Discrete Mathematical Structures

Discrete Mathematical Structures

Dan Yasaki
UNCG

Sebastian Pauli
UNCG

September 4, 2026

©2026 Dan Yasaki and Sebastian Pauli

This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License. To view a copy of this license, visit CreativeCommons.org[1]

[1] creativecommons.org/licenses/by-nc/4.0

Preface to the LaTeX version

This document grew from lecture notes following the seventh edition of **Discrete Mathematics and its Applications** by Kenneth Rosen [13].

I used various versions of the notes in conjunction with the book over the years whenever I taught MAT 253 Discrete Mathematical Structures. I will continue to develop this document incorporating feedback from readers.

MAT 253 core course in the mathematics curriculum designed for mathematics majors as an early introduction to discrete mathematical structures, rigorous proof techniques, and mathematical programming.

The catalog description of MAT 253 is:

> A rigorous introduction to discrete mathematical structures, proof techniques, and programming. Topics include sets, functions, sequences, relations, induction, propositional and predicate logic, modular arithmetic, and mathematical programming.

Upon successful completion of this course, students will be able to:

1. define the fundamental discrete mathematical structures.
2. identify and describe various types of relations.
3. explain how RSA encryption allows for secure message transcription.
4. translate pseudocode algorithms into Python scripts.
5. compute the number of solutions to several arrangement problems.
6. analyze simple algorithms and identify values of variables at various stages of completion.
7. combine definitions and results produced in class to create rigorous proofs of basic statements about discrete mathematical structures.
8. evaluate an argument for logical validity.

In writing this text, I wanted to produce an streamlined, yet inviting introduction to discrete structures. The text is interspersed with fun, yet (mostly) relevant xkcd comics. The prerequisites for the book are minimal; pre-calculus and an open mind. Foundations are built from definitions.

It is best to pair these notes with additional resources. Some free resources are identified below.

A major goal of this course is to teach you to communicate mathematics clearly. This involves learning precise statements of definitions, and learning to write clear, concise proofs. Here are two books that may help.

- Book of Proof[2] by Richard Hammock: This book is an introduction to the standard methods of proving mathematical theorems. It has been approved by the American Institute of Mathematics' Open Textbook Initiative.
- The Art of Proof[3] by Matthias Beck and Ross Geoghegan is an excellent introduction to writing good proofs. This book is not open source, but our library has a Springer subscription that includes this book. Go to the UNCG Library Catalog to find this book. You can download a free copy by entering your UNCG iSpartan credentials.

I thank Office of the Provost and the University Libraries for the Open Educational Resources Mini-Grant in summer 2018 that allowed me the extra time to adjust the course syllabus to accommodate this text.

Thanks to Cliff Smyth and Sebastian Pauli for piloting the use of these notes in their courses. Thanks to others that found typos and mistakes, including Heather Parlaman.

Dan Yasaki

[2] www.people.vcu.edu/~rhammack/BookOfProof/
[3] link.springer.com/book/10.1007%2F978-1-4419-7023-7

Contents

Back Matter

Chapter 0

Introduction

The *student learning outcomes* (SLOs) of a course based on these notes are:

Outcomes

1. Define the fundamental discrete mathematical structures.
2. Identify and describe various types of relations.
3. Explain how a public key cryptosystem works and apply it for encryption and decryption.
4. Translate pseudocode algorithms into Python scripts.
5. Compute the number of solutions to several arrangement problems.
6. Analyze simple algorithms and identify values of variables at various stages of completion.
7. Combine definitions and results produced in class to create rigorous proofs of basic statements about discrete mathematical structures.
8. Evaluate an argument for logical validity.

0.1 Preliminaries

We recall some terminology and notation that we use throughout the remainder of these notes. We will revisit these in later chapters.

0.1.1 Sets

We give basic terminology and notation for sets. We cover sets in more detail in Section 2.1.

A **set** is an unordered collection of objects, called **elements** or **members** of the set. We write $b \in A$ when b is an element of the set A. We read $b \in A$ as "b is an element of A" or "b is in A".

Sets can be specified in **roster form** by listing the elements explicitly between braces. For example, $\{1, 2, 7\}$ is the set containing 1, 2, and 7.

To indicate that the list of elements continues in a pattern that can be deduced from the preceding list of elements we use ellipsis (...). For example,

we have

$$\{1, 2, 3, \ldots, 10\} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}.$$

0.1.2 Functions

Functions describe a special relationship between the elements of two sets. We cover functions more thoroughly in Section 2.3.

Let A and B be nonempty sets. A **function** from A to B, denoted $f\colon A \to B$ is an assignment of exactly one element of B to each element of A.

The set A is called the **domain** of f and the set B is called the **codomain** of f.

0.1.3 Numbers

We introduce some sets that we use often in a somewhat intuitive manner. We revisit these definitions in Subsection 2.1.1.

Traditionally, the letters representing these sets were typeset in boldface. But, since on blackboards this is difficult to write in boldface, lecturers are writing bold symbols with certain doubled strokes. These double stroke characters are now also commonly used in print.

The set of **natural numbers** is

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

The set of **integers** is

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

The set of integers $\mathbb{Z}$ contains all elements of the set of natural numbers $\mathbb{N}$.

The set of **non-negative integers** is

$$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\}.$$

The set of **rational numbers** $\mathbb{Q}$ contains all numbers that can be written as fractions in the form $\frac{n}{d}$ where $n \in \mathbb{Z}$ and $d \in \mathbb{N}$.

The set of rational numbers $\mathbb{Q}$ contains all elements of the set of integers $\mathbb{Z}$.

The set of **real numbers** which we denote by $\mathbb{R}$ is actually fairly subtle to define. In a course in analysis they are defined as equivalence classes of Cauchy sequences over the rational numbers. For our purposes, the intuitive idea that they are all numbers that can be written in finite or infinite decimal representation is sufficient.

The set $\mathbb{R}$ of real numbers contains all elements of the set $\mathbb{Q}$ of rational numbers.

The real numbers that are not rational numbers are called **irrational**.

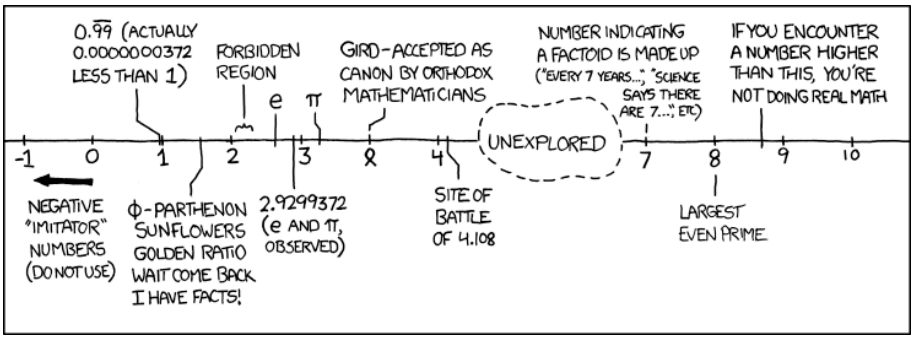


Figure 0.1.3.1 Number Line. (https://xkcd.com/899/) The Wikipedia page List of Numbers opens with "This list is incomplete; you can help by expanding it."

0.1.4 Arithmetic

We recall basic properties of the operations on real numbers. Since the natural numbers, integers, and rational numbers are in the set of real numbers, they also hold for these.

For all real numbers a and b we have:

Commutative laws	$a + b = b + a$ and $a \cdot b = b \cdot a$
Associative laws	$a + (b + c) = (b + a) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
Distributive law	$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$
Identity laws	$0 + a = a$ and $1 \cdot a = a$

We call 1 the multiplicative identity. Note that the additive identity 0 is an element of the real and rational numbers and the integers, but not the natural numbers.

The following applies only to integers, rational numbers, and real numbers a, since neither 0 nor negative numbers are in the set of natural number. It only holds for $a = 0$ in $\mathbb{N}_0$.

Inverse law for addition	$a + (-a) = 0$

We call $-a$ the additive inverse of a and write $a + (-b) = a - b$.

Because the sets of integers and natural numbers do not contain fractions, the following only holds for rational numbers and real numbers in general. The only integers for which it holds are $a = 1$ and $a = -1$. The only natural number for which it holds is $a = 1$.

Inverse law for multiplication	$a \cdot \frac{1}{a} = 1$

We call $\frac{1}{a}$ the multiplicative inverse of a.

We indicate the order of operations with parenthesis (evaluate the innermost expression first), in addition we follow the precedence of operations:

1. **exponentiation** and **roots**, since exponentiation is repeated multiplication (when the exponent is a natural number)

2. **multiplication** and **division**, since multiplication by natural numbers is repeated addition and division by natural numbers is repeated subtraction

3. **addition** and **subtraction**

Also see the first panel in Figure 0.1.4.2.

Recall that $a^2 = a \cdot a$. With the laws above we prove:

Theorem 0.1.4.1 *Let a and b be real numbers. Then*

1. $(a+b)^2 = a^2 + 2ab + b^2$
2. $(a-b)^2 = a^2 - 2ab + b^2$
3. $(a-b)(a+b) = a^2 - b^2$

Proof. We prove the first formula and leave the proof of the second formula as an exercise. For the third formula see Theorem 1.3.2.13.

By the distributive law we have

$$(a+b)(a+b) = a(a+b) + b(a+b).$$

Applying the distributive law twice we get

$$a(a+b) + b(a+b) = (a^2 + ab) + (ba + b^2).$$

By the commutative law of addition we have

$$(a^2 + ab) + (ba + b^2) = (a^2 + ab) + (ab + b^2).$$

With the associative law of addition we get

$$(a^2 + ab) + (ab + b^2) = a^2 + (ab + ab) + b^2 = a^2 + 2ab + b^2.$$

Thus we have proven that

$$(a+b)(a+b) = a^2 + 2ab + b^2.$$

■

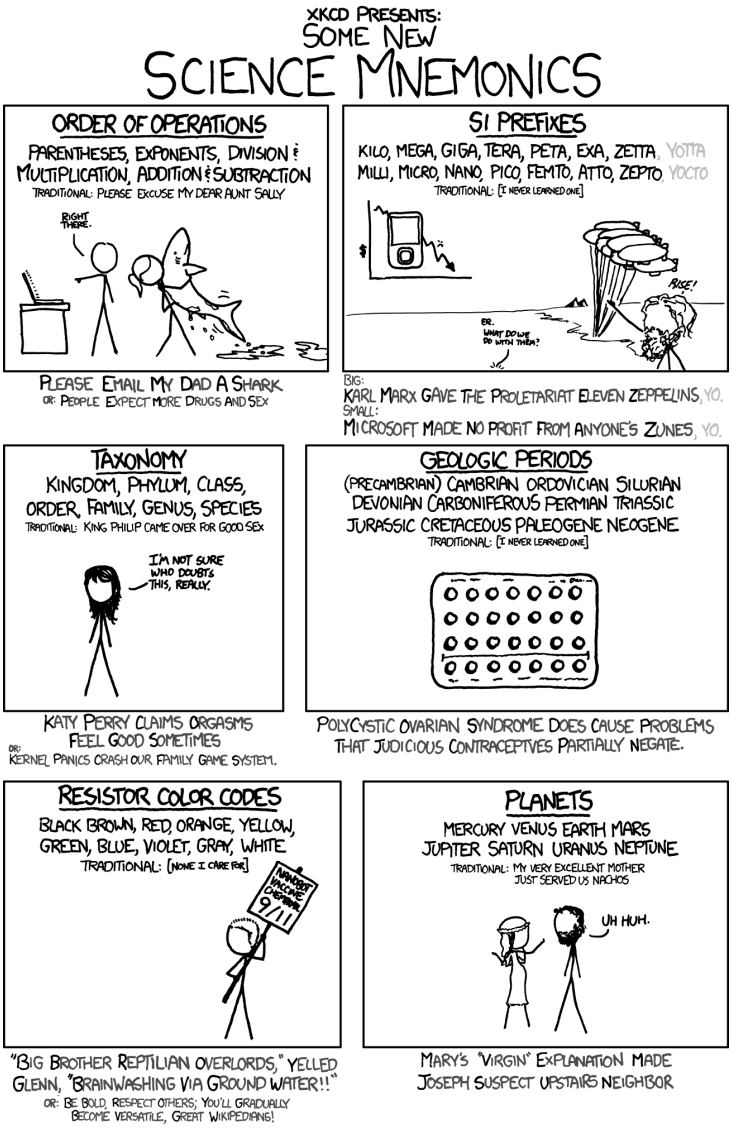


Figure 0.1.4.2 Mnemonics. (https://xkcd.com/992/) 'Sailor Moon's head exploded once' and 'Some men have explosive orgasms' both work for the Great Lakes from west to east (Paddle-to-the-Sea order).

0.1.5 Integer Division

Multiplication of two natural number can be viewed as repeated addition. In particular, when $n \in \mathbb{N}$ and $b \in \mathbb{Z}$ we have

$$n \cdot b = \underbrace{b + b + \ldots + b}_{n \text{ copies of } b}.$$

Dividing a natural number a by a natural number b answers the question

How often does b go into a ?

We can answer this question by starting with a and repeatedly subtracting b until we obtain a number less than b. We call the number of times q that we have subtracted b the quotient of the division of a by b and the value r with

$0 \leq r < b$ that is left over the remainder. We get

$$a \underbrace{-b-b-\ldots-b}_{q \text{ copies of } b} = a - q \cdot b = r < b.$$

If we rearrange this we get:

$$a = \underbrace{b+b+\ldots+b}_{q \text{ copies of } b} + r = q \cdot b + r.$$

The above proves a theorem called **Division Algorithm** for non-negative integers, see Theorem 5.1.1.9 for the general version.

Theorem 0.1.5.1 *Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Then there are unique $q \in \mathbb{N}_0$ and $r \in \mathbb{N}_0$ such that*

$$a = b \cdot q + r \text{ and } 0 \leq r < b.$$

We write $q = a \operatorname{div} b$ and $r = a \bmod b$.

We cover integer division greater detail in Divisibility.

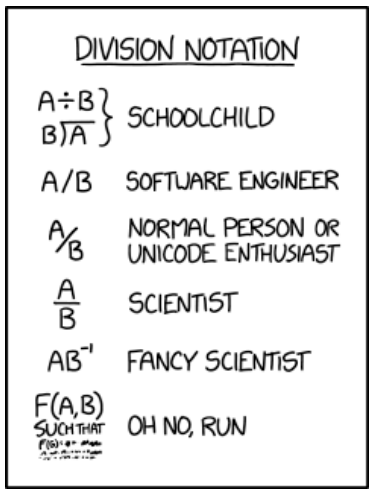


Figure 0.1.5.2 Division Notation. (https://xkcd.com/2687) Science tip: Scientists hardly ever use the two-dot division sign, and when they do it often doesn't even mean division, but they still get REALLY mad when you repurpose it to write stuff like SALE! ALL SHOES 30÷ OFF!

0.1.6 Exponentiation

We define exponentiation as repeated multiplication:

Definition 0.1.6.1 For a real number b and a natural number n the n-th power of b is

$$b^n = \underbrace{b \cdot b \cdot \ldots \cdot b}_{n \text{ copies of } b}.$$

We call b the base of b^n and n its exponent. ◊

Theorem 0.1.6.2 *For all real numbers a and b and natural numbers m and n we have:*

1. $(b^m) \cdot (b^n) = b^{(m+n)}$
2. $(a^n) \cdot (b^n) = (ab)^n$
3. $(b^m)^n = b^{(m \cdot n)}$

Proof.

1. With the definition of exponentiation 0.1.6.1 and then counting the numbers bs we get

$$\begin{aligned}(b^m)\cdot(b^n) &= \underbrace{b\cdot b\cdot \ldots \cdot b}_{m \text{ copies of}} \cdot \underbrace{b\cdot b\cdot \ldots \cdot b}_{n \text{ copies of } b} \\ &= \underbrace{b\cdot b\cdot \ldots \cdot b}_{(m+n) \text{ copies of } b} \\ &= b^{(m+n)}.\end{aligned}$$

2. In addition, applying the commutative law and the associative law of multiplication we get

$$\begin{aligned}a^n \cdot b^n &= \underbrace{a\cdot a\cdot \ldots \cdot a}_{n \text{ copies of } a} \cdot \underbrace{b\cdot b\cdot \ldots \cdot b}_{n \text{ copies of } b} \\ &= \underbrace{(a\cdot b)\cdot(a\cdot b)\cdot \ldots \cdot (a\cdot b)}_{n \text{ copies of } a\cdot b} \\ &= (a\cdot b)^n.\end{aligned}$$

3. Again, with the definition of exponentiation 0.1.6.1 and then counting the numbers bs we get

$$\begin{aligned}(b^m)^n &= (\underbrace{b\cdot b\cdot \ldots \cdot b}_{m \text{ copies of } b})^n \\ &= \underbrace{\underbrace{b\cdot b\cdot \ldots \cdot b}_{m \text{ copies of } b} \cdot \ldots \cdot \underbrace{b\cdot b\cdot \ldots \cdot b}_{m \text{ copies of } b}}_{n \text{ copies of } \underbrace{b\cdot b\cdot \ldots \cdot b}_{m \text{ copies of } b}} \\ &= \underbrace{b\cdot b\cdot \ldots \cdot b}_{(m\cdot n) \text{ copies of } b} \\ &= b^{(m\cdot n)}.\end{aligned}$$

■

So far we only have considered exponents that are natural numbers. In a generalization we would like Theorem 0.1.6.2 to still hold.

If we set $m = 0$ in Theorem 0.1.6.2Item 1 we have $(b^0)\cdot(b^n) = b^{(0+n)} = b^n$. With the identity law for multiplication we get

$$b^0 = 1.$$

For the negative exponent $n = -m$ in Theorem 0.1.6.2 2. we have $b^{(m+(-m))} = (b^m)\cdot\left(b^{(-m)}\right) = b^0 = 1$. Then by the Inverse law for multiplication we have

$$b^{(-m)} = \frac{1}{b^m}.$$

Definition 0.1.6.3 Let $a \in \mathbb{R}$. The square root of a is the positive solution of $x^2 = a$. ◊

0.2 Further reading

- The original version of these lecture notes were written for a course that followed the book **Discrete Mathematics and its Applications** by Kenneth Rosen [13].
- The material in Integers and Modern Applications for the Uninitiated[1] overlaps with the material in these notes, but lacks the emphasis on proofs.
- Oscar Levin's Discrete Mathematics: An Open Introduction[2] is another good representation to some of the material covered in this course.
- Book of Proof[3] by Richard Hammock is an introduction to the standard methods of proving mathematical theorems. It has been approved by the American Institute of Mathematics' Open Textbook Initiative.
- The Art of Proof[4] by Matthias Beck and Ross Geoghegan is an excellent introduction to writing good proofs. This book is not open source, but our library has a Springer subscription that includes this book. Go to the UNCG Library Catalog to find this book. You can download a free copy by entering your UNCG iSpartan credentials.

0.3 Python

One of the components of this course is mathematical programming. The language chosen for this course is Python 3.

You can use Python on the central Linux server of UNCG. It is also installed on the computers in all ITS computer labs. You can download it for free at `http://www.python.org/download/`.

There are several additional resources you may find helpful. Choose based on your programming background and desired difficulty level.

- The Python 3 Tutorial[1] does not attempt to be comprehensive and cover every single feature, or even every commonly used feature. Instead, it introduces many of Python's most noteworthy features, and will give you a good idea of the language's flavor and style. After reading it, you will be able to read and write Python modules and programs, and you will be ready to learn more about the various Python library modules described in The Python Standard Library.

  This will be the main source of information for the programming assignments. In a typical semester, we cover most of the sections 1 through 5.
- The Non-Programmer's Tutorial for Python 3[2] is a tutorial designed to be an introduction to the Python programming language. This guide is for someone with no programming experience.

[1] `go.uncg.edu/mat112`
[2] `discrete.openmathbooks.org`
[3] `www.people.vcu.edu/~rhammack/BookOfProof/`
[4] `link.springer.com/book/10.1007%2F978-1-4419-7023-7`
[1] `docs.python.org/3/tutorial/`
[2] `en.wikibooks.org/wiki/Non-Programmer's_Tutorial_for_Python_3`

- Python for Non-Programmers[3] is aimed at readers with no previous programming experience. If you've never programmed before, the tutorials on this page are recommended for you.
- Python for Programmers[4] is aimed at people who have previous experience with other programming languages (C, Perl, Lisp, Visual Basic, etc.).
- The Python Wiki[5] is a community place to gather and organize all things about Python. Feel free to exercise your editorial skills and expertise to make it a useful knowledge base and up-to-date reference on all Python-related topics.
- Learn Python the Hard Way[6] instructs you in Python by slowly building and establishing skills through techniques like practice and memorization, then applying them to increasingly difficult problems. By the end of the book you will have the tools needed to begin learning more complex programming topics.

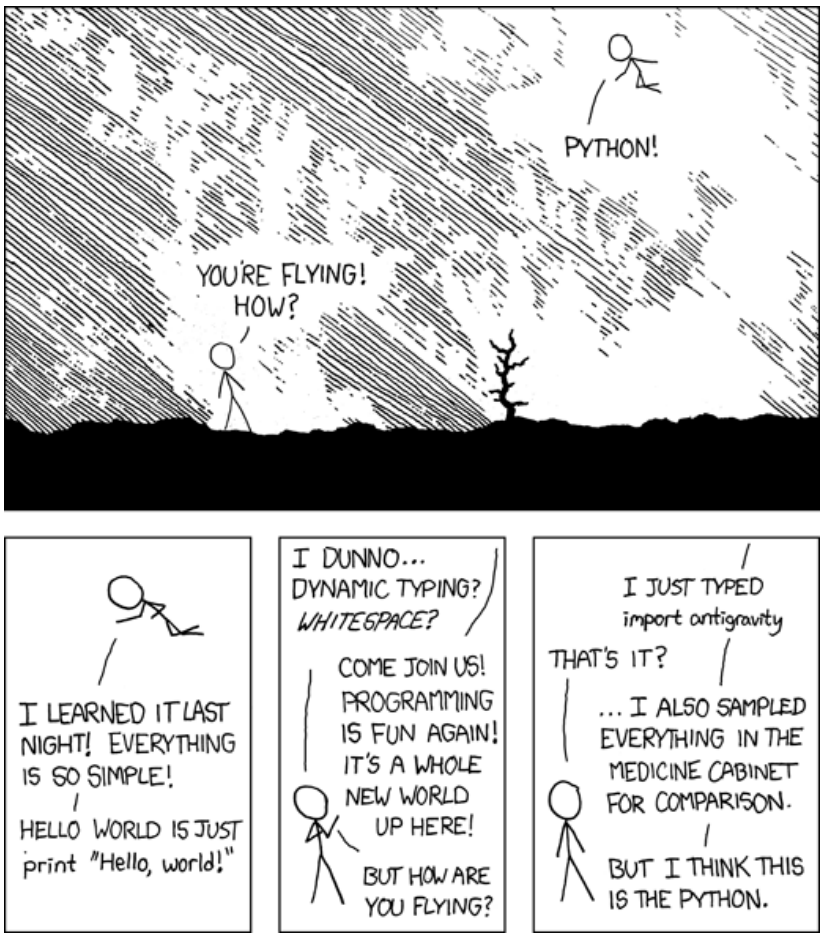


Figure 0.3.0.1 Python. (https://xkcd.com/353/) I wrote 20 short programs in Python yesterday. It was wonderful. Perl, I'm leaving you.

[3] wiki.python.org/moin/BeginnersGuide/NonProgrammers
[4] wiki.python.org/moin/BeginnersGuide/Programmers
[5] wiki.python.org/moin/FrontPage
[6] learnpythonthehardway.org/book/

Chapter 1

Logic and Proofs

> Contrariwise, 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 Carroll (1832--1898)

In this chapter, we discuss logic and proofs. The rules of logic specify the meaning of mathematical statements. These rules help us understand and reason with these statements to construct arguments to justify the truth of certain statements. Once we prove a mathematical statement true, we call it a **theorem** . The argument of justification is called a **proof** .

Clear reasoning and communication of ideas are important in all disciplines. We restrict ourselves to a small, but important corner of mathematics where we can completely describe the theory with minimal prerequisites. The skills you gain are applicable to many other situations.

1.1 Propositional logic

Outcomes

1. Reproduce definitions of the basic terminology of propositional logic
2. Express propositions in English
3. Formulate propositions using propositional logic
4. Evaluate compound proposition
5. Construct truth tables

While we may wish to understand all statements, this is too daunting of a task. Human language is just too subtle and nuanced. Instead, we develop a theory for understanding certain types of simpler statements known as *propositions*. There is a overlap between these simple statements and statements that are used outside of this context. For the most part the meanings agree, but there are instances where the intended meanings are different. Because of this, we will be precise about what we mean mathematically.

In this section, we develop a framework to study propositions. The rules of logic give precise meaning to propositions and help us understand the validity of arguments.

1.1.1 Propositions

Definition 1.1.1.1 A **proposition** is a declarative statement (a statement that declares a fact) that is either true or false but not both. The **truth value** of a proposition is often denoted T for true and F for false. ◊

Let's first look at some statements that are not propositions. These fall outside the scope of our study of propositional logic.

Example 1.1.1.2 Question. Consider the statement "Where is the bookstore?" This is an interrogative, not declarative, statement. Questions are not propositions. □

Example 1.1.1.3 Command. Consider the statement "Tie your shoe."

This is an imperative, not declarative, statement. Commands are not propositions. □

Example 1.1.1.4 Paradox. Consider the statement "This statement is false."

This is a declarative statement, but it is not a proposition. This is more subtle. Suppose the statement is true. The statement is asserting it is false. The only way for that to be true would be for the statement to be false. In that case, the assertion that it is false would be true\dots. This sort of self-referential statement that does not allow for a single truth value is not a proposition. □

The last example is a bit more subtle. Here, the statement is not a proposition because there are variables in the statement that are not quantified. We will see how to deal with such things in detail in Subsection 1.3.1.

Example 1.1.1.5 Unquantified. Consider the statement "$x + 1 = 3$."

This is a declarative statement, but it is not a proposition. Why? The truthfulness of the sentence depends on the value of the variable x , so it does not have a well-defined truth value. For example, it is true when $x = 2$, but it is false when $x = 3$. This sort of ambiguous statement is not a proposition. □

Now, let's look at some statements that are propositions.

Example 1.1.1.6 Washington. Consider the statement "Washington DC is the capital of the United States."

This is a proposition since it is a declarative statement that is true. □

Note that propositions are allowed to be false.

Example 1.1.1.7 Charlotte. The statement "Charlotte is the capital of NC," is a proposition since it is a declarative statement that is that is false. □

Example 1.1.1.8 Equation. Consider the statement "$3 \cdot 5 = 8$."

This is a proposition since it is a declarative statement that is false. □

We also allow statements that declare facts for which we may not know the truth value.

Example 1.1.1.9 Bob is taller than Alics. Consider the statement "Bob is taller than Alice." This is a proposition since it is a declarative statement that is either true or false, but not both. It does not bother us that we don't know if it is true or false. The statement is a proposition because it declares a fact that is either true or false. □

We will often use **propositional variables** to represent propositions. For example, let p = "2 + 3 = 5," and let q be "My name is Dan." In this case, the proposition p is true. We also have that the proposition q is true, provided we agree on the convention that first person pronouns refer to me, the author.

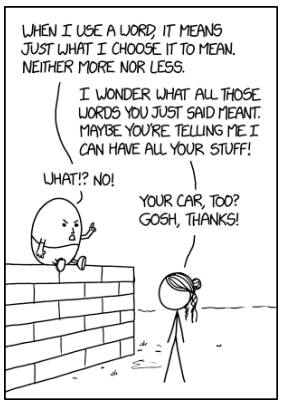


Figure 1.1.1.10 *Communicating* by Randal Munroe (https://xkcd.com/1860/). You're saying that the responsibility for avoiding miscommunication lies entirely with the listener, not the speaker, which explains why you haven't been able to convince anyone to help you down from that wall.

1.1.2 Compound propositions

Using logical operators, we will be able to use existing propositions to create **compound propositions**. . .

Definition 1.1.2.1 Let p be a proposition. The **negation** of p , denoted $\neg p$, is the proposition "It is not the case that p." It is read as "Not p" and has the opposite truth value from p. ♢

Figure 1.1.2.2 gives the truth table for negation. If p is true, then $\neg p$ is false. If p is false, then $\neg p$ is true.

p	$\neg p$
T	F
F	T

Figure 1.1.2.2 Truth table for negation (not p).

Example 1.1.2.3 Let p be the proposition "$2 + 3 = 5$". The negation $\neg p$ is the proposition "It is not the case that $2 + 3 = 5$." More directly, $\neg p$ can be expressed as "$2 + 3 \neq 5$." The original proposition p is true, so the negation is false. □

Example 1.1.2.4 Negation of a false proposition. Let q be the proposition "$2^3 = 5$." The negation $\neg q$ is the proposition "It is not the case that $2^3 = 5$." More directly, $\neg p$ can be expressed as "$2^3 \neq 5$." The original proposition is false, so the negation is true. □

Definition 1.1.2.5 Let p and q be propositions. The **conjunction** of p and q, denoted $p \wedge q$ is the proposition "p and q." It is true when p and q are both true and false otherwise. ♢

The conjunction is often referred to as **and** . Figure 1.1.2.6 gives the truth table for conjunction. If p is true and q is true, then $p \wedge q$ is true. If p or q are false (including the case where both p and q are false), then the conjunction $p \wedge q$ is false.

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Figure 1.1.2.6 Truth table for conjunction (p and q).

Example 1.1.2.7 Conjunction is true. Let p be the proposition "1+2 = 3," and let q be the proposition "$3 + 4 = 7$."

Since p and q are both true, the conjunction $p \wedge q$ is true. The conjunction $p \wedge q$ can be expressed as "$1 + 2 = 3$ and $3 + 4 = 7$." □

Example 1.1.2.8 Conjuction is false. Let p be the proposition "$5 < 3$" and let q be the proposition "$4 = 7$." Both p and q are false.

The conjunction $p \wedge q$ is false. The conjunction $p \wedge q$ can be expressed as "$5 < 3$ and $4 = 7$." □

Example 1.1.2.9 Conjunction is true. Let p be the proposition "$5 < 5$" and let q be the proposition "$10 = 5 \cdot 2$."

The conjunction $p \wedge q$ is false because p is false. The fact that q is true is not enough to make the conjunction true. The conjunction $p \wedge q$ can be expressed as "$5 < 5$ and $10 = 5 \cdot 2$." □

Example 1.1.2.10 Alice and Bob are tall. Let a be the propositions "Alice is tall," and let b be "Bob is tall."

The conjunction $a \wedge b$ of a and b is the proposition "Alice is tall, and Bob is tall." More directly, $a \wedge b$ can be expressed as "Alice and Bob are tall." The conjunction $a \wedge b$ is true if Alice and Bob are both tall; it is false if either one is not tall. □

Definition 1.1.2.11 Let p and q be propositions. The **disjunction** of p and q, denoted $p \vee q$, is the proposition "p or q." It is false when p and q are both false and true otherwise. ◊

The disjunction is often referred to as **or** . Figure 1.1.2.12 gives the truth table for disjunction. If at least one of p or q is true, then the disjunction $p \vee q$ is true. If p and q are both false, then $p \vee q$ is false.

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Figure 1.1.2.12 Truth table for disjunction (p or q).

A disjunction is true if at least one of the propositions is true.

Example 1.1.2.13 Squares and triangles. Let s be the proposition "Squares have four sides," and let t be "Triangles have five sides."

The disjunction $s \vee t$ of s and t is the proposition "Squares have four sides, or triangles have five sides." The disjunction is true because squares have four sides. It does not matter that triangles don't have five sides. Disjunctions are only false when both propositions being combined are false. □

The disjunction of two true propositions is true.

Example 1.1.2.14 Greensboro and Japan. Let g be the proposition "I live in Greensboro," and let j be the proposition "I was born in Japan." Then g and j are both true. The disjunction $g \vee j$ is true and can be expressed as "I live in Greensboro, or I was born in Japan.". □

The only way the disjunction of two propositions is false is if both propositions are false.

Example 1.1.2.15 Mascot and motto. Let s be the proposition "UNCG's mascot is Rudy the Rodeo Clown," and let m be the the proposition "The UNCG motto is *Sleep*." The disjunction can be expressed as $s \vee m =$ "UNCG's mascot is Rudy the Rodeo Clown and motto is *Sleep*." Since the mascot is Spiro the Spartan and the motto is *Service*, both s and m are false. Thus the disjunction $s \vee m$ is also false. □

Definition 1.1.2.16 Let p and q be propositions. The **exclusive disjunction** of p and q, denoted $p \oplus q$, is the proposition "p or q but not both." It is true when exactly one of p or q is true and false otherwise. ◊

The exclusive disjunction is often referred to as **exclusive or** or **xor**. Figure 1.1.2.17 gives the truth table for exclusive disjunction.

The exclusive disjunction of one true proposition and one false proposition is true.

p	q	$p \oplus q$
T	T	F
T	F	T
F	T	T
F	F	F

Figure 1.1.2.17 Truth table for exclusive disjunction (p xor q).

Example 1.1.2.18 Ostrich and beaver. Let o be the statement "An ostrich is a bird," and let b be the statement "A beaver is a bird."

The exclusive disjunction is $o \oplus b$ is "An ostrich is a bird, or a beaver is a bird but not both.". The exclusive disjunction is true, because an ostrich is a bird and a beaver is not a bird. □

The exclusive disjunction of two false propositions is false.

Example 1.1.2.19 False xor False. The exclusive disjunction of "Dogs are reptiles," and "Snakes are mammals," is "Dogs are reptiles or snakes are mammals, but not both." The exclusive disjunction is false, because dogs are not reptiles, and snakes are not mammals. Exclusive disjunctions are true when exactly one of the propositions being combined is true. □

The exclusive disjunction of two true propositions is false.

Example 1.1.2.20 True xor True. The exclusive disjunction of "Dogs are mammals," and "Snakes are reptiles," is "Dogs are mammals or snakes are reptiles, but not both." The exclusive disjunction is false, because dogs are mammals and snakes are reptiles. Exclusive disjunctions are true when exactly one of the propositions being combined is true. □

Definition 1.1.2.21 Let p and q be propositions. The **conditional** of p and q, denoted $p \to q$, is the proposition "if p then q ." It is false when p is true and q is false and true otherwise. In the conditional $p \to q$, p is called the **hypothesis** and q is called the **conclusion** . ◊

The conditional is often referred to as **if then**. Figure 1.1.2.28 gives the truth table for the conditional.

Example 1.1.2.22 True conditional. The conditional of "$5^2 = 10$" and "$3 \cdot 4 = 12$" is "if $5^2 = 10$, then $3 \cdot 4 = 12$." The conditional is true because the the hypothesis is false. Similarly, the conditional "if $5^2 = 10$, then $3 \cdot 4 = 13$," is also true. □

Example 1.1.2.23 False conditional. The conditional "if $3 \cdot 4 = 12$, then $5^2 = 10$," is false because the the hypothesis is true, but the conclusion is false. □

Definition 1.1.2.24 Let p and q be propositions. The **biconditional** of p and q, denoted $p \leftrightarrow q$, is the proposition "p if and only if q ." In other words, $p \leftrightarrow q$ is the proposition $(p \rightarrow q) \wedge (q \rightarrow p)$. It is true when p and q have the same truth values and false otherwise. ◊

The biconditional is often referred to as **if and only if** or **iff** . Figure 1.1.2.28 gives the truth table for the conditional.

The biconditional $p \leftrightarrow q$ is true when p and q are both true.

Example 1.1.2.25 True iff True. The biconditional "$2+4 = 6$ if and only if $2 \cdot 4 = 8$." is true because both component propositions are true. Biconditionals are true exactly when the component propositions have the same truth value. □

The biconditional $p \leftrightarrow q$ is also true when p and q are both false.

Example 1.1.2.26 False iff False. The biconditional "$2 + 4 = 7$ if and only if $2 \cdot 4 = 9$" is is true because both component propositions are false. Biconditionals are true exactly when the component propositions have the same truth value. □

The biconditional $p \leftrightarrow q$ is false when p is true and q is false, or when p is false and q is true.

Example 1.1.2.27 False iff True. The biconditional "$1 = 2$ if and only if $3 = 3$" is false, because "$1 = 2$" is false and "$3 = 3$" is true. □

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

Figure 1.1.2.28 Truth table for conditional (if p then q).

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

Figure 1.1.2.29 Truth tables for biconditional (p if and only if q).

Example 1.1.2.30 Singing in a store. Suppose p be the proposition "I'm in a store," and let q be the proposition "I'm singing." Then we have the following.

