{31}{10} + \frac{5}{4}s - \frac{11}{20}t \\ x_4 &= s \\ x_5 &= t \end{aligned}$$

In vector/matrix notation, the solution reads

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -7/10 \\ 12/5 \\ 31/10 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 4 \\ 5/4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1/10 \\ -1/5 \\ 11/20 \\ 0 \\ 1 \end{bmatrix}$$

■

Example 6.161 Solve the following system of linear equations:

$$x_1 + x_2 - 2x_3 + 3x_4 = 4$$
$$2x_1 + 3x_2 + 3x_3 - x_4 = 3$$
$$5x_1 + 7x_2 + 4x_3 + x_4 = 5$$

Solution The augmented matrix is

$$\overline{A} = \left[\begin{array}{cccc|c} 1 & 1 & -2 & 3 & 4 \\ 2 & 3 & 3 & -1 & 3 \\ 5 & 7 & 4 & 1 & 5 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 2 & 14 & -14 & -15 \end{array}\right]$$
$$\sim \left[\begin{array}{cccc|c} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{array}\right]$$

We can stop here and conclude that our system has no solution, since the third row corresponds to a degenerate equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -5$$

which, obviously, is inconsistent. ■

Continuing our discussion preceding the four aforementioned examples, let's make a few more observations: Looking closely at the general solution (6.30) of our equation $AX = B$, one immediately recognizes that we could express it as

$$X = C_0 + H$$

where C_0 is one particular solution and H is a general solution of the corresponding homogeneous system. Indeed, if we take all d parameters to be equal to zero, that is, $t_1 = t_2 = \cdots = t_d = 0$ then

$$C_0 = \begin{bmatrix} \beta_1' \\ \vdots \\ \beta_r' \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

is one particular solution of (6.30). We can continue choosing solutions by picking for the first, $t_1 = 1$, $t_2 = t_3 = \cdots = t_d = 0$; for the second, $t_1 = 0$, $t_2 = 1$, $t_3 = t_4 = \cdots = 0$; and so on, until finally for the dth solution, $t_1 = t_2 = \cdots = t_{d-1} = 0$, $t_d = 1$

The vectors C_i for $i = 1, \ldots, d$ are

$$C_1 = \begin{bmatrix} -\alpha_{1r+1} \\ \vdots \\ -\alpha_{rr+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} -\alpha_{1r+2} \\ \vdots \\ -\alpha_{rr+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots .., \quad C_d = \begin{bmatrix} -\alpha_{1n} \\ \vdots \\ -\alpha_{rn} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

and they are linearly independent. Since there are d of them, they form a fundamental set of solutions for a homogeneous system corresponding to (6.29). Evidently, then

$$X = C_0 + t_1 C_1 + \cdots + t_d C_d \tag{6.31}$$

is a general solution of (6.29). At the same time, (6.31) is also a **general** solution of (6.30). Needless to say, since systems (6.29), (6.28), and (6.25) are equivalent, (6.30) is the general solution of (6.25) that we have been searching for.

Example 6.162 Solve

$$\begin{aligned} x_1 + 2x_2 + 2x_3 + \quad\quad x_5 &= 3 \\ 2x_1 \quad\quad - x_3 - x_4 + 5x_5 &= 2 \\ x_1 + 2x_2 + 6x_3 - x_4 + 5x_5 &= 3 \\ x_1 - 2x_2 + 5x_3 - 12x_4 + 12x_5 &= -1 \end{aligned}$$

Solution

$$\overline{A} = \left[\begin{array}{ccccc|c} 1 & 2 & 2 & 3 & 1 & 3 \\ 2 & 0 & -1 & -1 & 5 & 2 \\ 1 & 2 & 6 & -1 & 5 & 3 \\ 1 & -2 & 5 & -12 & 12 & -1 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 2 & 2 & 3 & 1 & 3 \\ 0 & -4 & -5 & -7 & 3 & -4 \\ 0 & 0 & 4 & -4 & 4 & 0 \\ 0 & -4 & 3 & -15 & 11 & -4 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 2 & 2 & 3 & 1 & 3 \\ 0 & 1 & 5/4 & 7/4 & -3/4 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & -4 & 3 & -15 & 11 & -4 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & -1/2 & -1/2 & 5/2 & 1 \\ 0 & 1 & 5/4 & 7/4 & -3/4 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 8 & -8 & 8 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} 1 & 0 & 0 & -1 & 3 & 1 \\ 0 & 1 & 0 & 3 & -2 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

So we have

$$\begin{aligned} x_1 \qquad & - \; x_4 + 3x_5 = 1 \\ x_2 \quad & + 3x_4 - 2x_5 = 1 \\ x_3 - \; & x_4 + \; x_5 = 0 \end{aligned}$$

And the general solution is

$$\begin{aligned} x_1 &= 1 + s - 3t \\ x_2 &= 1 - 3s + 2t \\ x_3 &= s - t \\ x_4 &= s \\ x_5 &= t \end{aligned}$$

Written in vector/matrix notation, the solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ -3 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 2 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

■

6.18 THE BASICS OF EIGENVALUE AND EIGENVECTOR THEORY

On our journey through linear algebra, as in the studies of other mathematical disciplines (remember group theory), one especially welcomes the occurrence of **invariants**, that is, features that are uniquely characteristic to the particular mathematical object one is studying. You remember that in our studies of linear operators, for instance, the rank and defect were some of those unique characteristics that we found especially useful. Also, remember we called $\boldsymbol{L} < X$ an invariant subspace of a space $\boldsymbol{X}$, if under the action of an operator $F \in \text{Hom}\,(\boldsymbol{X}, \boldsymbol{X})$, $F\boldsymbol{l} \in \boldsymbol{L}$, $\forall \boldsymbol{l} \in \boldsymbol{L}$. Can we find the others? In particular, can we find an *invariant* by examining the matrix representation of an operator, independent of the selection of coordinate basis? And, of course, if we could, what would that tell us about the corresponding vector space? Well, suppose we start by asking how an operator behaves on the simplest invariant subspace of $\boldsymbol{X}$, that is, on a subspace of dimension 1. That space is easy to construct. Take any nonzero vector $\boldsymbol{x} \in \boldsymbol{X}$ and consider the set of all scalar multiples of $\boldsymbol{x}$, that is, $\boldsymbol{L} = \{\boldsymbol{l} = \lambda \boldsymbol{x} \mid \lambda \in \Phi\}$. $\boldsymbol{L}$ is obviously a subspace of $\boldsymbol{X}$, with respect to the same operations that make

$\boldsymbol{X}$ a vector space. Now, if $F \in \text{Hom}(\boldsymbol{X}, \boldsymbol{X})$ is such that $F\boldsymbol{l} \in \boldsymbol{L}$, then there exists $\lambda \in \Phi$ such that $F\boldsymbol{l} = \lambda \boldsymbol{l}$. With these ideas in mind, why not move forward and be more general?

Definition 6.75 Let $F : \boldsymbol{X} \to \boldsymbol{X}$ be an endomorphism of vector spaces, and let $A = A_F$ be a matrix representation of F, such that

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \tag{*}$$

where $\boldsymbol{x} \in \boldsymbol{X}$ is some *nonzero vector*,[11] and $\lambda \in \Phi$ some scalar. We say that $\boldsymbol{x}$ is an **eigenvector** of A, and λ is an **eigenvalue** of A. That is, an operator has an eigenvalue iff there exists a nonzero vector that gets sent by the operator to a scalar multiple of itself.

Example 6.163 Consider a vector space $\mathbf{R}^2$, and a vector $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \in \mathbf{R}^2$. Then $\boldsymbol{x}$ is an eigenvector of $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$ with eigenvalue $\lambda = 3$.