1. $\neg p$ is "I'm not in a store."
2. $p \wedge q$ is "I'm in a store, and I'm singing."
3. $p \vee q$ is "I'm in a store, or I'm singing."

4. $p \oplus q$ is "I'm in a store, or I'm singing but not both."
5. $p \to q$ is "If I'm in a store, then I'm singing."
6. $p \leftrightarrow q$ is "I'm in a store if and only if I'm singing."

□

Example 1.1.2.31 Swimming with sharks. Let p be the proposition "Swimming is allowed," and let q be the proposition "Sharks have been spotted."

- "Swimming is not allowed," can be written symbolically as $\neg p$.
- "Sharks have been spotted, but swimming is allowed," can be written as $q \wedge p$. Note that "but" has been converted logically to "and" in this context. We will see more trickiness of the English language in Subsection 1.1.6.
- "If sharks have not been spotted, then swimming is allowed," can be written as $\neg q \to p$.

□

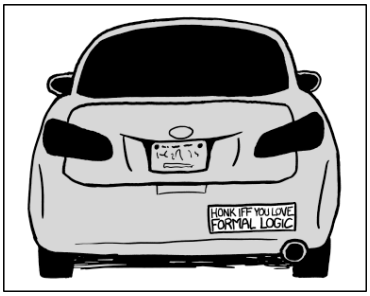


Figure 1.1.2.32 *Formal Logic* by Randal Munroe (https://xkcd.com/1033/. Note that this implies you should NOT honk solely because I stopped for a pedestrian and you're behind me.

1.1.3 Truth tables

The **truth table** for a compound proposition gives the possible truth values of a compound proposition in terms of the truth values of the original propositions.

Remark 1.1.3.1 In general, if a compound proposition involves k propositional variables, the truth table will have 2^k rows.

The truth tables of the standard compound propositions are given in Figure 1.1.2.2, Figure 1.1.2.6, Figure 1.1.2.12, Figure 1.1.2.17, Figure 1.1.2.28, and Figure 1.1.2.29.

Example 1.1.3.2 Truth table for $(p \vee \neg q) \to (p \wedge q)$. We construct the truth table of the compound proposition

$$(p \vee \neg q) \to (p \wedge q).$$

Solution. There are two propositional variables, so the table will have $2^2 = 4$

rows.

p	q	$\neg q$	$p \vee \neg q$	$p \wedge q$	$(p \vee \neg q)n \rightarrow (p \wedge q)$
T	T	F	T	T	T
T	F	T	T	F	F
F	T	F	F	F	T
F	F	T	T	F	F

□

1.1.4 Converse, contrapositive, and inverse

Definition 1.1.4.1 Let p and q be propositions, and consider the conditional $p \rightarrow q$. The **converse** of $p \rightarrow q$ is $q \rightarrow p$. The **contrapositive** of $p \rightarrow q$ is $\neg q \rightarrow \neg p$. The **inverse** of $p \rightarrow q$ is $\neg p \rightarrow \neg q$. ◊

The truth tables for the converse, contrapositive, and inverse of a conditional statement are given in Figure 1.1.4.3.

Example 1.1.4.2 Converse, contrapositive, and inverse. Consider the conditional "If I studied, then I passed the course."

- The converse is the conditional "If I passed the course, then I studied."
- The contrapositive is the conditional "If I did not pass the course, then I did not study."
- The inverse is the conditional "If I did not study, then I did not pass the course."

□

				Conditional	Converse	Contrapositive	Inverse
p	q	$\neg p$	$\neg q$	$p \rightarrow q$	$q \rightarrow p$	$\neg q \rightarrow \neg p$	$\neg p \rightarrow \neg q$
T	T	F	F	T	T	T	T
T	F	F	T	F	T	F	T
F	T	T	F	T	F	T	F
F	F	T	T	T	T	T	T

Figure 1.1.4.3 Truth tables for converse, contrapositive, and inverse of the conditional $p \rightarrow q$

Remark 1.1.4.4 Figure 1.1.4.3 shows that a conditional and its contrapositive have the same truth values. It also shows that the inverse of a conditional has the same truth values as the converse.

1.1.5 Logical operators

Negation can be viewed as an operator that takes a proposition and returns a proposition. Similarly, conjunction, disjunction, conditional, and biconditional can be viewed as binary operators that take two propositions and returns a proposition. Similarily to the precedence of operations for arithmetic operators, there is a prescribed order of precedence for the logical operators. It is given in Figure 1.1.5.1.

Precedence	Operator
1	$\neg$
2	$\wedge$
3	$\vee$
4	$\rightarrow$
5	$\leftrightarrow$

Figure 1.1.5.1 Precedence Order for Logical Operators

For example, the precedence order means that $\neg p \wedge q$ means the same thing as $(\neg p) \wedge q$. Omitting the parentheses in this case is standard and acceptable. We will allow this.

The precedence also means that $p \wedge q \vee r$ means $(p \wedge q) \vee r$. The omission of parentheses in this case is confusing. We will not allow this in this course.

Another example that will not be allowed is the following. The precedence gives that $p \wedge q \rightarrow r$ means $(p \wedge q) \rightarrow r$. In this case also, parentheses should be added for clarity.

1.1.6 English is hard

The English language is tricky and subtle. There are many ways to say the same thing, and there are many preconceptions that people have that must be let go when studying logic. In this section, we give lots of examples and highlight some of the common misconceptions.

Example 1.1.6.1 Happy in the rain. Let p be the proposition "It is raining," and let q be the proposition "I am happy." Then the following are some of the correct ways to express $p \wedge q$.

- It is raining, and I am happy.
- It is raining, but I am happy.
- It is raining, yet I am happy.
- Although it is raining, I am happy.

□

For propositions p and q, the following are correct ways to express $p \rightarrow q$.

- If p , then q.
- If p , q.
- p implies q.
- p only if q.
- p is sufficient for q.
- A necessary condition for p is q.
- q whenever p.
- q if p.
- q follows from p.
- q unless $\neg p$.

- A sufficient condition for q is p.
- q is necessary for p.

Note that "p only if q" says that p cannot be true when q is not true. Similarly, "q unless $\neg p$" says that if $\neg p$ is false, then q must be true. Looking at the truth tables, we see that they are restatements of if p , then q.

Warning 1.1.6.2 Conditionals are independent of cause and effect. The hypothesis and conclusion may not be related. There is no implied cause and effect or any other type of relationship between the statements in a true conditional proposition.

Example 1.1.6.3 Conditional and causation. "If the sun comes up in the morning, then $2 + 2 = 4$." This conditional is true since the hypothesis and conclusion are true. The conditional does not say anything about causation. □

Let's look at some English conditionals. We will rewrite them in the "if p then q" form to make them easier to analyze. Recall the truth table for $p \to q$:

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Example 1.1.6.4 Pass the course with effort. Consider the proposition "You will pass this course only if you put in the effort." Recall the conditional $p \to q$ can be phrased as "p only if q." We can rewrite this as "If you pass this course, then you must have put in the effort."

Note: It is possible that you fail this course, even if you put in the effort. This conditional says that if you pass this course, then we can deduce that you put in the effort. □

Example 1.1.6.5 Pass the course unless. Consider the proposition "You will pass this course unless you don't put in the effort." Recall the conditional $p \to q$ can be phrased "q unless not p." Thus we can rewrite the given proposition as "If you put in the effort, then you will pass this course." Note: It is possible to pass the course even if you do not put in the effort, but if you put in the effort, we can deduce that you will pass this course. □

Example 1.1.6.6 Wash the boss's car. Consider the proposition "It is necessary to wash the boss's car to get promoted."

Recall that the conditional $p \to q$ can be phrased as "q is necessary for p". Namely, in order for p to be true, q must be true. The given proposition is that washing the car is necessary for the promotion. We can rephrase that as "If you get promoted, then you must have washed the boss's car."

A useful way to think of these things is in terms of an obligation or contract. Note: This proposition makes no claims about what will happen if you wash the boss's car. Specifically, if you wash the boss's car, you should not expect to be promoted because of that. □

Example 1.1.6.7 Roller coaster. Express the following sentence symbolically:

"You cannot ride the roller coaster if you are under 4 feet tall unless you are older than 16 years old."

Let r , f, and s be the propositions

1. r is "You can ride the roller coaster,"

2. f is "You are under 4 feet tall,"

3. s is "You are older than 16 years old."

Then we the sentence can be expressed symbolically as

$$\neg s \to (f \to \neg r).$$

Can you convince yourself that the statement can also be expressed as

$$(f \wedge \neg s) \to \neg r?$$

We will see how to reconcile this in Section 1.2. □

Example 1.1.6.8 Truth-tellers and liars. An island has two tribes, truth-tellers and liars. You encounter two people, Alice and Bob, on the island. Alice says "Bob is a truth-teller," and Bob says, "We are from different tribes." Let's figure out who is from which tribe.

Let a be the proposition "Alice is a truth-teller," and let b be "Bob is a truth-teller.".

Then Alice's statement is b . Bob's statement is that a and b have opposite truth values, so his statement is $a \leftrightarrow \neg b$. Note that Alice's statement must have the same truth value as a because her statement is true if and only if she is a truth-teller. That means $a \leftrightarrow b$ must be true. Analogously, Bob's statement has the same truth value as b , so $(a \leftrightarrow \neg b) \leftrightarrow b$ must be true.

Let's work out the truth table. We are looking for a row where $a \leftrightarrow b$ and $(a \leftrightarrow \neg b) \leftrightarrow b$ are both true.

a	b	$\neg b$	$a \leftrightarrow \neg b$	$a \leftrightarrow b$	$(a \leftrightarrow \neg b) \leftrightarrow b$
T	T	F	F	T	F
T	F	T	T	F	F
F	T	F	T	F	T
F	F	T	F	T	T

From the last two columns, we see that $b \leftrightarrow a$ and $(a \leftrightarrow \neg b) \leftrightarrow b$ are both true in the row where a and b are false. In other words, we have that Alice and Bob are both liars. □

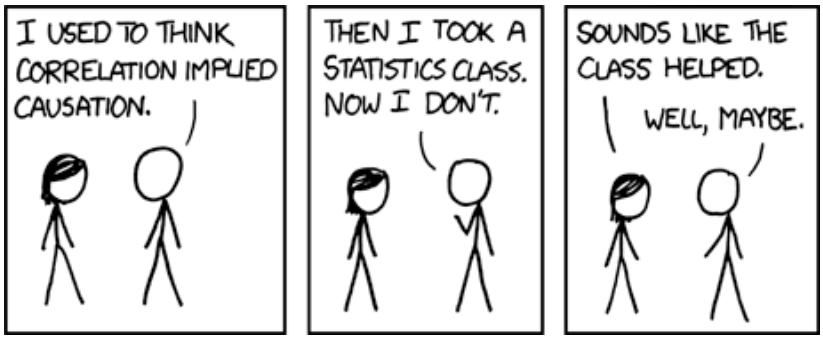


Figure 1.1.6.9 *Correlation* by Randal Munroe (https://xkcd.com/552/). Correlation doesn't imply causation, but it does waggle its eyebrows suggestively and gesture furtively while mouthing `look over there'.

1.1.7 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) proposition

(b) negation of a proposition

(c) conjunction of two propositions

(d) disjunction of two propositions

(e) exclusive disjunction of two propositions

(f) conditional of two propositions

(g) biconditional of two propositions

(h) converse of a conditional

(i) contrapositive of a conditional

(j) inverse of a conditional

2. Complete the following truth table.

p	q	$p \vee q$	$p \oplus q$	$p \wedge q$	$p \rightarrow q$	$p \leftrightarrow q$
T	T					
T	F					
F	T					
F	F					

3. Identify each of the following as a proposition or not. For each proposition, give the truth value.

(a) Greensboro is the capital of North Carolina.

(b) Let's meet at the dining hall at 7pm.

(c) Squares have three sides.

(d) $2 \cdot 5 = 11$.

(e) $a^2 + b^2 = c^2$.

4. Identify each of the following as a proposition or not. For each proposition, give the truth value.

(a) $1 + 1 = 2$

(b) Buy a bag of chips.

(c) If $2 + 3 = 7$, then $11 - 4 = 7$.

(d) If $11 - 4 = 7$, then $2 + 3 = 7$.

(e) 3 is odd or 253 is odd.

(f) If the sun came up today, then squares have four

5. Identify each of the following as a proposition or not. For each proposition, give the truth value.

(a) Cats are reptiles.

(b) Dogs are mammals.

(c) If cats are reptiles, then dogs are mammals.

(d) If dogs are mammals, then cats are reptiles.

(e) Cats are reptiles if and only if dogs are mammals.

6. What is the negation of each of these propositions?

(a) There are 60 minutes in an hour.

(b) 169 is a perfect square.

(c) Alice is less than 5 feet tall.

(d) Bob has more than 20 books.

(e) $5 \cdot 6 = 30$

7. Let r be the proposition "It is raining." and let w be "The ground is wet."
Express each of the propositions as an English sentence.

(a) $\neg r$

(b) $r \wedge w$

(c) $r \rightarrow w$

(d) $\neg r \rightarrow \neg w$

(e) $\neg r \wedge \neg w$

8. Let p be the propositions "There are 8 pints in a gallon." and let q be "There are 4 quarts in a gallon.".
Write each of these propositions using p and q and logical operators (including negations).

(a) There are 8 pints in a gallon, or there are 4 quarts in a gallon.

(b) If there are 8 pints in a gallon, then there are not 4 quarts in a gallon.

(c) There are not 8 pints in a gallon, or there are 8 pints in a gallon and 4 quarts in a gallon.

(d) There are 8 pints in a gallon if and only if are 4 quarts in a gallon.

(e) There are not 8 pints in a gallon, but there are 4 quarts in a gallon.

9. Let p be the proposition "You have the flu." and let q be "You miss the final exam." and let r be "You pass the course."
Express each proposition as an English sentence.

(a) $p \rightarrow (q \vee r)$

(b) $(\neg q \wedge r) \vee r$

(c) $p \rightarrow (q \wedge r)$

(d) $(\neg p \vee \neg q) \rightarrow r$

(e) $(p \wedge q) \vee (\neg q \wedge r)$

10. Write each of these statements in the form "if p then q."

(a) It is necessary to study every day to pass MAT 253.

(b) Alice gets caught whenever she cheats.

(c) To pass MAT 253, it is sufficient to attend class every day.

(d) The warranty is good only if you bought the computer less than 1 year ago.

(e) Bob will get a good job unless he does not learn discrete mathematics.

11. State the converse, contrapositive, and inverse of each of these conditional statements.

(a) If it is snowing tonight, I will stay home.

(b) The home team wins whenever it rains.

(c) Alice eats a desert only if she eats her vegetables.

(d) When Bob stays up late, it is necessary for him to sleep until noon.

(e) A positive integer is prime only if it has no positive divisors other than 1 and itself.

12. Construct a truth table for each of these compound propositions.

(a) $(p \wedge q) \rightarrow r$

(b) $(p \vee q) \rightarrow \neg r$

(c) $p \rightarrow (q \wedge r)$

(d) $\neg q \rightarrow (r \vee \neg p)$

(e) $p \wedge (q \vee \neg r)$

1.2 Propositional equivalence

Outcomes

1. Prove propositional equivalences
2. Reproduce the most important such equivalences.

1.2.1 Terminology

Some compound propositions are true for a silly reason. Basically, the compound proposition is true, regardless of the truth values of the component propositions. For example, "I'll get to it when I get to it." If g is the proposition "I get to it,", then this compound proposition is the conditional $g \rightarrow g$. Whether g is true or false, the conditional is true. This is an example of a tautology.

Definition 1.2.1.1 A **tautology** is a compound proposition that is true no matter what the truth values of the propositional variables that occur in it. $\Diamond$

Example 1.2.1.2 Tautology. Let p be a proposition. Then $p \to p$ is a tautology. To see this, we compute the truth table for $p \to p$.

p	$p \to p$
T	T
F	T

Since $p \to p$ is always true, $p \to p$ is a tautology. □

Definition 1.2.1.3 A **contradiction** is a compound proposition that is always false no matter what the truth values of the propositional variables that occur in it. ◊

Example 1.2.1.4 Contradiction. Let p be a proposition. Then $\neg(p \to p)$ is a contradiction. To see this, we compute the truth table for $\neg(p \to p)$.

p	$p \to p$	$\neg(p \to p)$
T	T	F
F	T	F

Since $\neg(p \to p)$ is always false, $\neg(p \to p)$ is a contradiction. □

Remark 1.2.1.5 In general, if P is a tautology, then $\neg P$ is a contradiction.

Definition 1.2.1.6 A **contingency** is a compound proposition that is neither a tautology nor a contradiction. ◊

Example 1.2.1.7 Contingency. Let p and q be propositions. The conditional $p \to q$ is a contingency. To see this, compute the truth table for $p \to q$.

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

Since $p \to q$ is true for some choices of p and q and false for some other choices, $p \to q$ is a contingency. □

Example 1.2.1.8 Contradiction and Tautology. Let p be a proposition. Then $p \wedge \neg p$ is a contradiction and $p \vee \neg p$ is a tautology as shown in the table below. Not that the column for $p \wedge \neg p$ is all false, and the column for $p \vee \neg p$ is all true.

p	$\neg p$	$p \wedge \neg p$	$p \vee \neg p$
T	F	F	T
F	T	F	T

□

Definition 1.2.1.9 Two propositions p and q are **logically equivalent** denoted $p \equiv q$,when $p \leftrightarrow q$ is a tautology. ◊

Theorem 1.2.1.10 Contrapositive. *Let p and q be propositions then the conditional $p \to q$ is logically equivalent to its contrapositive:*

$$p \to q \equiv \neg q \to \neg q$$

and its converse is equivalent to its inverse:

$$\neg q \to \neg p \equiv \neg p \to \neg q.$$

Proof. By looking at the truth table in Figure 1.1.4.3, we see that a conditional $p \to q$ is logically equivalent to its contrapositive $\neg q \to \neg q$, because the truth values in the respective columns are the same, which proves the first equivalence.

Also the truth values in the columns for the converse $\neg q \to \neg p$ are the same as the truth values in the column of the inverse $\neg p \to \neg q$, which proves the second equivalence. ■

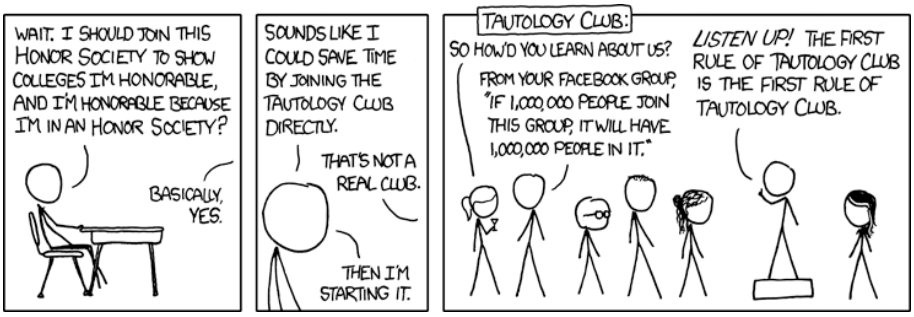


Figure 1.2.1.11 *Honor Societies* by Randal Muroe (https://xkcd.com/703/). Hey, why do YOU get to be the president of Tautology Clu--- wait, I can guess.

1.2.2 Important propositional equivalences

The following theorem can be thought of as telling us how to distribute a "not" across an "and" or an "or".

Theorem 1.2.2.1 De Morgan's laws for propositions. *Let p and q be propositions.*

1. $\neg(p \wedge q) \equiv \neg p \vee \neg q$
2. $\neg(p \vee q) \equiv \neg p \wedge \neg q$

Proof. We prove $\neg(p \vee q) \equiv \neg p \wedge \neg q$ and leave the other as an exercise. First compute the truth table.

p	q	$\neg p$	$\neg q$	$p \vee q$	$\neg(p \vee q)$	$\neg p \wedge \neg q$
T	T	F	F	T	F	F
T	F	F	T	T	F	F
F	T	T	F	T	F	F
F	F	T	T	F	T	T

Since the last two columns are the same, we have $\neg(p \vee q) \equiv \neg p \wedge \neg q$ as desired. ■

Example 1.2.2.2 Jake is wearing khakis. Let's negate the statement "Jake is wearing khakis and sounds hideous."

Let j be "Jake is wearing khakis," and let h be "Jake sounds hideous." Then the original statement is $j \wedge h$. By De Morgan's law, the negation is

$$\neg(j \wedge h) = \neg j \vee \neg h.$$

This is, "Jake is not wearing khakis, or he does not sound hideous." □

There are several other important propositional equivalences that we should know.

Theorem 1.2.2.3 Proposition Identities I. *Let p be a proposition.*

Identity laws	$p \wedge T \equiv p$ *and* $p \vee F \equiv p$.
Domination laws	$p \vee T \equiv T$ *and* $p \wedge F \equiv F$.
Idempotent laws	$p \vee p \equiv p$ *and* $p \wedge p \equiv p$.
Double negation law	$\neg(\neg p) \equiv p$
Negation law	$p \vee \neg p \equiv T$ *and* $p \wedge \neg p \equiv F$.

Proof. Exercise. Just compute the truth tables and verify the corresponding columns are the same. ■

Theorem 1.2.2.4 Proposition Identities II. *Let p and q be propositions.*

Commutative laws	$p \vee q \equiv q \vee p$ *and* $p \wedge q \equiv q \wedge p$.
Absorption laws	$p \vee (p \wedge q) \equiv p$ *and* $p \wedge (p \vee q) \equiv p$.
Contrapositive law	$p \to q \equiv \neg q \to \neg p$

Proof. Exercise. Just compute the truth tables and verify the corresponding columns are the same. ■

Theorem 1.2.2.5 Proposition Identities III. *Let p , q, and r be propositions.*

Associative laws	$p \vee (q \vee r) \equiv (p \vee q) \vee r$; $p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$.
Distributive laws	$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$; $p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$.

Proof. Exercise. Just compute the truth tables and verify the corresponding columns are the same. ■

We can also use logical equivalences to turn conditionals into disjunctions.

Theorem 1.2.2.6 *Let p and q be propositions. Then*

$$p \to q \equiv \neg p \vee q.$$

Proof. We compute the relevant truth table.

p	q	$\neg p$	$\neg p \vee q$	$p \to q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Since the last two columns are equal, we have $p \to q \equiv \neg p \vee q$. ■

Once we have proven all of these important logical equivalences, we can use them to prove additional equivalences and simplify compound propositions without resorting to truth tables.

Example 1.2.2.7 Roller coaster (continued). Recall Example 1.1.6.7, where we translated "You cannot ride the roller coaster if you are under 4 feet tall unless you are older than 16 years old." Let

- r be "You can ride the roller coaster,"
- f be "You are under 4 feet tall,"
- s be "You are older than 16 years old."

Then we had two seemingly correct translations

$$\neg s \to (f \to \neg r)$$

and

$$(f \wedge \neg s) \to \neg r.$$

Let's see that these statements are logically equivalent by simplifying each one. We change each conditional to a disjunction and simplify.

$$\begin{aligned} \neg s \to (f \to \neg r) &\equiv s \vee (f \to \neg r) \\ &\equiv s \vee (\neg f \vee \neg r) \end{aligned}$$

$$\begin{aligned} (f \wedge \neg s) \to \neg r &\equiv \neg(f \wedge \neg s) \vee \neg r \\ &\equiv (\neg f \vee s) \vee \neg r \\ &\equiv (s \vee \neg f) \vee \neg r \\ &\equiv s \vee (\neg f \vee \neg r) \end{aligned}$$

Since both statements are equivalent to $s \vee (\neg f \vee \neg r)$, the statements are equivalent to each other. □

1.2.3 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) tautology

(b) contradiction

(c) contingency

(d) logically equivalent propositions

2. State precisely De Morgan's laws for propositions. Be sure to set up any notation that is required.

3. Use De Morgan's laws to find the negation of these statements.

(a) Alice will go to graduate school or get a job in industry.

(b) Bob majored in math and computer science.

(c) Carl is tall and thin.

(d) Dan has a laptop and a desktop.

(e) Eve or Frank will pick you up at the airport.

4. Show that each of these conditional statements is a tautology by using truth tables.

(a) $p \to (p \vee q)$

(b) $(p \wedge q) \to p$

(c) $(p \wedge q) \to (p \to q)$

(d) $(\neg p \wedge (p \vee q)) \to q$

(e) $\neg p \to (p \to q)$

5. Complete the following truth table. Is $p \to q$ is logically equivalent to $\neg p \vee q$? Justify. Be sure to say what portion of the computation explains your response.

p	q	$\neg p$	$\neg p \vee q$	$p \to q$
T	T			
T	F			
F	T			
F	F			

6. Verify these associative law using a truth table. Which columns show the logical equivalence?

(a) $(p \vee q) \vee r \equiv p \vee (q \vee r)$

(b) $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$

7. Verify these absorptions law using a truth table. Which columns show the logical equivalence?

(a) $p \vee (p \wedge q) \equiv p$

(b) $p \wedge (p \vee q) \equiv p$

8. Show that $\neg(p \oplus q)$ and $p \leftrightarrow q$ are logically equivalent.

9. Show that $(p \to q) \vee (p \to r)$ and $p \to (q \vee r)$ are logically equivalent.

10. Show that $(p \wedge q) \to r$ and $(p \to r) \wedge (q \wedge r)$ are not logically equivalent.

11. Show that $p \leftrightarrow q$ and $\neg p \leftrightarrow \neg q$ are logically equivalent.

12. Prove the proposition identities in Theorem 1.2.2.3.

13. Prove the proposition identities in Theorem 1.2.2.4.

14. Prove the proposition identities in Theorem 1.2.2.5.

1.3 Predicates and quantifiers

Outcomes

1. Formulate propositions with predicate logic.
2. Translate between English sentences (or mathematical statements) and logical expressions.
3. Prove existential quantification with witnesses
4. Disprove universal quantification with witnesses

1.3.1 Predicates

Statements involving variables, such as "x is greater than 3" are often found in mathematical assertions. Such statements are not propositions as they are neither true nor false when the value of the variable is not specified. The statement has two parts. The first part, the variable x , is the **subject** . The second part, the **predicate** , is the property that the subject can have. In this case, the predicate is "is greater than 3". We denote such statements as $P(x)$, which we read as the **propositional function** P evaluated at x.

Definition 1.3.1.1 A **propositional function** P is a function which for each input outputs a proposition. ♢

Example 1.3.1.2 Propositional function $x > 3$. Let $P(x)$ be the statement "$x > 3$".

Then $P(2)$ is the proposition "$2 > 3$," which is false.

The proposition $P(4)$ is the statement "$4 > 3$", which is true. □

Propositional functions can have many variables:

Example 1.3.1.3 Propositional function with three variables. Let $P(x, y, z)$ be "x, y, and z live in the same dorm.".

Then $P(\text{Alice}, \text{Bob}, \text{Carl})$ is the proposition "Alice, Bob, and Carl live in the same dorm." □

Definition 1.3.1.4 The collection of all inputs that we allow for P is the **domain of the propositional function** P .

Let x be in the domain of a propositional function P. Then $P(x)$ is a proposition. The property of x given by the proposition $P(x)$ is called the **predicate** of $P(x)$. ♢

Definition 1.3.1.5 The **integers** are the numbers:

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

♢

Example 1.3.1.6 Predicate. Let $P(x) =$ "$x < 5$" with domain the integers.

The predicate of P is "is less than five" or "< 5".

1. $P(4)$ is "$4 < 5$" which is true.
2. $P(10)$ is "$10 < 5$" which is false.

□

Example 1.3.1.7 Determine whether $Q(1)$ and $Q(2)$ are true. Let $Q(x)$ be "$(x + 3)^3 - 25 = (5x)^2$".

Determine whether

1. $Q(1)$ is true or false and
2. $Q(2)$ is true or false.

Answer. $Q(1)$ is false and $Q(2)$ is true.

Solution.

1. We first determine whether $Q(1)$ is true or false.

   For $x = 1$ we have that the left hand side of "$Q(x)$" is

$$(x + 3)^3 - 25 = (1 + 3)^3 - 25 = 4^3 - 25 = 64 - 25 = 39$$

and the right hand side of "$Q(x)$" is

$$(5x)^2 = 5^2 = 25.$$

Because $39 \neq 25$ we conclude that $Q(1)$ is false.

2. We now determine whether $Q(2)$ is true or false.
   For $x = 2$ we have that the left hand side of "$Q(x)$" is

$$(x+3)^3 - 25 = (2+3)^3 - 25 = 5^3 - 25 = 125 - 25 = 100$$

and the right hand side of "$Q(x)$" is

$$(5x)^2 = 10^2 = 100.$$

Because the left hand side and the right hand side of $Q(2)$ are equal we conclude that $Q(2)$ is true.

□

Example 1.3.1.8 Prove $P(1)$ and $P(2)$ are true. Let $P(n)$ be the proposition "$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$."

Prove $P(1)$ and $P(2)$ are true.

Answer. $P(1)$ is the statement "$1 = \frac{1(1+1)}{2}$." This is true since the right side is

$$\frac{1(1+1)}{2} = \frac{2}{2} = 1,$$

which is equal to the left side.

Similarly, $P(2)$ is the statement "$1 + 2 = \frac{2(2+1)}{2}$." This is true since the left side is

$$1 + 2 = 3,$$

and the right side is

$$\frac{2(2+1)}{2} = \frac{6}{2} = 3.$$

□

1.3.2 Universal quantification

Quantification expresses the extent to which a propositional function is true; for example, all, some, none, many, and so on.

Definition 1.3.2.1 The **universal quantification** of P , denoted $\forall x\, P(x)$, is the proposition "$P(x)$ for all x in the domain." The symbol $\forall$ is the **universal quantifier** The domain of P specifies the possible values for x . ◊

The domain of a propositional function matters. In Example 1.3.2.2 and Example 1.3.2.4 we consider propositional functions with the same predicate but with a different domain.

Example 1.3.2.2 For all integers. Let $S(x)$ be "$x > -2$" with domain the integers. Consider the proposition

$$\forall x\, S(x)$$

Because -5 is an integer and "$-5 > -2$" is false $\forall x\, S(x)$ is false. □

The propositional function in the following example has the domain the natural numbers. We recall their definition:

Definition 1.3.2.3 The **natural numbers** are the numbers:

$$1, 2, 3, 4, 5, 6, \ldots$$

$\Diamond$

Example 1.3.2.4 For all natural numbers. Let $R(x)$ be "$x > -2$" with domain the natural numbers. Consider the proposition

$$\forall x\, R(x)$$

Because all natural numbers are positive and $x > -2$ is true for all positive numbers we have that $\forall x\, R(x)$ is true. $\square$

Remark 1.3.2.5 The universal quantification of a propositional function is a proposition, since it is a declarative sentence that is either true or false, but not both.

Definition 1.3.2.6 Let $P(x)$ be a propositional function. A **counterexample** of $\forall x\, P(x)$ is an element x_0 such that $P(x_0)$ is false. $\Diamond$

Remark 1.3.2.7 To prove a universal quantification is false, it suffices to exhibit a single counterexample.

In Example 1.3.2.2 we haven given the counterexample $x = -5$ to prove that the proposition "all integers are greater than -2" is false.

Example 1.3.2.8 Single counterexample. Let $P(x)$ be "$x^3 \geq 0$" with domain the integers. Then $\forall x\, P(x)$ is false. To show this, it is enough to provide a single counterexample. Because $(-2)^3 = 8$ and "$-8 > 0$" is false, so $P(-2)$ is false. Thus $\forall x\, P(x)$ is false. $\square$

Proof Technique 1.3.2.9 To show $\forall x\, P(x)$ is false. *Suppose P is a propositional function, and we want to prove the universal quantification $\forall x\, P(x)$ is false.*

1. *Find a counterexample. Specifically, find x_0 in the domain of P such that $P(x_0)$ is false.*
2. *Conclude $\forall x\, P(x)$ is false.*

Example 1.3.2.10 All dogs have brown fur. Consider the statement "All dogs have brown fur." This universally quantified statement is false. My dog Duey provides a counterexample.

Let's examine this in more detail. Let $B(x)$ be "x has brown fur," with domain the set of all dogs. The universally quantified statement "All dogs have brown fur," can be written as $\forall x\, B(x)$. This is false, because we can produce a counterexample, namely my dog Duey. He is a dog, so he is in the domain of B . He does not have brown fur, so B(Duey) is false. Thus $\forall x\, B(x)$ is false. In other words, not all dogs have brown fur. $\square$

Proof Technique 1.3.2.11 To show $\forall x\, P(x)$ is true. *Suppose P is a propositional function, and we want to prove the universal quantification $\forall x\, P(x)$ is true.*

1. *Fix a generic element x in the domain of P.*
2. *Show $P(x)$ is true for this fixed generic element.*
3. *Conclude $\forall x\, P(x)$ is true.*

Example 1.3.2.12 Squares of integers are positive. Let $P(x)$ be "$x^2 \geq 0$" with domain the integers. Then $\forall x\, P(x)$ is true. To see this, fix a generic integer x . Then x^2 is non-negative since the square of any real number is non-negative. Thus $\forall x\, P(x)$ is true. □

Theorem 1.3.2.13 *Let $S(a, b)$ be "$(a + b)(a - b) = a^2 - b^2$" with domain the integers for a and b*

Then $\forall a, b\, S(a, b)$ (is true).

Proof. Let a and b be integers. Starting on the left hand side of $S(a, b)$ we get

$$\begin{aligned}(a + b)(a - b) &= a(a - b) + b(a - b) \text{by the distirbutive law} \\ &= a^2 - ab + ba - b^2 \\ &= a^2 - b^2.\end{aligned}$$

Thus $S(a, b)$ is true for all integers a and b. ■

In the proof of Theorem 1.3.2.15 we will apply the following result:

Theorem 1.3.2.14 *The product of two negative integers is positive and the product of two positive integers is positive.*

Theorem 1.3.2.15 *Let $P(x) =$ "$x^2 \geq 0$" with domain the integers. Then $\forall x\, P(x)$ is true.*

Proof. We use a generic element to prove Theorem 1.3.2.15.

1. Let b be an integer.
2. If $b = 0$ then $b^2 = a \cdot b = 0 \cdot 0 = 0 \geq 0$ thus $P(b) = P(0)$ is true.

If b is negative, by Theorem 1.3.2.14, we have that $b^2 = b \cdot b$ is positive. Thus for negative b we have $P(b)$ is true.

If b is positive, by Theorem 1.3.2.14, we have that $b^2 = b \cdot b$ is positive. Thus for positive b we have $P(b)$ is true.

Because each integer b is either equal to zero or negative or positive and $b^2 \geq 0$ in each of the three cases we have proven that $b^2 \geq 0$ for all integers b. ■

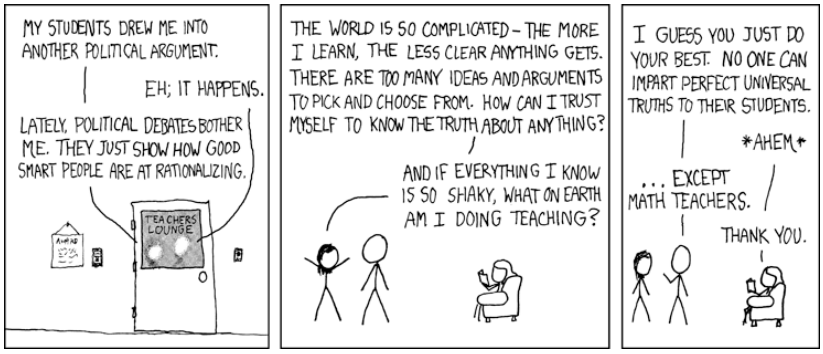


Figure 1.3.2.16 *Certainty* by Randal Muroe (https://xkcd.com/263/). $a(b + c) = (ab) + ac$. Polticize that bitches.

1.3.3 Existential quantification

Definition 1.3.3.1 The **existential quantification** of P , denoted $\exists x\, P(x)$ is the proposition "There exists an element x in the domain such that $P(x)$." The symbol $\exists$ is the **existential quantifier**. ◊

Example 1.3.3.2 Existential quantification is true. Let $Q(x)$ be "$x < 0$" with domain the integers. Consider the proposition

$$\exists x\, Q(x)$$

Because $Q(-3) =$ "$-3 < 0$" is true and -3 is an integer, we have found an element in the domain for which $Q(x)$ is true. Thus the proposition $\exists x\, Q(x)$ is true. □

Example 1.3.3.3 Existential quantification is false. Let $R(x) =$ "$x < 0$" with domain the natural numbers. Consider the proposition

$$\exists x\, R(x)$$

Because all natural numbers are positive there is no natural number b such that $b < 0$. Thus $\exists x\, R(x)$ is false. □

Definition 1.3.3.4 A **witness** of $\exists x\, P(x)$ is an element x_0 in the domain of P such that $P(x_0)$. ◊

Remark 1.3.3.5 To prove an existential quantification is true, it suffices to exhibit a single witness.