Indeed,

$$\begin{aligned} A\boldsymbol{x} &= \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ &= 3\boldsymbol{x} \end{aligned}$$

■

Example 6.164 Let $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ be a matrix representing some endomorphism $\mathbf{R}^2 \to \mathbf{R}^2$. Then, $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbf{R}^2$ is an eigenvector of A with eigenvalue $\lambda = 4$. Indeed,

$$\begin{aligned} A\boldsymbol{x} &= \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= 4\boldsymbol{x} \end{aligned}$$

■

[11] Note that we must require $\boldsymbol{x} \neq 0$, otherwise every scalar would satisfy equation (*).

Example 6.165 Consider $\Lambda = \begin{bmatrix} \lambda & 0 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} = \lambda I,\ \lambda \in \Phi$. Then, for any $\boldsymbol{x} \in \boldsymbol{X},\ \Lambda \boldsymbol{x} = \lambda \boldsymbol{x}$.

In other words, any vector is an eigenvector of Λ. ■

Definition 6.76 Let $A = (\alpha_{ik})$ be an (n, n) matrix over a field Φ. A matrix

$$C = A - \lambda I = \begin{bmatrix} \alpha_{11} - \lambda & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} - \lambda & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \lambda \end{bmatrix}$$

is called the **characteristic matrix** of A.

Definition 6.77 The determinant

$$\begin{aligned} \det C &= \det(A - \lambda I) \\ &= a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0 \\ &= k_A(\lambda), \qquad a_i \in \Phi \end{aligned}$$

is called the **characteristic polynomial** of A.

The equation

$$\begin{aligned} \det C &= \det(A - \lambda I) \\ &= a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0 = 0 \end{aligned}$$

is called the **characteristic equation** of A.

Example 6.166 Find the characteristic polynomial $k_A(\lambda)$ for

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

Solution

$$\begin{aligned} k_A(\lambda) = \det(A - \lambda I) &= \begin{vmatrix} 1 - \lambda & 2 \\ 3 & 4 - \lambda \end{vmatrix} \\ &= (1 - \lambda)(4 - \lambda) - 6 \\ &= \lambda^2 - 5\lambda - 2 \end{aligned}$$

■

Example 6.167 Find the characteristic polynomial for

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

Solution

$$k_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 4 & -17 & 8-\lambda \end{vmatrix}$$

$$= -\lambda^3 + 8\lambda^2 + 17\lambda + 4$$ ■

Example/Exercise 6.168 Find the characteristic polynomial $k_A(\lambda)$ if

$$A = \begin{bmatrix} 3 & 0 & 0 & 13 \\ -25 & 7 & 11 & -6 \\ 18 & 0 & 1 & 5 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

In general, as the aforementioned examples show, the coefficients of the characteristic polynomial are succinctly given by the following

Theorem 6.96 The coefficients of the characteristic polynomial of a matrix A are as follows:

$$a_n = (-1)^n, \quad a_{n-1} = (-1)^{n-1}\mathrm{tr}A, \quad a_0 = \det(A)$$

Theorem 6.97 Let $F : \boldsymbol{X} \to \boldsymbol{X}$ be a linear operator whose matrix representation is $A = A_F$. Then the following are equivalent:

(i) A scalar $\lambda \in \Phi$ is an eigenvalue of A (eigenvalue of F).
(ii) The matrix $A - \lambda I$ is singular (the linear operator $F - \lambda I$ is singular).
(iii) $\lambda \in \Phi$ is a root of the characteristic polynomial $k_A(\lambda)$ ($\lambda \in \Phi$ is a root of the characteristic polynomial $k_F(\lambda)$).

Theorem 6.98 An (n, n) matrix A is invertible iff $\lambda = 0$ is not an eigenvalue of A.

Proof Let A be an (n, n) matrix whose characteristic polynomial is

$$k_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 \qquad (*)$$

We immediately realize that $\lambda = 0$ is an eigenvalue of A if $a_0 = 0$.

On the other hand, recall that A is invertible if $\det A \neq 0$. So, consider

$$\det(A - \lambda I) = k_A(\lambda)$$

which, with $\lambda = 0$ becomes

$$\begin{aligned}\det(A - \lambda I) &= \det A \\ &= a_0\end{aligned}$$

Thus, $\det A = 0$ iff $a_0 = 0$, in which case A is not invertible. Conversely, if $\lambda = 0$ is not an eigenvalue of A, then $\det A = a_0 \neq 0$, and A is an invertible matrix. ■

Theorem 6.99 If λ is an eigenvalue of an invertible linear operator F, then λ^{-1} is an eigenvalue of F^{-1}.

Proof Since F is invertible, by Theorem 6.98 $\lambda \neq 0$. Then there exists an eigenvector $\boldsymbol{x} \neq 0$, such that $F\boldsymbol{x} = \lambda\boldsymbol{x}$. But then

$$\begin{aligned}F^{-1}F\boldsymbol{x} &= F^{-1}(\lambda\boldsymbol{x}) \\ &= \lambda F^{-1}\boldsymbol{x}\end{aligned}$$

Thus,

$$\boldsymbol{x} = \lambda F^{-1}\boldsymbol{x}$$

and so,

$$\begin{aligned}\lambda^{-1}\boldsymbol{x} &= (\lambda^{-1}\lambda)\, F^{-1}\boldsymbol{x} \\ &= F^{-1}\boldsymbol{x}\end{aligned}$$

Therefore, λ^{-1} is an eigenvalue of F^{-1}. ■

Theorem 6.100 Similar matrices have the same characteristic polynomial.

Proof Recall: two (n, n)-matrices A and B are said to be similar, $A \simeq B$, if there exists a regular (n, n) matrix T (a "*change-of-basis*" *matrix*), such that

$$B = T^{-1}AT$$

So,

$$\begin{aligned}
\det(B - \lambda I) &= \det(T^{-1}AT - \lambda T^{-1}T) \\
&= \det(T^{-1}AT - T^{-1}(\lambda I)T) \\
&= \det(T^{-1}(A - \lambda I)T) \\
&= \det(T^{-1}) \cdot \det(A - \lambda I) \cdot \det(T) \\
&= \det(T^{-1}) \cdot \det(T) \cdot \det(A - \lambda I) \\
&= \det(T^{-1}T) \cdot \det(A - \lambda I) \\
&= 1 \cdot \det(A - \lambda I) \\
&= \det(A - \lambda I)
\end{aligned}$$

■

The important conclusion is this:

If $F : \boldsymbol{X} \to \boldsymbol{X}$ is any linear operator and if A_F and A'_F are two matrix representations of F in two different bases, then, since $A_F \simeq A'_F$, the characteristic polynomials are basis independent, that is,

$$k_A(\lambda) = k_{A'}(\lambda)$$

Finally, we need to answer the naturally occurring questions of finding the eigenvalues and eigenvectors for a given operator/matrix.

Example 6.169 Let's find the eigenvalues and eigenvectors of an operator whose matrix representation is

$$A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$$

We need to solve the characteristic equation $k_A(\lambda) = 0$, that is,

$$\begin{aligned}
k_A(\lambda) &= \det(A - \lambda I) = 0 \\
&= \begin{vmatrix} -4 - \lambda & -6 \\ 3 & 5 - \lambda \end{vmatrix} \\
&= \lambda^2 - \lambda - 2 = 0 \\
&= (\lambda - 2)(\lambda + 1) = 0
\end{aligned}$$

Thus, $\lambda_1 = 2$ and $\lambda_2 = -1$ are the eigenvalues of A. To find the eigenvectors, we need to solve

$$(A - \lambda_1 I)\boldsymbol{x} = \boldsymbol{0}$$

and

$$(A - \lambda_2 I)\boldsymbol{x} = \mathbf{0}$$

So, for $\lambda_1 = 2$, we have

$$(A - \lambda_1 I)\boldsymbol{x} = \begin{bmatrix} -4-2 & -6 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

that is,

$$\begin{aligned} -6x_1 - 6x_2 &= 0 \\ 3x_1 + 3x_2 &= 0 \end{aligned}$$

Thus, $x_1 = -x_2 = r$, where r is some scalar, and the eigenvector of A, corresponding to eigenvalue $\lambda_1 = 2$, is $\boldsymbol{x}_{\lambda_1} = r \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Similarly, for $\lambda_2 = -1$, we have

$$(A - \lambda_2 I)\boldsymbol{x} = \begin{bmatrix} -4+1 & -6 \\ 3 & 5+1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

that is,

$$\begin{aligned} -3x_1 - 6x_2 &= 0 \\ 3x_1 + 6x_2 &= 0 \end{aligned}$$

Solving this we get $x_1 = -2x_2 = -2s$, where $s \in \mathbf{R}$ is some scalar. The corresponding eigenvector of A is $\boldsymbol{x}_{\lambda_2} = s \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.[12] ■

Example 6.170 Let's consider again $A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$, from Example 6.163, and let's find the corresponding eigenvalues and eigenvectors.

$$\begin{aligned} k_A(\lambda) = \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 0 \\ 8 & -1-\lambda \end{vmatrix} \\ &= (3 - \lambda)(-1 - \lambda) \end{aligned}$$

[12]From now on, in order to avoid a clutter of subscripts, instead of writing $\boldsymbol{x}_{\lambda_i}$, we will simply write $\boldsymbol{x}_i$. I hope this won't cause any confusion.

Thus,

$$(3-\lambda)(-1-\lambda)=0$$

gives the eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -1$.

Next, we calculate

$$(A-\lambda I)\boldsymbol{x}=\mathbf{0}$$

$$\begin{bmatrix} 3-\lambda & 0 \\ 8 & -1-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

For $\lambda_1 = 3$, we get

$$\begin{bmatrix} 0 & 0 \\ 8 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

that is, $2x_1 = x_2 = r$, where $r \in \boldsymbol{R}$ is some scalar. Hence, the corresponding eigenvector is $\boldsymbol{x}_1 = r\begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$. With $r = 2$, we get $\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. ■

Example 6.171 Let

$$A=\begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix}$$

be the matrix representation of some linear operator $F : \mathbf{R}^3 \to \mathbf{R}^3$. Find the corresponding eigenvectors and eigenvalues.

Solution

$$\begin{aligned} k_A(\lambda) &= \det(A-\lambda I) \\ &= \begin{vmatrix} -1-\lambda & 2 & 2 \\ 2 & 2-\lambda & 2 \\ -3 & -6 & -6-\lambda \end{vmatrix} \\ &= -\lambda(\lambda+2)(\lambda+3) \end{aligned}$$

Solving $k_A(\lambda) = 0$ we get $\lambda_1 = 0, \quad \lambda_2 = -2, \quad \lambda_3 = -3$. Now we find the corresponding eigenvectors.

$$(A-\lambda I)\boldsymbol{x}=\mathbf{0}$$

For $\lambda_1 = 0$,

$$\begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

that is,

$$\begin{aligned} -x_1 + 2x_2 + 2x_3 &= 0 \\ 2x_1 + 2x_2 + 2x_3 &= 0 \\ -3x_1 - 6x_2 - 6x_3 &= 0 \end{aligned}$$

This gives us

$$x_1 = 0, \quad x_2 = r, \quad x_3 = -r, \qquad r \in \mathbf{R}$$

The corresponding eigenvector is $\boldsymbol{x}_1 = r\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

Similarly, for $\lambda_2 = -2$ we get $\boldsymbol{x}_2 = s\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, and for λ_3, $\boldsymbol{x}_3 = t\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $s, t \in \mathbf{R}$.

With this, we have obtained three characteristic subspaces $\boldsymbol{L}_1, \boldsymbol{L}_2, \boldsymbol{L}_3 < \mathbf{R}^3$ corresponding to eigenvalues λ_1, λ_2, and λ_3:

$$\boldsymbol{L}_1(\lambda_1) = \left\{ r\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \mid \quad r \in \mathbf{R} \right\}$$

$$\boldsymbol{L}_2(\lambda_2) = \left\{ s\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \mid \quad s \in \mathbf{R} \right\}$$

$$\boldsymbol{L}_3(\lambda_3) = \left\{ t\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \mid \quad t \in \mathbf{R} \right\}$$

■

As you might have expected, we appropriately call the subspaces occurring in the aforementioned examples the *eigenspaces* corresponding to eigenvalues λ_1, λ_2, and λ_3. Thus,

Definition 6.78 Let $F : \boldsymbol{X} \to \boldsymbol{X}$ be a linear operator, such that $\lambda_i \in \Phi$ are eigenvalues corresponding to eigenvectors $\boldsymbol{x}_i \in \boldsymbol{X}$. We say that the set

$$\boldsymbol{L}_i = \{\boldsymbol{x}_i \mid F(\boldsymbol{x}_i) = \lambda_i \boldsymbol{x}_i\}$$

is an **eigenspace** of F corresponding to eigenvalue λ_i.

We end this section with two motivating theorems.

Theorem 6.101 Every matrix A (every linear operator F) is a zero of its characteristic polynomial.

Theorem 6.102 If $F : \boldsymbol{X} \to \boldsymbol{X}$ is a linear operator whose matrix representation is A with different eigenvalues $\lambda_1, \ldots, \lambda_k$ corresponding to eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \boldsymbol{X}$, respectively, then the set $\boldsymbol{L} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ is the set of linearly independent vectors.

Proof Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \boldsymbol{X}$ be the eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_k$. Suppose the vectors in $\boldsymbol{L} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ are not linearly independent. Any eigenvector by definition is linearly independent, thus $\boldsymbol{L}_1 = \{\boldsymbol{x}_1\}$ is linearly independent. Let $r \in \mathbf{N}$, $1 \leq r < k$, be the largest number, such that $\boldsymbol{L}_r = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\}$ is linearly independent. Then, $\boldsymbol{L}_{r+1} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r, \boldsymbol{x}_{r+1}\}$ is certainly linearly dependent and there are $r+1$ scalars $\alpha_i \in \Phi$, not all equal to zero, such that

$$\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_r \boldsymbol{x}_r + \alpha_{r+1} \boldsymbol{x}_{r+1} = \boldsymbol{0} \qquad (*)$$

Multiplying both sides by A we get

$$\alpha_1 A\boldsymbol{x}_1 + \cdots + \alpha_r A\boldsymbol{x}_r + \alpha_{r+1} A\boldsymbol{x}_{r+1} = \boldsymbol{0}$$

and, keeping in mind that

$$A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1, \ \ldots, \ A\boldsymbol{x}_{r+1} = \lambda_{r+1}\, \boldsymbol{x}_{r+1}$$

we obtain

$$\alpha_1 \lambda_1 \boldsymbol{x}_1 + \cdots + \alpha_{r+1} \lambda_{r+1} \boldsymbol{x}_{r+1} = \boldsymbol{0} \qquad (**)$$

Now, let's multiply (*) by λ_{r+1} and subtract the resulting equation from (**), that is,

$$\alpha_1(\lambda_1 - \lambda_{1+r})\boldsymbol{x}_1 + \alpha_2(\lambda_2 - \lambda_{r+1})\boldsymbol{x}_2 + \cdots + \alpha_r(\lambda_r - \lambda_{r+1})\boldsymbol{x}_r = \boldsymbol{0}$$

But, remember, $\boldsymbol{L}_r = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\}$ is a set of linearly independent vectors, so

$$\alpha_1(\lambda_1 - \lambda_{1+r}) = \alpha_2(\lambda_2 - \lambda_{r+1}) = \cdots = \alpha_r(\lambda_r - \lambda_{r+1}) = 0$$

And, since all λ_i's are assumed distinct, it follows that

$$\alpha_1 = \alpha_2 = \cdots = \alpha_r = 0$$

Substituting these values in (*) yields

$$\alpha_{r+1}\boldsymbol{x}_{r+1} = \mathbf{0}$$

Since eigenvector $\boldsymbol{x}_{r+1} \neq 0$, we conclude that $\alpha_{r+1} = 0$. This contradicts our assumption that not all $\alpha_1, \ldots, \alpha_{r+1}$ are zero, and with this we complete our proof. ■

Do all matrices $A = (\alpha_{ij})$, $\alpha_{ij} \in \mathbf{R}$ have (real) eigenvalues? Unfortunately not! However, if A is a real symmetric matrix, then the following proposition is true.