Proof Technique 1.3.3.6 To show $\exists x\, P(x)$ is true. *Suppose P is a propositional function, and we want to prove the existential quantification $\exists x\, P(x)$ is true.*

1. *Find a witness. Specifically, find an element x_0 in the domain of P such that $P(x_0)$ is true.*
2. *Conclude $\exists x\, P(x)$ is true.*

Example 1.3.3.7 Witness. Let $P(x)$ be "$x^4 = 16$" with domain the integers.

1. $\exists x\, P(x)$ is true since $2^4 = 16$ and 2 is in the domain of P. We say $x = 2$ is a witness for $\exists x\, P(x)$.
2. $\forall x\, P(x)$ is false since $3^4 = 81 \neq 16$ and 3 is in the domain of P. We say $x = 3$ is a counterexample for $\forall x\, P(x)$.

□

To show an existential quantification false, we need to show that a witness does not exist. Specifically, to show $\exists x\, P(x)$ is false, we need to show that $P(x)$ is false for every x in the domain of P .

Proof Technique 1.3.3.8 To show $\exists x\, P(x)$ is false. *Suppose P is a propositional function, and we want to prove the existential quantification $\exists x\, P(x)$ is false.*

1. *Fix a generic element x in the domain of P.*
2. *Show $P(x)$ is false.*
3. *Conclude $\exists x\, P(x)$ is false.*

Remark 1.3.3.9 To prove the existential quantification of P is true, it is enough to exhibit a single witness. To prove the existential quantification of P is false, we need to show that the universal quantification of $\neg P$ is true.

We wrap up this section with a comparison of the proof techniques for universal and existential quantification.

Proposition	When true?	When false?
$\forall x\, P(x)$	$P(x)$ is true for every x	There is a counterexample x_0 for which $P(x_0)$ is false.
$\exists x\, P(x)$	There is a witness x_0 for which $P(x_0)$ is true	$P(x)$ is false for every x

Figure 1.3.3.10 Quantifiers and proof techniques.

Example 1.3.3.11 MAT 253 students born in NC. Let $N(x)$ be the propositional function "x was born in NC," with domain the MAT 253 students.

1. $\forall x\, N(x)$
2. $\exists x\, N(x)$

What are the truth values of the two propositions above? Do you need to discuss this with your classmates to determine the truth values?

1. $\forall x\, N(x)$ says "Every student in MAT 253 was born in NC." To prove this statement false, it is enough to produce a counterexample. We need to just find one student in MAT 253 that was not born in NC. To prove this statement true, we need to check that every MAT 253 student was born in NC.
2. $\exists x\, N(x)$ says "There is a student in MAT 253 that was born in NC." To prove this statement true, it is enough to produce a witness. We need to find just find one student in MAT 253 that was not born in NC.

Now suppose instead we take the domain of N to be all people in the US, but we want to have the same propositions about our MAT 253 class. Then we need to introduce a new propositional function $S(x)$ which is "x is a MAT 253 student."

1. We can write "Every student in MAT 253 was born in NC," as $\forall x(S(x) \to N(x))$.
2. We can write "There is a student in MAT 253 that was born in NC," as $\exists x\,(S(x) \wedge N(x))$

Note that we have shown that $p \to q$ is equivalent to $\neg p \vee q$. That means that $\forall x\,(\neg S(x) \vee N(x))$ is a valid (though convoluted) way to write that every student in MAT 253 was born in NC. □

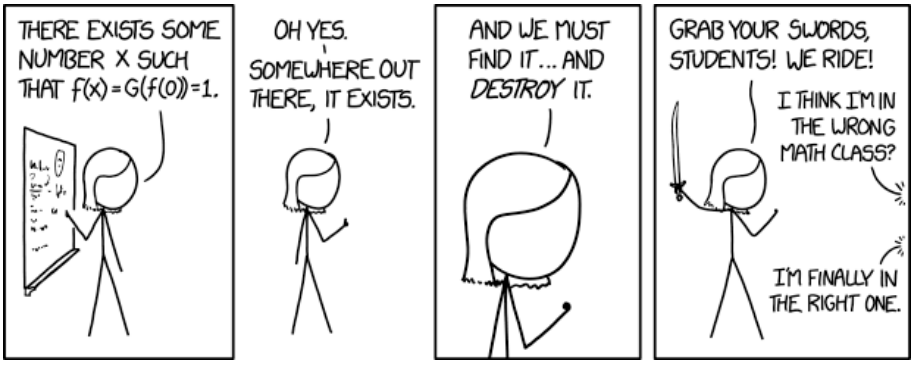


Figure 1.3.3.12 *Existence Proof* by Randal Munroe (https://xkcd.com/1856/). Real analysis is way realer than I expected.

1.3.4 Logical equivalence

When the domain of a propositional function only consists of a couple of elements we can rewrite the universal and existential quantification with the logical operations $\wedge$ and $\vee$.

Example 1.3.4.1 Quantification with domain $\{1,2,3\}$. Consider a propositional function P with domain $\{1,2,3\}$. Then

1. $\forall x\, P(x) = P(1) \wedge P(2) \wedge P(3)$
2. $\exists x\, P(x) = P(1) \vee P(2) \vee P(3)$

□

Example 1.3.4.2 De Morgan with domain $\{1,2\}$. Let $Q(x)$ be a propositional function with domain $\{1,2\}$. Then, by De Morgan's laws 1.2.2.1, we have

$$\neg\big(\forall x\, P(x)\big) = \neg\big(P(1) \wedge P(2)\big) \equiv \neg P(1) \vee \neg P(2) = \exists x\, \neg P(x)$$

and

$$\neg\big(\exists x\, P(x)\big) = \neg\big(P(1) \vee P(2)\big) \equiv \neg P(1) \wedge \neg P(2) = \forall x\, \neg P(x)$$

□

Before we state and prove the theorems analogous to De Morgan's laws for propositional functions we generalize the notion of logical equivalence.

Definition 1.3.4.3 Two statements S and T involving predicates and quantifiers are *logically equivalent*, denoted $S \equiv T$, if they have the same truth value no matter which predicates are substituted into these statements and which domain is used. ◊

Theorem 1.3.4.4 De Morgan's laws for quantifiers. *Let P be a propositional function. Then the following logical equivalences hold.*

1. $\neg\forall x\, P(x) \equiv \exists x \neg P(x)$.
2. $\neg\exists x\, P(x) \equiv \forall x \neg P(x)$.

Proof. We prove the first statement and leave the proof of the second as an exercise.

We want to show

$$\neg\forall x\, P(x) \equiv \exists x \neg P(x).$$

We just need to show that the left side and the right side have the same truth values, independent of what P or x actually is.

We first prove that if $\neg\,(\forall x\, P(x))$ is true then $\exists x\, \neg P(x)$ is true

If $\neg\,(\forall x\, P(x))$ is true then $\forall x\, P(x)$ is false.

If $\forall x\, P(x)$ is false there is a counterexample c in the domain of P such that $P(c)$ is false.

If $P(c)$ is false then $\neg P(c)$ is true.

So c is a witness for $\exists x\, \neg P(x)$, which means that $\exists x\, \neg P(x)$ is true.

We now prove that if $\neg\,(\forall x\, P(x))$ is false then $\exists x\, \neg P(x)$ is false.}

If $\neg\,(\forall x\, P(x))$ is false then $\forall x\, P(x)$ is true.

So $P(x)$ is true for all x in the domain of P.

So $\neg P(x)$ is false for all x in the domain of P.

That is, there is no x in the domain of P such that $\neg P(x)$ is true.

So $\exists x\, \neg P(x)$ is false.

As the two propositions can only be true or false this proves that $\neg(\forall x\, P(x)) \equiv \exists x\, \neg P(x)$. ■

Example 1.3.4.5 No one is perfect. Consider the statement "No one is perfect." Translate it into a logical expression using predicates and quantifiers. Then use De Morgan's laws to rewrite it.

The statement is that there does not exist a person that is perfect. Let $P(x)$ be "x is perfect," with domain the set of all people. Then the statement is $\neg\exists x\, P(x)$. By De Morgan's laws

$$\neg\exists x\, P(x) \equiv \forall x \neg P(x),$$

which says that everyone is imperfect. □

Example 1.3.4.6 Some raccoons have rabies. Consider the statement "Some raccoons have rabies."

1. Express the statement with a propositional function and a quantifier.
2. Give the negation of the expression.
3. Formulate the negation such that there is no negation to the left of the quantifier.
4. Express the negation in English.

Solution.

1. Let the domain of our propositional function be the set of all raccoons.
   Let our propositional function $F(x)$ be "x has rabies".
   Then "Some raccoons have rabies" can be expressed as $\exists x\, F(x)$
2. The negation of $\exists x\, F(x)$ is $\neg\exists x\, F(x)$
3. By De Morgan's laws we have $\neg\exists x\, F(x) \equiv \forall x\, \neg F(x)$.
4. $\forall x\, \neg F(x)$ in English is "No raccoon has rabies."

□

Example 1.3.4.7 Lewis Carroll. Consider the three statements

1. All lions are fierce.
2. Some lions do not drink coffee.
3. Some fierce creatures do not drink coffee.

We write the three statements into logical expression using predicates and quantifiers.

Let L , F, and C be propositional functions with domain all creatures, where

- $L(x)$ is given by "x is a lion,"
- $F(x)$ is given by "x is fierce," and
- $C(x)$ is given by "x drinks coffee,"

Then the three statements become:

1. $\forall x\, (L(x) \to F(x))$

2. $\exists x\,(L(x) \wedge \neg C(x))$
3. $\exists x\,(F(x) \wedge \neg C(x))$

Suppose we know that the first two statements are true. Does this allow us to deduce the third statement ?

Suppose the first two statements are true. The second statement guarantees the existence of a creature, we'll call this creature Bob, such that $L(\text{Bob})$ and $\neg C(\text{Bob})$ are both true. Since $L(\text{Bob})$ is true, the conditional in the first statement tells us that $F(\text{Bob})$ is true. Since $F(\text{Bob})$ and $\neg C(\text{Bob})$ are both true, Bob provides the example to show the third statement is true.

In other words, since some lions do not drink coffee, there must be a lion that does not drink coffee. Let's call this lion Bob. Then since all lions are fierce, we have that Bob is fierce. Since Bob is fierce and does not drink coffee, we know that some fierce creatures do not drink coffee. □

1.3.5 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) universal quantification of a propositional function

(b) existential quantification of a proposition function

(c) logically equivalent statements involving predicates and quantifiers

2. State precisely De Morgan's laws for quantifiers. Be sure to set up any notation that is required.

3. Let $P(x)$ denote the propositional function "$x \leq 10$," with domain $\mathbb{R}$. What are these truth values?

(a) $P(0)$

(b) $P(-2)$

(c) $P(253)$

(d) $\forall x\,P(x)$

(e) $\exists x\,P(x)$

4. Let $S(x)$ be the statement "x studies more than 30 hours per week," where the domain for x consists of all students in MAT 253. Express each of these quantifications in English.

(a) $\exists x\,S(x)$

(b) $\exists x \neg S(x)$

(c) $\forall x\,S(x)$

(d) $\forall x \neg S(x)$

5. Let $B(x)$ be the statement "x has a bird," let $C(x)$ be the statement "x has a cat," and let $D(x)$ be the statement "x has a dog," where the domain for x is all MAT 253 students. Express each of these statements in terms of $B(x)$, $C(x)$, $D(x)$, quantifiers, and logical operators.

(a) There is a student in MAT 253 that has a bird, a cat, and a dog.

(b) Every student in MAT 253 has a dog.

(c) No student in MAT 253 has a bird.

(d) Every student in MAT 253 that has a dog also has a bird.

(e) Some student in MAT 253 has a bird, a cat, or a dog.

6. Determine the truth value of each of these statements if the domain of all variables consists of all integers.

(a) $\forall n\,(n^2 \geq 0)$

(b) $\forall n\,(n^2 > 0)$

(c) $\exists n\,(n^2 > 0)$

(d) $\forall n\,(n \geq 0 \vee n < 0)$

(e) $\exists n\,(n \geq 0 \wedge n < 0)$

7. Translate each of these statements into logical expressions using predicates, quantifiers, and logical operators.

(a) Something is not right.

(b) Everything is fine.

(c) Every bird can fly.

(d) Some old dogs can learn new tricks.

(e) No one is perfect.

8. Prove each of these universally quantified statements false by providing a counterexample, where the domain for all the variables consists of all real numbers.

(a) $\forall x\,(x^2 > 0)$

(b) $\forall x\,(x^2 \geq x)$

(c) $\forall x\,((x+3)^2 = x^2 + 3^2)$

(d) $\forall x\,(\frac{x}{2}$ is rational)

(e) $\forall x\,(\sqrt{x^2} = x)$

9. Prove each of these existential statements true by providing a witness, where the domain for all the variables consists of all real numbers.

(a) $\exists x\,(3x$ is irrational)

(b) $\exists x\,(x^2 \leq 0)$

(c) $\exists x\,(1 < x < 2)$

(d) $\exists x\,((x+3)^2 = x^2 + 3^2)$

(e) $\exists x\,(\pi x^2$ is an integer)

10. Let $D(x)$ be "x is an odd duck", and let $M(x)$ be "x is a mathematician", with domain consisting of all people.

Express the statement "Every mathematician is an odd duck" in terms of $D(x)$, $M(x)$, quantifiers, and logical operators.

11. Let $Q(x)$ be the statement "$2x < 0.001$" with domain the set of all real numbers. What is the truth value of $\exists x\, Q(x)$? Explain.

12. Express each of these statements using quantifiers. Use De Morgan's laws for quantifiers to find the negation of each of these statements. Form the negation of the statement so that no negation is to the left of the quantifier. Finally, express the negation in simple English. (Do not simply use the phrase "It is not the case that \dots")

(a) All dogs are nice.

(b) Some students do not study every day.

(c) All birds can fly.

(d) Some dogs have fleas.

(e) Some integers are rational numbers.

(f) All real numbers are integers.

13. Consider the propositional functions $M(x)$ = "x is a mathematician", $S(x)$ = "x is silly", and $C(x)$ = "x drinks coffee", with domain the set of all people.

Express each of the following sentences in terms of $M(x)$, $S(x)$, $C(x)$, quantifiers, and logical operators.

(a) No mathematician is silly.

(b) All mathematicians drink coffee.

(c) Mathematicians that do not drink coffee are silly.

(d) Some people that drink coffee are not silly.

(e) Not all silly people are mathematicians.

14. Alice was overheard saying, "I will go out with Bob when pigs fly." Rewrite this in the form "if p , then q." Assuming that Alice speaks the truth, use propositional logic to explain what, if any, implications this has for Alice and Bob.

15. Suppose we have the following rules for Alice.

Taylor's rule	Alice eats her veggies, or she can't have dessert.
Leslie's rule	If Alice eats her veggies, then she can have dessert.
Alex's rule	Alice can have her dessert, when she eats her veggies.
Cameron's rule	Alice can have her dessert, only if she eats her veggies.
Alice's rule	Alice can have her dessert.

If Alice's parents agree on a rule, what are their names ? (They do not have the same name. There may be more than one correct answer.)

1.4 Introduction to proofs

Outcomes

1. Prove basic results with a direct proof, a proof by contraposition, a proof by contradiction, or an existence proof.
2. Distinguish between correct and incorrect arguments

1.4.1 Even and odd integers

We give definitions of certain types of integers so that we have some objects to play with.

Definition 1.4.1.1 An integer n is **even** if there exists an integer k such that $n = 2k$. ◊

Proof Technique 1.4.1.2 Show n is even. *Suppose n is an integer, and we want to prove n is even.*

1. *Find k such that $n = 2k$.*
2. *Show k is an integer.*
3. *Conclude n is even, by definition.*

Example 1.4.1.3 Twelve is even. The integer 12 is even because $12 = 2 \cdot 6$, and 6 is an integer. □

Example 1.4.1.4 Zero is even. The integer 0 is even because $0 = 2 \cdot 0$, and 0 is an integer. □

Definition 1.4.1.5 An integer n is **odd** if there exists an integer k such that $n = 2k + 1$. ◊

Proof Technique 1.4.1.6 Show n is odd. *Suppose n is an integer, and we want to prove n is odd.*

1. *Find k such that $n = 2k + 1$.*
2. *Show k is an integer.*
3. *Conclude n is odd, by definition.*

Example 1.4.1.7 Twenty three is odd. The integer 23 is odd because $23 = 2 \cdot 11 + 1$, and 11 is an integer. □

Example 1.4.1.8 Negative twenty seven is odd. The integer -27 is odd because $-27 = 2 \cdot (-14) + 1$, and -14 is an integer. □

In Theorem 5.1.1.15 we show that every integer is even or odd and that no integer is both. Thus we have:

Remark 1.4.1.9 Let a be an integer then

1. ¬"a is odd"="a is even"
2. ¬"a is even"="a is odd"

1.4.2 Direct proof

Recall that the conditional $p \to q$ has the following truth table.

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

It follows that if we want to prove that $p \to q$ is true, we need to show that the second row ($p = \text{T}$, $q = \text{F}$) does not occur.

In the approach called the **direct proof** we start by assuming p is true and then show that q is true.

Proof Technique 1.4.2.1 To show $p \to q$ with direct proof. *Suppose p and q are propositions, and we want to prove the conditional "if p then q" by direct proof.*

1. *Assume p is true.*
2. *Deduce q is true under the assumption.*
3. *Conclude $p \to q$.*

Example 1.4.2.2 The sum of two odd integers is even. We prove that the sum of two odd integers is even.

Answer. Let a and b be odd integers.
Since a is odd, there exists an integer k such that $a = 2k + 1$.
Since b is odd, there exists an integer ℓ such that $b = 2\ell + 1$.
Compute $a + b$ and simplify.

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

Since k and ℓ are integers, we have $k + \ell + 1$ is an integer, and so $a + b$ is even.

Solution. We first develop a proof and the write down the actual proof as the answer to this problem.

1. First, we rewrite what we want to show in the form "if p , then q." Additionally, we will give friendly names to objects.[1] Let's name our odd integers a and b. Then the statement we wish to prove is:

   "If a and b are odd integers, then $a + b$ is even."

2. When we want to prove a statement of the form "if p , then q" directly, we assume p is true and try to show q. This is commonly where we set some notation as well. Let a and b be odd integers.

3. Next we need to recall what an *odd* integer is. An integer n is *odd* if there exists an integer k such that $n = 2k + 1$.

4. Now apply the definition to our situation. Since a is odd, there exists an integer k such that $a = 2k + 1$.

5. Since b is odd, there exists an integer ℓ such that $b = 2\ell + 1$.

6. Check back above to be sure that the two integers whose existence is guaranteed have different names. They need different names because they need not be the same integer.

7. Now look back to the goal that we set in 1. Since we want to say something about $a + b$, it makes sense to compute $a + b$ and simplify.

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

8. The line above should prove $a + b$ is even, provided the bit in the parentheses is an integer. Make some remark noting that, and we are done. Since k and ℓ are integers, we have $k + \ell + 1$ is an integer, and so $a + b$ is even.

Putting all that together yields the following. □

Example 1.4.2.3 The product of two even integers is even. Prove that the product of two even integers is even.

Answer. Let m and n be even integers.
Then there exist integers k and ℓ such that $m = 2k$ and $n = 2\ell$.
With the above we get

$$mn = 2k \cdot 2\ell = 4k\ell = 2(2k\ell).$$

Since k and ℓ are integers, $2k\ell$ is an integer. Thus mn is even. □

Example 1.4.2.4 Squares. Let m and n be real numbers. Prove that $m^2 = n^2$ if and only if $m = n$ or $m = -n$.

Answer. To prove

"$m^2 = n^2$ if and only if $m = n$ or $m = -n$."

we first prove

"If $m^2 = n^2$ then $m = n$ or $m = -n$."

and then

"If $m = n$ or $m = -n$ then $m^2 = n^2$."

Suppose m and n are integers such that $m^2 = n^2$. Then

$$0 = m^2 - n^2 = (m + n)(m - n).$$

It follows that $m + n = 0$ or $m - n = 0$. Thus $m = -n$ or $m = n$, which proves that "if $m^2 = n^2$ then $m = n$ or $m = -n$" is true.
Clearly, if $m = n$ then $m^2 = n^2$.
Also, if $m = -n$ then $m^2 = (-n)^2 = n^2$.
Thus "if $m = n$ or $m = -n$ then $m^2 = n^2$" is true. □

[1]Fear of a name only increases fear of the thing itself. --Dumbledore

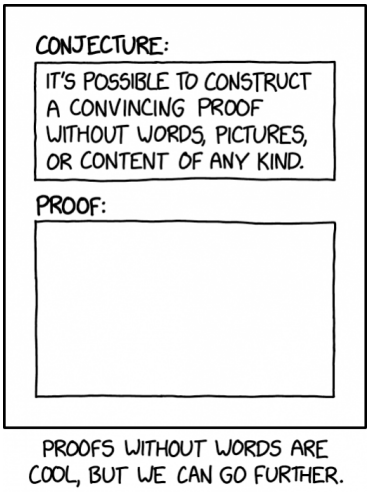


Figure 1.4.2.5 *Proof without Content* by Randal Muroe (https://xkcd.com/3201). There's also a proof without content of a conjecture without content, but it's left as an exercise for the reader

1.4.3 Proof by contraposition

In Theorem 1.2.1.10 we have shown that the contrapositive of $p \to q$ is logically equivalent to $\neg q \to \neg p$.

Thus proving $p \to q$ is true is the same as proving $\neg q \to \neg p$ is true. This is called **proof by contraposition** .

Proof Technique 1.4.3.1 To show $p \to q$ by contraposition. *Let p and q be propositions. To prove the conditional "if p then q" by contraposition, we:*

1. *State that it is sufficient to prove the contrapositive.*
2. *Assume q is false, (or equivalently $\neg q$ is true).*
3. *Deduce p is false, (or equivalently $\neg p$ is true) under the assumption.*
4. *Conclude $p \to q$.*

Example 1.4.3.2 Even square. Let n be an integer. Prove that if n^2 is even, then n is even.

Answer. To prove

"If n^2 is even, then n is even."

it is sufficient to prove the contrapositive

"If n is odd, then n^2 is odd."

Suppose n is odd.

Then there exists an integer k such that $n = 2k + 1$. With this we get

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$$

Since k is an integer, we have that $m = 2k^2 + 2k$ is an integer, and so

$$n^2 = 2m + 1$$

is odd.

We have proven "If n is odd, then n^2 is odd" is true and thus the original proposition "If n^2 is even, then n is even." also is true. □

Example 1.4.3.3 If $a \cdot b + 6$ is odd. Prove that if $a \cdot b + 6$ is odd then both a and b are odd.

Answer. We prove

"If $a \cdot b + 6$ is odd then both a and b are odd."

by contrapositive, that is we prove

"If not both a and b are odd then $a \cdot b + 6$ is even."

If not both a and b are odd we are in one of the following cases:

1. a is even and b is even, or
2. a is even and b is odd, or
3. a is odd and b is even.

We consider each of the three cases separately.

1. Suppose a is even and b is even.

   Then there are integers k and ℓ such that $a = 2k$ and $b = 2\ell$. We get

   $$a \cdot b + 6 = (2k)(2\ell) + 6 = 2(2k\ell) + 6 = 2(2k\ell + 3).$$

   Since k and ℓ are integers $(2k\ell + 3)$ is also an integer. Thus by Definition 1.4.1.1 we have that $a \cdot b + 6$ is even.

2. Suppose a is even and b is odd.

   Then there are integers k and ℓ such that $a = 2k$ and $b = 2\ell + 1$. We get

   $$a \cdot b + 6 = (2k)(2\ell + 1) + 6 = 2k(\ell \cdot 1) + 6 = 2(k(\ell + 1) + 3).$$

   Since k and ℓ are integers $(k(\ell + 1) + 3)$ is also an integer. Thus by Definition 1.4.1.1 we have that $a \cdot b + 6$ is even.

3. Suppose a is odd and b is even.

   Then there are integers k and ℓ such that $a = 2k + 1$ and $b = 2\ell$. We get

   $$a \cdot b + 6 = (2k + 1)(2\ell) + 6 = 2(\ell \cdot (2k + 1) + 3).$$

   Since k and ℓ are integers $(\ell \cdot (2k + 1) + 3)$ is also an integer. Thus by Definition 1.4.1.1 we have that $a \cdot b + 6$ is even.

Because in each of the three cases we found that $a \cdot b + 6$ is even, we have proven that "if not both a and b are odd $a \cdot b + 6$ is even", which concludes our prove. □

Example 1.4.3.4 If $x^2 + 2x + 1$ is odd then x is even by contraposition. Prove by contraposition that for an integer x we have if $x^2 + 2x + 1$ is odd then x is even.

Answer. The contraposition of the statement is

If x is odd then $x^2 + 2x + 1$ is even.

Suppose x is odd. Then by Definition 1.4.1.5 there is an integer k such that $x = 2k + 1$. We get

$$x^2+2x+1 = (2k+1)^2+2(k+1)+1 = 4k^2+4k+1+2(k+1)+1 = 4k^2+6k+4 = 2(2k^2+3k+2).$$

Because $2k^2 + 3k + 2$ is the sum of products of integers it is an integer and thus by Definition 1.4.1.1 we have that $x^2 + 2x + 1$ is even.

Because we have proven that the contrapositive of the original statement is true we can conclude that the original statement is true. □

1.4.4 Proof by contradiction

Let s and r be propositions. Recall that $r \wedge \neg r$ is a contradiction. In particular, $r \wedge \neg r$ is always false. If we can prove that $\neg s \to (r \wedge \neg r)$ is true, then $\neg s$ must be false. In this case, s must be true.

Unwinding that, we get a technique known as **proof by contradiction** . Namely, we can prove s is true by producing a direct proof of $\neg s \to (r \wedge \neg r)$. That is, to prove s is true, we start by assuming $\neg s$ is true and try to reach a contradiction.

Proof Technique 1.4.4.1 To prove s is true by contradiction. *Suppose s is a proposition, and we want to prove s by contradiction.*

1. *State that we will use proof by contradiction.*
2. *Assume s is false, (or equivalently $\neg s$ is true).*
3. *Deduce an auxiliary proposition r and its negation $\neg r$.*
4. *Announce the contradiction.*
5. *Conclude s is true.*

For the next example we need the following definition.

Definition 1.4.4.2 A number r is **rational** if there exists integers p and q with $q \neq 0$ such that $r = \frac{p}{q}$.

We say that a rational number $\frac{p}{q}$ is written in lowest terms when there exists no integer d with $d > 1$, $d \neq p$, and $d \neq q$ and no integers k and l such that $p = d \cdot k$ and $q = d \cdot l$.

A number that is not rational is called **irrational**. ◊

Example 1.4.4.3 The square root of two is irrational. Prove that $\sqrt{2}$ is irrational.

Answer. We proceed by contradiction. Suppose $\sqrt{2}$ is rational. Then there exists integers p and q , with $q \neq 0$ such that $\sqrt{2} = \frac{p}{q}$. Note that we can arrange that the fraction is in lowest terms by canceling common factors, so without loss of generality assume that p and q have no common factors.

Squaring both sides of $\sqrt{2} = \frac{p}{q}$ yields $2 = \frac{p^2}{q^2}$.

Multiplying both sides by q^2 we get $2q^2 = p^2$. Because q^2 is an integer this means that p^2 is even. It follows with Example 1.4.3.2 that p is even. So there exists an integer k such that $p = 2k$.

With $p = 2k$ we get $2q^2 = (2k)^2 = 4k^2$. Dividing by 2, we obtain $q^2 = 2k^2$ which implies that q^2 is even. Again, it follows with Example 1.4.3.2 that q is even.

From our assumption that $\sqrt{2}$ is a rational number we have deduced that p and q must be even. This contradicts our assumption that p and q have no common factor.

Thus $\sqrt{2}$ cannot be a rational number, which means that $\sqrt{2}$ is irrational. □

Example 1.4.4.4 No rational number. Use a proof by contradiction to show that there is no rational number r for which $r^3 + r + 1 = 0$.

Answer. Suppose there is a rational number $r = \frac{a}{b}$ such that $r^3 + r + 1 = 0$, where we write $\frac{a}{b}$ is lowest terms so that a and b have no common factors. Then

$$\left(\frac{a}{b}\right)^3 + \frac{a}{b} + 1 = 0.$$

Multiplying both sides by b^3 , we get

$$a^3 + ab^2 + b^3 = 0.$$

There are four cases to consider, depending on the parity (even/odd) of a and b.

a **and** b **both even**	This case contradicts that $\frac{a}{b}$ is in lowest terms, because they have the common factor 2.
a **and** b **both odd**	Then a^3, ab^2, and b^3 are all odd. The sum of three odd numbers is odd and thus cannot be equal to 0 which is even.
a **is even** b **is odd**	Then a^3 and ab^2 are even, and b^3 is odd. The sum of two even integers and an odd integer is odd thus cannot be equal to 0 which is even.
a **is odd** b **is even**	Then a^3 is odd, and ab^2 and b^3 are even. The sum of an odd integer and two even integers is odd thus cannot be equal to 0 which is even.

Since all four cases yield a contradiction, there is not rational number r such that $r^3 + r + 1 = 0$. □

We give an alternative proof of the result from Example 1.4.3.4.

Example 1.4.4.5 If $x^2 + 2x + 1$ is odd then x is even by contradiction. Prove by contradiction that if for an integer x we have $x^2 + 2x + 1$ is odd then x is even.

Answer. We assume that $x^2 + 2x + 1$ is even then x is even and show that this leads to a contradiction.

As x is even by Definition 1.4.1.1 there exists an integer k such that $x = 2k$. Then

$$x^2 + 2x + 1 = (2k)^2 + 2(2k) + 1 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$$

Because $2k^2 + 2k$ is the sum of products of integers it is an integer. Thus by Definition 1.4.1.5 we have that $x^2 + 2x + 1$ is odd. This contradicts our assumption that $x^2 + 2x + 1$ is even. So $x^2 + 2x + 1$ must be odd which concludes our proof. □

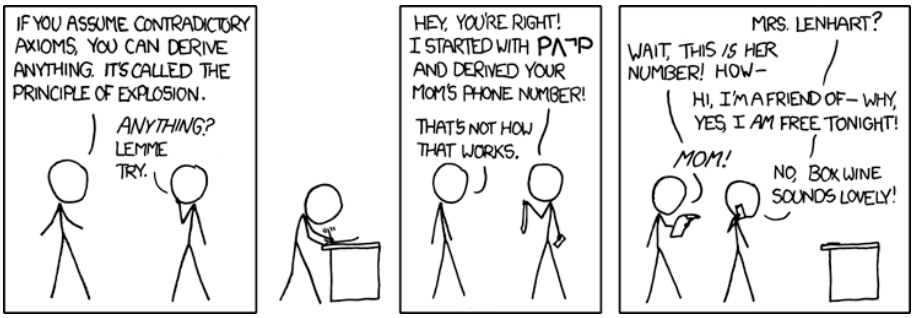


Figure 1.4.4.6 *Principle of Explosion* by Randal Munroe (https://xkcd.com/704/). You want me to pick up waffle cones? Oh, right, for the wine. One sec, let me just derive your son's credit card number and I'll be on my way.

1.4.5 Existence proofs

Many theorems assert the existence of an object with certain properties. They are of the form $\exists x\, P(x)$. One way to prove such a statement is to find a **witness** a such that $P(a)$ is true. This is called an **constructive existence proof** . . See Proof Technique 1.3.3.6 for details.

Example 1.4.5.1 Sum of cubes. Prove there is an integer that can be written as the sum of cubes of positive integers in two different ways.

Answer. There is an integer that can be written as the sum of cubes of positive integers in two different ways, for example,

$$1729 = 10^3 + 9^3 = 12^3 + 1^3.$$

Solution. It suffices to provide a witness. We can find the a witness with Python script Listing 1.4.5.2, for example.

We look for an integer that can be expressed as the sum of cubes of two distinct positive integers in two different ways. □

Listing 1.4.5.2 Python script to find witness for Example 1.4.5.1.

```
1  def witness_sum_of_cubes(bound):^^I
2    ijs = {} # dictionary to store possible witnesses
3    for i in range(bound):
4      for j in range(i+1,bound):
5        s = i**3 + j**3
6        if s in ijs.keys():
7          print("found␣witness")
8          print(s,"=",i,"**3␣+",j,"**3␣␣and")
9          print(s,"=",ijs[s][0],"**3␣+",ijs[s][1],"**3")
10         return True
11       else:
12         ijs[s] = [i,j]
13   print("no␣witness␣found")
14   return False
15
16 witness_sum_of_cubes(13)
```

There are some proofs of existence that do not produce a witness. These are called **non-constructive proofs**.

Example 1.4.5.3 Irrational power. Prove there exist irrational numbers a and b such that a^b is rational.

Solution. In our proof we distinguish two cases and use that we have already

proven that $\sqrt{2}$ is irrational in Example 1.4.4.3.

Case I: $\sqrt{2}^{\sqrt{2}}$ is rational If $\sqrt{2}^{\sqrt{2}}$ is rational, then $a = \sqrt{2}$, $b = \sqrt{2}$ is our witness, and we are done.

Case II: $\sqrt{2}^{\sqrt{2}}$ is irrational If $\sqrt{2}^{\sqrt{2}}$ is irrational then

$$(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^2 = 2,$$

which is rational, so $a = \sqrt{2}^{\sqrt{2}}$, $b = \sqrt{2}$ is our witness.

□

Note that the proof of Example 1.4.5.3 does not tell us which case occurs, so it is non-constructive.

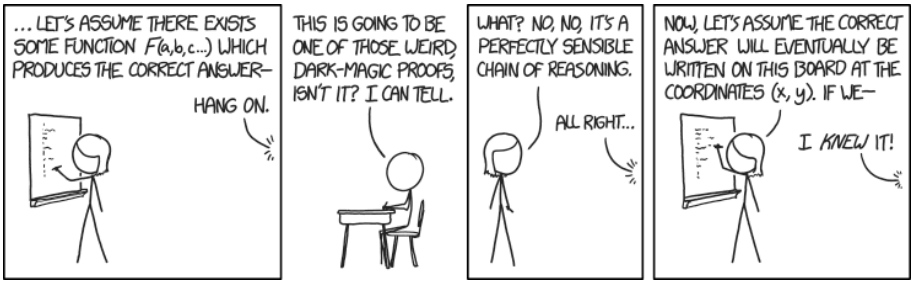


Figure 1.4.5.4 *Proofs* by Randal Munroe (https://xkcd.com/1724/). Next, let's assume the decision of whether to take the Axiom of Choice is made by a deterministic process ...

1.4.6 Summary guidelines for proofs

1. Use complete sentences.
2. Each sentence should set notation or be a true statement.
3. Each true statement should be a conclusion that can be drawn from the previous statements using a definition, computation, or result proved in class.
4. Do not assert the truth of statement to be proven at the beginning of a proof. Preface such statements with "We wish to prove" or something similar.
5. Oftentimes, a good first step is just unwinding the definitions.
6. To prove "if p , then q" directly, start the proof by assuming p is true. Then deduce that q must be true. See Proof Technique 1.4.2.1.
7. To prove "if p , then q" by contraposition, start the proof by assuming q is false. Then deduce that p must be false. See Proof Technique 1.4.3.1.
8. To prove p by contradiction, start the proof by assuming p is false. Then deduce a contradiction. See Proof Technique 1.4.4.1.

Here are some examples of what is meant by Item 2 above.

- "ax" is not a sentence.
- "$ax = b$ has a solution." is a sentence, but it is not true or false. We need to know more about a and b.

- "Let $a \in \mathbb{R}$, $a \neq 0$. Then $ax = b$ has a solution." is a bit better. The first sentence sets notation, but the second sentence is still neither true nor false since we have not specified the universe for b .
- "Let $a \in \mathbb{R}$, $a \neq 0$. Then $ax = b$ has a solution for every $b \in \mathbb{R}$." The first sentence sets notation. All of the notation is defined. The second sentence is true.

1.4.7 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) even integer

(b) odd integer

(c) rational number

2. Use a direct proof to show that the sum of two odd integers is even.

3. Use a direct proof to show that the sum of two even integers is even.

4. Use a direct proof to show that the product of two odd integers is odd.

5. Use a proof by contradiction to prove that the sum of an irrational number and a rational number is irrational.

6. Use a direct proof to show the product of two rational numbers is a rational number.

7. Prove each of these statements is false by providing a counterexample.

(a) The product of two irrational numbers is irrational.