Theorem 6.103 If $A = (\alpha_{ij})$, $\alpha_{ij} \in \mathbf{R}$, is a symmetric matrix, then all eigenvalues are real (the characteristic polynomial has all real zeros).

Theorem 6.104 If $A = (\alpha_{ij})$, $\alpha_{ij} \in \mathbf{R}$, is a matrix with n distinct eigenvalues, then the corresponding eigenvectors constitute a basis for $\mathbf{R}^n$.

Theorem 6.105 If $\lambda \in \Phi$ is an eigenvalue of a matrix A and $\boldsymbol{x}$ is the corresponding eigenvector, then for every $n \in \mathbf{N}$, λ^n is an eigenvalue of A^n for the corresponding eigenvector $\boldsymbol{x}$.

Proof If $A\boldsymbol{x} = \lambda\boldsymbol{x}$, then clearly for $n = 2$

$$\begin{aligned} A^2\boldsymbol{x} &= A(A\,\boldsymbol{x}) \\ &= A\lambda\boldsymbol{x} \\ &= \lambda(A\boldsymbol{x}) \\ &= \lambda(\lambda\boldsymbol{x}) \\ &= \lambda^2\boldsymbol{x} \end{aligned}$$

It follows that, assuming that the same holds for $n = k$,

$$\begin{aligned} A^{k+1}\boldsymbol{x} &= A(A^k\boldsymbol{x}) \\ &= A(\lambda^k\boldsymbol{x}) \\ &= \lambda^k(A\boldsymbol{x}) \\ &= \lambda^k(\lambda\boldsymbol{x}) \\ &= \lambda^{k+1}\boldsymbol{x} \end{aligned}$$

Hence, $A^n\boldsymbol{x} = \lambda^n\boldsymbol{x}$ for every $n \in \mathbf{N}$, as claimed. ■

Diagonalization, Eigenvalues, and Eigenvectors

Let $F : \boldsymbol{X} \rightarrow \boldsymbol{X}$ be a linear operator with several matrix representations. One would like to find a basis of $\boldsymbol{X}$, with respect to which the matrix representation of F is the simplest possible. As we have seen so far the simplest matrix representation of an operator is the one in which the matrix is diagonal.

Definition 6.79 A linear operator $F : \boldsymbol{X} \rightarrow \boldsymbol{X}$ is said to be **diagonalizable** if there exists a basis $\boldsymbol{B_X}$ with respect to which the matrix representation $[F]_{\boldsymbol{B_X}}$ is diagonal. Diagonalizable operators are sometimes also called **semi-simple operators**.

Keeping in mind our definition of similar matrices, we also have

Definition 6.80 An (n, n) matrix A is said to be **diagonalizable** if there exists a matrix T, such that

$$B = T^{-1}AT$$

is a diagonal matrix.

Equivalently, a matrix A is diagonalizable if there exists an invertible matrix T and a diagonal matrix B, such that $A = TBT^{-1}$.

Theorem 6.106 A linear operator $F : \boldsymbol{X} \rightarrow \boldsymbol{X}$ is diagonalizable iff there exists a basis

$$\boldsymbol{B_X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} < \boldsymbol{X}$$

such that all $\boldsymbol{x}_i \in \boldsymbol{B_X}$ are eigenvectors of F.

Proof Let $\boldsymbol{B_X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ be a basis, such that every $\boldsymbol{x}_i$ is an eigenvector of F corresponding to eigenvalue λ_i. Then,

$$F(\boldsymbol{x}_i) = \lambda_i \boldsymbol{x}_i$$

The matrix representation of F in this basis is clearly

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}, \quad \lambda_i \in \Phi$$

that is, a diagonal matrix. Conversely, if there is a basis $\boldsymbol{B_X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$, such that the matrix representation of F in this basis is a diagonal matrix, that is,

$$[F]_{\boldsymbol{B_X}} = \begin{bmatrix} \beta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \beta_n \end{bmatrix}, \quad \beta_i \in \Phi$$

then,

$$F(\boldsymbol{x}_i) = \beta_i \boldsymbol{x}_i, \qquad i = 1, \ldots, n$$

We see that the vectors $\boldsymbol{x}_i \in \boldsymbol{B}_X$ are eigenvectors of F with corresponding eigenvalues β_i. ■

In other words, if a linear operator $F : \boldsymbol{X} \to \boldsymbol{X}$ is diagonalizable, that is, its matrix representation is diagonalizable, then the original system of linear equations is reduced to a simple system

$$\begin{aligned} F(\boldsymbol{x}_1) &= \beta_1 \boldsymbol{x}_1 \\ F(\boldsymbol{x}_2) &= \qquad \beta_2 \boldsymbol{x}_2 \\ \vdots \qquad & \qquad\qquad \ddots \\ F(\boldsymbol{x}_n) &= \qquad\qquad\qquad \beta_n \boldsymbol{x}_n \end{aligned}$$

Formally, we state this as

Theorem 6.107 If an operator $F : \boldsymbol{X} \to \boldsymbol{X}$ is diagonalizable, then its diagonal matrix representation is a matrix whose (diagonal) elements are eigenvalues corresponding to eigenvectors of F.

Equivalently,

Theorem 6.108 If a matrix A has n distinct eigenvalues, then A is diagonalizable.

Example 6.172 Let $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$. Find a matrix T so that $T^{-1}AT$ is diagonal.

Solution First, we find the eigenvalues λ.

Solving

$$\det(A - \lambda I) = 0$$

$$\begin{vmatrix} 1-\lambda & 1 \\ -2 & 4-\lambda \end{vmatrix} = 0$$

$$= (1 - \lambda)(4 - \lambda) = 0$$

that is,

$$\lambda^2 - 5\lambda + 6 = 0$$

so. $\lambda_1 = 2, \quad \lambda_2 = 3.$

Next, we need to find the corresponding eigenvectors:
For $\lambda_1 = 2$:

$$\begin{bmatrix} 1-2 & 1 \\ -2 & 4-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus,

$$-x_1 + x_2 = 0$$
$$-2x_1 + 2x_2 = 0$$

and we have $x_1 = x_2$. So, the eigenvector corresponding to $\lambda_1 = 2$ is

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Similarly for $\lambda_2 = 3$:

$$\begin{bmatrix} 1-3 & 1 \\ -2 & 4-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

we get

$$-2x_1 + x_2 = 0$$

that is, $2x_1 = x_2$, and the eigenvector corresponding to the eigenvalue $\lambda_2 = 3$ is

$$\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Evidently $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ are linearly independent, and thus we can diagonalize our matrix A. Let T be a matrix whose columns are $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, that is,

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

T^{-1} turns out to be (you should check this)

$$T^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

Finally, we get

$$T^{-1}AT = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Of course, we could have picked $\lambda_1 = 3$ and $\lambda_2 = 2$ and, following the same procedure (you should do this too), we would have obtained

$$T^{-1}AT = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$ ■

Example 6.173 Let's consider the matrix from Example 6.171 again.

$$A = \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix}$$

We have already found three linearly independent vectors

$$\boldsymbol{x}_1 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

So, we expect to diagonalize matrix A with matrix $T\,T = \begin{bmatrix} 0 & 2 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$. We also need $T^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ -1 & -2 & -2 \end{bmatrix}$ so that we can diagonalize matrix A:

$$T^{-1}AT = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ -1 & -2 & -2 \end{bmatrix} \begin{bmatrix} -1 & 2 & 2 \\ 2 & 2 & 2 \\ -3 & -6 & -6 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

where the eigenvalues $\lambda_1 = 0$, $\lambda_2 = -2$, and $\lambda_3 = -3$ are exactly those obtained in Example 6.18.9. ■

The previous example leads us to the following conclusion.