(b) The sum of two irrational numbers is irrational.

(c) Every positive integer can be written as the sum of the squares of three integers.

(d) If n is an integer, then $n^2 > n$.

(e) If a and b are rational numbers, then a^b is also rational.

8. Let $P(n)$ be the proposition "$n^2 \geq n$." Prove $P(1)$ is true.

9. Use a proof by contradiction to prove that if n is an integer and $n^3 + 5$ is odd, then n is even.

10. Prove there is a right triangle with all three sides having rational lengths.

11. Prove there exists a pair of consecutive integers such that one of these integers is a perfect square and the other is a perfect cube.

12. Use a proof by contraposition to prove that if n is an integer and $n^3 + 7$ is even, then n is odd.

13. Go to http://link.springer.com/book/10.1007%2F978-1-4419-7023-7 while on campus and download *The Art of Proof* by Matthias Beck and Ross Geoghegan. Read carefully Chapters 1--7.

Chapter 2

Basic Structures

We look at some fundamental structures in mathematics: sets, functions, and sequences.

2.1 Sets

Outcomes

- Reproduce the basic terminology of set theory.
- Describe sets in roster form and set builder notation.
- Determine the cardinality of a set, the powers set of a set, and Cartesian products
- Compute the power set of a set.
- Construct Cartesian products.
- Prove equality and subset relations of sets.

2.1.1 Basic terminology

Definition 2.1.1.1 A **set** is an unordered collection of objects, called **elements** or **members** of the set.

We write $b \in A$ which we read "b is an element of A" to denote that b is an element of the set A. ◊

Definition 2.1.1.2 The **universal set** is the set that contains all objects that are relevant for a specific discussion. ◊

In our discussions of sets and set operations the domain of all propositional functions and thus for all quantification will be the universal set, unless specified explicitly.

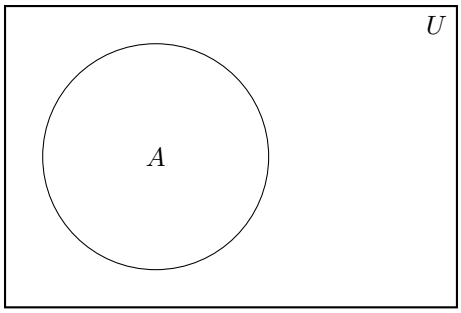


Figure 2.1.1.3 The box in the Venn diagram represents the universal set U that contains all objects under consideration. Other sets, like the set A pictured, are represented by circles or ovals inside this box.

Definition 2.1.1.4 Two sets A and B are **equal** , denoted $A = B$, if they have the same elements, that is, $\forall x(x \in A \leftrightarrow x \in B)$. ♢

By the convention stated above $\forall x(x \in A \leftrightarrow x \in B)$ is read

> "For all elements x of the universal set we have that x is an element of the set A if and only if x is an element of the set B."

There are several ways to describe a set.

We specify a set in **roster form** by listing the elements explicitly between braces, for example, $\{1, 2, 7\}$. To indicate that the list of continues in a pattern that can be deduced from the preceding list of elements we use ellipsis $(\ldots)$, for example

$$\{1, 2, 3, \ldots, 10\} = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}\,.$$

When we specify a set in **set builder notation** we characterize the elements of the set by stating properties they must have to be members, for example,

$$\{x : x \text{ is an odd positive integer less than } 10\} = \{1, 3, 5, 7, 9\}\,.$$

Example 2.1.1.5 Set or element. Is $\{3\}$ the same thing as 3 ?

Answer. No. Sets are like bags. A bag with an apple in it is different from an apple. The set $\{3\}$ is a set with one element 3 in it. The number 3 is not the same as the set containing just 3. □

We formulate the sets of numbers that we often use in set builder notation.

Recall that set of **integers** is $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

The set of **natural numbers** or **positive integers** is $\mathbb{N} = \{x : x \in \mathbb{Z} \wedge x > 0\} = \{1, 2, 3, \ldots\}$.

The set of **non-negative integers** is $\mathbb{N}_0 = \{x : x \in \mathbb{Z} \wedge x \geq 0\} = \{0, 1, 2, \ldots\}$.

The set of **rational numbers** is $\mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z} \wedge q \in \mathbb{Z} \wedge q \neq 0 \right\}$.

We denote the set of **real numbers** by $\mathbb{R}$. , For our purposes, the intuitive idea that they are all numbers that can be written in finite or infinite decimal representation is sufficient.

Example 2.1.1.6 Intervals. We have special notation for intervals of real numbers.

$$\begin{aligned}
[a, b] &= \{x : x \in \mathbb{R}, a \leq x \leq b\} \\
(a, b] &= \{x : x \in \mathbb{R}, a < x \leq b\} \\
[a, b) &= \{x : x \in \mathbb{R}, a \leq x < b\} \\
(a, b) &= \{x : x \in \mathbb{R}, a < x < b\}
\end{aligned}$$

□

Example 2.1.1.7 Equality. Let $A = \{1, 2, 2, 5\}$, and let $B = \{1, 2, 5\}$. Why is $A = B$?

Solution. We have $A = B$, because for every element x, the truth value of $x \in A$ is the same as the truth value of $x \in B$. Thus $A = B$ by definition. □

To show two sets are equal, we use subsets as described later in Proof Technique Proof Technique 2.1.2.7. It is more straightforward to show two sets are not equal. Since set equality is defined as the universal quantification ("for all") of a biconditional ("if and only if"), to show two sets are not equal, it is enough to find any counterexample to the biconditional. In particular, we just need to produce any element that they do not share.

Proof Technique 2.1.1.8 To show $A \neq B$. *Suppose A and B are sets, and we want to prove A is not equal to B.*

1. *Find a particular element a in A that is not a member of B; or find a particular element b in B that is not a member of A.*
2. *Conclude $A \neq B$.*

Definition 2.1.1.9 The **empty set**, denoted $\emptyset$, is the set with no elements. ◊

Example 2.1.1.10 Two formulations for the empty set. The empty set is $\{\}$.

That is, $\emptyset = \{\}$ □

2.1.2 Subsets

Definition 2.1.2.1 A set A is a **subset** of a set B, denoted $A \subseteq B$, if every element of A is also an element of B. that is , $\forall x(x \in A \rightarrow x \in B)$.

If $A \subseteq B$, and $A \neq B$, we say A is a **proper subset** of B, denoted $A \subset B$. ◊

Example 2.1.2.2 Subset and proper subset. Let $A = \{a, b, c\}$, and let $B = \{a, b, c, d, e, f\}$. Then A is a subset of B. Why? For every x, whenever x is in A we have x is in B. Thus $A \subseteq B$, by definition.

Furthermore, we have A is a proper subset of B, since we additionally have that A is not equal to B. Thus $A \subset B$, by definition. □

Remark 2.1.2.3 This notation is reminiscent of the notation for inequality and strict inequality.

Note that subsets are defined in terms of a universal quantification ("for all"). To prove universal statements, we use generic elements.

Proof Technique 2.1.2.4 To show $A \subseteq B$. *Suppose A and B are sets, and we want to prove A is a subset of B.*

1. *Let x be a generic element of A.*
2. *Show that x is in B.*
3. *Conclude $A \subseteq B$.*

Theorem 2.1.2.5 *For any set A,*

1. *$\emptyset \subseteq A$, and*
2. *$A \subseteq A$.*

Proof. We apply Proof Technique Proof Technique 2.1.2.4 to show $\emptyset \subseteq A$. Since $\emptyset$ has no elements, there is nothing to check. In particular, we need to check that for each a in $\emptyset$, the statement $a \in A$ is true. Since there are no elements in $\emptyset$, there is nothing to check.

Next, we apply Proof Technique Proof Technique 2.1.2.4 to show $A \subseteq A$. Let x be a generic element in A. Then x is in A by construction. Thus $A \subseteq A$. ■

It seems like we just ran around in circles in the proof above. Read it over again, paying attention to what it is we need to show.

For a set A to be a proper subset of set B, we need A to be a subset of B that is not equal to B.

Proof Technique 2.1.2.6 To show $A \subset B$. *Suppose A and B are sets, and we want to prove A is a proper subset of B.*

1. *Use Proof Technique Proof Technique 2.1.1.8 to show $A \neq B$.*
2. *Use Proof Technique Proof Technique 2.1.2.4 to show*
3. *Conclude $A \subset B$.*

By definition, two sets are equal if they have the same elements. This implies that each one is a subset of the other. We can turn this into a proof technique for showing two sets are equal.

Proof Technique 2.1.2.7 To show $A = B$. *Suppose A and B are sets, and we want to show A is equal to B.*

1. *Use Proof Technique Proof Technique 2.1.2.4 to show $A \subseteq B$.*
2. *Use Proof Technique Proof Technique 2.1.2.4 to show $B \subseteq A$.*
3. *Conclude $A = B$.*

2.1.3 Size of a set

We give an intuitive definition of the number of elements in a set.

Definition 2.1.3.1 Let A be a set. If there are exactly n distinct elements in A, where n is a non-negative integer, we say A is a **finite set** and that n is the **cardinality** of A, denoted $\#A$. ♢

We give a precise definition of cardinality in Definition 4.1.1.2. In Section 4.1 we also discuss sets that are not finite and their cardinality The notion if infinity is subtle; there are different sizes of infinity.

Example 2.1.3.2 The cardinality of the empty set is zero. The empty set $\emptyset$ has size 0

$$\#\emptyset = 0$$

since $\{\}$ has no elements. □

Example 2.1.3.3 Cardinality of a set in roster form. Let E be the even integers between -5 and 5, not including -5 and 5. Then we can compute the size of E by writing E using the roster method and counting.

$$E = \{-4, -2, 0, 2, 4\},$$

so $\#E = 5$. □

Example 2.1.3.4 Number of consonants. Let C be the set of consonants in the English alphabet. Instead of listing the elements of C in roster method,

we note that there are 26 letters in the English alphabet. Of these, 5 are vowels and the rest are consonants. Thus

$$\#C = 26 - 5 = 21.$$

□

2.1.4 Power sets

Definition 2.1.4.1 Let A be a set. The **power set** of A, denoted $\wp(A)$, is the set of all subsets of A. ◊

Example 2.1.4.2 Powerset of a set with cardinality three. Suppose $A = \{2, x, \pi\}$. The power set of A is the set of all subsets of A. Let's work out the subsets systematically, ordered by size.

Size 0	There is only 1 subset of size 0. It is the empty set: $\emptyset$.
Size 1	These singleton sets each contain one element of A. Since there are 3 elements in A, there are 3 subsets of size 1: $\{2\}$, $\{x\}$, and $\{\pi\}$.
Size 2	Subsets of A of size 2 can be constructed by omitting one element. Since there are 3 elements in A, there are 3 subsets of size 2: $\{2, x\}$, $\{2, \pi\}$, and $\{x, \pi\}$.
Size 3	There is only 1 subset of size 3. It is the set A itself: $\{2, x, \pi\}$.
Size > 3	Since $\#A = 3$, there are no subsets of A that have size larger than 3.

Thus the power set of A is

$$\wp(A) = \{\emptyset, \{2\}, \{x\}, \{\pi\}, \{2, x\}, \{2, \pi\}, \{x, \pi\}, \{2, x, \pi\}\}.$$

□

Example 2.1.4.3 What adding one element does to the power set. The only subset of $\emptyset$ is $\emptyset$, thus

$\wp(\emptyset) = \{\emptyset\}$ with with $\#\wp(\emptyset) = 1$.

The subsets of $\{1\}$ are $\emptyset$ and the set $\{1\}$ itself, thus

$\wp(\{1\}) = \{\emptyset, \{1\}\}$ with $\#\wp(\{1\}) = 2$.

Because $\{1\} \subseteq \{1, 2\}$ the subsets of $\{1\}$ are also subsets of $\{1, 2\}$.

In addition $\{1, 2\}$ also has subsets that contain 2. We obtain these by adding the element 2 to the subsets of $\{1\}$ and obtain $\{2\}$ and $\{1, 2\}$. Together the subsets that do not contain 2 and the subsets that contain 2 are all the subsets of $\{1, 2\}$. We get

$\wp(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ with $\#\wp(\{1, 2\}) = 4$.

Because $\{1, 2\} \subseteq \{1, 2, 3\}$ the subsets of $\{1, 2\}$ are also subsets of $\{1, 2, 3\}$. In addition $\{1, 2, 3\}$ also has subsets that contain 3. We obtain these by adding the element 3 to $\emptyset$ and $\{1\}$ and $\{1, 2\}$. We get

$\wp(\{1, 2, 3\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ with $\#\wp(\{1, 2, 3\}) = 8$.

□

In Example 2.1.4.3 we observe that adding an element to a set, doubles the number of sets in the powerset. Since $\#\wp(\emptyset) = 1$ this yields the following theorem.

Theorem 2.1.4.4 *If a set A has cardinality n, then the cardinality of the power set of A is $\#\wp(A) = 2^n$.*

We revisit this result as Theorem 4.1.2.8 in Chapter 4 after we develop techniques on counting and give another proof.

2.1.5 Cartesian products

Definition 2.1.5.1 Let A and B be sets. The **Cartesian product** of A and B, denoted $A \times B$, is the set of ordered pairs (a, b), where $a \in A$ and $b \in B$. That is,

$$A \times B = \{(a, b) : a \in A \wedge b \in B\}.$$

◊

Example 2.1.5.2 A Cartesian product. Let $A = \{0, 1\}$, and let $B = \{4, 5, 6\}$. Then the Cartesian product $A \times B$ is the set of all ordered pairs (a, b), where a is an element from A and b is an element of B. Thus

$$A \times B = \{(0, 4), (0, 5), (0, 6), (1, 4), (1, 5), (1, 6)\}.$$

Analogously, the Cartesian product $B \times A$ is the set of all ordered pairs (b, a), where b is an element of B and a is an element of A. Thus

$$B \times A = \{(4, 0), (4, 1), (5, 0), (5, 1), (6, 0), (6, 1)\}.$$

□

Theorem 2.1.5.3 $AB\#A \times B = \#A \cdot \#B$

More generally, we can define the Cartesian product of more than two set.

Definition 2.1.5.4 The **Cartesian product** of n sets $A_1, A_2, \ldots, A_n$, denoted $A_1 \times A_2 \times \cdots \times A_n$ is the set of ordered n-tuples $(a_1, a_2, \ldots, a_n)$ where $a_i \in A_i$ for $i = 1, 2, \ldots, n$. That is,

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_i \in A_i \text{ for } i = 1, 2, \ldots, n\}.$$

◊

Definition 2.1.5.5 Let A be a set. We use the notation A^2 to denote the Cartesian product $A \times A$. Similarly we write A^3 for the Cartesian product $A \times A \times A$, and so on. ◊

Example 2.1.5.6 The Cartesian plane. $\mathbb{R}^2$ is the familiar Cartesian plane. □

2.1.6 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) set

(b) equal sets

(c) empty set

(d) subset

(e) proper subset

(f) cardinality of a set

(g) power set of a set

(h) Cartesian product of sets

2. Write each set in roster notation.

(a) $\{n \in \mathbb{Z} : n \text{ is odd and } n \leq 10\}$

(b) $\{n \in \mathbb{Z} : |n| \leq 3\}$

(c) the set of positive, even integers less than 6

(d) $\left\{x \in \mathbb{R} : x^2 - x - 1 = 0\right\}$

(e) $\{x \in \mathbb{R} : x \text{ is the square of an integer and } x < 100\}$

(f) $\left\{x \in \mathbb{R} : x^2 = -1\right\}$

3. Use set builder notation to give a description of each of these sets.

(a) $\{2, 3, 4, 5, 6, 7, 8\}$

(b) $\{0, 3, 6, 9, 12\}$

(c) $\{-2, -1, 0, 1, 2\}$

(d) $\{0, 4, 9, 16, 25, 36\}$

4. For each of these pairs of sets, determine whether the first is a subset of the second, the second is a subset of the first, or neither is a subset of the other.

(a) the set of dogs; the set of mammals

(b) the set of people who speak English; the set of people who

(c) the set of math majors; the set of computer science majors

(d) the set of people over 6 feet tall; the set of people over 5

(e) the set of animals; the set of alligators

5. For each of these sets, determine whether 3 is an element of that set.

(a) $\{1, 3, 5, 7, 9\}$

(b) $\{\{3, 5\}, \{2, 4\}\}$

(c) $\{(0, 3), (3, 0), (3, 3)\}$

(d) $\{\{3\}, \{5\}\}$

(e) the set of positive, odd integers less than 20

6. Find $A \times B$, where $A = \{a, b\}$ and $B = \{0, 4, 8\}$. What is the cardinality of each of these sets?

(a) $\{1, 3, 3, 5, 5, 6, 10, 10\}$

(b) $\{\{1, 3, 5, 6\}, \{10, 12, 15\}\}$

(c) $\emptyset$

(d) $\{0, \emptyset\}$

(e) $\{\emptyset, \{\}\}$

(f) $\wp(\{0, 1\})$

7. Suppose A, B, and C are sets such that $A \subseteq B$ and $B \subseteq C$. Prove that $A \subseteq C$.

8. Let A, B, C, and D be sets. Show that if $A \subseteq C$ and $B \subseteq D$, then $A \times B \subseteq C \times D$.

9. Let $A = \{a, b, c\}$, and let $B = \{x, y\}$. Find the following sets.

(a) $A \times B$

(b) $B \times A$

(c) A^2

(d) B^3.

(e) $\wp(A)$

10. Determine whether these statements are true or false.

(a) $x \in \{x\}$

(b) $\{x\} \in \{x\}$

(c) $x \subseteq \{x\}$

(d) $\{x\} \subseteq \{x\}$

(e) $\emptyset \subseteq \{x\}$

2.2 Set operations

Outcomes

1. Evaluate expressions involving set operations
2. Prove set identities
3. Reproduce the most important such identities.

2.2.1 Basic operations

Definition 2.2.1.1 Let A and B be sets. The **union** of sets A and B, denoted $A \cup B$ is the set that contains those elements that are in A or in B or in both. That is,

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

$\Diamond$

Example 2.2.1.2 Union. Let $A = \{1, 2, 3, 5, 10\}$, and let $B = \{2, 3, 4, 5\}$. To compute the union $A \cup B$, we combine all of the elements of A together with all of the elements of B into one set. The union of A and B is

$$A \cup B = \{1, 2, 3, 4, 5, 10\}.$$

$\square$

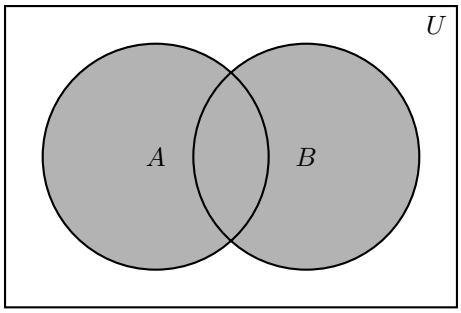


Figure 2.2.1.3 The shaded area in the Venn diagram corresponds to the union $A \cup B$ of two sets A and B.

Definition 2.2.1.4 Let A and B be sets. The **intersection** of A and B, denoted $A \cap B$ is the set that contains those elements that are in A and also in B. That is,

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

$\Diamond$

Example 2.2.1.5 Intersection. Let $A = \{1, 2, 3, 5, 10\}$, and let $B = \{2, 3, 4, 5\}$. To compute the intersection $A \cap B$, we take the elements of A that are also elements of B. In other words, we only keep elements that A and B have in common. The intersection of A and B is

$$A \cap B = \{2, 3, 5\}.$$

$\square$

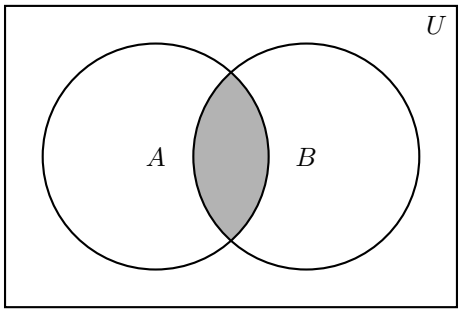


Figure 2.2.1.6 The shaded area in the Venn diagram corresponds to the intersection $A \cap B$ of two sets A and B.

Definition 2.2.1.7 The **difference** of A and B, denoted $A - B$, is the set containing elements that are in A but not in B. That is,

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

$\Diamond$

Example 2.2.1.8 Two differences. Let $A = \{1, 2, 3, 5, 10\}$, and let $B = \{2, 3, 4, 5\}$. The difference of A and B consists of members of A that are not also members of B, so we take all the elements of A and throw out the ones that are in B. That gives

$$A - B = \{1, 10\}.$$

Analogously, the difference of B and A consists of members of B that are not also members of A, so

$$B - A = \{4\}.$$

$\square$

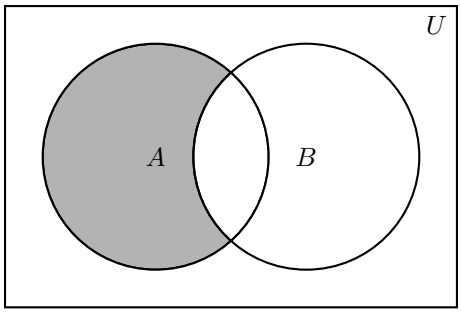


Figure 2.2.1.9 The shaded area in the Venn diagram corresponds to the set difference $A - B$ of two sets A and B.

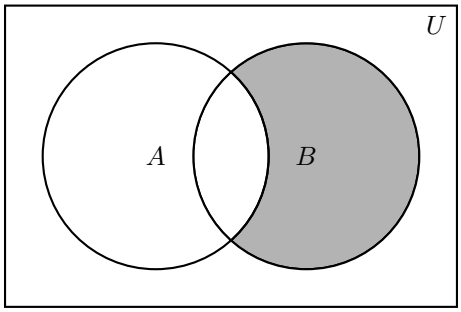


Figure 2.2.1.10 The shaded area in the Venn diagram corresponds to the set difference $B - A$ of two sets A and B.

Definition 2.2.1.11 The **symmetric difference** of A and B, denoted by $A \oplus B$, is the set containing those elements in either A or B, but not in both A and B. That is

$$A \oplus B = \{x : x \in A \oplus x \in B\}.$$

◊

Example 2.2.1.12 Symmetric difference. Let $A = \{1, 2, 3, 5, 10\}$, and let $B = \{2, 3, 4, 5\}$. To compute the symmetric difference of A and B, we combine the elements of A and B, and them remove the elements that are members of both A and B. Thus

$$A \oplus B = \{1, 4\} \text{ and } B \oplus A = \{1, 4\}.$$

□

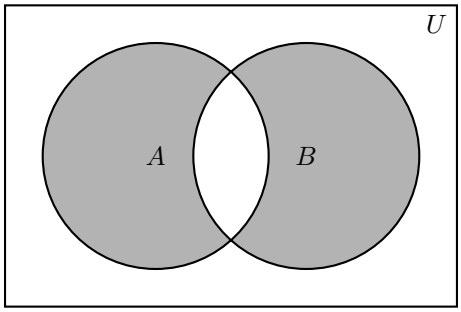


Figure 2.2.1.13 The shaded area in the Venn diagram corresponds to the symmetric difference $A \oplus B$ of the two sets A and B.

Definition 2.2.1.14 Let U be the universal set. The **complement** of A,

denoted $\overline{A}$, is the difference of U and A. That is

$$\overline{A} = \{x : x \notin A\} = U - A.$$

$\Diamond$

Example 2.2.1.15 Complement. Let $A = \{1, 2, 3, 5, 10\}$, and let $B = \{2, 3, 4, 5\}$. Suppose $U = \{1, 2, \ldots, 10\}$. Then the complement of A consists of elements in U that are not in A. We can compute it by taking all of the elements of U and throwing out the elements of A. That gives

$$\overline{A} = \{4, 6, 7, 8, 9\}.$$

Analogously, the complement of B is consists of members of U that are not in B, so

$$\overline{B} = \{1, 6, 7, 8, 9, 10\}.$$

□

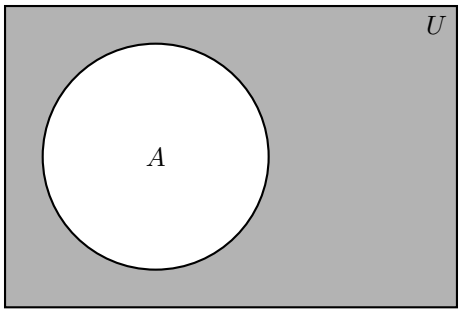


Figure 2.2.1.16 The shaded area in the Venn diagram corresponds to the complement $\overline{A}$ of the set A.

Definition 2.2.1.17 Two sets A and B are **disjoint** if $A \cap B = \emptyset$. $\Diamond$

Example 2.2.1.18 Disjoint. Let $A = \{1, 3, 5, 7\}$, and let $B = \{2, 4, 6\}$. Then A and B have no elements in common, so $A \cap B = \emptyset$. Thus A and B are disjoint. □

Example 2.2.1.19 Intersection with a subset. Prove if $A \subseteq B$, then $A \cap B = A$.

Proof. We need to show

1. $A \cup B \subseteq B$, and
2. $B \subseteq A \cup B$.

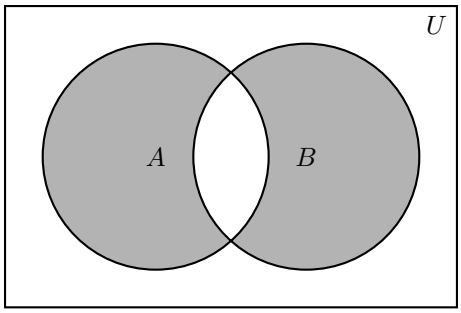
The second statement $B \subseteq A \cup B$ is clear from the the definition. Namely, for any element b in B, we must have $b \in A \cup B$ by definition of union. It remains to show $A \cup B \subseteq B$. Let x be a generic element in $A \cup B$. Then by definition, $x \in A$ or $x \in B$. Since $A \subseteq B$, in either case we have $x \in B$. Thus $A \cup B \subseteq B$.

Therefore $A \cup B = B$ as desired. ■

□

Example 2.2.1.20 Venn diagram for $(A - B) \cup (B - A)$.. Draw a Venn diagram representing $(A - B) \cup (B - A)$.

Answer.

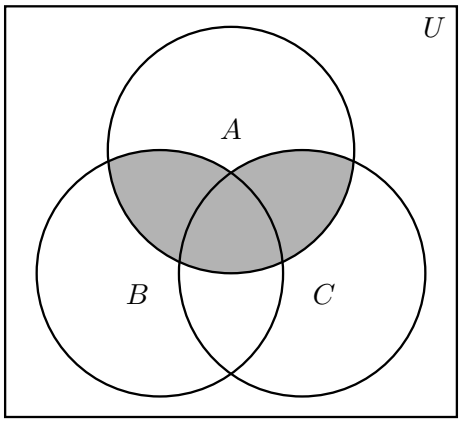


Comparing this Venn diagram with Figure 2.2.1.13 we notice that $(A-B)\cup(B-A) = A \oplus B$.

Solution. From the definition, $(A-B)\cup(B-A)$ consists of elements of A that are not in B together with elements of B that are not in A. We shade the part of A that is not in B and the part of B that is not in A. □

Example 2.2.1.21 Venn diagram for $A\cap(B\cup C)$.. Draw a Venn diagram representing $A\cap(B\cup C)$.

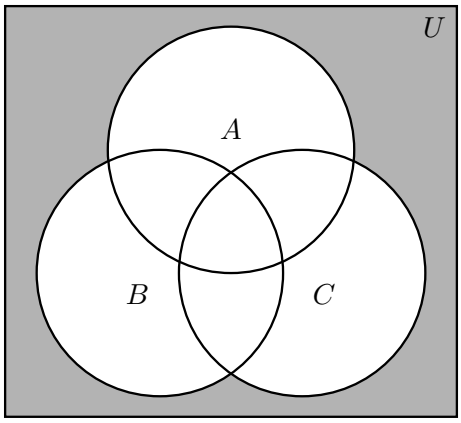
Answer.



Solution. From the definition, we have that $A\cap(B\cup C)$ consists of elements that are in A and in $B\cup C$. □

Example 2.2.1.22 Venn diagram for $\overline{A}\cap\overline{B}\cap\overline{C}$.. Draw a Venn diagram representing $\overline{A}\cap\overline{B}\cap\overline{C}$.

Answer.



Solution. From the definition, we have that $\overline{A}\cap\overline{B}\cap\overline{C}$ consists of elements

that are not in A and not in B and not in C. □

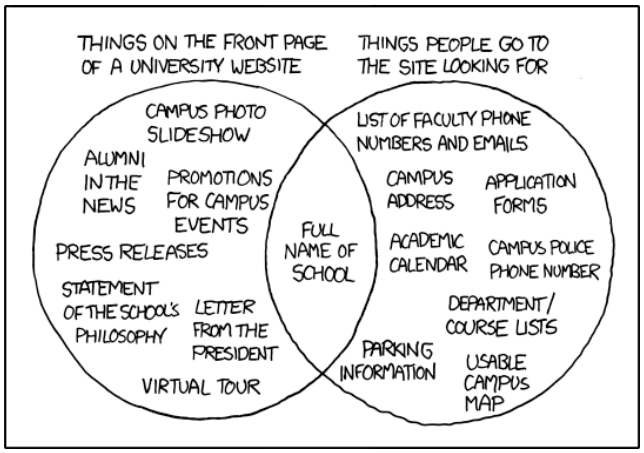


Figure 2.2.1.23 *University Website* by Randall Munroe (https://xkcd.com/773/). People go to the website because they can't wait for the next alumni magazine, right? What do you mean, you want a campus map? One of our students made one as a CS class project back in '01! You can click to zoom and everything.

2.2.2 Set identities

Theorem 2.2.2.1 *Let A and B be sets. Then*

$$A - B = A \cap \overline{B}.$$

Proof with generic elements. One approach is to show $A - B \subseteq A \cap \overline{B}$ and $A \cap \overline{B} \subseteq A - B$. We can do this with generic elements.

Let $x \in A - B$. Then by definition, $x \in A$ and $x \notin B$. Then $x \in \overline{B}$. Since $x \in A$ and $x \in \overline{B}$, we have $x \in A \cap \overline{B}$. Therefore $A - B \subseteq A \cap \overline{B}$.

Let $x \in A \cap \overline{B}$. Then by definition $x \in A$ and $x \in \overline{B}$. Then $x \notin B$. Since $x \in A$ and $x \notin B$, we have $x \in A - B$. Thus $A \cap \overline{B} \subseteq A - B$.

Therefore $A - B = A \cap \overline{B}$. ■

Proof with set builder notation and logical equivalences. Another approach is to use set builder notation and logical equivalences.

$$\begin{aligned} A - B &= \{x : x \in A \wedge x \notin B\} \\ &= \left\{x : x \in A \wedge x \in \overline{B}\right\} \\ &= \left\{x : x \in A \cap \overline{B}\right\} \\ &= A \cap \overline{B}. \end{aligned}$$

■

Proof with a membership table. Another approach is to compute a **membership table** , a table displaying the membership of elements in sets. The concept is similar to using truth tables to display the truth value of propositions. Use 1 to indicate membership in the set and 0 to denote the element is not in the set.

Below, we use a membership table to prove Theorem 2.2.2.1.

$x \in A$	$x \in B$	$x \in \overline{B}$	$x \in A \cap \overline{B}$	$x \in A - B$
1	1	0	0	0
1	0	1	1	1
0	1	0	0	0
0	0	1	0	0

Since the last two columns are the same, we have $A \cap \overline{B} = A - B$. ■

Remark 2.2.2.2 When using a membership table to prove a set identity, remember to specify which columns give the desired result in the conclusion.

Theorem 2.2.2.3 Set identities I. *Let A be a set with universal set U.*

Identity laws	$A \cap U = A$ *and* $A \cup \emptyset = A$
Domination laws	$A \cup U = U$ *and* $A \cap \emptyset = \emptyset$
Idempotent laws	$A \cup A = A$ *and* $A \cap A = A$
Complementation law	$\overline{\overline{A}} = A$
Complement laws	$A \cup \overline{A} = U$, $A \cap \overline{A} = \emptyset$

Proof. Exercise. Either compute membership tables and compare columns or use generic elements to show each set is a subset of the other. ■

Theorem 2.2.2.4 Set identities II. *Let A and B be sets.*

Commutative laws	$A \cup B = B \cup A$ *and* $A \cap B = B \cap A$.
Absorption laws	$A \cup (A \cap B) = A$ *and* $A \cap (A \cup B) = A$.

Proof. Exercise. Either compute membership tables and compare columns or use generic elements to show each set is a subset of the other. ■

Theorem 2.2.2.5 Set identities III. *Let A, B, and C be sets.*

Associative laws	$A \cup (B \cup C) = (A \cup B) \cup C$ *and* $A \cap (B \cap C) = (A \cap B) \cap C$.
distributive laws	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *and* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

We prove the associative law for intersection in several ways and leave the rest as an exercise. Let A and B be sets. We want to show

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

Proof with generic elements. We need to show

1. $A \cap (B \cap C) \subseteq (A \cap B) \cap C$, and
2. $(A \cap B) \cap C \subseteq A \cap (B \cap C)$.

Let $x \in A \cap (B \cap C)$. Then x must be in A and also in $B \cap C$. Hence x must be in A and also in B and in C. Since x is in both A and B, we conclude that

$x \in A \cap B$. This, together with the fact that $x \in C$ tells us that $x \in (A \cap B) \cap C$, as desired.

Now we prove the other direction. (The argument is nearly identical.) Let $x \in (A \cap B) \cap C$. Then $x \in A \cap B$ and also in C. Hence x is in A and B and C. Since x is in B and in C, we have that $x \in B \cap C$. This, together with the fact that $x \in A$ gives $x \in A \cap (B \cap C)$, as desired. Therefore $A \cap (B \cap C) = (A \cap B) \cap C$. ■

Proof with a membership table. An alternate approach is to use a membership table. Use 1 to indicate membership in the set and 0 to denote the element is not in the set. Remember to specify which columns give the desired result in the conclusion.

We compute the membership table.

$x \in A$	$x \in B$	$x \in C$	$x \in A \cap B$	$x \in (A \cap B) \cap C$	$x \in B \cap C$	$x \in A \cap (B \cap C)$
1	1	1	1	1	1	1
1	1	0	1	0	0	0
1	0	1	0	0	0	0
1	0	0	0	0	0	0
0	1	1	0	0	1	0
0	1	0	0	0	0	0
0	0	1	0	0	0	0
0	0	0	0	0	0	0

Since the 5th and 7th columns are identical, we have

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

■

Theorem 2.2.2.6 De Morgan's laws for sets. *Let A and B be sets.*

1. $\overline{A \cup B} = \overline{A} \cap \overline{B}$
2. $\overline{A \cap B} = \overline{A} \cup \overline{B}$

Proof. Exercise. Either compute membership tables and compare columns or use generic elements to show each set is a subset of the other. ■

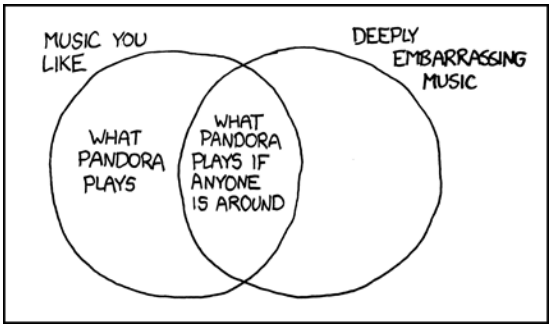


Figure 2.2.2.7 *Pandora* by Randall Munroe (https://xkcd.com/668/). What? Oh, no, the 'Enchanted' soundtrack was just playing because Pandora's algorithms are terrible. [silence] … (quietly) That's how you knooooooow …

2.2.3 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) union of sets

(b) intersection of sets

(c) difference of sets

(d) complement of a set

(e) disjoint sets

2. State precisely De Morgan's laws for sets. Be sure to set up any notation that is required.

3. Let $C = \{r, s, t, l, n, e\}$, and let $W = \{f, o, r, t, u, n, e\}$. Compute each of these sets.

(a) $C \cap W$

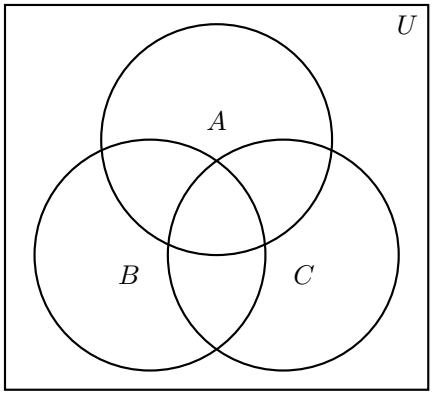
(b) $C \cup W$

(c) $C \oplus W$

(d) $C - W$

(e) $W - C$

4. Let A, B, C be sets. Shade the portion of the Venn diagram corresponding to each of these sets.



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

(b) $A \cap (B \cap C)$

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

(d) $(\overline{A} \cap \overline{B}) \cup \overline{C}$

(e) $(A - B) - C$

(f) $(A - B) \cup (A - C) \cup (B - C)$

5. Let $A = \{1, 3, 5, 7, 9\}$, and let $B = \{1, 2, 3, 4, 5\}$. Compute the following sets.

(a) $A \cap B$

(b) $A - B$

(c) $A \cup B$

(d) $B - A$

(e) $B \oplus A$

6. Let A and B be sets. Prove that if $A \subseteq B$, then $A \cup B = B$.

7. Prove the set identities in Theorem 2.2.2.3.

8. Prove the set identities in Theorem 2.2.2.4.

9. Prove the set identities in Theorem 2.2.2.5.

10. Prove De Morgan's laws for sets.

11. Let A and B be sets. Prove $A \oplus B = (A - B) \cup (B - A)$.

12. Let A and B be sets.

(a) Prove if $A \subseteq B$ then $\overline{B} \subseteq \overline{A}$.

(b) Prove if $\overline{B} \subseteq \overline{A}$ then $A \subseteq B$.

2.3 Functions

Outcomes

1. Reproduce the definition of a function
2. Proof that a function is injective
3. Proof that a function is surjective
4. Proof that a function is bijective
5. Evaluate the floor and ceiling functions.

2.3.1 Basic terminology

We have seen examples of functions in previous math classes. Most were likely given by formulas, such as $f(x) = x^2 + 2$. Functions can given in other ways and are not always given by formulas. The key property of a function is that it accepts inputs and provides a corresponding output value for each possible input.

Definition 2.3.1.1 Let A and B be nonempty sets. A **function** or **mapping** or **transformation** from A to B, denoted $f\colon A \to B$ is an assignment of exactly one element of B to each element of A.

The set A is called the **domain** and B is called the **codomain** of f. ◊

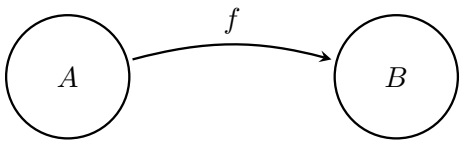


Figure 2.3.1.2 Function from A to B.

Definition 2.3.1.3 Let $f\colon A \to B$ be a function, and let a be an element in A. The **image** of a is the unique element of B that is assigned by the function f to the element a.

In this case, write $f(a) = b$, where b is the image of a. We say a is a **preimage** of b. ◊

By Definition 2.3.1.1 each element of the domain of a function can only have one image. But an element of the codomain of a function can have several preimages.

Example 2.3.1.4 More than one preimage. Under the function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3\}$ in Figure 2.3.5.7 the element 2 of the codomain $\{1, 2, 3\}$ has the preimages 2 and 4. □

A function must output a value for *every* input from the domain. Not every element in the codomain needs to be an image, though.

Example 2.3.1.5 Square and square root functions. Consider $f(x) = x^2 + 3$. Then f defines a function from $\mathbb{R}$ to $\mathbb{R}$. We write $f\colon \mathbb{R} \to \mathbb{R}$. The domain is $\mathbb{R}$. The codomain is $\mathbb{R}$.

On the other hand, the square root 0.1.6.3 function $g(x) = \sqrt{x}$ is not a function from $\mathbb{R}$ to $\mathbb{R}$. This is because, for example, -2 is an element of $\mathbb{R}$, but the formula does not define a value for $g(-2)$. We can view g as a function from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}$. □

Functions do not have to be given by formulas, They do not have to be from sets of numbers to sets of numbers.

Example 2.3.1.6 Temperature function. Each point on the surface of the earth has a particular temperature right now, and the temperature (in degrees Celsius) is a real number. Thus, temperature defines a function T from the surface of the earth to $\mathbb{R}$, where $T(x)$ is the temperature at the point x. □

Definition 2.3.1.7 Let $f\colon A \to B$ be a function. The **image** of f, denoted $\mathrm{im}(f)$ or $f(A)$, is the set of all images of elements of A,

$$\mathrm{im}(f) = \{f(a) : a \in A\}.$$

More generally, if S is a subset of A, the **image** of S under f, denoted $f(S)$ is the set

$$f(S) = \{f(s) : s \in S\}.$$

◊

The image of a function f is also called the range of f.

Functions are often given by a table of values.

Example 2.3.1.8 Gem stones. Consider the prices of some expensive gem stones.

Gem stone	price per carat
Blue Diamond	$3930000
Pink Star Diamond	$1200000
Musgravite	$35000
Jadeite	$20000
Alexandrite	$12000
Red Beryl	$10000
Padparadscha Sapphire	$8000

This table describes a function p which gives the price per carat of these gem stones. The domain of p is { Blue Diamond, Pink Star Diamond, Musgravite, Jadeite, Alexandrite, Red Beryl, Padparadscha Sapphire}. We have that p(Jadeite) = $20000, but p(Opal) does not exist, since Opal is not in the domain of p. □

For a function $f\colon A \to B$, we need each element a in A to be assigned to a unique element $f(a)$ in B. We do not require every element b in B to have a corresponding element in A. Specifically, an element of B may be the image of one, none, or several elements of A.

Example 2.3.1.9 Father function. Let P denote the set of all people (alive or dead). Let $F\colon P \to P$ be the function defined by the rule $F(x)$ is the father of x. For example, F(Ben Stiller) = Jerry Stiller, since Jerry is Ben's dad.

On the other hand, if we would run into trouble if we tried to define a function G for grandfathers. This is because people have two grandfathers (a paternal grandfather and a maternal grandfather), so G(Ben Stiller) is ambiguous. Functions are not allowed to have this ambiguity. Thus G does not define a function from P to P, though it *does* define a relation that we discuss in Section 6.1. □

Example 2.3.1.10 A function with domain $\{1,2,3\}$. Let $A = \{1,2,3\}$, and let $B = \{1,2,3,4\}$. Let $f\colon A \to B$ be the function defined by $f(1) = 4$, $f(2) = 3$, and $f(3) = 1$. Let's see what some of the terms mean for this function.

- The domain is $\{1,2,3\}$.
- The codomain} is $\{1,2,3,4\}$.
- The image of 3 is 1, since $f(3) = 1$.
- A preimage of 3 is 2, since $f(2) = 3$.
- The element 2 has no preimage, since there is no element in the domain that f sends to 2.
- The image of $\{1,2\}$ under f is $\{3,4\}$, since

$$f(\{1,2\}) = \{f(1), f(2)\} = \{4,3\} = \{3,4\}\,.$$

- The image of the function f the same as the image of the domain A under f. This is $\{1,3,4\}$, since

$$f(A) = f(\{1,2,3\}) = \{f(1), f(2), f(3)\} = \{4,3,1\} = \{1,3,4\}\,.$$

□

We can think about functions in terms of objects and arrows, where each arrow joins an element a in A to the corresponding element $f(a)$ in B as show in the next example.

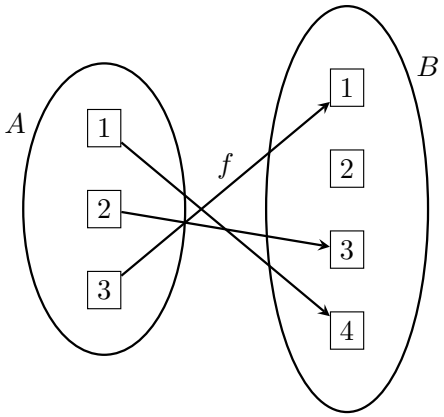


Figure 2.3.1.11 We represent the function $f : A \to B$ where $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$ from Example 2.3.1.10 graphically using objects and arrows.

Example 2.3.1.12 Volume. The volume of a box is a function of its length, width, and height. Let's break this example down and recognize the components described above.

What is the *domain* ?

Consider the the set of triples of numbers representing the length, width, and height of a box. Then each number must be nonnegative (if we allow degenerate boxes). For concreteness, let's consider ordered triples (ℓ, w, h) for length, width, and height. The set of all such triples is the domain of the volume function.

The volume of a box is a real number, so we can take the *codomain* to be $\mathbb{R}$.

If V is the name of our volume function, then we have $V \colon \mathbb{R}^3_{\geq 0} \to \mathbb{R}$. The *image* of V is the set of realizable outputs of the function V. Thus the image is the set of nonnegative real numbers,

$$\mathrm{im}(V) = \mathbb{R}_{\geq 0}.$$

□

2.3.2 Graphs of functions

We can represent a function $f : A \to B$ as a subset of $A \times B$.

Definition 2.3.2.1 Let $f : A \to B$ be a function. The graph of f is

$$\{(a, f(a)) : a \in A\} \subset A \times B.$$

◊

Example 2.3.2.2 Function given by its graph. Let $g : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3\}$ be given by its graph

$$\{(1, 1), (2, 3), (3, 1), (4, 2)\} \subset \{1, 2, 3, 4\} \times \{1, 2, 3\}.$$

Then $g(1) = 1$, $g(2) = 3$, $g(3) = 1$, and $g(4) = 2$. □

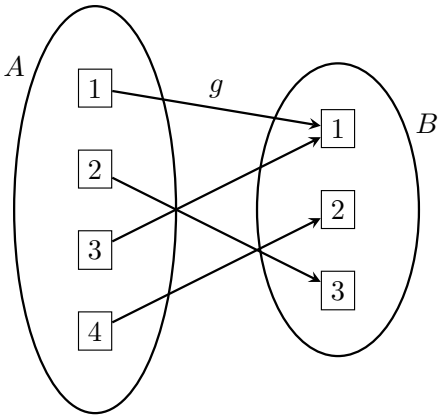
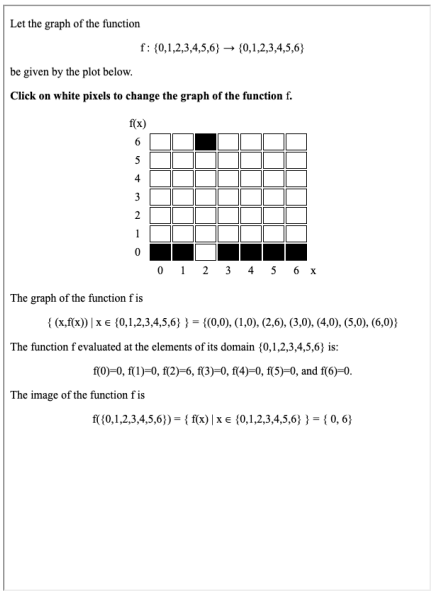


Figure 2.3.2.3 Graphical representation of the function $g : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3\}$ with graph $\{(1, 1), (2, 3), (3, 1), (4, 2)\}$ from Example 2.3.2.2.

Note that in both Figure 2.3.1.11 and Figure 2.3.2.3 each element of A is assigned to a unique element of B. An element of B may be the image of 1, none, or several elements of A.

We plot a function $f : A \to B$ by representing the elements of its graph $G \subset A \times B$ by black pixels and the elements of $(A \times B) - G$ by white pixels on the $A \times B$ plane. In Example 2.3.2.4 observe how changing the plot of the graph of a function changes its graph.

Example 2.3.2.4 Interactive plot of a function and its graph.



□

Functions need to assign exactly one element of the codomain to each element in the domain. Relations are generalizations of functions that do not have this restriction. Just like the graph of a functions relations can be represented as a subset of a Cartesian product.

We cover relations in more detail in Chapter 6.

Definition 2.3.2.5 Let A and B be sets. A subset R of a $A \times B$ is called a **relation** from A to B. ◊

In Figure 2.3.2.6 and Figure 2.3.2.7 we give examples of relations that are

not functions.

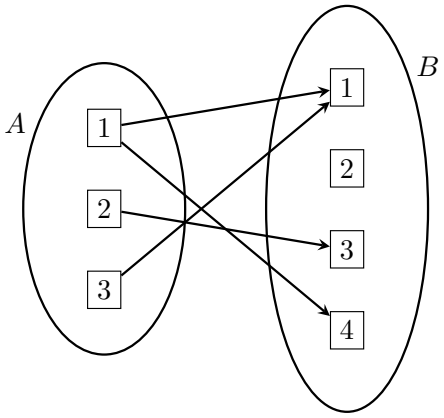


Figure 2.3.2.6 The graphical representations of the relation $\{(1,1),(1,4),(2,3),(3,1)\} \subseteq A \times B$ where $A = \{1,2,3\}$ and $B = \{1,2,3,4\}$. This relation is not a function, because the element $1 \in A$ is assigned to $1 \in B$ and $4 \in B$.

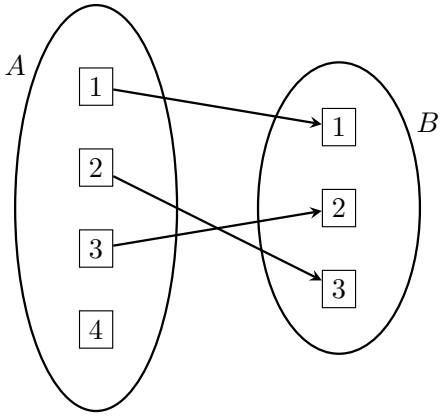


Figure 2.3.2.7 The graphical representations of the relation $\{(1,1),(2,3),(3,2)\} \subseteq A \times B$ where $A = \{1,2,3,4\}$ and $B = \{1,2,3\}$. This relation is not a function, because the element $4 \in A$ is not assigned to any element in B and functions need to assign exactly one element of the codomain to each element in the domain.

Remark 2.3.2.8 A function is a special kind of relation from A to B. Specifically, each element a in A is related to the element $f(a)$ in B. We discuss relations in detail in Section 6.1.

2.3.3 Floor and Ceiling

Definition 2.3.3.1 The **floor** of a real number x, denoted $\lfloor x \rfloor$, is the largest integer that is less than or equal to x. ◊

Example 2.3.3.2 Floor. The floor $\lfloor 2.7 \rfloor = 2$, since 2 is the largest integer that is less than or equal to 2.7. □

Warning 2.3.3.3 Computing the floor of a real number is different from just throwing away everything after the decimal. In the next example, we compute the floor of a negative number.

Example 2.3.3.4 Floor of a negative number. The floor $\lfloor -2.7 \rfloor = -3$, since -3 is the largest integer that is less than or equal to -2.7. □

Example 2.3.3.5 Floor of an integer. The floor of an integer is the integer itself, so $\lfloor 3 \rfloor = 3$, $\lfloor -3 \rfloor = -3$, and $\lfloor 0 \rfloor = 0$. □

Definition 2.3.3.6 The **ceiling** of a real number x, denoted $\lceil x \rceil$, is the smallest integer that is greater than or equal to x. ◊

Example 2.3.3.7 Ceiling. The ceiling $\lceil 2.7 \rceil = 3$, since 3 is the smallest integer that is greater than or equal to 2.7. □

Warning 2.3.3.8 Computing the ceiling of a real number is different from just throwing away everything after the decimal and adding 1. In the next example, we compute the ceiling of a negative number.

Example 2.3.3.9 Ceiling of a negative number. The ceiling $\lceil -2.7 \rceil = -2$, since -2 is the smallest integer that is greater than or equal to -2.7. □

Example 2.3.3.10 Ceiling of an integer. The ceiling of an integer is the integer itself, so $\lceil 3 \rceil = 3$, $\lceil -3 \rceil = -3$, and $\lceil 0 \rceil = 0$. □

Remark 2.3.3.11 If we think of x as a point on the usual real number line, then the floor of x is the integer that is directly to the left of x, unless x itself is an integer. The ceiling of x is the integer that is directly to the right of x, unless x itself is an integer.

2.3.4 Injective functions

Some functions "lose information" in the sense that the output does not uniquely determine the input. For example, consider the father function F given in Example 2.3.1.9. Knowing that the father of x is George Foreman does uniquely determine x since George Foreman has twelve children (including five sons all named George). Other functions do not lose information in this way. This property is known as **injectivity**. Namely, injective functions are the functions that do not lose information.

Definition 2.3.4.1 Let $f\colon A \to B$ be a function. The function f is **injective** or **one-to-one** if for all a and $\hat{a}$ in A, $f(a) = f(\hat{a})$ implies $a = \hat{a}$. ◊

Example 2.3.4.2 Squaring function is not injective. The squaring function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not injective because $f(-2) = f(2)$, but $-2 \neq 2$. □

Example 2.3.4.2 shows the general technique for showing a function is not injective.

Proof Technique 2.3.4.3 Show f is not injective. *Suppose $f\colon A \to B$ is a function, and we want to show f is not injective.*

1. *Find two different elements $a \neq \hat{a}$ in the domain A*
2. *Conclude f is not injective.*

Example 2.3.4.4 Students and ID. The function U denote the set of UNCG students, and consider the function $\mathrm{ID}\colon U \to \mathbb{Z}$ that assigns to each student their university ID. In other words, $\mathrm{ID}(x)$ is the university ID of x. These ID numbers are unique by design so that so that no two students have the same university ID. Specifically, the only way to have $\mathrm{ID}(\text{student } a) = \mathrm{ID}(\text{student } \hat{a})$ is to have student a be the same student as $\hat{a}$. The ID function is injective. □

For an injective function, every element in the codomain can have *at most one* preimage.

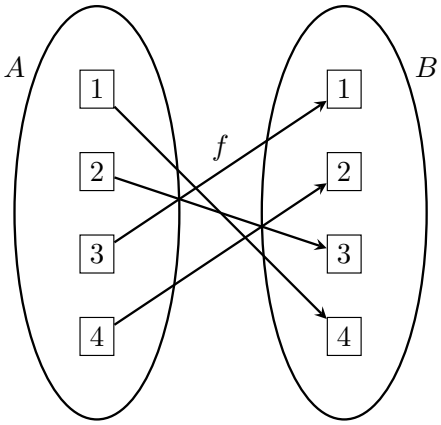


Figure 2.3.4.5 Graphical representations of an injective function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3, 4\}$. Each element in the codomain B has at most one preimage in the domain A.

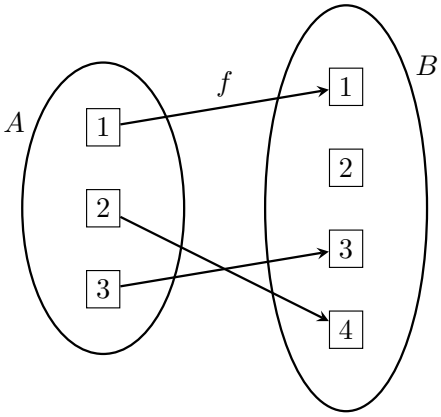


Figure 2.3.4.6 Graphical representations of an injective function $f : A \to B$ where $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$. Each element in the codomain B has at most one preimage in the domain A.

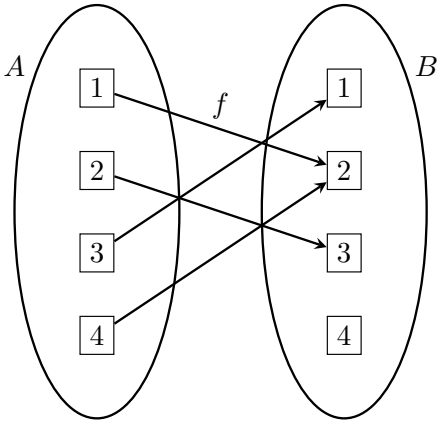


Figure 2.3.4.7 Graphical representation of a function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3, 4\}$ that is not injective. There is an element in the codomain has more than one preimage. Both 1 and 4 map to 2, so 2 has two preimages.

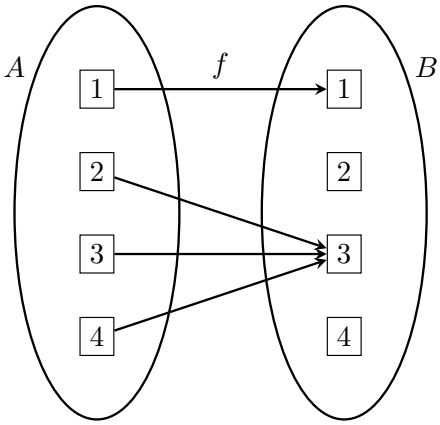


Figure 2.3.4.8 Graphical representation of a function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3, 4\}$ that is not injective. There is an element in the codomain has more than one preimage. Because 2, 3, and 4 map to 3, the element 3 of the codomain has three preimages.

Example 2.3.4.9 Prove injectivity. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = 3x - 2$. Show that f is injective

Solution. We show f is injective using the definition.

1. First, we rewrite what we want to show in the form "if p, then q." This will involve giving names to objects, and recalling the definition of *injective*. We want to prove f is injective, which means we want to prove: "If $f(a_1) = f(a_2)$, then $a_1 = a_2$."

2. As before, to prove a statement of the form "if p, then q" directly, we assume p is true and try to show q. This is commonly where we set some notation as well. It is also a good place to remind the reader what we want to show, for example, "Let a_1 and a_2 be real numbers such that $f(a_1) = f(a_2)$. We want to show that $a_1 = a_2$".

3. Now look back at the goal. It should be some relationship between a_1 and a_2. It makes sense to write out what we know and simplify to see if we get what we want.

$$\begin{aligned} f(a_1) &= f(a_2) \\ 3a_1 - 2 &= 3a_2 - 2 \\ 3a_1 &= 3a_2 \\ a_1 &= a_2 \end{aligned}$$

4. The part above should complete the proof. Since we chose generic a_1 and a_2 from the domain of f, the argument covers all a_1 and a_2 in the domain of f.

□

The technique in Example 2.3.4.9 can be generalized to prove a function is injective.

Proof Technique 2.3.4.10 Show f is injective. *Suppose $f : A \to B$ is a function, and we want to show f is injective.*

1. *Fix generic elements a and $\hat{a}$ in the domain A such that $f(a) = f(\hat{a})$.*

2. *Deduce $a = \hat{a}$.*

3. *Conclude f is injective.*

2.3.5 Surjective functions

Some functions have a codomain that is larger than it needs to be. Namely, they have a codomain that is larger than the image of the function. For example, consider the absolute value function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|$. Then the codomain of f is $\mathbb{R}$, but the image of f is just the set of nonnegative real numbers, $\text{im}(f) = \mathbb{R}_{\geq 0}$.

Negative numbers are in the codomain and have no preimage. Other functions have the property that the codomain is as small as possible. Namely, the codomain is exactly equal to the image. For example, we can modify f to define a new function $g\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ given by $g(x) = |x|$. Notice now that the codomain is equal to the image. Another way of saying that is every element of the codomain has at least one preimage. This property is known as **surjectivity**. Namely, **surjective functions** codomains equal to their image.

Definition 2.3.5.1 Let $f\colon A \to B$ be a function. The function f is **surjective** or **onto** if for each b in the codomain B, there exists an element a in the domain A such that $f(a) = b$; equivalently, if $f(A) = B$. A surjective function is called a **surjection** . ♢

Example 2.3.5.2 Squaring function is not surjective. The squaring function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not surjective because $-2 \in \mathbb{R}$, but there is no $a \in \mathbb{R}$ such that $f(a) = -2$. □

Example 2.3.5.2 shows the general technique for proving a function is not surjective.

Proof Technique 2.3.5.3 Show f is not surjective. *Suppose $f\colon A \to B$ is a function, and we want to prove f is not surjective.*

1. *Find particular b in the codomain B that is not in the image of f, that is , b such that $f(a) \neq b$ for all a in the domain A.*

2. *Conclude f is not surjective.*

We can make a non-surjective function into a surjective one by shrinking the codomain.

Example 2.3.5.4 The squaring function is not surjective. The squaring function is surjective onto $\mathbb{R}_{\geq 0}$. Specifically, the function $g\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ given by $g(x) = x^2$ is surjective. □

Warning 2.3.5.5 A function is a triple of data (f, A, B). In particular, the domain and codomain are part of the function. Because of this, when we change the domain or codomain, we are really creating a *new* function. For example, the squaring function from $\mathbb{R}$ to $\mathbb{R}$ is *different* from the squaring function from $\mathbb{R}$ to $\mathbb{R}_{\geq 0}$.

For a function to be surjective, every element in the codomain must have *at least one* preimage. For example, the functions in Figure 2.3.5.6 and Figure 2.3.5.7 are surjective.

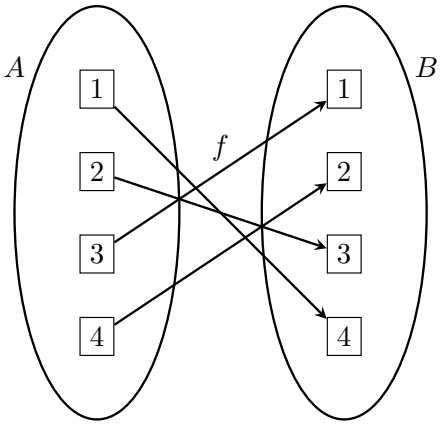


Figure 2.3.5.6 Graphical representation of a surjective function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3, 4\}$. Each element in the codomain B has at least one preimage in the domain A.

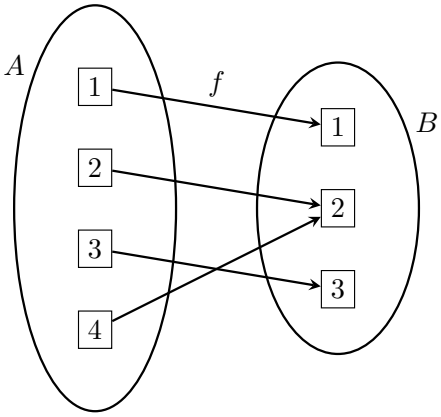


Figure 2.3.5.7 Graphical representation of a surjective function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3\}$. Each element in the codomain B has at least one preimage in the domain A.

In Figure 2.3.5.8 and Figure 2.3.5.9 we give examples of functions that are not surjective. In each instance, there is an element in the codomain has has no preimage.

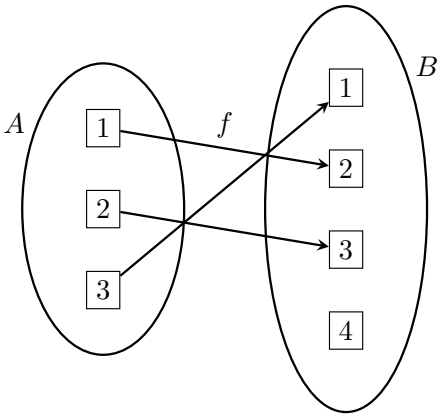


Figure 2.3.5.8 Graphical representations of a function $f : A \to B$ where $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$ that is not surjective. There is no element in the domain A that maps to the element 4 of the codomain B, so 4 has no preimage.

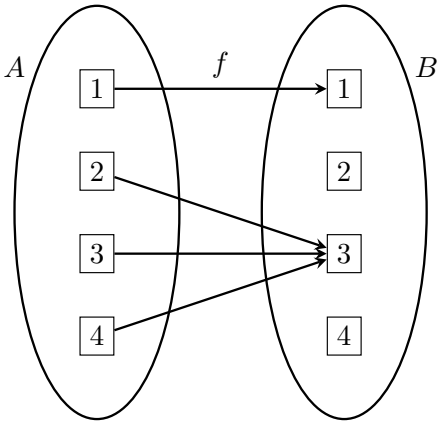


Figure 2.3.5.9 Graphical representations of a function $f : A \to B$ where $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3, 4\}$ that is not surjective, because there are no elements that map to 2 or 4, so both 2 and 4 have no preimages.

Example 2.3.5.10 Prove surjectivity. Let $f(x) = 3x - 2$. Show that f is surjective.

Solution. We prove that f is surjective as follows.

1. First we recall the definition of surjective. At the same time,
2. Now this still does not look like a "if p, then q" type of
3. As before, to prove a statement of the form "if p, then q"
4. Since we want to find a, we should write out the conditions
5. We have found a as desired. Since b is

□

The technique in Example 2.3.5.10 an be generalized to prove a function is surjective.

Proof Technique 2.3.5.11 Show f is surjective. *Suppose $f : A \to B$ is a function, and we want to prove f is surjective.*

1. *Fix a generic element b in the codomain B.*

2. *Find a in the domain A such that $f(a) = b$.*

3. *Conclude f is surjective.*

2.3.6 Bijective functions

Definition 2.3.6.1 Let $f\colon A \to B$. The function f is **bijective** if f is injective and surjective (or one-to-one and onto). A bijective function is called a **bijection** or **one-to-one correspondence**. ◊

Example 2.3.6.2 Identity function is bijective. The identity function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x$ is bijective. □

Proof Technique 2.3.6.3 Show f is bijective. *Suppose $f\colon A \to B$ is a function, and we want to prove f is bijective.*

1. *Use Proof Technique Proof Technique 2.3.4.10 to show f is*

2. *Use Proof Technique Proof Technique 2.3.5.11 to show f is*

3. *Conclude f is bijective.*

Remark 2.3.6.4 If A and B are finite sets of the same size, a function $f\colon A \to B$ is surjective if and only if f is injective. Furthermore, if $\#A < \#B$ then f cannot be surjective. If $\#B < \#A$ then f cannot be injective.

Example 2.3.6.5 Area of a circle. The area of a circle is a function of its radius. This function is injective, because if the area of a circle of radius r is equal to the area of a circle of radius r', we must have $r = r'$. This function is not surjective, when viewed as a function from $\mathbb{R}_{>0}$ to $\mathbb{R}$, since there is no radius that would give rise to area -1. When we restrict the codomain to $\mathbb{R}_{>0}$, this function is surjective since we can construct a circle for any given area. □

Proof Technique 2.3.6.6 Show f is not bijective. *Suppose $f\colon A \to B$ is a function, and we want to prove f is no bijective.*

Complete one of the following.

1. *Use Proof Technique Proof Technique 2.3.4.3 to show f is not injective.*

2. *Use Proof Technique Proof Technique 2.3.5.3 to show f is not surjective.*

3. *Conclude f is not bijective.*

Definition 2.3.6.7 Let $f\colon A \to B$ be a bijection. The **inverse** of f, denoted f^{-1} , is a function $f^{-1}\colon B \to A$ that assigns to an element $b \in B$, the unique element $a \in A$ such that $f(a) = b$. ◊

From Definition 2.3.6.7, it immediately follows that f is an invertible function if and only if

$$f(f^{-1}(b)) = b \quad \text{for all } b \in B; \text{ and}$$
$$f^{-1}(f(a)) = a \quad \text{for all } a \in A.$$

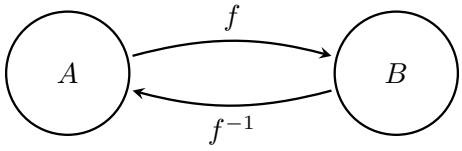


Figure 2.3.6.8 A function f from A to B and its inverse from B to A.

Remark 2.3.6.9

1. The definition of inverse makes sense since f is a
2. f^{-1} does not mean $\frac{1}{f}$.
3. Non-bijective functions do not have an inverse.

Example 2.3.6.10 Inverse of a function. Consider the function $f(x) = 3x + 2$. Find the inverse of f.

Solution. Let $y = f(x)$. Then $f^{-1}(y) = x$. That means we can just take $y = 3x + 2$, and solve for x. Doing so, we see that $x = \frac{y-2}{3}$. That means $f^{-1}(y) = \frac{y-2}{3}$. Convention often dictates the input variable be named x, so we write it as

$$f^{-1}(x) = \frac{x-2}{3}.$$

□

Finding the inverse of a function is more difficult than checking if given functions are inverses of each other.

Example 2.3.6.11 Check the inverse. Let f and g be functions from $\mathbb{R}$ to $\mathbb{R}$ given by $f(x) = x^3 - 4$ and $g(x) = \sqrt[3]{x+4}$. Determine whether g is the inverse of f.

Solution. We need to check that and both hold. We compute each one separately. For all $x \in \mathbb{R}$, we have

$$\begin{aligned} f(g(x)) &= f(\sqrt[3]{x+4}) \\ &= (\sqrt[3]{x+4})^3 - 4 \\ &= (x+4) - 4 \\ &= x; \end{aligned}$$

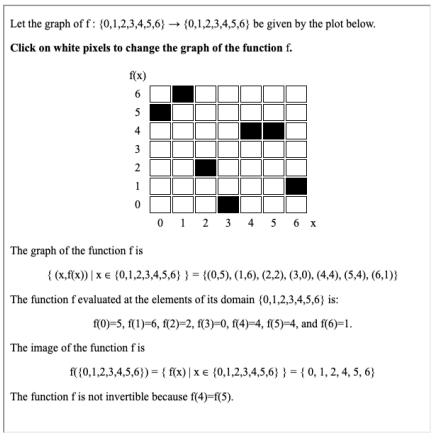
$$\begin{aligned} g(f(x)) &= g(x^3 - 4) \\ &= \sqrt[3]{(x^3-4)+4} \\ &= \sqrt[3]{x^3} \\ &= x. \end{aligned}$$

Thus g is the inverse of f.

Note: This also shows that f is the inverse of g. □

In Example 2.3.6.12 observe how changing the plot of the graph of a function changes its graph and the properties of the function.

Example 2.3.6.12 Interactive plot of a function with invertibility.



□

2.3.7 Composition of functions

Definition 2.3.7.1 Let $f\colon A \to B$, and let $g\colon B \to C$. The **composition**, denoted $g \circ f$, is the function $g \circ f\colon A \to C$ defined by

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

◊

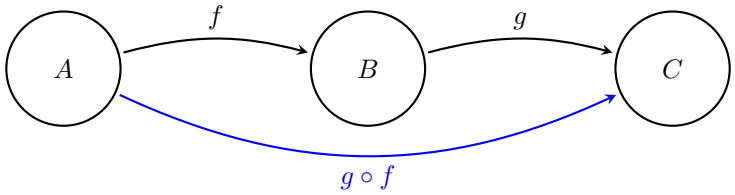


Figure 2.3.7.2 Composition $g \circ f$ of two functions $f : A \to B$ and $g : B \to C$.

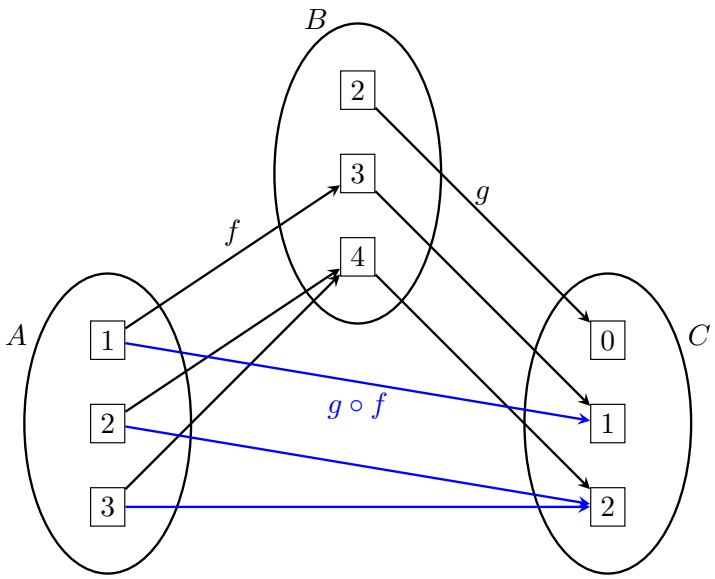


Figure 2.3.7.3 Two functions $f : A \to B$ and $g : B \to C$ and their composition $g \circ f$ where $A = \{1, 2, 3\}$, $B = \{2, 3, 4\}$, and $C = \{0, 1, 2\}$.

Example 2.3.7.4 Composition. Let $f(x) = x^2 + 2x - 1$ and $g(x) = 3x - 5$ be functions from $\mathbb{R}$ to $\mathbb{R}$.