Theorem 6.109 Let A be a matrix representation of a linear operator $F : \boldsymbol{X} \to \boldsymbol{X}$, that is,

$$A = [F]_{\boldsymbol{B}_{X'}}$$

and let

$$A \simeq D = T^{-1}AT$$

Then, T is a matrix whose columns are the coordinates of the eigenvectors in the same basis $\boldsymbol{B}$.

Example 6.174 Let $F : \boldsymbol{X} \to \boldsymbol{X}$ be a linear operator whose matrix representation is

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$

Let's find the bases for the eigenspaces $\boldsymbol{B}_{\boldsymbol{X}}$, whose vectors are columns of the matrix that diagonalizes A, and let's find

$$D = T^{-1}AT$$

Solution

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 0-\lambda & 0 & -2 \\ 1 & 2-\lambda & 1 \\ 1 & 0 & 3-\lambda \end{vmatrix} \\ &= \lambda^3 - 5\lambda^2 + 8\lambda - 4 \\ &= (\lambda - 1)(\lambda - 2)^2 = 0 \end{aligned}$$

So, the only two distinct eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = 2$. The corresponding eigenvectors follow:

For $\lambda_1 = 1$, we have

$$\begin{bmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and the system of linear equations is

$$\begin{aligned} -x_1 - 2x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_3 &= 0 \end{aligned}$$

Solving this system we get

$$x_1 = -2t, \quad x_2 = t, \quad x_3 = t$$

Thus, the eigenvectors corresponding to $\lambda_1 = 1$ are vectors of the form $t\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$.

So, the basis vector is $\boldsymbol{x}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$.

For $\lambda_2 = 2$, we have

$$\begin{bmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

that is,

$$-2x_1 - 2x_3 = 0$$
$$x_1 + x_3 = 0$$

Solving this system yields

$$x_1 = -s, \quad x_2 = t, \quad x_3 = s$$

The eigenvectors corresponding to $\lambda_2 = 2$ are vectors of the form

$$\begin{bmatrix} -s \\ t \\ s \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

It is easy to verify that vectors $\boldsymbol{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ and $\boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ are linearly independent and thus they form a basis for the eigenspace corresponding to $\lambda_2 = 2$. Summing up, the bases for the eigenspaces are formed by

$$\boldsymbol{x}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

and we expect that the matrix that diagonalizes A is

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

Let's see! We need

$$T^{-1} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}$$

to calculate

$$\begin{aligned} T^{-1}AT &= \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \\ &= D \end{aligned}$$

■

The technique of diagonalization can sometimes be the best way to compute the powers of a matrix. Consider the following:

Let $A = (\alpha_{ik})$ be a diagonalizable (n, n) matrix, and let T be the matrix that diagonalizes A. Suppose we want to compute A^p where $p \in \mathbf{N}$. Well, if A is diagonalizable by T, then

$$D = T^{-1}AT = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

and certainly

$$D^2 = (T^{-1}AT)^2 = \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix}$$

But also,

$$\begin{aligned} (T^{-1}AT)^2 &= (T^{-1}AT)(T^{-1}AT) \\ &= T^{-1}(ATT^{-1})AT \\ &= T^{-1}AAT \\ &= T^{-1}A^2T \end{aligned}$$

So, $D^2 = T^{-1}A^2T$. Proceeding in the same way, we can obtain $D^p = T^{-1}A^pT$.

Hence,

$$A^p = TD^pT^{-1}$$
$$= T\begin{bmatrix} \lambda_1^p & 0 & \cdots & 0 \\ 0 & \lambda_2^p & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n^p \end{bmatrix} T^{-1}$$

Example 6.175 Compute A^9 if

$$A = \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix}$$

Solution

$$\det(A - \lambda I) = \begin{vmatrix} -4-\lambda & -6 \\ 3 & 5-\lambda \end{vmatrix}$$
$$= \lambda^2 - \lambda - 2$$

Solving the characteristic equation

$$(\lambda - 2)(\lambda + 1) = 0$$

we get the eigenvalues of A: $\lambda_1 = 2$, and $\lambda_2 = -1$.

Next, we compute the eigenvectors:

For $\lambda_1 = 2$:

$$\begin{bmatrix} -4-2 & -6 \\ 3 & 5-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

implies

$$-6x_1 - 6x_2 = 0$$
$$3x_1 + 3x_2 = 0$$

and thus $x_1 = -x_2$, and the eigenvectors corresponding to $\lambda_1 = 2$ are $r\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Similarly, for $\lambda_2 = -1$:

$$\begin{bmatrix} -4+1 & -6 \\ 3 & 5+1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

yields

$$-3x_1 - 6x_2 = 0$$
$$3x_1 + 6x_2 = 0$$

and we have $x_1 = -2x_2$. The corresponding eigenvectors are $s\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Now, we know that a matrix

$$T = \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$$

diagonalizes matrix A, that is,

$$\begin{aligned} T^{-1}AT &= \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -4 & -6 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} = D \end{aligned}$$

To complete our calculation, we will need $D^9 = \begin{bmatrix} 512 & 0 \\ 0 & -1 \end{bmatrix}$. So, finally, we have

$$\begin{aligned} A^9 &= TD^9T^{-1} \\ &= \begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 512 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -514 & -1026 \\ 513 & 1025 \end{bmatrix} \end{aligned}$$

■

I admit this is rather boring, but one has to do it at least a couple of times in one's lifetime to get some proficiency and understanding of the underlying principles. Consequently, here is another example, with some steps omitted, that you should complete yourself.

Example 6.176 Find A^{13} if

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Solution A matrix

$$T = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

diagonalizes A and its inverse is

$$T^{-1} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix}$$

So,

$$D = T^{-1}AT = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus,

$$\begin{aligned} A^{13} &= TD^{13}T^{-1} \\ &= \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2^{13} & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 1^{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix} \end{aligned}$$

■

Theorem 6.110 Let $F : \boldsymbol{X} \to \boldsymbol{X}$ be a linear operator and A its (n, n) matrix representation. Then, the following statements are equivalent:

(i) A is diagonalizable.
(ii) There are n linearly independent eigenvectors corresponding to A.