Let's compute some compositions. First, consider $f \circ g$.

$$\begin{aligned}(f \circ g)(x) &= f(g(x)) \\ &= f(3x - 5) \\ &= (3x - 5)^2 + 2(3x - 5) - 1 \\ &= (9x^2 - 30x + 25) + (6x - 10) - 1 \\ &= 9x^2 - 24x + 14.\end{aligned}$$

Next, let's compose them in the opposite order.

$$\begin{aligned}(g \circ f)(x) &= g(f(x)) \\ &= g(x^2 + 2x - 1) \\ &= 3(x^2 + 2x - 1) - 5 \\ &= (3x^2 + 6x - 3) - 5 \\ &= 3x^2 + 6x - 8.\end{aligned}$$

Finally, let's compose g with itself.

$$\begin{aligned}(g \circ g)(x) &= g(g(x)) \\ &= g(3x - 5) \\ &= 3(3x - 5) - 5 \\ &= (9x - 15) - 5 \\ &= 9x - 20.\end{aligned}$$

□

Definition 2.3.7.5 Let A be a nonempty set. We call $\mathrm{id}_A : A \to A$ with $\mathrm{id}_A(x) = x$ the identity function on A. ◊

It follows from Definition 2.3.6.7 and Definition 2.3.7.1 that for a bijective function $f : A \to B$ and its inverse $f^{-1} : B \to A$ that

$$f \circ f^{-1} = \mathrm{id}_B \quad \text{and} \quad f^{-1} \circ f = \mathrm{id}_A.$$

2.3.8 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) function

(b) image of an element

(c) image or range of a function

(d) image of a set under a function

(e) preimage of an element

(f) graph of a function

(g) relation

(h) floor of a real number

(i) ceiling of a real number

(j) injective function

(k) surjective function

(l) bijective function

(m) inverse of a function

(n) composition of functions

2. Let $f(x) = x^2 + 1$ and $g(x) = x + 2$ be functions from $\mathbb{R}$ to $\mathbb{R}$.

(a) Find the composition $h = f \circ g$. What is the image of h?

(b) Find the composition $i = g \circ f$. What is the image of i?

3. Construct a function that is

(a) injective but not surjective.

(b) surjective but not injective.

(c) neither surjective nor injective.

(d) both injective and surjective.

4. Let $f\colon \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = 3x + 1$ Prove that

(a) f is surjective.

(b) f is injective.

(c) f is bijective.

5. Let $\mathcal{W}$ be the set of words in the Oxford English Dictionary, and let $\mathcal{A}$ be the English alphabet. Let $f\colon \mathcal{W} \to \mathcal{A}$ be the function defined by

$$f(x) = \text{first letter of } x.$$

(a) What is the image of f?

(b) Is f is surjective? Justify.

(c) Is f is injective? Justify.

6. Give an explicit example of a function from $\mathbb{Z}$ to $\mathbb{Z}$ that is

(a) injective, but not surjective.

(b) surjective, but not injective.

(c) injective and surjective.

(d) neither injective nor surjective.

Justify your answers.

7. Let $S = \{-2, -1, 0, 3, 4, 5\}$ Find $f(S)$ for each of the following. Be sure to express the answer as a set.

(a) $f(x) = 1$

(b) $f(x) = 2x + 1$

(c) $f(x) = x^2$

(d) $f(x) = \left\lceil \frac{x}{2} \right\rceil$

(e) $f(x) = \left\lfloor \frac{x^2+1}{3} \right\rfloor$

8. Let A and B be sets, and let $f\colon A \to B$ be a function. Let S and T be subsets of A. Prove $f(S \cup T) = f(S) \cup f(T)$.

9. Let A and B be sets, and let $f\colon A \to B$ be a function. Let S and T be subsets of A.

(a) Prove $f(S \cap T) \subseteq f(S) \cap f(T)$.

(b) Give an explicit example that shows the inclusion above can be proper.

10. Let A and B be sets, and let $f\colon A \to B$ be a function. Let S and T be subsets of A. Prove that if f is injective, then $f(S \cap T) = f(S) \cap f(T)$.

2.4 Sequences and summations

Outcomes

1. Determine whether sequences are arithmetic or geometric progressions
2. Evaluate summations.
3. Evaluate recurrence relations.
4. Solve recurrence relations.

2.4.1 Basic Terminology

Definition 2.4.1.1 A **sequence** is a function from a subset of the integers (usually $\{0, 1, 2, \ldots\}$ or $\{1, 2, 3, \ldots\}$) to a set S.

We use a_n to denote the image of n, and we call a_n a **term** of the sequence.

We use the notation $(a_n)_n$ to describe the whole sequence, and a_n represents an individual term in the sequence. ◊

Remark 2.4.1.2 Note the notation conflicts with our notation for sets. The context will make it clear which we are discussing.

Example 2.4.1.3 Sequence. The sequence $(a_n)_n$, where $a_n = \frac{1}{n^2}$ for $n = 1, 2, 3, \ldots$ has terms

$$1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \frac{1}{25}, \ldots, \frac{1}{n^2}, \frac{1}{(n+1)^2}, \ldots.$$

□

2.4.2 Geometric and arithmetic sequences

Definition 2.4.2.1 A **geometric progression** is a sequence of the form

$$a, ar, ar^2, ar^3, \ldots, ar^n, ar^{n+1}, \ldots,$$

where the **initial term** a and **common ratio** r are real numbers. ◊

Suppose $(a_n)_n$ is a geometric progression. How do we compute its initial term and common ratio? The initial term is the first term in the sequence. The common ratio is the ratio of consecutive terms:

$$\frac{a_{n+1}}{a_n} = \frac{ar^{n+1}}{ar^n} = r,$$

assuming the sequence is not the zero sequence.

This is also how we can tell if a sequence is a geometric progression. A geometric progression must have the ratio of consecutive terms constant, no matter where we look in the sequence.

Example 2.4.2.2 Geometric progression. The sequence $3, 6, 12, 24, \ldots$ is a geometric progression with initial term $a = 3$. The common ratio is $r = 2$, since

$$\frac{6}{3} = \frac{12}{6} = \frac{24}{12} = \cdots = 2.$$

□

Example 2.4.2.3 Geometric progression. The sequence $(a_n)_n = 2 \cdot 5^n$ for $n = 0, 1, 2, \ldots$ is a geometric progression with initial term $a = 2$, since $a_0 = 2 \cdot 5^0 = 2$. The common ratio $r = 5$, since

$$\frac{a_{n+1}}{a_n} = \frac{2 \cdot 5^{n+1}}{2 \cdot 5^n} = 5.$$

□

Example 2.4.2.4 Not a geometric progression. The sequence $1, 3, 5, 7, 9, \ldots$ is not a geometric progression since the ratio of consecutive terms is not constant. For example,

$$\frac{3}{1} \neq \frac{5}{3}.$$

□

Definition 2.4.2.5 An **arithmetic progression** is a sequence of the form

$$a, a + d, a + 2d, \ldots, a + nd, \ldots,$$

where the **initial term** a and **common difference** d are real numbers. ◊

Suppose $(a_n)_n$ is an arithmetic progression. How do we compute its initial term and common difference? The initial term is the first term in the sequence. The common difference is the difference of consecutive terms:

$$a_{n+1} - a_n = (a + (n + 1)d) - (a + nd) = d.$$

This is also how we can tell if a sequence is an arithmetic progression. An arithmetic progression must have the difference of consecutive terms constant, no matter where we look in the sequence.

Example 2.4.2.6 Arithmetic progression. The sequence $2, 5, 8, 11, 14, \ldots$ is an arithmetic progression with initial term 2. The common difference is 3, since

$$5 - 2 = 8 - 5 = 11 - 8 = 14 - 11 = \cdots = 3.$$

□

Example 2.4.2.7 Arithmetic progression. The sequence $(a_n)_n$ defined by $a_n = 3 + 5n$ for $n = 0, 1, 2, \ldots$ is an arithmetic progression with initial term $a = 3$, since $a_0 = 3 + 5 \cdot 0 = 3$. The common difference is $d = 5$, since

$$a_{n+1} - a_n = (3 + 5(n + 1)) - (3 + 5n) = 5.$$

□

Example 2.4.2.8 Not an arithmetic progression. The sequence $1, 2, 4, 8, 16, 32, \ldots$ is not an arithmetic progression, since the difference of consecutive terms is not constant. For example,

$$2 - 1 \neq 4 - 2.$$

□

2.4.3 Recurrence relations

Definition 2.4.3.1 A **recurrence relation** for a sequence $(a_n)_n$ is an equation that expresses a_n in terms of one or more of the previous terms of the sequence. ◊

Example 2.4.3.2 Compound interest. Suppose we deposit P_0 dollars in a savings account that yields 3% interest per year, compounded annually. Find a formula for the amount in the account after n years.

Solution. Let $(P_n)_n$ be the sequence where the n th term is the amount in the account after n years. Then

$$\begin{aligned}
P_0 &= P_0 \\
P_1 &= P_0 + 0.03P_0 = 1.03P_0 \\
P_2 &= P_1 + 0.03P_1 = 1.03P_1 = (1.03)^2 P_0 \\
P_3 &= P_2 + 0.03P_2 = 1.03P_2 = (1.03)^3 P_0 \\
&\vdots \quad \vdots \\
P_n &= 1.03P_{n-1} = (1.03)^n P_0.
\end{aligned}$$

□

Example 2.4.3.3 Recurrence relation. Suppose $(a_n)_n$ is a sequence that satisfies the recurrence relation $a_n = -a_{n-1} + 1$ for $n = 1, 2, 3, \ldots$, and suppose $a_0 = 2$.

Solution. Then

$$\begin{aligned}
a_0 &= 2 \\
a_1 &= -a_0 + 1 = -2 + 1 = -1 \\
a_2 &= -a_1 + 1 = -(-1) + 1 = 2 \\
a_3 &= -a_2 + 1 = -2 + 1 = -1 \\
a_4 &= -a_3 + 1 = -(-1) + 1 = 2 \\
&\vdots \quad \vdots \\
a_n &= \begin{cases} 2 & \text{if } n \text{ is even,} \\ -1 & \text{if } n \text{ is odd.} \end{cases}
\end{aligned}$$

□

Definition 2.4.3.4 A **solution** of a recurrence relation is a sequence whose terms satisfy the recurrence relation. ◊

We call a solution $(a_n)_n$ to a recurrence relation a **closed formula** if the expression for a_n can be evaluated in a fixed finite number of operations on n.

Example 2.4.3.5 Solution of a recurrence relation. Is $(a_n)_n$, where $a_n = 3n$ for $n = 0, 1, 2, \ldots$ a solution of the recurrence relation $a_n = 2a_{n-1} - a_{n-2}$ for $n = 2, 3, \ldots$?

Solution. We just need to check if the terms of the sequence satisfy the recurrence relation. We compute for $n \geq 2$,

$$\begin{aligned} 2a_{n-1} - a_{n-2} &= 2(3(n-1)) - (3(n-2)) \\ &= (6n-6) - (3n-6) = 3n = a_n. \end{aligned}$$

Thus $a_n = 3n$ is a solution of the recurrence relation.

Note: It is easy to check that the sequence $(b_n)_n$, defined by $b_n = 5$ for $n = 0, 1, 2, \ldots$ is also a solution of the recurrence relation. In particular, a recurrence relation can have more than one solution. □

Example 2.4.3.6 Factorial as the solution of a recurrence relation. The sequence $(a_n)_n$ defined by $a_n = n!$ for $n = 1, 2, 3, \ldots$ is a solution to the recurrence relation $a_n = na_{n-1}$, since $n! = n((n-1)!)$. □

Next, we define a famous sequence definied by a recurrence relation and two base cases which is named for a twelfth century Italian mathematician.

Definition 2.4.3.7 The **Fibonacci sequence** $f_0, f_1, f_2, \ldots$ is defined by the initial conditions $f_0 = 0$, $f_1 = 1$, and the recurrence relation

$$f_n = f_{n-1} + f_{n-2},$$

for $n = 2, 3, \ldots$. ◊

Example 2.4.3.8 First terms of the Fibonacci sequence. It is straightforward to compute more terms of the Fibonacci sequence.

$$\begin{aligned} f_0 &= 0 \\ f_1 &= 1 \\ f_2 &= f_1 + f_0 = 1 + 0 = 1 \\ f_3 &= f_2 + f_1 = 1 + 1 = 2 \\ f_4 &= f_3 + f_2 = 2 + 1 = 3 \\ f_5 &= f_4 + f_3 = 3 + 2 = 5 \end{aligned}$$

□

Using linear algebra, we can find a closed formula for the terms of the Fibonacci sequence. We will not derive this formula in this course, but you will be asked to prove it in Exercise 3.2.3.5.

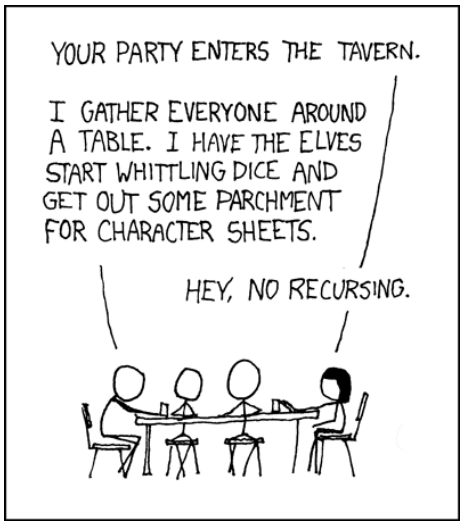


Figure 2.4.3.9 *Tabletop Roleplaying* by Randall Munroe (https://xkcd.com/244/). I may have also tossed one of a pair of teleportation rings into the ocean, with interesting results.

2.4.4 Recursion

Recall that a sequence $(a_n)_n$ can be defined for all $n \in \mathbb{N}_0$ by giving

1. an **initial value** a_0 and
2. a **recurrence relation** that for $k \geq 2$ gives a_k as a formula in a_{k-1}.

Recurrence relations example of a more general concept called **recursion**, where a concept or process depends on a simpler version of itself.

In Python a **recursive function** is a function that calls itself. To assure termination we need to have a base case for which the function does not call itself.

We give examples of recurrence relations and their implementation with recursive functions in Python..

Example 2.4.4.1 Factorial. Let the sequence $(f_n)_n$ be given by the recurrence relation $f_n = n \cdot f_{n-1}$ and the initial value $f_0 = 1$.

We compute f_5.

$$\begin{aligned} f_5 &= 5 \cdot f_4 \\ &= 5 \cdot (4 \cdot f_3) \\ &= 5 \cdot (4 \cdot (3 \cdot f_2)) \\ &= 5 \cdot (4 \cdot (3 \cdot (2 \cdot f_1))) \\ &= 5 \cdot (4 \cdot (3 \cdot (2 \cdot (1 \cdot f_0)))) \\ &= 5 \cdot (4 \cdot (3 \cdot (2 \cdot (1 \cdot 1)))) \\ &= 5 \cdot (4 \cdot (3 \cdot (2 \cdot 1))) \\ &= 5 \cdot (4 \cdot (3 \cdot 2)) \\ &= 5 \cdot (4 \cdot 6) \\ &= 5 \cdot 24 \\ &= 120 \end{aligned}$$

We have computed $f_5 = 120$ using 4 multiplications.

We can directly transfer the definition of the recurrence relation into Python

by letting a function call itself. Our Python function evaluates $f(n)$ for a non-negative integer n recursively following the recurrence relation $f(n) = n \cdot f(n-1)$ and the initial value $f(0) = 1$.

```
1  def f(n):
2    if n==0:
3      return 1
4    return n*f(n-1)
```

The function $f(n)$ can also be evaluated using a `for` loop.

```
1  def f(n):
2    f = 1
3    for n in range(1,n):
4      f=f*n
5    return f
```

□

Example 2.4.4.2 Compound interest recursive. Alice deposits \$1000 in an account that pays 5% interest yearly and her grandmother transfers \$100 to the account every year right after the interest t has been paid.

The amount A_t in Alice's account t years after her initial deposit is given by initial value

$$A_0 = 1000$$

and the recurrence relation

$$A_t = A_{t-1} + A_{t-1} \cdot 5\% + 100 = A_{t-1} \cdot 1.05 + 100.$$

We can use the following Python function for recursively evaluating A. When the input is a non-negative integer t the function `A` returns the amount in Alice's account after t years.

```
1  def A(t):
2    if t==0:
3      return 1000
4    return A(t-1)*1.05+100
```

□

Example 2.4.4.3 Exponentiation recursive. Let $b \in \mathbb{R}$ and let the sequence $(p_n)_n$ be given by $p_n = b^n$. Then p_n satisfies the recurrence relation

$$p_n = b \cdot p_{n-1}.$$

which together with the initial value $p_0 = 1$ uniquely defines p_n. We formulate this in Python:

```
1  def p(b,n):
2    if n==0:
3      return 1
4    return b*p(b,n)
```

□

Example 2.4.4.4 Fast exponentiation recursive. Let $b \in \mathbb{R}$ and let the sequence $(p_n)_n$ be given by $p_n = b^n$.

Recall that by Theorem 0.1.5.1 for $n \in \mathbb{N}_0$ there exits $q \in \mathbb{N}_0$ and $r \in \{0, 1\}$ such that $n = 2 \cdot q + r$. Thus

$$b^n = b^{2 \cdot q + r} = b^{q \cdot 2} \cdot b^r = (b^q)^2 \cdot b^r.$$

As in Theorem 0.1.5.1 we write $q = n \operatorname{div} 2$ and $r = n \bmod 2$ and get the

recurrence relation

$$p_n = (p_q)^2 \cdot b^r = (p_{(n \operatorname{div} 2)})^2 \cdot b^{(n \bmod 2)}$$

In Python this is considerably more efficient than the function in Example 2.4.4.3.

```
1  def p(b,n):
2    if n==0:
3      return 1
4    q,r = divmod(n,2)
5    return p(b,q)**2 * b**r
```

□

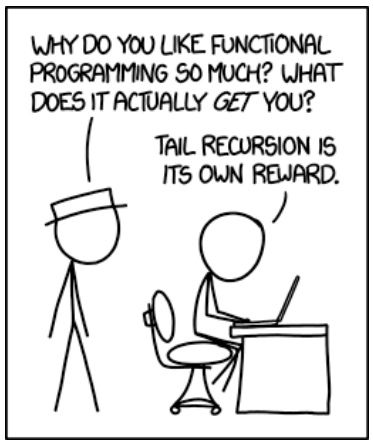


Figure 2.4.4.5 *Functional* by Randall Munroe (https://xkcd.com/1270). Functional programming combines the flexibility and power of abstract mathematics with the intuitive clarity of abstract mathematics.

2.4.5 Summations

We now consider the addition of terms of a sequence.

Definition 2.4.5.1 Let $(a_n)_n$ be a sequence of numbers. For a positive integer n, the **sum** from 1 to n of a_n , is

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n.$$

More generally, for $m \leq n$, we have

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + \cdots + a_n.$$

Even more generally, for a finite set I, we $\sum_{i \in I} a_i$ be the sum of the a_i with i in I. ◊

Remark 2.4.5.2 We are using the fact that addition is an associative operator. In particular,

$$\sum_{i=1}^{n} a_i = ((\ldots(((a_1 + a_2) + a_3) + a_4) + \ldots) + a_n)$$

For summing infinitely many things, one must define the sum to be the limit of partial sums as described above. This will be covered in more detail in

Calculus, since the notion of limit is required.

Given a sequence $(a_i)_i$ as a function in Python `b`, the following the function in Listing 2.4.5.3 computes the sum from m to n of $(a_i)_i$.

Listing 2.4.5.3

```
def summation(b,m,n):
  """
␣␣Input:␣function␣b,␣integers␣m,n␣in␣the␣domain␣of␣b
␣␣Return␣sum␣of␣b(i)␣for␣i␣from␣m␣to␣n.
␣␣"""
  s = 0  # initialize sum
  for i in range(m,n + 1):
    s += a(i)
  return s
```

Example 2.4.5.4 Sum of squares.

$$\sum_{i=1}^{3} i^2 = 1^2 + 2^2 + 3^2 = 14.$$

□

Theorem 2.4.5.5 *Let $(a_n)_n$ and $\{b_n\}$ be defined on a finite set S. Suppose $\sum_{i\in S} a_i$ and $\sum_{i\in S} b_i$ exist. Then*

$$\sum_{i\in S} a_i + \sum_{i\in S} b_i = \sum_{i\in S} (a_i + b_i),$$

Proof. With the associativity and commutativity of addition we get

$$\begin{aligned}\sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i &= (a_m + a_{m+1} + \cdots + a_n) + (b_m + b_{m+1} + \cdots + b_n)\\ &= (a_m + b_m) + (a_{m+1} + b_{m+1}) + \cdots + (a_n + b_n)\\ &= \sum_{i=m}^{n} (a_i + b_i)\end{aligned}$$

■

Theorem 2.4.5.6 *Let $(a_n)_n$ be defined on a finite set S, and let c be a real number. Suppose $\sum_{i\in S} a_i$ exists. Then*

$$c \sum_{i\in S} a_i = \sum_{i\in S} c a_i.$$

Proof. With the distributivity of multiplication over addition we get

$$\begin{aligned}c \cdot \sum_{i=m}^{n} a_i &= c \cdot (a_m + a_{m+1} + \cdots + a_n)\\ &= (c \cdot a_m) + (c \cdot a_{m+1}) + \cdots + (c \cdot a_n)\\ &= \sum_{i=m}^{n} (c \cdot a_i).\end{aligned}$$

■

We can use these properties to break down complicated sums into simpler ones.

Example 2.4.5.7 Simplify a summation. Break down the summation

$$\sum_{k=0}^{100}(3k^2 - 5k + 2)$$

to a sum of simpler summations.

Solution. With the properties in Theorem 2.4.5.5 and Theorem 2.4.5.6 we get:

$$\begin{aligned}\sum_{k=0}^{100}(3k^2 - 5k + 2) &= \sum_{k=0}^{100} 3k^2 - \sum_{k=0}^{100} 5k + \sum_{k=0}^{100} 2 \\ &= 3\sum_{k=0}^{100} k^2 - 5\sum_{k=0}^{100} k + 2\sum_{k=0}^{100} 1.\end{aligned}$$

It remains to work out formulas for these summations:

$$\sum_{k=0}^{100} k^2, \quad \sum_{k=0}^{100} k, \quad \text{and} \quad \sum_{k=0}^{100} 1,$$

which is done in Example 2.4.5.13. □

Theorem 2.4.5.8 *If a and r are real numbers with $r \neq 0$, then*

$$\sum_{i=0}^{n} ar^i = \begin{cases} \frac{ar^{n+1}-a}{r-1} & \text{if } r \neq 1, \\ (n+1)a & \text{if } r = 1. \end{cases}$$

Proof. If $r = 1$, then

$$\sum_{i=0}^{n} ar^i = \sum_{i=0}^{n} a = \underbrace{a + a + \cdots + a}_{n+1 \text{ times}} = (n+1)a.$$

Now suppose $r \neq 1$. Let $s = \sum_{i=0}^{n} ar^i$. Then

$$rs = \sum_{i=0}^{n} ar^{i+1} = \sum_{i=1}^{n+1} ar^i.$$

Then

$$s(r-1) = rs - s = \sum_{i=1}^{n+1} ar^i - \sum_{i=0}^{n} ar^i = ar^{n+1} - a.$$

Dividing by $r - 1$ gives the result. ■

Example 2.4.5.9 Compute a sum. Compute the sum

$$1 + \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{100}}.$$

Solution. We rewrite the sum as

$$1 + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^{100} = \sum_{i=0}^{100} \left(\frac{1}{2}\right)^i.$$

This is the sum of a geometric progression with initial term $a = 1$ and common ratio $r = \frac{1}{2}$. Using Theorem 2.4.5.8 with $a = 1$, $r = \frac{1}{2}$, and $n = 100$, we have

that the sum is equal to

$$\frac{ar^{n+1}-a}{r-1} = \frac{\left(\frac{1}{2}\right)^{101}-1}{\frac{1}{2}-1} = 2\left(1-\frac{1}{2^{101}}\right),$$

which is just a bit less than 2. □

We revisit Example 2.4.4.2 and derive a closed formula for the compound interest recurrence relation.

Example 2.4.5.10 Compound interest closed formula. Alice deposits \$1000 in an account that pays 5% interest yearly and her grandmother transfers \$100 to the account every year right after the interest t has been paid.

The amount A_t in Alice's account t years after her initial deposit is given by initial value

$$A_0 = 1000$$

and the recurrence relation

$$A_t = A_{t-1} + A_{t-1} \cdot 5\% + 100 = A_{t-1} \cdot 1.05 + 100.$$

Derive a closed formula for A_n.

Solution. The first 3 terms of the sequence $(A_n)_n$ are:

$$\begin{aligned}
A_0 &= 1000\\
A_1 &= 1000 \cdot 1.05 + 100\\
A_2 &= A_1 \cdot 1.05 + 100 = (1000 \cdot 1.05 + 100) \cdot 1.05 + 100\\
&= 1000 \cdot 1.05^2 + 100 \cdot 1.05 + 100\\
A_3 &= A_2 \cdot 1.05 + 100\\
&= (1000 \cdot 1.05^2 + 100 \cdot 1.05 + 100) \cdot 1.05 + 100\\
&= 1000 \cdot 1.05^3 + 100 \cdot 1.05^2 + 100 \cdot 1.05 + 100
\end{aligned}$$

For the general case we get:

$$\begin{aligned}
A_n &= 1000 \cdot 1.05^n + 100 \cdot 1.05^{n-1} + \cdots + 100 \cdot 1.05 + 100\\
&= 1000 \cdot 1.05^n + 100 \cdot \sum_{i=0}^{n-1} 1.05^i
\end{aligned}$$

Now we apply one of the summation formulae from Section Section 2.4 to simplify our formula for A_n:

$$\begin{aligned}
A_n &= 1000 \cdot 1.05^n + 100 \cdot \frac{1.05^n - 1}{1.05 - 1}\\
&= 1000 \cdot 1.05^n + 100 \cdot \frac{1.05^n - 1}{0.05}\\
&= 1000 \cdot 1.05^n + 2000 \cdot (1.05^n - 1)\\
&= 3000 \cdot 1.05^n - 2000
\end{aligned}$$

□

Before we prove the next formulas, we consider a special case.

Example 2.4.5.11 Sum of natural numbers up to 100. Evaluate $\sum_{k=1}^{100} k$

Solution. Let $S = \sum_{k=1}^{100} k$. We write the summands of S twice in opposing

order and then add the two sums summand by summand:

$$\begin{aligned} S &= \quad 1 + \quad 2 + \quad 3 + \cdots + \quad 98 + \quad 99 + 100 \\ S &= 100 + \quad 99 + \quad 98 + \cdots + \quad 3 + \quad 2 + \quad 1 \\ 2S &= \underbrace{101 + 101 + 101 + \cdots + 101 + 101 + 101}_{100 \text{ copies of } 101} = 101 \cdot 100 \end{aligned}$$

Division by 2 yields

$$\sum_{k=1}^{100} k = S = \frac{101 \cdot 100}{2} = 5050.$$

□

Theorem 2.4.5.12 Summation formulae.

1. $\sum_{k=1}^{n} 1 = n$
2. $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$
3. $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$
4. $\sum_{k=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}$.

Proof. We prove the first two formulas here. We reprove the second and prove the third and fourth in Section 3.1 on mathematical induction.

1. We have
$$\sum_{k=1}^{n} 1 = \underbrace{1 + 1 + \cdots + 1}_{n \text{ copies of } 1} = n.$$

2. Let $S = \sum_{k=1}^{n} k$. We write the summands of S twice in opposing order and then add the two sums summand by summand:
$$\begin{aligned} S &= \quad 1 \; + \quad 2 \; + \cdots + (n-1) + \quad n \\ S &= \quad n \; + (n-1) + \cdots + \quad 2 \; + \quad 1 \\ 2S &= \underbrace{(n+1) + (n+1) + \cdots + (n+1) + (n+1)}_{n \text{ copies of } (n+1)} = n(n+1) \end{aligned}$$
Division by 2 yields
$$\sum_{k=1}^{n} k = S = \frac{n(n+1)}{2}.$$

■

Example 2.4.5.13 Sums from 0 to 100. With Theorem 2.4.5.12 we compute the missing sums

$$\sum_{k=0}^{100} 1 \text{ and } \sum_{k=0}^{100} k \text{ and } \sum_{k=0}^{100} k^2$$

from Example 2.4.5.7.

Solution.

1. First, we compute $\sum_{k=0}^{100} 1$. We break off piece that starts at $k = 1$ so we can use Theorem 2.4.5.12, Finally,

$$\sum_{k=0}^{100} 1 = 1 + \sum_{k=1}^{100} 1,$$

so by Theorem 2.4.5.12,

$$\sum_{k=0}^{100} 1 = 1 + 100 = 101.$$

2. Similarly, we write

$$\sum_{k=0}^{100} k = 0 + \sum_{k=1}^{100} k.$$

We use Theorem 2.4.5.12 with $n = 100$ to get

$$\sum_{k=0}^{100} k = 0 + \frac{100(100+1)}{2} = 5050.$$

3. Finally, we get

$$\sum_{k=0}^{100} k^2 = 0^2 + \sum_{k=1}^{100} k^2.$$

Since $0^2 = 0$, this is just $\sum_{k=1}^{100} k^2$, which matches the formula in the theorem using $n = 100$. Thus

$$\sum_{k=0}^{100} k^2 = 0 + \frac{100(100+1)(2(100)+1)}{6} = 338350.$$

□

Example 2.4.5.14 Evaluation of a sum.

$$\begin{aligned}\sum_{k=1}^{500}(7k+3) &= 7\sum_{k=1}^{500} k + \sum_{k=1}^{500} 3 \\ &= 7 \cdot \frac{500(501)}{2} + 3(500) \\ &= 878250.\end{aligned}$$

□

Example 2.4.5.15 Evaluation of a sum. Evaluate $\sum_{k=0}^{100} (4 \cdot k + 7)$.

Solution. We have

$$\begin{aligned}\sum_{k=0}^{100} (4 \cdot k + 7) &= \sum_{k=0}^{100} (4 \cdot k) + \sum_{k=0}^{100} 7 \\ &= \sum_{k=0}^{100} (4 \cdot k) + 707\end{aligned}$$

$$
\begin{aligned}
&= 4 \cdot \left(\sum_{k=1}^{100} k \right) + 707 \\
&= 4 \cdot \left(\sum_{k=0}^{100} k \right) + 707 \\
&= 4 \cdot \left(\frac{100 \cdot (100+1)}{2} \right) + 707 \\
&= 2 \cdot 100 \cdot (100+1) + 707 \\
&= 2 \cdot 10100 + 707 = 20200 + 707 = 20907
\end{aligned}
$$

□

2.4.6 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) sequence

(b) geometric progression, common ratio, initial term

(c) arithmetic progression, common difference, initial term

(d) recurrence relation

(e) solution of a recurrence relation

(f) Fibonacci sequence

(g) sum

2. Compute the initial term and common ratio of the following geometric progression.

$$16, 8, 4, 2, 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$$

3. Compute $\sum_{i=1}^{5} i$. Simplify.

4. Compute $\sum_{i=1}^{100} (3i + 7)$. Simplify.

5. Compute $\sum_{k=50}^{100} k^2$. Simplify.

6. Compute $\sum_{k=0}^{10} 3 \cdot 2^k$. Simplify.

7. What are the terms a_1, a_2, a_3, and a_4 of each of these sequences $(a_n)_n$?

(a) $a_n = 1$

(b) $a_n = \left\lfloor \frac{n}{2} \right\rfloor$

(c) $a_n = (-1)^n$

(d) $a_n = 2(-3)^{n+1} + 4$

(e) $a_n = (-\frac{1}{2})^n$

8. What are the terms a_1, a_2, a_3, and a_4 of each of these sequences $(a_n)_n$?

(a) $a_n = 1 + (-1)^n$

(b) $a_n = -(-3)^n$

(c) $a_n = 3n + 4$

(d) $a_n = 3^n - 2^n$

(e) $a_n = \sqrt{\lfloor n \rfloor}$

9. Show that each of these sequences $(a_n)_n$ is a solution of the recurrence relation

$$a_n = -3a_{n-1} + 4a_{n-2}.$$

(a) $a_n = 0$

(b) $a_n = 1$

(c) $a_n = (-4)^n$

(d) $a_n = 2(-4)^n + 3$

10. Suppose you are hired at company in this year with a starting salary of \$50000. Every year, you receive a 5% raise plus an additional \$1000.

(a) Set up a recurrence relation for your salary after n years.

(b) What is your salary in 10 years?

(c) Find an explicit formula for your salary in n years.

Chapter 3

Induction

Many times, a universal quantification can be reinterpreted as a nice family of propositions, indexed by positive integers. For example, consider the statement "the sum of the first n positive integers is $\frac{n(n+1)}{2}$." This can be rephrased as the universally quantified statement "for every positive integer n, $1+2+\cdots+n = \frac{n(n+1)}{2}$." This can be seen as a family of infinitely many simpler propositions:

$$\begin{aligned} 1 &= 1 \\ 1+2 &= \frac{2(2+1)}{2} \\ 1+2+3 &= \frac{3(3+1)}{2} \\ &\vdots \end{aligned}$$

Mathematical induction is a powerful proof technique that allows us to prove certain universally quantified statements by examining closely the related family of more basic propositions. Despite its name, mathematical induction is an example of deductive, not inductive, reasoning.

Outcomes

1. Reproduce the principle of mathematical induction
2. Prove theorems using mathematical induction.

3.1 Mathematical induction

3.1.1 Propositional functions revisited

We often use propositional functions in our induction proofs. Recall that a propositional function is a function that outputs a proposition for each input. It is important to distinguish the parts of the proposition that is represented by a propositional functions and the values of a propositional function at the elements of its domain. The value of a propositional function evaluated at an element of its domain is always true or false.

Example 3.1.1.1 Propositional function. Let $P(n)$ be the proposition "$\sum_{i=1}^{n} i = \frac{n\cdot(n+1)}{2}$" where $n \in \mathbb{N}$.

1. $P(2)$ is $\sum_{i=1}^{2} i = \frac{2\cdot(2+1)}{2}$. We have that $P(2)$ is true as both sides of the

equality evaluate to 3.

2. `$P(2) = \sum_{i=1}^{2} i$' and `$P(2) = \frac{2\cdot(2+1)}{2}$' do not make sense because the right hand sides both evaluate to 3,

□

Mathematical induction can be used to prove statements that assert $P(n)$ is true for all positive integers n.

Example 3.1.1.2 Inductive reasoning. Let P be a propositional function with domain $\mathbb{N}$. We are told that

1. $P(1)$ is true.
2. if $k \in \mathbb{N}$ and $P(k)$ is true then $P(k+1)$ is true.

For which natural numbers n do you know that $P(n)$ is true ? We have

- $P(1)$ is true by (Item 1).
- Because $P(1)$ is true also $P(1+1) = P(2)$ is true by (Item 2).
- Because $P(2)$ is true also $P(2+1) = P(3)$ is true by (Item 2).
- Because $P(3)$ is true also $P(3+1) = P(4)$ is true by (Item 2).

Continuing this pattern we find that by (Item 2) $P(k)$ is true for all following natural numbers. That is, $P(n)$ is true for all $n \in \mathbb{N}$. □

Example 3.1.1.3 Inductive reasoning. Let $Q(n)$ be a propositional function with domain $\mathbb{N}$. Suppose

1. $Q(0)$ is true.
2. Let $k \in \mathbb{N}$. If $Q(k)$ is true then $Q(k+2)$ is true.

For which $n \in \mathbb{N}$ do we know that $Q(n)$ is true ? We have

- $Q(0)$ is true by (Item 1).
- Because $Q(0)$ is true also $Q(0+2) = Q(2)$ is true by (Item 2).
- Because $Q(2)$ is true also $Q(2+2) = Q(4)$ is true by (Item 2).
- Because $Q(4)$ is true also $Q(4+2) = Q(6)$ is true by (Item 2).

Continuing this pattern we find that by (Item 2) $Q(8)$ is true, $Q(10)$ is true, $Q(12)$ is true, and so on. So, with the information given we can conclude that $Q(n)$ is also true for all following even natural numbers. Thus $Q(n)$ is true for all $n \in \{0, 2, 4, 6, 8, 10, 12, \dots\}$. □

3.1.2 The principle of mathematical induction

In its most basic form, mathematical induction has two parts:

1. Prove $P(1)$ is true.
2. Prove for all positive integers k, if $P(k)$ is true, then

Then symbolically, mathematical induction says

$$(P(1) \wedge \forall k(P(k) \to P(k+1))) \to \forall n P(n).$$

A useful analogy to keep in mind is stacking dominos to topple. How can we prove that the n th domino topples for all n? First, we need to know that the first domino falls. A bit of thought shows that this is not enough. We need to additionally know that the dominos are stacked expertly. Namely, they are stacked in such a way that if the k th domino falls, then the $(k+1)$ st domino falls.

Theorem 3.1.2.1 Principle of Mathematical Induction. *Let* $\{P(n)\colon n \geq n_0\}$ *be a family of propositions such that*

1. $P(n_0)$ *is true.*
2. $P(k)$ *implies* $P(k+1)$*, for* $k \geq n_0$*.*

Then $P(n)$ *is true for all* $n \geq n_0$*.*

Follow these steps to write a proof by induction.

Introduction State that we are using induction. To be extra clear, tell the reader on what we are doing induction. A common way of saying this is "We proof the theorem by induction on n".
Define $P(n)$.

Basis step Prove $P(n_0)$ is true.

Inductive step For fixed (generic) $k \geq n_0$, assume $P(k)$ is true. This is known as assuming the **inductive hypothesis**.
Prove $P(k+1)$ is true. In this step you apply that $P(k)$ is true.

Conclusion State that by the principle of mathematical induction you have proven that $P(n)$ is true for all $n \in \{n_0, n_0+1, n_0+2, \dots\}$.

We proof Theorem the following theorem by induction.

Theorem 3.1.2.2 *Let* $n \in \mathbb{N}$ *then* $\displaystyle\sum_{i=1}^{n} i = \frac{n \cdot (n+1)}{2}$.

Proof.

Introduction We proof the theorem by induction on n.
Let $P(n)$ be "$\sum_{i=1}^{n} i = \frac{n\cdot(n+1)}{2}$".

Basis step We prove that $P(1)$ is true by evaluating both sides of the equation at $n = 1$ and concluding that they are equal.
For the left hand side of $P(1)$ we have

$$\sum_{i=1}^{1} i = 1$$

and for the right hand side of $P(1)$ we have

$$\frac{1 \cdot (1+1)}{2} = \frac{2}{2} = 1.$$

Thus $P(1)$ is true.

Inductive step Suppose that for a fixed $k \geq 1$ we have $P(k)$ is true, that is "$\sum_{i=1}^{k} i = \frac{k\cdot(k+1)}{2}$" is true.

We now want to prove that $P(k+1)$ is true, that is that "$\sum_{i=1}^{k+1} i = \frac{(k+1)\cdot(k+2)}{2}$" is true. With the induction hypothesis that $P(k)$ is true we get

$$\begin{aligned}\sum_{i=1}^{k+1} i &= \left(\sum_{i=1}^{k} i\right) + (k+1)\\ &= \frac{k\cdot(k+1)}{2} + (k+1)\\ &= \frac{k\cdot(k+1)}{2} + \frac{2k+2}{2}\\ &= \frac{k\cdot(k+1)+(2k+2)}{2}\\ &= \frac{k^2+3k+2}{2}\\ &= \frac{(k+1)\cdot(k+2)}{2}\end{aligned}$$

Thus $P(k+1)$ is true.

Conclusion By the principle of mathematical induction we have proven that $P(n)$ is true for all $n \in \mathbb{N}$.

■

For the next example of a proof by induction we need the notion of divisibility by 5.

Definition 3.1.2.3 Let $m \in \mathbb{N} \cup \{0\}$ and $d \in \mathbb{N}$. We say m is **divisible** by d when there exists $s \in \mathbb{N} \cup \{0\}$ such that $m = d \cdot s$. ♢

Theorem 3.1.2.4 *Let $n \in \mathbb{N} \cup \{0\}$. Then $n^5 - n$ is divisible by 5.*

Proof. ***Introduction.*** We proof the theorem by induction on n.

Let $P(n)$ be "$n^5 - n$ is divisible by 5" with domain $\mathbb{N} \cup \{0\}$.

Basis step. We prove that $P(0)$ is true.

For $n = 0$ we have $0^5 - 0 = 0$. Because $0 = 5 \cdot 0$ and because $0 \in \mathbb{N} \cup \{0\}$ we have that 5 divides 0, which proves that $P(0)$ is true.

Inductive step For fixed $k \geq 0$, assume $P(k)$ is true, that is "$k^5 - k$ is divisible by 5" is true.

We now want to show that $P(k+1)$ is true, that is "$(k+1)^5 - (k+1)$ is divisible by 5" is true. We get

$$\begin{aligned}(k+1)^5-(k+1) &= k^5+5\cdot k^4+10\cdot k^3+10\cdot k^2+5\cdot k+1-k-1\\ &= k^5+5\cdot(k^4+2\cdot k^3+2\cdot k^2+k)-k\\ &= k^5-k+5\cdot(k^4+2\cdot k^3+2\cdot k^2+k)\end{aligned}$$

Because of the assumption that $P(k)$ is true, there exists $s \in \mathbb{N} \cup \{0\}$ such that $k^5 - k = 5 \cdot s$. With this we get

$$\begin{aligned}(k+1)^5-(k+1) &= k^5-k+5\cdot(k^4+2\cdot k^3+2\cdot k^2+k)\\ &= 5\cdot s+5\cdot(k^4+2\cdot k^3+2\cdot k^2+k)\\ &= 5\cdot(s+k^4+2\cdot k^3+2\cdot k^2+k).\end{aligned}$$

Thus $(k+1)^5 - (k+1)$ is divisible by 5 which means that $P(k+1)$ is true.

Conclusion. By the principle of mathematical induction we have proven that $P(n)$ is true for all $n \in \mathbb{N} \cup \{0\}$. ■

Definition 3.1.2.5 Let the sequence $\{f_n\}$ be given by the recurrence relation $f_n = n \cdot f_{n-1}$ and the initial value $f_0 = 1$. We call f_n the factorial of n and denote it by $n!$. ◊

Remark 3.1.2.6 For $n \in \mathbb{N}$ we have $n! = \prod_{i=1}^{n} i = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-1) \cdot n$.

Theorem 3.1.2.7 *For all $n \in \mathbb{N}$ we have* $\sum_{i=1}^{n} (i \cdot (i!)) = (n+1)! - 1$.

Proof.

Introduction We proof the theorem by induction on n.

Let $P(n)$ be "$\sum_{i=1}^{n} (i \cdot (i!)) = (n+1)! - 1$."

Basis step We prove that $P(1)$ is true.

For the left hand side of $P(n)$ with $n = 1$ we have

$$\sum_{i=1}^{1} (i \cdot (i!)) = (1 \cdot (1!)) = 1 \cdot 1 = 1$$

For the right hand side we have

$$(1+1)! - 1 = 2! - 1 = 2 - 1 = 1$$

As the left hand side and the right hand side are equal $P(1)$ is true.

Inductive step For fixed $k \geq 0$, assume $P(k)$ is true, that is "$\sum_{i=1}^{k} (i \cdot (i!)) = (k+1)! - 1$" is true.

We now want to show that $P(k+1)$ is true, that is

$$\sum_{i=1}^{k+1} (i \cdot (i!)) = (k+2)! - 1.$$

We have

$$\sum_{i=1}^{k+1} (i \cdot (i!)) = \sum_{i=1}^{k} (i \cdot (i!)) + (k+1) \cdot (k+1)!$$

With the assumption that $P(k)$ is true we obtain

$$\begin{aligned} \sum_{i=1}^{k+1} (i \cdot (i!)) &= \sum_{i=1}^{k} (i \cdot (i!)) + (k+1) \cdot (k+1)! \\ &= (k+1)! - 1 + (k+1) \cdot (k+1)! \\ &= (1+k+1) \cdot (k+1)! - 1 \\ &= (k+2) \cdot (k+1)! - 1 \\ &= (k+2)! - 1 \end{aligned}$$

Thus if $P(k)$ is true then $P(k+1)$ is true.

Conclusion By the principle of mathematical induction we have proven that $P(n)$ is true for all $n \in \mathbb{N}$.

■

Now that you have seen a couple of example, we only label the basis and the inductive steps.

Theorem 3.1.2.8 *Let $n \in \mathbb{N} \cup \{0\}$. Then $\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$.*

Proof. We proof the theorem by induction on n. Let $P(n)$ be "$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$".

Basis step. We prove that $P(0)$ is true. For the left hand side of $P(0)$ we have

$$\sum_{i=0}^{0} 2^i = 2^0 = 1$$

and fro the right hand side of $P(1)$ we have

$$2^{0+1} - 1 = 2^1 - 1 = 2 - 1 = 1.$$

Thus $P(0)$ is true.

Inductive step. We assume $P(k)$ is true for a fixed $k \geq 0$, that is "$\sum_{i=0}^{k} 2^i = 2^{k+1} - 1$" is true. We now want to show that $P(k+1)$ is true, that is we want to show that "$\sum_{i=0}^{k+1} 2^i = 2^{k+2} - 1$" is true. We have

$$\begin{aligned}
\sum_{i=0}^{k+1} 2^i &= \left(\sum_{i=0}^{k} 2^i\right) + 2^{k+1} \\
&= \left(2^{k+1} - 1\right) + 2^{k+1} \text{ because we assume } P(k) \text{ is true} \\
&= 2 \cdot 2^{k+1} - 1 \\
&= 2^{k+2} - 1
\end{aligned}$$

Thus $P(k+1)$ is true.

By the principle of mathematical induction we have proven that $P(n)$ is true for all $n \in \mathbb{N} \cup \{0\}$. ■

Mathematical induction can be used to prove statements other than formulas.

Example 3.1.2.9 Tower of Hanoi. The Tower of Hanoi is a puzzle game consisting of three rods, and a number of disks of different sizes which can slide onto any rod. The puzzle starts with the disks in a neat stack in ascending order of size on one rod, the smallest at the top, thus making a conical shape.

The objective of the puzzle is to move the entire stack to another rod, obeying the following simple rules:

1. Only one disk can be moved at a time.
2. Each move consists of taking the upper disk from one of the stacks and placing it on top of another stack that is , a disk can only be moved if it is the uppermost disk on a stack.
3. No disk may be placed on top of a smaller disk.

See the *Wikipedia* article for more details (https://en.wikipedia.org/wiki/

Tower_of_Hanoi)

Answer. Let $P(n)$ be the proposition "$m_n = 2^n - 1$."

We proceed by induction on n.

Basis step. It is clear that $m_1 = 1$. Thus $P(1)$ is true.

Inductive step. Fix $k \geq 1$. Assume $P(k)$ is true so that $m_k = 2^k + 1$. Then

$$\begin{aligned} m_{k+1} &= 2m_k + 1 \\ &= 2(2^k - 1) + 1 \text{ (by inductive hypothesis)} \\ &= 2^{k+1} - 1. \end{aligned}$$

Thus $P(k+1)$ is true.

Therefore, by mathematical induction, $m_n = 2^n - 1$ for all $n \geq 1$.

Solution. Let hanoi(n) denote the Tower of Hanoi game with n disks. Once we have a winning strategy for hanoi(k) for some positive integer k, we can get a winning strategy for hanoi($k+1$) as follows.

1. Use strategy to move the smallest k disks to an empty rod.
2. Move largest disk to the remaining empty rod.
3. Use strategy to move the smallest k disks on top of the

If there are m_k moves in the hanoi(k) winning strategy, then there are $m_k + 1 + m_k = 2m_k + 1$ moves in the hanoi($k+1$) strategy described above. It is clear that $m_1 = 1$. Then its easy to see that $m_2 = 3$, $m_3 = 7$, $m_4 = 15$, and so on. It looks like $m_k = 2^k - 1$. Let's prove it by induction. □

Remark 3.1.2.10 A proof almost identical to the one above can be used to prove that the winning strategy for hanoi(n) with the minimum number of moves has $2^n - 1$ moves.

Example 3.1.2.11 All people are the same age. Find the flaw in the following proof that any set with n people are all the same age.

We proceed by induction. Let $P(n)$ be the proposition "In any set of n people are all the same age."

We proceed by induction on n.

Basis step. It is clear that in any set with 1 person, they are the same age. Thus $P(1)$ is true.

Inductive step Fix $k \geq 1$. Assume $P(k)$ is true so that any set of k people are the same age. Let S be a set of $k+1$ people. We have to show that everyone in S is the same age. Pick one person X in S. Let $T = S - \{X\}$. Then $\#T = k$, so by the inductive hypothesis everyone in T is the same age, call it a. Pick another person X' in S, $X \neq X'$. Let $T' = S - \{X'\}$. Then $\#T' = k$, so by the inductive hypothesis everyone in T' is the same age, call it a'.

Notice that we have $a = a'$ since T and T' have some members in common. Furthermore, $T \cup T' = S$. Thus everyone in S is the same age.

Therefore by mathematical induction, any set of n people are all the same age.

Solution. What is the error in the "proof"? When we choose $X' \neq X$, we are assuming $k \geq 2$. Furthermore, when we assert T and T' have members in common, we assume $k \geq 3$. This shows that we need to be really careful that the basis step is correctly chosen so that the inductive step can work. □

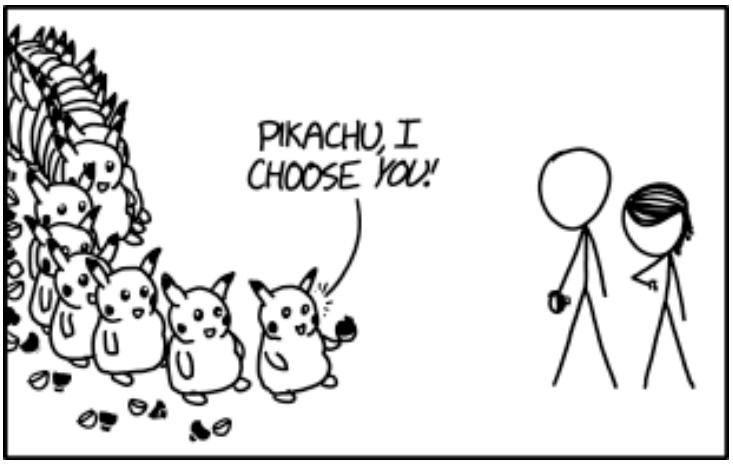


Figure 3.1.2.12 *Win By Induction*? by Randall Munroe (https://xkcd.com/1516/). This would be bad enough, but every 30th or 40th pokéball has TWO of them inside.

3.1.3 Exercises

1. State precisely the Principle of Mathematical Induction. Be sure to set up any notation that is required.

2. Complete the steps below to prove that for every positive integer n,

$$1 \cdot 2 + 2 \cdot 3 + \cdots + n(n+1) = \frac{n(n+1)(n+2)}{3}.$$

I have written in quotes part of what you need to write. You should complete the proof.

Introduction What is the proposition $P(n)$? Write:

$P(n)$...

Basis step What is the statement $P(1)$? Write:

$P(1) = \ldots$

Prove $P(1)$ is true.

Inductive step State the inductive hypothesis:

$k \geq 1$...

What is it we want to prove, assuming the inductive hypothesis? Write:

...

Prove it. Be clear where you are using the inductive hypothesis.
Write:

Conclusion State the conclusion:

...

3. For each positive integer n, let $P(n)$ be the proposition

$$2^0 + 2^1 + 2^2 + \cdots + 2^n = 2^{n+1} - 1.$$

Prove $P(n)$ is true for all positive integers n by induction.

4. Prove this extension of De Morgan's law. Let $p_1, p_2, \cdots, p_n$ be propositions. Prove that

$$\neg(p_1 \vee p_2 \vee \cdots \vee p_n) \equiv \neg p_1 \wedge \neg p_2 \wedge \cdots \wedge \neg p_n.$$

5. Find the flaw in the following proof that $n^2 + 5n + 1$ is even for all $n \geq 0$.

Let $P(n)$ be the proposition "$n^2 + 5n + 1$ is even." Fix $k \geq 0$. Assume $P(k)$ is true so that k^2+5k+1 is even. We want to show $(k+1)^2+5(k+1)+1$ is even. We compute

$$\begin{aligned}(k+1)^2 + 5(k+1) + 1 &= k^2 + 2k + 1 + 5k + 5 + 1 \\ &= (k^2 + 5k + 1) + 2(k+3).\end{aligned}$$

Since $k^2 + 5k + 1$ is even by inductive hypothesis and $2(k + 3)$ is visibly even, $(k + 1)^2 + 5(k + 1) + 1$ is even.

Thus by induction, $n^2 + 5n + 1$ is even for all $n \geq 0$.

6. Find the flaw in the following proof that $2^n \leq n + 1$ for all $n \geq 1$.

Let $P(n)$ be the proposition "$2^n \leq n + 1$".

We proceed by induction on n.

Base case: $2^1 = 2$ and $1 + 1 = 2$, so $2^1 \leq 1 + 1$.

Inductive step: Fix $k \geq 1$. Assume $P(k)$ is true. that is , assume $2^k \leq k + 1$.

$$\begin{aligned}2^{k+1} &\leq (k+1) + 1 \\ 2 \cdot 2^k &\leq k + 2 \\ 2^k &\leq \frac{k}{2} + 1 \leq k + 1,\end{aligned}$$

which is true by inductive hypothesis. Thus by induction, $2^n \leq n + 1$ for all $n \geq 1$.

7. Fix real numbers a and r, with $r \neq 1$. Prove that for every positive integer n,

$$a + ar^1 + ar^2 + \cdots + ar^n = \frac{ar^{n+1} - a}{r - 1}.$$

8. If you have had calculus, prove the power rule for positive exponents. Specifically, prove that for every positive integer n,

$$\frac{d}{dx}(x^n) = nx^{n-1}.$$

(Hint: Use induction on n and the Product rule, writing $x^n = x \cdot x^{n-1}$.)

9. Prove that for every positive integer n,

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

10. Prove that for every positive integer n,

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

11. Prove for every positive integer n,

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

12. Prove that for every positive integer n,

$$1 \cdot 1! + 2 \cdot 2! + \cdots + n \cdot n! = (n+1)! - 1.$$

13. Prove that $n < 2^n$, for every positive integer n.

14. Prove that $11^n - 4^n$ is divisible by 7, for all $n \geq 0$. Redo the proof using congruences instead of induction.

15. Prove that for every positive integer n,

$$\frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n \cdot (n+1)} = \frac{n}{n+1}.$$

3.2 Strong induction

Outcomes

1. Reproduce the principle of mathematical induction.
2. Reproduce the well-ordering property.
3. Prove theorems using strong induction and the well-ordering property.

3.2.1 Principle of strong induction

Strong induction is another form of mathematical induction that can often be used when we cannot easily use mathematical induction. For **strong induction** , replace the inductive step in the Principle of Mathematical Induction by a stronger assumption. In its most basic form, there are two parts:

1. Prove $P(1)$ is true.
2. Prove that for all positive integers k, if $P(j)$ is true for all $1 \leq j \leq k$, then $P(k+1)$ is true.

Then symbolically, mathematical induction says

$$(P(1) \wedge \forall k((P(1) \wedge P(2) \wedge \cdots \wedge P(k)) \rightarrow P(k+1))) \rightarrow \forall n P(n).$$

Theorem 3.2.1.1 Principle of Strong Induction. *Let $\{P(n)\colon n \geq n_0\}$ be a family of propositions such that*

1. *$P(n_0)$ is true.*
2. *$P(n_0) \wedge P(n_0+1) \wedge \cdots \wedge P(k)$ implies $P(k+1)$, for $k \geq n_0$.*

Then $P(n)$ is true for all $n \geq n_0$.

For the inductive step, we need to show "$P(n_0) \wedge P(n_0+1) \wedge \cdots \wedge P(k)$ implies $P(k+1)$." For fixed (generic) $k \geq n_0$, assume $P(j)$ is true for all $n_0 \leq j \leq k$. Prove $P(k+1)$ is true.

Example 3.2.1.2 Natural numbers are products of primes. Prove that every positive integer greater than 1 can be written as a product of primes.

Proof. Let $P(n)$ be the proposition "n can be written as a product of primes." We want to show $P(n)$ is true for all $n \geq 2$. We proceed by induction on n.

Basis step: $P(2)$ is true since 2 is prime.

Inductive step: Fix $k \geq 2$. Suppose $P(j)$ is true for all $0 \leq j \leq k$ so that j can be written as a product of primes when $0 \leq j \leq k$. We want to show $k+1$ can be written as a product of primes. There are two cases to consider.

$k+1$ **is prime**	In this case, $P(k+1)$ is trivially true.
$k+1$ **is composite**	In this case, there exist integers a and

Therefore by strong induction, every integer greater than or equal to 2 can be written as a product of primes. ■

□

When we use the inductive hypothesis, it is important to verify that we are in the range where we can use it. See the example below.

Example 3.2.1.3 Find the flaw. Find the flaw in the following proof that $3^n = 1$ for every nonnegative integer n.

Let $P(n)$ be the proposition $3^n = 1$.

Basis step: $P(0)$ is true since $3^0 = 1$.

Inductive step: (Uses strong induction.) Fix $k \geq 0$. Assume $P(j)$ is true for all $0 \leq j \leq k$. In particular, we assume that $3^j = 1$ for all $0 \leq j \leq k$. We wish to show that $3^{k+1} = 1$. We compute

$$3^{k+1} = 3^k \cdot 3^1 = \frac{3^k \cdot 3^k}{3^{k-1}} = \frac{1 \cdot 1}{1} = 1,$$

since by the inductive hypothesis, $3^k = 1$ and $3^{k-1} = 1$. Thus, by induction, $3^n = 1$ for all $n \geq 0$. □

Example 3.2.1.4 Paying with 3 Ruble and 5 Ruble banknotes. Every amount of 8 Rubles or more can be paid using 3 Ruble and 5 Ruble banknotes.

Solution. Let $R(n)$ be the proposition "an amount of n Rubles can be paid with 3 Ruble and 5 Ruble banknotes."

We want to prove $R(n)$ is true for $n \geq 8$. We proceed by induction on n.

Basis steps:

$$8 = 1 \cdot 3 + 1 \cdot 5$$
$$9 = 3 \cdot 3 + 0 \cdot 5$$
$$10 = 0 \cdot 3 + 2 \cdot 5$$

Inductive step: Fix $k \geq 10$. Suppose $P(j)$ is true for all $j \in \mathbb{N}$ with $8 \leq j \leq k$. That is, j Rubles can be formed using 3 Ruble and 5 Ruble bank notes. We want to form $k+1$ Rubles.

Let $j = (k+1) - 3 = k - 2$. Then $j \leq k$. Since $k \geq 10$, we have $j = k - 2 \geq 10 - 2 = 8$. By the inductive hypothesis, we can form j Rubles using 3 Ruble and 5 Ruble banknotes. Just add another 3 Ruble banknote to get $j + 3 = k - 2 + 3 = k + 1$ Rubles. Thus $R(k+1)$ is true.

Therefore by strong induction, $R(n)$ is true for all $n \geq 8$. □

Example 3.2.1.5 Fibonacci sequence bound. Recall the Fibonacci sequence is defined by $f_0 = 0$, $f_1 = 1$, and

$$f_{n+1} = f_n + f_{n-1}, \quad \text{for } n \geq 1.$$

Prove that $f_n \leq 2^n$ for all $n \geq 0$.

Proof. Let $P(n)$ be the proposition "$f_n \leq 2^n$." We proceed by induction on n.

Basis step: It is clear that $f_0 = 0 \leq 2^0$ and $f_1 = 1 \leq 2^1$, so $P(0)$ and $P(1)$ are true.

Inductive step: Fix $k \geq 1$. Assume $P(j)$ is true for $0 \leq j \leq k$ so that $f_j \leq 2^j$ for $0 \leq j \leq k$. We want to prove $f_{k+1} \leq 2^{k+1}$. We have

$$\begin{aligned} f_{k+1} &= f_k + f_{k-1} \\ &\leq 2^k + 2^{k-1} \\ &\leq 2 \cdot 2^k \\ &= 2^{k+1}. \end{aligned}$$

Thus $P(k+1)$ is true.

Therefore by strong induction, $f_n \leq 2^n$ for all $n \geq 0$. ■

□

3.2.2 Well-ordering Property

The validity of mathematical induction and strong induction follows from a fundamental result about the integers known as the **well-ordering property** . Note that while we list the well-ordering property below as a theorem, it is in fact an axiom (self-evident truth) that we assume in the precise construction of the set of integers.

Axiom 3.2.2.1 Well-Ordering Property. *Every nonempty set of nonnegative integers has a least element.*

Why does the well-ordering property imply that induction is valid ?

Proof of Theorem 3.1.2.1. For simplicity, just think about the most basic form, as described before Theorem Theorem 3.1.2.1. We proceed by contradiction.

Suppose the principle of mathematical induction is false. Then the set S of positive integers for which $P(n)$ is false is nonempty.

By the well-ordering property, S must have a least element. Let m be the least element in S. Then $P(m)$ is false.

Because $P(1)$ is true by the basis step, we have $m \neq 1$. Because m is positive and not 1, we must have that m is at least 2. Then $m-1$ is a positive integer less than m, and $P(m-1)$ is true. By the inductive step, we know that $P(m-1)$ is true implies $P(m)$ is true. But $P(m)$ is true contradicts the proposition that $P(m)$ is true which we followed from the existence of the nonempty set S. Thus such a set S cannot exists which implies that induction is valid.

The proof the more general case as well as strong-induction is similar. ■

Remark 3.2.2.2 This well-ordering property is not true for subsets of the real numbers, so we cannot use these induction techniques to prove things like $P(x)$ is true for all real positive numbers x.

Figure 3.2.2.3 *Set Theory* by Randall Munroe (https://xkcd.com/982). Proof of Zermelo's well-ordering theorem given the Axiom of Choice: 1: Take S to be any set. 2: When I reach step three, if S hasn't managed to find a well-ordering relation for itself, I'll feed it into this wood chipper. 3: Hey, look, S is well-ordered.

3.2.3 Exercises

1. State precisely the following. Be sure to set up any notation that is required.

(a) The Principle of Strong Induction.

(b) The Well-Ordering Property.

2. Complete the steps below to prove that every amount of postage of 12 cents or more can be formed using 4-cent and 5-cent stamps.

(a) What is the proposition $P(n)$? "Let $P(n)$ be the proposition …."

(b) "basis steps:"

i. What is the statement $P(12)$?

ii. Prove $P(12)$ is true.

iii. In this case, the proof is simpler if we use strong induction. State and prove $P(13)$, $P(14)$, and $P(15)$.

(c) "Inductive step:"

i. State the inductive hypothesis: "Fix $k \geq \ldots$. Assume … for all $\cdots \leq j \leq k$."
Be careful here. We gain the additional advantage of additional cases in the basis step by forcing k to be larger, and allowing j to go down as much as possible.

ii. What is it we want to prove, assuming the inductive hypothesis? "We want to show …."

iii. Prove it. Be clear where we are using the inductive hypothesis. Verify that we are working in a range where the inductive hypothesis is assumed to hold.

iv. Write "This completes the inductive step."

(d) State the conclusion: "By mathematical induction,..."

3. Suppose $a_1 = 10$, $a_2 = 5$, and $a_n = 2a_{n-1} + 3a_{n-2}$ for $n \geq 3$. Use strong induction to prove that 5 divides a_n for $n \geq 1$.

4. Let n be a positive integer. Show that any $2^n \times 2^n$ chess board with 1 square removed can be tiled using L-shaped pieces, where these pieces cover three squares at a time, as shown below.

5. Let f_n denote the n-th Fibonacci number. Prove that

$$f_n = \frac{\phi^n - (1-\phi)^n}{\sqrt{5}} \text{ where } \phi = \frac{1+\sqrt{5}}{2}.$$

Before starting the proof by induction show:

(a) $\phi^2 - \phi - 1 = 0$

(b) $\phi^2 = \phi + 1$

(c) $(1-\phi)^2 = 2 - \phi$.

(d) $\phi(\phi + 1) = \phi^3$

6. Prove that the power set $\wp(S)$ of a finite set S has cardinality

$$|\wp(S)| = 2^{|S|}.$$

(Hint: Use induction on the size $|S|$ of S. Fix an element a in S, and count the subsets of S containing a and the subsets not containing a. Note: A direct proof of this is given in Theorem Theorem 2.1.4.4.)

7. Prove that if a sequence $\{a_n\}$ satisfies

$$a_{n+1} = \frac{a_n}{a_n + 1}$$

then

$$a_n = \frac{a_0}{na_0 + 1}.$$

Chapter 4

Counting

> *When angry count to ten before you speak. If very angry, count to one hundred.*
>
> —Thomas Jefferson (1743--1826)

We learn to count at a young age. It is something that is fundamental and basic. There is archaeological evidence suggesting humans have been counting for over 50000 years. In this chapter, we define precisely what it means to *count* something. Then we develop advanced counting techniques and look at some of the implications.

4.1 Basics of counting

Outcomes

1. Reproduce counting rules and to show how
2. Solve counting problems.

Combinatorics is the study of arrangements of objects. We want to count the number of ways to do certain things.

4.1.1 Cardinality

Definition 4.1.1.1 Two sets A and B have the same **cardinality**, denoted $\#A = \#B$ if there is a bijection from A to B. ◊

The above lets us state precisely what it means to **count** something.

Definition 4.1.1.2 Let A be a set. If there exists a natural number n such that there is a bijection $c\colon A \to \{1, 2, \ldots, n\}$ then we call A **finite**.

To **count** the elements of A means to construct such a bijection c. When we have such a bijection c then $n = \#A$ is the cardinality of A. ◊

Recall that the cardinality of a finite set is defined earlier in Definition 2.1.3.1. This definition of counting is exactly what we show children when we teach them to count to answer "how many" type questions, though not in this language. Specifically, we teach them to construct a bijection in order to find the size of a set. For example, imagine teaching a child to count the number of hearts (♡) shown below:

♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡

Most likely, you pointed to each heart, one-by-one, and recited: "one, two, three, four, \dots, thirteen, fourteen." In doing so, you created a bijection between the set of hearts and $\{1, 2, \dots, 14\}$. From this, you cam deduce that there are 14 hearts.

Notice that we do not need to have constructed the same bijection in order to get the right number for the cardinality. The key point is that we create a function that is both injective and surjective. Do you see why?

We can generalize the definition for counting and cardinality to include some infinite sets. An infinite set is any set that is does not have a finite cardinality. It turns out, there are different sizes of infinity.

Definition 4.1.1.3 A **countable** set is either finite or has the same cardinality as the set of positive integers. An **uncountable** set is any set that is not countable. ♢

Notice that this gives at least two different "sizes" for infinite sets. Some sets are infinite, but we can count them by constructing a bijection to the positive integers. Other infinite sets are so large that we cannot even construct such a bijection to the positive integers.

Example 4.1.1.4 The set of integers is countable. The set of integers $\mathbb{Z}$ is a countable set. To show this, we need to construct a bijection from $\mathbb{Z}$ to $\{1, 2, \dots\}$. Consider the function $f\colon \mathbb{Z} \to \{1, 2, \dots\}$ defined by

$$f(x) = \begin{cases} 2x + 1 & \text{if } x \geq 0, \\ -2x & \text{if } x < 0. \end{cases}$$

So $f(0) = 1$, $f(1) = 3$, $f(3) = 7$, \dots, and $f(-1) = 2$, $f(-2) = 4$, $f(-3) = 6$, \dots. In other words, f sends the nonnegative integers to the odd integers, and the negative integers are sent to the even integers.

Let's show that f is bijective. First, to see that f is surjective, fix a generic element n in $\{1, 2, \dots\}$. If n is even, then $-\frac{n}{2}$ is a negative integer, and so

$$f(-\frac{n}{2}) = -2(-\frac{n}{2}) = n.$$

If n is odd, then $\frac{n-1}{2}$ is a non-negative integer, and so

$$f(\frac{n-1}{2}) = 2(\frac{n-1}{2}) + 1 = n.$$

Thus f is surjective.

To see f is injective, fix generic integers a and $\hat{a}$ such that $f(a) = f(\hat{a})$. If $f(a)$ is even, then $f(\hat{a})$ is also even, and so both a and $\hat{a}$ are negative. Then $f(a) = -2a$, and $f(\hat{a}) = -2\hat{a}$. Since $f(a) = f(\hat{a})$, we have $-2a = -2\hat{a}$ so $a = \hat{a}$. Now suppose $f(a)$ is odd. Then $f(\hat{a})$ is also odd, and so both a and $\hat{a}$ are nonnegative. Then $f(a) = 2a + 1$, and $f(\hat{a}) = 2\hat{a} + 1$. Since $f(a) = f(\hat{a})$, we have $2a + 1 = 2\hat{a} + 1$ so $a = \hat{a}$. Thus f is injective.

Since f is injective and surjective, f is bijective and so $\mathbb{Z}$ is countable. □

Warning 4.1.1.5 Example 4.1.1.4 is a bit counterintuitive. It shows that $\mathbb{Z}$ can be put in bijection with a proper subset of itself. In particular,

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

has the same size as

$$\mathbb{N} = \{1, 2, 3, \dots\}.$$

This hints at some of the strange things that can happen with counting infinite

sets.

Since the inverse of a bijective function is bijective, we can show an infinite set is countable by constructing a bijection from the set of positive integers to the set. Such a bijection is a sequence that includes all of the elements of A, listed with no repeats. In other words, an infinite set is countable if and only if it is possible to list all of the elements in a sequence with no repeats. In fact, we can relax the "no repeats" part, since if we do have a sequence with repeats, we can easily extract a subsequence without repeats.

For example, if we are given the sequence

$$1, 2, 3, 4, 5, 5, 6, 7, 7, 8, 9, 9, 10, \dots$$

we instead consider the sequence

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \dots .$$

In Example 4.1.1.4 that shows the set of integers is countable, the sequence that is inverse the bijection given is

$$0, -1, 1, -2, 2, -3, 3, \dots .$$

Sometimes it is easier to think of a sequence instead of a formula for the function. Since the composition of bijective functions is bijective, to show an infinite set A is countable it is enough to find a bijection from A to $\mathbb{Z}$. If we have such a bijection, we compose it with the bijection from $\mathbb{Z}$ to $\{1, 2, \dots\}$ given in Example 4.1.1.4 to get the desired bijection from A to $\{1, 2, \dots\}$. Using this idea, it is clear that the set of even integers $2\mathbb{Z}$ is countable, since $f\colon 2\mathbb{Z} \to \mathbb{Z}$ given by $f(x) = \frac{x}{2}$ is a bijection from $2\mathbb{Z}$ to $\mathbb{Z}$. A similar argument shows the set of odd integers is countable.

What about the set of rational numbers? Surely there are more rational numbers than integers? It turns out, the set of rational numbers is also countable.

Theorem 4.1.1.6 *The set of rational numbers is countable.*

Proof. It suffices to show the set of positive rational numbers is countable. (Why? Exercise.)

Every positive rational number is the quotient $\frac{p}{q}$ of two positive integers p and q. We arrange the positive rational numbers by listing those with denominator $q = 1$ in the first row, those with denominator 2 in the second row, and so on. The rational numbers with denominator k are listed in the k th row, as shown in Figure 4.1.1.7. From the figure, we create a sequence by looking at the diagonals: list $\frac{p}{q}$ with $p + q = 2$, followed by those with $p + q = 3$, followed by with $p + q = 4$, and so on as shown in Figure 4.1.1.7. The numbers we keep for the sequence are circled. The uncircled numbers are those we leave out because they are already listed. Because all the positive rational numbers are listed once. we hove show the set of positive integers is countable. The exercise completes the proof by extending this bijection to a bijection from $\mathbb{Z}$ to $\mathbb{Q}$. ■

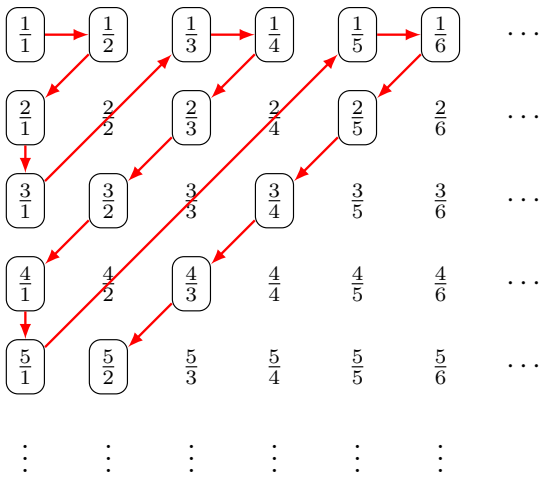


Figure 4.1.1.7 The positive rationals are countable.

One may start to think that every infinite set is countable. In fact, there are sets that are larger. In 1879, Georg Cantor produced a proof using a technique now known as the **Cantor diagonalization argument**, that proves the set of real numbers is not countable.