Proof (i)⇒(ii):

If A is diagonalizable, then there exists a matrix T such that

$$D = T^{-1}AT$$

is a diagonal matrix. Then, of course,

$$TD = AT$$

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ be column vectors of T. Note that, since T is regular, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent. Furthermore, let $\lambda_1, \ldots, \lambda_n$ be diagonal entries of D, then

$$\begin{aligned} TD &= [\boldsymbol{x}_1\lambda_1 \quad \boldsymbol{x}_2\lambda_2 \quad \cdots \quad \boldsymbol{x}_n\lambda_n] \\ &= [\lambda_1\boldsymbol{x}_1 \quad \lambda_2\boldsymbol{x}_2 \quad \cdots \quad \lambda_n\boldsymbol{x}_n] \end{aligned}$$

and

$$\begin{aligned} AT &= A[\boldsymbol{x}_1 \;\; \boldsymbol{x}_2 \;\; \cdots \;\; \boldsymbol{x}_n] \\ &= [A\boldsymbol{x}_1 \;\; A\boldsymbol{x}_2 \;\; \cdots \; A\boldsymbol{x}_n] \end{aligned}$$

So, we have

$$A\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1, \quad A\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2, \; \ldots \; , A\boldsymbol{x}_n = \lambda_n \boldsymbol{x}_n$$

And thus, being linearly independent vectors, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are the eigenvectors of A.

(ii)⇒(i):

Suppose $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are linearly independent eigenvectors of A corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Let

$$T = [\boldsymbol{x}_1 \;\; \boldsymbol{x}_2 \;\; \cdots \;\; \boldsymbol{x}_n]$$

and let D be the diagonal matrix whose entries are $\lambda_1, \ldots, \lambda_n$. Then,

$$\begin{aligned} AT &= A[\, \boldsymbol{x}_1 \;\; \boldsymbol{x}_2 \;\; \cdots \;\; \boldsymbol{x}_n] \\ &= [A\boldsymbol{x}_1 \;\; A\boldsymbol{x}_2 \;\; \cdots \; A\boldsymbol{x}_n] \\ &= [\lambda_1 \boldsymbol{x}_1 \;\; \lambda_2 \boldsymbol{x}_2 \;\; \cdots \;\; \lambda_n \boldsymbol{x}_n] \\ &= TD \end{aligned}$$

But, T is a regular matrix and therefore,

$$T^{-1}AT = D$$

■

Finally, we would like to address the question of linear operators whose matrix representation is not diagonalizable. For example, even a very simple matrix $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ is not diagonalizable (as you can easily check).

Exercise 6.1 Determine whether the following matrix:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}$$

is diagonalizable or not.

Solution As usual, we first calculate the characteristic polynomial

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & 2-\lambda & 0 \\ -3 & 5 & 2-\lambda \end{vmatrix}$$
$$= (1-\lambda)(2-\lambda)^2$$

The solutions of the characteristic equation are $\lambda_1 = 1$ and $\lambda_2 = 2$. These eigenvalues, as you can easily check, correspond to eigenvectors,

$$\boldsymbol{x}_1 = \begin{bmatrix} 1/8 \\ -1/8 \\ 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

respectively. Since our matrix A is of type (n, n), and we have only two basis vectors, we conclude that A is not diagonalizable.

Example/Exercise 6.177 Convince yourself that the following matrices:

$$A = \begin{bmatrix} 5 & -3 \\ 3 & -1 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 3 & -1 \\ -1 & 2 & 0 \end{bmatrix}$$

are not diagonalizable.

6.19 SUPPLEMENTARY PROBLEMS

1. Let $\boldsymbol{X} = (\mathbf{R}^+;\ +,\ \cdot)$ be a set where operations on $\mathbf{R}^+$ are defined as follows:

$$x + y = x \cdot y, \quad \forall\, x, y \in \mathbf{R}^+$$
$$\lambda \cdot x = x^\lambda, \quad \forall\, x \in \mathbf{R}^+, \quad \lambda \in \mathbf{R}$$

Determine whether $\boldsymbol{X}$ is a vector space or not.

2. Let $\boldsymbol{X} = (\mathbf{R}^2;\ +,\ \cdot)$ where

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1,\ x_2 + y_2), \quad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$$

and

$$\lambda \boldsymbol{x} = (\lambda x_1, 0), \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \forall \lambda \in \mathbf{R}$$

Determine whether $\boldsymbol{X}$ is a vector space or not.

3. Determine whether the following set of all pairs of real numbers $\boldsymbol{X} = \{\,(1, x) \mid x \in \mathbf{R}\,\}$ endowed with operations $(1, x) + (1, y) = (1, x + y)$ and $\lambda(1, x) = (1, \lambda x)$, $\forall\, x, \lambda \in \mathbf{R}$ is a vector space.

4. Let Φ be any field. Show that $\boldsymbol{X} = \{\mathbf{0}\}$ is a (trivial) vector space.

5. Let S be any nonempty set, and Φ any field. Define

$$\boldsymbol{F} = \Phi^S = \{f \mid f : S \to \Phi\}$$

to be the set of all mappings from S to Φ. Show that, endowed with the usual operations of addition of functions and multiplication of functions by a scalar, $\boldsymbol{F}$ becomes a vector space.

6. Consider a set

$$\mathbf{R}^{\mathbf{R}} = \{f : \mathbf{R} \to \mathbf{R}\}$$

of all real functions with the usual addition of functions and multiplication of functions by a scalar. Show that such a set is a vector space.

7. Let $\mathbf{R}_{[0,1]} = \{f : [0, 1] \to \mathbf{R}\}$ be the set of all real functions from $[0, 1]$ to $\mathbf{R}$. Show that, with the usual operations, $\mathbf{R}_{[0,1]}$ is a vector space.

8. Consider a space Φ^4 and a set

$$\boldsymbol{F} = \{\,(x_1,\, x_2,\, x_3,\, x_4) \in \Phi^4 \mid x_3 = 5x_4 + a, \quad a \in \Phi\,\}$$

Show that $\boldsymbol{F}$ is a subspace of Φ^4 iff $a = 0$.

9. Consider a space Φ^3 and the following subsets of Φ^3:

$$\boldsymbol{A} = \{\,(x_1,\, x_2,\, x_3) \in \Phi^3 \mid x_1 + 2x_2 + 3x_3 = 0\,\},$$
$$\boldsymbol{B} = \{\,(\boldsymbol{x}_1,\, \boldsymbol{x}_2,\, \boldsymbol{x}_3) \in \Phi^3 \mid x_1 + 2x_2 + 3x_3 = 4\},$$
$$\boldsymbol{C} = \{(x_1,\, x_2,\, x_3) \in \Phi^3 \mid x_1 = 3x_3\,\}$$

Determine whether $\boldsymbol{A} < \Phi^3$, $\boldsymbol{B} < \Phi^3$, $\boldsymbol{C} < \Phi^3$.

10. Determine whether

(i) $\boldsymbol{L} = \{\boldsymbol{x} = (x, 0, x, 0, \ldots) \mid x \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^\infty$.

(ii) $\boldsymbol{L} = \{\boldsymbol{x} = (x, 1, x, 1, \ldots) \mid x \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^\infty$.

11. Let $\mathbf{R}^{\mathbf{R}}$ be the set of all real functions. Determine whether the following are subspaces of $\mathbf{R}^{\mathbf{R}}$:

(i) The set of all even functions.

(ii) The set of all odd functions.

12. Let $L(X)$ be the set of all subspaces of a space X. Show that "*to be a subspace*" is a relation of equivalence.

13. Determine whether or not the following sets of vectors are independent in $\mathbf{R}^3$:

(i) $S = \{(1, 1, 1), (2, 2, 2), (0, 1, 5)\}$

(ii) $T = \{(1, -2, 3), (7, 4, -2), (3, -6, 9)\}$

(iii) $U = \{(8, -1, 3), (4, 0, 1)\}$

14. Show that the following set of vectors is independent in $\mathbf{R}^4$:

$$S = \{(1, 2, 2, -1), ((4, 9, 9, -4), (5, 8, 9, -5)\}$$

15. Show that the following set of vectors is dependent in $\mathbf{R}^4$:

$$S = \{(0, 1, 0, -1), (1, 3, 3, 3), (1, 2, 3, 4)\}$$

16. Let X be a finite-dimensional vector space, and let $S = \{x_1, \ldots, x_n\}$ be a set of linearly independent vectors from X. Show that a set

$$\overline{S} = \{x_1 - x_2, x_2 - x_3, x_3 - x_4, \ldots, x_{n-1} - x_n, x_n\}$$

is also linearly independent.

17. Determine whether or not the following set of vectors are independent in $P_2(t)$:

(i) $x_1 = 2t^2 + 1, \quad x_2 = t^2 + 4t, \quad x_3 = t^2 - 4t + 1$

(ii) $x_1 = t^2 + 3, \quad x_2 = t + 1, \quad x_3 = 2t^2 - 3t + 3$

18. Consider the space $\mathbf{R}^3$. Determine whether the following vectors $x_1 = (1, -1, 0)$, $x_2 = (1, 1, 0)$ and $x_3 = (0, 1, 1)$ form a basis for $\mathbf{R}^3$.

19. Let $L < R^5$ be a subspace defined by

$$L = \{(x_1, \ldots, x_5) \in \mathbf{R}^5 \mid x_1 = 3x_2, \; x_3 = 7x_4\}$$

Find a basis of L.

20. Let $\boldsymbol{L}_1,\ \boldsymbol{L}_2,\ \boldsymbol{L}_3 < \mathbf{R}^3$ be three subspaces defined by

$$\boldsymbol{L}_1 = \{(x, y,\ z) \mid x = z\},\ \ \boldsymbol{L}_2 = \{(x, y,\ z) \mid x + y + z = 0\},$$
$$\boldsymbol{L}_3 = \{(0,\ 0,\ z)\}.$$

Show that $\mathbf{R}^3 = \boldsymbol{L}_1 + \boldsymbol{L}_2 = \boldsymbol{L}_2 + \boldsymbol{L}_3 = \boldsymbol{L}_1 + \boldsymbol{L}_3$.

21. Let $\boldsymbol{X}$ be a finite-dimensional vector space and let $\boldsymbol{L}$ be a subspace of $\boldsymbol{X}$. Show that if $\dim \boldsymbol{L} = \dim \boldsymbol{X}$ then $\boldsymbol{L} = \boldsymbol{X}$.

22. Let $\boldsymbol{D} = \left\{ D = \begin{bmatrix} d_{11} & 0 \\ 0 & d_{22} \end{bmatrix} \mid d_{ii} \in \Phi \right\}$ be the set of all diagonal matrices over a field Φ. Show that $\boldsymbol{D}$ is a subspace of the space $\boldsymbol{M}_{22}$ of all $(2, 2)$-matrices.

23. Let $\boldsymbol{L}, \boldsymbol{M} < \mathbf{R}^3$ be two subspaces. Show that $\boldsymbol{L} \cap \boldsymbol{M} \neq \varnothing$.

24. Suppose $F : \mathbf{R}^3 \to \mathbf{R}^2$ is defined by

$$F(\boldsymbol{x}) = F \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + y + z \\ 2x - 3y + 4z \end{bmatrix}$$

Show that F is a linear operator.

25. Show that

(i) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(\boldsymbol{x}) = F \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ x \end{bmatrix}$ is a linear operator, and

(ii) $F : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $F(\boldsymbol{x}) = F \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xy \\ x \end{bmatrix}$ is not.

26. Let $\boldsymbol{M}_{nn}$ be the space of all (n, n)-matrices, and let $M \in \boldsymbol{M}_{nn}$ be some arbitrary but fixed matrix. Show that $F : \boldsymbol{M}_{nn} \to \boldsymbol{M}_{nn}$, defined by $F(A) = AM + MA$, is a linear operator.

27. Let $F : \mathbf{R}^4 \to \mathbf{R}^2$ be a linear operator, such that

$$\operatorname{Ker} F = \{(x_1, x_2, x_3, x_4) \mid x_1 = 5x_2;\ x_3 = 7x_4\}$$

Show that F is surjective.

28. Let $\boldsymbol{X}, \boldsymbol{L}$ be two finite-dimensional vector spaces. Show that if $\boldsymbol{L} < X$, and $F : \boldsymbol{L} \to \boldsymbol{X}$ is an isomorphism, then $\boldsymbol{L} = \boldsymbol{X}$.

29. Let A be an (n, n) matrix. Show that

(i) $A + A^{\tau}$ is symmetric.

(ii) $A - A^{\tau}$ is antisymmetric.

30. Show that if A is a matrix such that $AA^{\tau} = O$, then $A = O$.

31. Show that if a matrix A is symmetric(antisymmetric), and B is any square matrix, then $B^{\tau}AB$ is symmetric (antisymmetric).

32. Show that the following matrix:

$$\mathrm{A} = \begin{bmatrix} 3/7 & 2/7 & 6/7 \\ -6/7 & 3/7 & 2/7 \\ 2/7 & 6/7 & -3/7 \end{bmatrix}$$

is orthogonal.

33. Show that if a matrix A satisfies one of the following two properties, then it also satisfies the third.
(i) A is symmetric, (ii) A is orthogonal, and (iii) A is involutory.

34. Reduce the following matrices to echelon form:

(i) $A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \\ 2 & 4 & -2 & 3 \end{bmatrix}$

(ii) $B = \begin{bmatrix} 1 & -7 & 10 & 2 \\ 1 & 2 & 2 & 1 \\ 3 & 3 & 4 & 2 \\ 1 & 5 & 3 & 5 \end{bmatrix}$

35. Reduce the following matrices to row canonical form:

(i) $A = \begin{bmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 8 & 8 & -1 & 26 & 23 \end{bmatrix}$

36. Let $A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}, \quad a_{ii} \neq 0$. Show that

$$\widetilde{A} = \prod_{i=1}^{n} a_{ii} \begin{bmatrix} a_{11}^{-1} & 0 & \cdots & 0 \\ 0 & a_{22}^{-1} & 0 & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & a_{nn}^{-1} \end{bmatrix}.$$

37. Show that
(i) $(\widetilde{\lambda A}) = \lambda^{n-1}\widetilde{A}$
(ii) $(\widetilde{AB}) = \widetilde{B}\widetilde{A}$

38. Let $T : \mathbf{R}^2 \to \mathbf{R}^3$ be a linear operator defined by the matrix $T = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -1 & -2 \end{bmatrix}$, and let $\boldsymbol{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\boldsymbol{x}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\boldsymbol{x}_3 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Find $T(\boldsymbol{x}_1)$, $T(\boldsymbol{x}_2)$, and $T(\boldsymbol{x}_3)$.

39. Let $P : \Phi^3 \to \Phi^2$ be a linear operator defined by $P(\alpha, \beta, \gamma) = (\alpha, \beta)$. Convince yourself that its matrix representation in the standard basis (canonical basis) is

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

40. Let a matrix for the operator $F : \mathbf{R}^3 \to \mathbf{R}^3$ be defined in the standard basis by

$$\begin{aligned} x_1' &= 3x_1 + 5x_2 - x_3 \\ x_2' &= 4x_1 - x_2 + x_3 \\ x_3' &= 3x_1 + 2x_2 - x_3 \end{aligned}$$

Find the matrix for F and calculate $F \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

41. Show that the "rotation" matrix

$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

is orthogonal.

42. Matrices

(i) $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $C = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$ represent reflections in $\boldsymbol{R}^2$, and matrices

(ii) $D = \begin{bmatrix} d & 0 \\ 0 & 1 \end{bmatrix}$, $E = \begin{bmatrix} 1 & 0 \\ 0 & d \end{bmatrix}$, $0 < d < 1$ represent contraction in $\boldsymbol{R}^2$. Determine which particular reflection/contraction each matrix pertains to.

43. Let an operator $\mathcal{T}$ be a composition of reflection about the X-axis, followed by a rotation trough $\frac{\pi}{2}$, followed by a dilatation of factor d. If $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, find $\mathcal{T}(\boldsymbol{x})$.

44. Find the inverse of A by row reducing $[A|I]$ to $[I|A^{-1}]$.

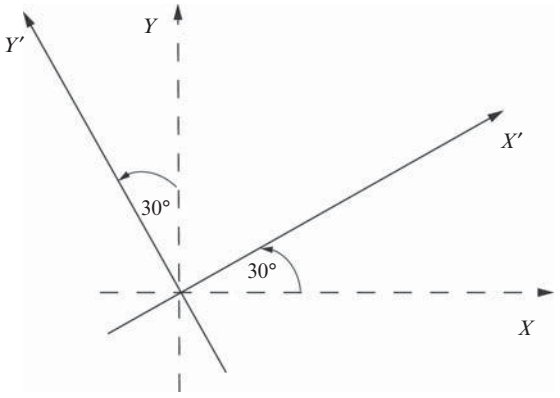
45. Consider the canonical basis of $\mathbf{R}^2$, $\boldsymbol{E} = \left\{ \boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and another basis, $\boldsymbol{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \ \boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \right\}$. Find the change of basis matrix $\boldsymbol{E} \xrightarrow{T} \boldsymbol{B}$

46. Let $F : \mathbf{R}^3 \to \mathbf{R}^3$ be a linear operator whose matrix representation in some basis is

$$F = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 4 & 3 \\ 2 & 7 & 4 \end{bmatrix}$$

Find the matrix of F in basis $\boldsymbol{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$.

47. Let $X'Y'$ be a coordinate system in $\mathbf{R}^2$, obtained by rotating a system XY 30° counterclockwise (see the figure below).



Find the change-of-base matrix for the $X'Y'$-system.

48. Let $\boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$ and $\boldsymbol{B}' = \{\boldsymbol{b}'_1, \boldsymbol{b}'_2\}$ be the basis of X and Y, respectively. Furthermore, let $T : \boldsymbol{X} \to \boldsymbol{Y}$ be a linear operator such that

$$T(\boldsymbol{b}_1) = 2\boldsymbol{b}'_1 - \boldsymbol{b}'_2$$
$$T(\boldsymbol{b}_2) = 3\boldsymbol{b}'_1 + 2\boldsymbol{b}'_2$$
$$T(\boldsymbol{b}_3) = \boldsymbol{b}'_1 - 4\boldsymbol{b}'_2$$

Find the matrix representation of T with respect to $\boldsymbol{B}, \boldsymbol{B}'$ basis and find $T(\boldsymbol{x})$ if $\boldsymbol{x} = 3\boldsymbol{b}_1 + 2\boldsymbol{b}_2 - \boldsymbol{b}_3$.

49. Find the matrix representation of the following linear operators relative to the canonical basis of $\mathbf{R}^n$:

(i) $F : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $F(x_1, x_2, x_3) = (2x_1 - 4x_2 + 9x_3,\ 5x_1 + 3x_2 - 2x_3)$

(ii) $F : \mathbf{R}^4 \to \mathbf{R}$ defined by $F(x_1, x_2, x_3, x_4) = 2x_1 + x_2 - 7x_3 - x_4$.

50. Let

(i) $A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 0 & 2 & 3 \end{bmatrix}$. Find A^{-1} by reducing $[A \mid I]$ to the form $[I \mid A^{-1}]$.

(ii) $B = \begin{bmatrix} 1 & 1 & 2 \\ 3 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}$. Find B^{-1} by reducing $[B \mid I]$ to the form $[I \mid B^{-1}]$.

(iii) $C = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$. Find C^{-1} by reducing $[C \mid I]$ to the form $[I \mid C^{-1}]$.

51. Let

(i) $A = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & -1 & 3 \end{bmatrix}$ and

(ii) $B = \begin{bmatrix} 1 & 2 & 1 \\ 3 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

Find A^{-1} and B^{-1} using Theorem 6.73.

52. Use Cramer's method to solve the following system:

$$\begin{aligned} x_1 + \quad & 2x_3 = 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

53. Use Cramer's method to solve the following system:

$$\begin{aligned} x_1 - 3x_2 + x_3 &= 11 \\ 2x_1 - x_2 + 3x_3 &= 7 \\ 3x_1 + x_2 - 2x_3 &= 5 \end{aligned}$$

54. Show that the following system of linear equations has no solution:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 3 \\ 2x_1 + x_2 + 3x_3 &= 3 \\ x_1 + x_2 + 2x_3 &= 0 \end{aligned}$$

55. Solve the following system of linear equations by the Gauss–Jordan method:

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + \quad x_5 \quad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= 0 \\ 5x_3 + 10x_4 + \quad 15x_6 &= 0 \\ 2x_1 + 6x_2 + \quad 8x_4 + 4x_5 - 18x_6 &= 0 \end{aligned}$$

56. Solve the following system of linear equations by the Gauss–Jordan method:

$$\begin{aligned} 3x_1 + x_2 - 3x_3 + x_4 - x_5 &= 1 \\ 2x_1 - x_2 + 7x_3 - 3x_4 + 5x_5 &= 2 \\ 3x_1 - 2x_2 + 7x_3 - 5x_4 - x_5 &= 3 \\ x_1 + 3x_2 - 2x_3 + 5x_4 + 8x_5 &= 3 \end{aligned}$$

57. Solve the following system of linear equations by the Gauss–Jordan method:

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 8 \\ -x_1 - 2x_2 + 3x_3 &= 1 \\ 3x_1 - 7x_2 + 4x_3 &= 10 \end{aligned}$$

58. Solve the following system of linear equations by the Gauss–Jordan method:

$$\begin{aligned} x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\ 3x_1 - 9x_2 + 7x_3 - x_4 + 3x_5 &= 7 \\ 2x_1 - 6x_2 + 7x_3 - 4x_4 - 5x_5 &= 7 \end{aligned}$$

59. Find eigenvalues and eigenvectors for the following matrices:

(i) $A = \begin{bmatrix} 0 & 3 \\ 8 & -1 \end{bmatrix}$;

(ii) $B = \begin{bmatrix} -2 & 6 \\ 3 & -0 \end{bmatrix}$;

(iii) $C = \begin{bmatrix} 3 & -1 & 1 \\ 7 & -5 & 1 \\ 6 & -6 & 2 \end{bmatrix}$;

(iv) $D = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$.

60. Let F be an (n, n) matrix. Show that the coefficient of λ^n in the characteristic polynomial is 1.

61. Let $F : \mathbf{R}^3 \to \mathbf{R}^3$ be an operator defined by

$$F\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 - 2x_3 \\ 2x_1 + 3x_2 - 4x_3 \\ x_1 + x_2 - x_3 \end{bmatrix}$$

(i) Determine whether F is diagonalizable and, if so, find a T that diagonalizes F.

(ii) Find the eigenvalues and eigenvectors of F.

62. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, with $a, b, c, d \in \mathbf{R}$. What are the necessary and sufficient conditions so that A is diagonalizable?

63. Show that $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}$ is not diagonalizable.

64. Let $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$. Find T that diagonalizes A and find A^7.

65. Let $A = \begin{bmatrix} 1 & 3 & 7 & 1 \\ 0 & \frac{1}{2} & 3 & 4 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Find the eigenvalues of A^9.

INDEX

Principles of Mathematics: A Primer, First Edition. Vladimir Lepetic.

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.