Theorem 4.1.1.8 *The set of real numbers is uncountable.*

Proof. It suffices to prove that the open interval $(0,1) = \{x : (0 < x) \wedge (x < 1)\}$ is uncountable. We prove this by contradiction. Suppose the interval $(0,1)$ is countable. Then there exists a bijective function \(c:\NN\to (0,1)\). Write the images of c as

$$\begin{aligned} c(1) &= 0.c_{1,1}c_{1,2}c_{1,3}c_{1,4}c_{1,5}\dots \\ c(2) &= 0.c_{2,1}c_{2,2}c_{2,3}c_{2,4}c_{2,5}\dots \\ c(3) &= 0.c_{3,1}c_{3,2}c_{3,3}c_{3,4}c_{3,5}\dots \\ &\vdots \quad \vdots \end{aligned}$$

where for \(i\in\NN\) the decimal expansion of \(c(i)\) is

$$c(i) = 0.c_{i,1}c_{i,2}c_{i,3}c_{i,4}c_{i,5}\dots \text{ with } c_{i,j} \in \{0,1,2,\dots,9\} \text{ for } j \in \mathbb{N}.$$

By assumption, the image $c(\mathbb{N})$ of c contains every real number in the interval $(0,1)$. To reach a contradiction, it suffices to produce a real number in the interval $(0,1)$ that is provably not in $c(\mathbb{N})$. Consider the number

$$d = 0.d_1d_2d_3d_4d_5\cdots \in (0,1),$$

where d_j is defined by

$$d_j = \begin{cases} 3 & \text{if } c_{j,j} \neq 3, \\ 4 & \text{if } c_{j,j} = 3. \end{cases}$$

We have that $d \neq c(1)$, since they differ in the first digit after the decimal, $d_1 \neq c_{1,1}$. Similarly, $d \neq c(2)$, since they differ in the second digit after the decimal, $d_2 \neq c_{2,2}$, and so on. In general, d cannot be equal yo $c(k)$ since $d_k \neq c_{k,k}$. See the boxed digits below:

$$\begin{aligned} c(1) &= 0.\boxed{c_{1,1}}c_{1,2}c_{1,3}c_{1,4}c_{1,5}\dots \\ c(2) &= 0.c_{2,1}\boxed{c_{2,2}}c_{2,3}c_{2,4}c_{2,5}\dots \end{aligned}$$

$$c(3) = 0.c_{3,1}c_{3,2}\boxed{c_{3,3}}c_{3,4}c_{3,5}\dots$$
$$\vdots \quad \vdots$$

In other words, d is not in the image of c. This contradicts the assumption that $c(\mathbb{N}) = (0, 1)$. Thus there is no such bijective function, and hence the interval $(0, 1)$ is uncountable. It follows that the set of real numbers is uncountable. ■

Warning 4.1.1.9 There is a subtle point that not every real number has a unique decimal representation. For example, $0.5000\dots$ is the same real number as $0.4999\dots$. We avoid this issue by choosing d to have digits only involving 3 and 4 to be sure we are producing a number that has only one representation, so that we can be sure it is not in the sequence.

4.1.2 Product Rule

Theorem 4.1.2.1 Product Rule. *Suppose a procedure can be broken down into a sequence of two tasks. If there are n_1 ways to do the first task, and for each of these ways there are n_2 ways to do the second task, then there are n_1n_2 ways to do the procedure.*

Example 4.1.2.2 Chairs. Suppose auditorium chairs are labeled with an uppercase letter followed by a positive integer not exceeding 100. How many different labels are possible?

Solution. We can view this problem as counting they ways to assign a label to a chair. We proceed by constructing tasks.

Assign a letter	There are 26 ways to assign the letter.
Assign an integer	For each choice of letter, there are 100 ways to assign the integer.

By the Product Rule, there are $26 \cdot 100 = 2600$ different ways to assign a label. Hence there are 2600 different labels. □

The Product Rule generalizes to m tasks. Suppose a procedure can be carried out by performing tasks $T_1, T_2, \dots, T_k$ tasks in sequence. If each task T_i can be done in n_i ways regardless of how the previous tasks were done, then there are $n_1n_2 \cdots n_m$ ways to carry out the procedure.

Example 4.1.2.3 Bit strings of length 8. How many bit strings of length 8 are there?

Solution. We just have to choose each bit, so there are 8 tasks. Each bit can be 0 or 1, so there are 2 ways to choose each bit. The Product Rule says there are $2^8 = 256$ such bit strings. □

Example 4.1.2.4 Bit strings with start 1. How many bit strings of length 8 that start with 1 are there?

Solution. The first bit must be 1. We are left with choosing the remaining bits, so there are 7 tasks. There are 2 ways to choose each of the remaining 7 bits. The Product Rule says there are $2^7 = 128$ such bit strings. □

Example 4.1.2.5 Bit strings with end 00. How many bit strings of length 8 that end with 00 are there?

Solution. The last two bits must be 00. We are left with choosing the remaining bits, so there are 6 tasks. There are 2 ways to choose each of the remaining 6 bits. The Product Rule says there are $2^6 = 64$ such bit strings. □

Example 4.1.2.6 Bit strings with start 1 and end 00. How many bit strings of length 8 that start with 1 and end with 00 are there?

Solution. The first bit must be 1 and the last two must be 00. We are left with choosing the remaining bits, so there are 5 tasks. There are 2 ways to choose each of the remaining 5 bits. The Product Rule says there are $2^5 = 32$ such bit strings. □

Example 4.1.2.7 Birthdays. Suppose there are 22 people in the class. How likely is it that two share a birthday?

Solution. First, let's count the number of ways the birthdays could be arranged, assuming no one was born on February 29. We order the 22 people, and break this own into 22 tasks. The i th task is choosing a birthday for the i th person. There are 365 ways to complete each task, regardless of how previous tasks were done. By the Product Rule, the number of birthday arrangements is

$$365^{22} \approx 2.346621351 \times 10^{56}.$$

Next, let's compute the number of ways the birthdays could be arranged, assuming no one was born on February 29 and no two people share a birthday. Again, we break this into 22 tasks, where the i th task is choosing the birthday for the i th person. The difference is in the number of ways to complete each task. There are 365 was to complete the first task. There are only 364 ways to complete the second task, since we cannot choose the same birthday as the first person. Notice that the number of ways does not depend on what day was chosen for the first person. Similarly, there are 363 ways to complete the second task, and so on. In general, there are $365 - (i - 1)$ ways to complete the i th task. By the Product Rule, the number of birthday arrangements so that no two share a birthday is

$$\underbrace{365 \cdot 364 \cdot 363 \ldots 344}_{22 \text{ terms}} \approx 1.230344586 \times 10^{56}.$$

The likelihood of an event happening (assuming all possible events are equally likely) is the ratio of the number of ways the particular event can occur divided by the total number of events. We can compute that

$$\frac{1.230344586 \times 10^{56}}{2.346621351 \times 10^{56}} \approx 0.5243.$$

That means if birthdays are arranged randomly, there is a 52% chance that no two people in the class share a birthday. This is the largest gathering where that is greater than 50%. If we do the same computation with twice as many people, there is only a 6.7% chance that no one shares a birthday. □

Theorem 4.1.2.8 *If a set A has cardinality n, then the cardinality of the power set of A is $\#\wp(A) = 2^n$.*

Proof. Label the elements of A,

$$A = \{a_1, a_2, \ldots, a_n\}.$$

Note that we uniquely determine a subset of A by specifying the elements that are in the subset.

We specify a subset of A by completing n tasks, where the i th task is deciding whether or not a_i is in the subset. There are two ways to complete each task, regardless of how previous tasks are completed. By the Product Rule, there are 2^n different subsets of A. ■

Example 4.1.2.9 License plates. How many license plates can be made, if the license plate consists of 3 upper case letters followed by 4 one digit numbers.

Solution. We view this as a sequence of 7 tasks: choose a letter, choose a letter, choose a letter, choose a digit, choose a digit, choose a digit, and choose a digit. There are 26 letters and 10 digits, so there are 26 ways to choose a letter and 10 ways to choose a digit. The number of ways to complete each task does not depend on how the previous tasks were completed. Thus by the Product Rule, there are

$$26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 \cdot 10 = 175760000$$

different license plates. □

Example 4.1.2.10 Number of functions. Let $A = \{a, b, c\}$, and let $B = \{1, 2, 3, 4, 5\}$.

1. How many functions $f\colon A \to B$ are there?
2. How many injective functions $f\colon A \to B$ are there?
3. How many surjective functions $f\colon A \to B$ are there?

Solution.

1. We need to choose a value for $f(x)$ for each x in A. There are 5 choices for $f(a)$, 5 choices for $f(b)$, and 5 choices for $f(c)$. By the Product Rule, there are $5 \cdot 5 \cdot 5 = 125$ functions from A to B.
2. We need to choose a value for $f(x)$ for each x in A, but we need to do so in a way that produces an injective function. There are 5 choices for $f(a)$. There are 4 choices for $f(b)$, since we cannot choose the same value as was chosen for $f(a)$. Similarly, there are 3 choices for $f(c)$, since we must avoid $f(a)$ and $f(b)$. By the Product Rule, there are $5 \cdot 4 \cdot 3 = 60$ injective functions from A to B.
3. Since $\#B > \#A$, there are no surjective functions from A to B.

□

4.1.3 Sum Rule

Theorem 4.1.3.1 Sum Rule. *If a task can be done in either one of n_1 ways or n_2 ways, where none of the set of n_1 ways is the same as any of the set of n_2 ways. Then there are $n_1 + n_2$ ways to do the task.*

The Sum Rule can be generalized. If a task can be done in one of n_1 ways, or in one of n_2 ways, \dots, or in one of n_m ways, where none of the set of n_i ways is the same as any of the set of n_j ways, for all $i \neq j$, then there are $n_1 + n_2 + \cdots + n_m$ ways to complete the task.

Example 4.1.3.2 Goats and Sheep. Suppose there are 10 goats and 15 sheep in a certain village.

1. Suppose a gift of one animal (goat or sheep) is to be given to a visiting dignitary. How many ways are there to select a gift?
2. Suppose it is decided instead that the gift should consist of one goat and one sheep. How many ways are there to select a gift?

Solution.

1. There are 10 ways to select a goat and 15 ways to select a sheep, and none of the ways that select a goat is the same as any of the ways that select a sheep. By the Sum Rule, there are $10 + 15 = 25$ ways to select the gift.

2. The procedure of selecting a gift can be broken down into two tasks: select the goat and select the sheep. The number of ways to select the sheep does not depend on which goat was selected. By the Product rule, there are $10 \cdot 15 = 150$ ways to select a gift.

□

4.1.4 Principle of Inclusion-Exclusion

Definition 4.1.4.1 Principle of Inclusion-Exclusion. If a task can be done in n_1 ways or n_2 ways, then the number of ways to do the task is n_1+n_2-k, where k is the number of ways to do the task that is common to the n_1 and n_2 ways. ◊

The Principle of Inclusion-Exclusion s sometimes called the **Subtraction Rule** .

Example 4.1.4.2 Bit strings with start 1 **or end** 00**.** How many bit strings of length 8 are there that start with 1 or end with 00?

Solution. From Example 4.1.2.3, there are 2^7 bit strings of length 8 that start with 1. From Example 4.1.2.5, there are 2^6 bit strings of length 8 that end with 00. From Example 4.1.2.6, there are 2^5 bit strings that start with 1 and end with 00. By the Principle of Inclusion-Exclusion, there are $2^7 + 2^6 - 2^5 = 160$ bit strings that start with 1 or end with 00. □

Example 4.1.4.3 Students in an auditorium. Suppose there are 350 undergraduates in an auditorium. Of these, 220 are computer science majors, 147 are math majors, and 51 are double majoring in computer science and math. How many are neither computer science nor math majors?

Solution. First count the number of computer science or math majors. By the Principle Inclusion-Exclusion, there are $220 + 147 - 51 = 316$ computer science or math majors. Then there are $350 - 316 = 34$ students that are neither computer science nor math majors. □

4.1.5 Division Rule

Theorem 4.1.5.1 Division Rule. *Suppose we can complete a task using a procedure that can be carried out in n ways, and for each of these ways exactly d of the ways corresponds to the same way. Then there are* $\frac{n}{d}$ *ways to do the task.*

Example 4.1.5.2 Seating arrangements. Seat four people at a circular table.

We consider two seating arrangements to be the same if each person has the same left and right neighbor. In other words, two seating arrangements are the same if one can be rotated into the other.

How many seating arrangements are there ?

Solution. By the Product Rule there are $4 \cdot 3 \cdot 2 \cdot 1 = 24$ ways to arrange the four people around the table. For each way, exactly four give the same seating arrangement. By the Division Rule, there are $\frac{24}{4} = 6$ different seating arrangements. □

4.1.6 Exercises

1. State precisely the following rules. Be sure to set up any notation that is required.

(a) Product Rule

(b) Sum Rule

(c) Principle of Inclusion-Exclusion

(d) Division Rule

2. How many bit strings of length eight contain exactly three 0s ? Hint: Think about choosing locations for the 0s.

3. How many answer keys are possible for a twenty question multiple choice test, where each question has exactly six choices?

4. Suppose there are 350 undergraduates in an auditorium. Of these, 220 are computer science majors, 147 are math majors, and 51 are double majoring in computer science and math. How many are neither computer science nor math majors?

5. Let $A = \{a, b, c, d\}$, and let $N = \{1, 2, 3, 4, 5, 6, 7\}$.

(a) How many functions are there from A to N? How many functions are there from N to A?

(b) How many injective functions are there from A to N? How many injective functions are there from N to A?

(c) How many surjective functions are there from A to N? It is a trickier problem to determine the number of surjective functions from N to A.

6. Suppose a multiple choice exam consists of 20 questions, each with choices A, B, C, D.

(a) How many possible answer keys are there?

(b) In how many ways can a student answer the questions on the test, if the student answers every question?

(c) In how many ways can a student answer the questions on the test, if the student student can leave answers blank?

7. A particular brand of shirt comes in a variety of colors (red, green, blue, black, pink, orange, purple) and sizes (XS, S, M, L, XL). How many different types of shirts are there?

8. How many license plates can be made using either three uppercase English letters (A--Z) followed by three digits (0--9) or four uppercase English letters (A--Z) followed by two digits (0--9)?

9. In how many ways can Alice arrange six of her dolls in a row if she has twenty-three dolls?

10. In how many ways can a photographer at a wedding arrange five people in a row from a group of twenty people, given that the bride must be in the picture?

11. In how many ways can Alice arrange twelve of her fifty-four books on a shelf, given that she must include *Strangers Have the Best Candy* and *The Joy of Chickens*?

12. Each user on a computer system has a password, where each character is an uppercase letter (A--Z) or a digit (0--9).

(a) How many possible 8 character passwords are there?

(b) How many possible 8 character passwords are there that do not use any digits?

(c) How many possible 8 character passwords are there, if each password must contain at least one digit?

13. Count the number of times that the letter F appears in the following sentence:

14. Let f be a bijection from the set of positive rational numbers $\mathbb{Q}_{>0}$ to the set of positive integers $\mathbb{N} = \mathbb{Z}_{>0}$. Use f to construct a bijection from $\mathbb{Q}$ to $\mathbb{Z}$. Use this to fill in the gap in the proof in Theorem 4.1.1.6.

Hint: Extend the domain of f to be $\mathbb{Q}$: For a positive rational number r you know $f(r)$. What is a reasonable choice to pick to define $f(-r)$?

15. Let A and B be sets and $n \in \mathbb{N} \cup \{0\}$. Complete the definitions.

(a) We say A and B have the **same cardinality** when ...

(b) We say the cardinality of A is n when ...

(c) We say A is **finite** when ...

(d) We say A is **infinite** when ...

(e) We say A is **countable and infinite** when...

16. Recall the Definition of Cartesian power.

(a) Prove that the set $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N} = \{(a, b) \mid a \in \mathbb{N} \wedge b \in \mathbb{N}\}$ is countable.

(b) Let A be a countable set. Prove that the set $\mathbb{N} \times A$ is countable.

(c) Let $n \in \mathbb{N}$. Prove that the set $\mathbb{N}^n$ is countable.

17. Prove that the set of sequences $(a_n)_{n \in \mathbb{N}}$ with $a_n \in \{0, 1\}$ is uncountable.

4.2 Pigeonhole Principle

Outcomes

1. Apply the Pigeonhole Principle in enumeration
2. Apply the Pigeonhole Principle in proofs.

4.2.1 The Pigeonhole Principle

The basic idea is really simple. For example, if we have three boxes and want to put away four or more toys, there must be at least one box containing two or more toys. The technique can be used cleverly to prove statements that are not so trivial.

Theorem 4.2.1.1 Pigeonhole Principle. *If k is a positive integer, and $k + 1$ or more objects are placed into k boxes, then there is at least one box containing two or more objects.*

Example 4.2.1.2 Words starting with the same letter. In any group of 27 English words, at least two must start with the same letter, since there are only 26 letters in the English alphabet. □

Example 4.2.1.3 People with the same number of hairs. Assume no human has more than 200000 hairs on his/her head. Since there are more than 200000 people in Greensboro, there are at least two people in Greensboro with exactly the same number of hairs on their head. □

Theorem 4.2.1.4 *If A and B are nonempty sets with $\#B = k$ and $\#A \geq k+1$, then there are no injective functions from A to B.*

Proof. We show that any function from A to B is not injective. Make a box for each b in B. For each a in A, place a in box labeled b if $f(a) = b$. By the Pigeonhole Principle, there is a box with at least two elements in it. Thus there is a b in B with at least two different preimages. Thus f is not injective. ■

Example 4.2.1.5 Multiples consisting only of zeros and ones. We find that these integers n, there is a positive multiple of n that has only zeros and ones in its decimal expansion.

n	multiple of n
2	$2 \cdot 5 = 10$
3	$3 \cdot 37 = 111$
4	$4 \cdot 25 = 100$
5	$5 \cdot 2 = 10$
6	$6 \cdot 185 = 1110$
7	$7 \cdot 143 = 1001$

Prove that for any integer n, there is a positive multiple of n that has only zeros and ones in its decimal expansion.

Solution. There are n congruence classes modulo n. Use these as labels on a box. Consider the $n+1$ integers

$$1, 11, 111, \ldots, \underbrace{111\cdots1}_{n+1 \text{ times}}.$$

By the Pigeonhole Principle, there must be a congruence class with at least two of these integers in it. The difference of these two integers is 0 modulo n, hence divisible by n. Additionally, the difference only has 0 s and 1 s in its decimal expansion. □

4.2.2 The Extended Pigeonhole Principle

Theorem 4.2.2.1 Extended Pigeonhole Principle. *Suppose n and k are positive integers. If n objects are placed into k boxes, then there is at least one box containing $\lceil \frac{n}{k} \rceil$ objects.*

Example 4.2.2.2 People born in the same month. In a group of 100 people, there are at least $\lceil \frac{100}{12} \rceil = 9$ born in the same month. □

Example 4.2.2.3 Bowl of fruit. Suppose a bowl contains apples, oranges, and bananas. How many random selections are required to ensure we have at least two of the same fruit?

Solution. Label boxes by fruit types. Place fruit in the box identifying its type. Then there are 3 boxes, so if we select $3 + 1 = 4$ fruit, the Pigeonhole

Principle says we will have at least two of the same type.

If we wanted at least five of the same type, the Extended Pigeonhole Principle says we need $\lceil \frac{n}{3} \rceil \geq 5$. That means we need $\frac{n}{3} > 4$, so $n > 12$. Thus if we select 13 fruit, we are guaranteed at least five of the same type. □

Example 4.2.2.4 Natural numbers adding up to seven. How many distinct random selections from $\{1, 2, 3, 4, 5, 6\}$ are required to guarantee at least one pair that adds up to 7?

Solution. Note that $1 + 6 = 2 + 5 = 3 + 4 = 7$ are the only ways to get a pair to add to 7. Label 3 boxes: $\{1, 6\}$, $\{2, 5\}$, and $\{3, 4\}$. By the Pigeonhole Principle, if we select $3 + 1 = 4$ numbers, one box must contain two or more elements. Thus we will have a pair that sums to 7 if we select 4 numbers. □

Example 4.2.2.5 Dealing cards. Suppose we deal cards from a standard deck of cards (jokers removed). How many cards must be dealt to guarantee that at least 3 of the same suit are chosen?

Solution. Imagine 4 boxes, labeled by the 4 suits: hearts (♡), diamonds (♢), clubs (♣), spades (♠). By the Extended Pigeonhole Principle, if we deal n cards, we are guaranteed at least one suit has $\lceil \frac{n}{4} \rceil$ cards in it. We want $\frac{n}{4} > 2$, so we need $n > 2 \cdot 4 = 8$. Thus 9 cards will guarantee 3 of the same suit. □

4.2.3 Exercises

1. State precisely the Extended Pigeonhole Principle. Be sure to set up any notation that is required.

2. Assume UNCG has 18502 students. How many students at UNCG *must* share a birthday? Explain using the Extended Pigeonhole Principle. (You may assume no one was born on February 29.)

3. Suppose the final exam is graded on a scale from 0 to 100 points. How many students must be in the class to guarantee that at least two students receive the same score on the final exam?

4. How many numbers must be selected from the set $\{1, 3, 5, 7, 9, 11, 13, 15\}$ to guarantee at least one pair of these numbers add up to 16?

5. What is the minimum number of people required to be sure that at least four will have birthdays in the same month?

6. A standard deck of cards consists of 52 cards. Each card is one of thirteen ranks (A, 2, 3, \dots, J, Q, K) and one of four suits (♣, ♠, ♡, ♢).

(a) How many cards must be selected to guarantee that at least three of the same suit are chosen?

(b) How many cards must be selected to guarantee that at least three of the same rank are selected?

7. A standard deck of cards consists of 52 cards. Each card is one of thirteen ranks (A, 2, 3, \dots, J, Q, K) and one of four suits (♣, ♠, ♡, ♢).

(a) How many cards must be selected to guarantee that at least two hearts (♡) are selected?

(b) How many cards must be selected to guarantee that at least three spades (♠) are selected?

(c) How many cards must be selected to guarantee that at least two hearts (♡) and three spades (♠) are selected?

8. Alice selects clips randomly from a bowl that contains ten large paper clips and thirty small paper clips.

(a) How many must she select to be sure of having at least three of the same size?

(b) How many must she select to be sure of having at least five of the same size?

9. Show that there are at least seventeen people in Greensboro (population 285000) with the same three initials, assuming everyone has a first, middle, and last initial.

10. There are 38 different time periods during which classes can be scheduled. If there are 650 different classes, how many different rooms will be needed?

11. Show that among any group of five integers, there are at least two with the same remainder when divided by 4.

12. There are 50 baskets of apples. Each basket contains at least one apple and no more than 24 apples. Show that there are at least 3 baskets containing the same number of apples. If you use the Pigeonhole Principle or its extension, be sure to tell me what are the pigeons and what are the boxes. Hint: The apples are not pigeons. The baskets are not boxes.

4.3 Permutations and combinations

Outcomes

1. Solve counting problems with permutations and combinations
2. Prove theorems by combinatorial arguments.

4.3.1 Permutations

Definition 4.3.1.1 A **permutation** of a set of distinct objects is an ordered arrangement of these objects. ◊

Example 4.3.1.2 Permutations of a set with three elements. Let $S = \{x, y, z\}$. The permutations of S are

$$\{xyz, xzy, yxz, yzx, zxy, zyx\}.$$

□

Definition 4.3.1.3 An ordered arrangement of r elements of a set is called an **r-permutation** . The number of r-permutations of a set of size n is denoted $P(n, r)$. ◊

Remark 4.3.1.4 If $\#S = n$, then an n-permutation is the same thing as a permutation.

Example 4.3.1.5 Permutations of $\{a, b, c, d\}$. Let $S = \{a, b, c, d\}$.

1.
2. The 2-permutations of S are
3. The 3-permutations of S are
4. As an exercise, try to list the permutations of S. There

□

Theorem 4.3.1.6 *Let n and r be integers such that $0 \leq r \leq n$. Then*

$$P(n,r) = \frac{n!}{(n-r)!} = n(n-1)(n-2)\cdots(n-(r-1)).$$

Proof. Constructing an r-permutation, can be broken down into a sequence of r tasks. First, we select the first term in the arrangement. There are n ways to do this. Then, we select the second term in the arrangement. There are $n-1$ ways to do this because one of the ways is no longer valid as it was already used in the previous step. Similarly, there are $n-2$ ways to select the third term, and so on. The result then follows by the Product Rule. ■

Remark 4.3.1.7 $0! = 1$, so $P(n,n) = n!$.

Example 4.3.1.8 Gold, Silver, and Bronze finishers. Suppose there are eight runners in a race. How many different ways can we have Gold, Silver, and Bronze finishers, assuming no ties?

Solution. We want to count the number of 3-permutations of the runners. Since there are eight runners, we have

$$P(8,3) = \frac{8!}{(8-3)!} = \frac{8!}{5!} = 8 \cdot 7 \cdot 6 = 336.$$

□

Remark 4.3.1.9 For n and k such that $0 \leq k \leq n$, cancellation gives

$$\frac{n!}{k!} = n(n-1)\cdots(k+1).$$

Example 4.3.1.10 Permutations of the alphabet. Let $S = \{a, b, c, \ldots, x, y, z\}$. The number of permutations of S is

$$P(26,26) = 26! = 403291461126605635584000000.$$

□

Example 4.3.1.11 Permutations of the alphabet with *abc*. How many permutations of $S = \{a, b, c, \ldots, x, y, z\}$ contain *abc*?

We can view this as a different problem. We want permutations of

$$S' = \{abc, d, e, f, \ldots, x, y, z\}.$$

Then $\#S' = 24$, and

$$P(24,24) = 24! = 620448401733239439360000.$$

□

Example 4.3.1.12 Seven people. How many ways are there to arrange seven people in from a group of ten (including me) in a row if I need to be one of the seven?

Solution. We break this into tasks. First select the position for me. There are 7 choices. The remaining spots are ordered. We need a 6-permutation of the set of 9 remaining people. There are

$$P(9,6) = \frac{9!}{(9-6)!} = 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 = 60480$$

such permutations. By the Product Rule, there are $7 \cdot 60480 = 423360$ arrangements. □

4.3.2 Combinations

Definition 4.3.2.1 A **combination** of a set of distinct objects is an unordered selection of these objects. An unordered selection of r-elements from a set is called an **r-combination** . The number of r-combinations of a set of size n is denoted $C(n, r)$. ♢

Example 4.3.2.2 Combinations and Permutations of $\{a, b, c, d\}$. Let $S = \{a, b, c, d\}$.

1. The 1-combinations of S are
2. The 2-permutations of S are
3. The 3-combinations of S are
4. The 4-combinations of S are

□

Remark 4.3.2.3 There is only one way to select all of the elements of S, so $C(n, n) = 1$, in general.

Theorem 4.3.2.4 *Let n and r be integers such that $0 \leq r \leq n$. Then*

$$C(n, r) = \frac{n!}{r!(n-r)!}.$$

Proof. We can carry out the last of listing the r-combinations by listing the r-permutations. By Theorem 4.3.1.6, there are $\frac{n!}{(n-r)!}$ ways to do this. For each of these ways, exactly $r!$ of the ways correspond to the same way. The result then follows by the Division Rule. ■

Example 4.3.2.5 Poker hands. How many poker hands (5 cards) are in a standard deck of 52 cards?

Solution. We want to count 5-combinations of cards. Since there are 52 cards, the number of 5-combinations is

$$C(52, 5) = \frac{52!}{5!(52-5)!} = \frac{52!}{5!47!} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2598960.$$

□

Selecting r elements from a set of size n to be in an r-combination is equivalent to selecting $n - r$ elements not to be in the r-combination. Thus the number of ways to do each must be the same. This gives the following theorem.

Theorem 4.3.2.6 *Let n and r be integers such that $0 \leq r \leq n$. Then*

$$C(n, r) = C(n, n-r).$$

Example 4.3.2.7 Exam committees. Suppose there are 9 freshmen and 11 sophomores in MAT 253. How many different final exam committees can be made if a committee consists of 3 freshmen and 4 sophomores ?

Solution. The procedure of choosing a committee can be broken down into two tasks, namely, selecting the freshmen and selecting the sophomores. The number of ways to select freshmen is

$$C(9, 3) = \frac{9!}{3!(9-3)!} = \frac{9!}{3!6!} = \frac{9 \cdot 8 \cdot 7}{3 \cdot 2 \cdot 1} = 84.$$

The number of ways to select sophomores is

$$C(11,4) = \frac{11!}{4!(11-4)!} = \frac{11!}{4!7!} = \frac{11 \cdot 10 \cdot 9 \cdot 8}{4 \cdot 3 \cdot 2 \cdot 1} = 330.$$

Then by the Product Rule, the number of different committees is

$$84 \cdot 330 = 27720.$$

□

Example 4.3.2.8 Bit strings with three zeros. How many bit strings of length ten have exactly three 0s?

Solution. We can describe a bit string with exactly three 0s by specifying the location of the three 0s. There are ten possible places for them to go, and we cannot distinguish the three 0s, so the order does not matter. Thus the number of bit strings with exactly three 0s is

$$C(10,3) = \frac{10!}{3!(10-3)!} = \frac{10!}{3!7!} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1} = 120.$$

□

Example 4.3.2.9 Soccer team. Thirteen students on a soccer team consisting of freshman and sophomores show up for a game. Of the thirteen that show up, three are sophomores. How many ways are there to choose ten players to take the field if at least one of these players must be a sophomore ?

Solution. Which of the following computations is correct?

1. There are $C(3,1) = 3$ ways to choose a sophomore. Then there are 12 remaining players, and we need to choose 9 of them. The number of ways to choose 9 from 12 is

   $$C(12,9) = \frac{12!}{9!(12-9)!} = \frac{12!}{9!3!} = \frac{12 \cdot 11 \cdot 10}{3 \cdot 2 \cdot 1} = 220.$$

   By the Product Rule, the number of ways to choose a team with at least one sophomore is the product

   $$C(3,1)C(12,9) = 3 \cdot 220 = 660.$$

2. There are 3 ways to choose a sophomore. Of the remaining 10 freshmen, we need to choose 9. The number of ways to choose 9 from 10 is $C(10,9) = 10$. By the Product Rule, the number of ways to choose a team with at least one sophomore is the product $3 \cdot 10 = 30$.

3. To pick a team, independent of class year, we need to choose 10 people from a group of 13. The number of ways to pick 10 from 13 is

   $$C(13,10) = \frac{13!}{10!(13-10)!} = \frac{13!}{10!3!} = \frac{13 \cdot 12 \cdot 11}{3 \cdot 2 \cdot 1} = 286.$$

   There is $C(10,10) = 1$ way to choose a team of all freshmen. Therefore there are $286 - 1 = 285$ ways to choose a team with at least one sophomore.

The third method is correct. The second method counts the teams with exactly 1 sophomore. It does not allow for more than one sophomore. The first method counts certain teams more than once. For example if Alice is

chosen first as part of the $C(3,1)$, and then Beverly is chosen as part of the $C(12,9)$, the team created is duplicated when Beverly is chosen first as part of the $C(3,1)$ and Alice is chosen as part of the $C(12,9)$. We can see furthermore that the value of 660 cannot be correct, as the total number of possible teams is 286, as shown in the third method. □

Example 4.3.2.10 People in a line. How many ways are there for eight freshmen and five sophomores to stand in a line so that no two sophomores stand next to each other?

Hint: First position the freshmen and then consider the possible positions for the sophomores.

Solution. There are $P(8,8) = 8!$ ways to arrange the freshmen in order. There are 9 slots where a sophomore can go so that the sophomores are separated by at least one freshman. There are $C(9,5) = \frac{9!}{5!(9-5)!}$ ways to choose the slots for the sophomores. There are $P(5,5) = 5!$ ways to arrange the sophomores. By the Product Rule, there are

$$8! \cdot \frac{9!}{5!(9-5)!} \cdot 5! = 609638400 \quad \text{ways.}$$

□

4.3.3 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) permutation

(b) r-permutation

(c) combination

(d) r-combination

2. A local pizza shop offers their pies in small, medium, large, or extra large. For toppings, they offer: pepperoni, sausage, bacon, olives, onions, peppers, and anchovies. How many different pizzas can they make that have exactly three (different) toppings?

Compute the following.

(a) $P(7,3)$

(b) $C(7,3)$

(c) $C(8,0)$

(d) $P(8,5)$

(e) $C(8,3)$

3. Fifteen people on a softball team show up for a game.

(a) How many ways are there to select 9 to take the field?

(b) How many ways are there to assign the 9 positions?

4. How many permutations of the letters ABCDEFGH contain the string ABC?

5. In how many different orders can ten runners finish a race if no ties are allowed?

6. List all the permutations of $\{1, 2, 3\}$

7. List all the 3-combinations of $\{a, e, i, o, u\}$.

8. List all the 3-permutations of $\{1, 2, 3, 4, 5, 6\}$.

9. In how many ways can a set of four letters be selected from the English alphabet?

10. How many ways are there for 10 sophomores and 6 freshmen to stand in a line if so that no two freshmen stand next to each other? (Hint: First position the sophomores and then consider the possible positions for the freshmen.)

11. Harry, Hermione, Ron, Fred, George, Ginny, Luna, Neville, Seamus, and Hagrid go to some pictures taken.

(a) How many ways are there to arrange four people from that group in a row for the picture?

(b) Suppose Harry is willing to pay for any picture that he is in. How many ways are there to arrange four people from that group in a row for the picture, if Harry must be one of the four?

Chapter 5

Number Theory and Applications

> Mathematics is the queen of the sciences and number theory is the queen of mathematics. She often condescends to render service to astronomy and other natural sciences, but in all relations she is entitled to the first rank.
>
> —Carl Friedrich Gauss (1777--1855)

At its core, **number theory** is the study of integers. In this chapter, we examine introduce modular arithmetic to explore divisibility. We discuss modern applications of number theory including ASCII encoding and encryption.

5.1 Divisibility and modular arithmetic

Objectives

Introduce fundamental concepts from number theory, including the division algorithm, congruences, and the rules of modular arithmetic.

5.1.1 Divisibility

Definition 5.1.1.1 Let a and b be integers, with $a \neq 0$. We say a **divides** b, denoted $a \mid b$, if there exists an integer k such that $b = ak$. In this case, we call a a **factor** of **divisor** of b, and b is a **multiple** of a. We write $a \nmid b$ when a does not divide b. ◊

Example 5.1.1.2 Divisibility.

3 **divides** 12	$3 \mid 12$ since $12 = 3 \cdot 4$.
3 **does not divide** 7	$3 \nmid 7$ since there is no integer k such that $7 = 3k$.

□

Theorem 5.1.1.3 *Let a, b, and c be integers with $a \neq 0$.*

1. *If $a \mid b$ and $a \mid c$, then $a \mid (b + c)$.*

2. *If $a \mid b$, then $a \mid bc$.*

3. *Suppose in addition $b \neq 0$. If $a \mid b$ and $b \mid c$, then $a \mid c$.*

Proof.

1. Suppose $a \mid b$ and $a \mid c$. Then there exists integers m and n such that $b = am$ and $c = an$. We want to show $a \mid (b + c)$. We compute

$$b + c = am + an = a(m + n).$$

   Since m and n are integers, $m + n$ is an integer. Thus $a \mid (b + c)$.

2. Suppose $a \mid b$. Then there exists an integer n such that $b = an$. We want to show $a \mid bc$. We compute

$$bc = (an)c = a(nc).$$

   Since n and c are integers, nc is an integer. Thus $a \mid bc$.

3. Suppose $a \mid b$ and $b \mid c$. Then there exist integers m and n such that $b = am$ and $c = bn$. We want to show $a \mid c$. We compute

$$c = bn = (am)n = a(mn).$$

   Since m and n are integers, mn is an integer. Thus $a \mid c$.

■

Corollary 5.1.1.4 *Let $a, b, c \in \mathbb{Z}$ with $a \neq 0$. If $a \mid b$ and $a \mid c$, then $a \mid (sb+tc)$ for all $s, t \in \mathbb{Z}$.*

Theorem 5.1.1.5 Division algorithm for positive integers. *Let a and b be integers with $b > 0$. There exists unique integers q and r with $0 \leq r < b$ such that $a = b \cdot q + r$.*

Proof. The existence of q and r follows from Listing 5.1.1.6.

We now prove the uniqueness of q and r. Suppose that there are $q_1 \in \mathbb{Z}$ and $q_2 \in \mathbb{Z}$ and $r_1 \in \mathbb{Z}$ and $r_2 \in \mathbb{Z}$ such that

1. $a = b \cdot q_1 + r_1$ and $0 \leq r_1 < b$ and

2. $a = b \cdot q_2 + r_2$ and $0 \leq r_2 < b$.

Then $b \cdot q_1 + r_1 = b \cdot q_2 + r_2$. We need to show that $q_1 = q_2$ and $r_1 = r_2$.

Case 1: $r_1 = r_2$ If $r_1 = r_2$ then

$$\begin{aligned} b \cdot q_1 + r_1 &= b \cdot q_2 + r1 && \text{subtract } r_1 \\ b \cdot q_1 &= b \cdot q_2 && \text{divide by } b \\ q_1 &= q_2 \end{aligned}$$

Case 2: $q_1 = q_2$ If $q_1 = q_2$ then $bq_1 = bq_2$. Thus $b \cdot q_1 + r_1 = b \cdot q_2 + r2$ becomes $b \cdot q_1 + r_1 = b \cdot q_1 + r2$. Subtracting $b \cdot q_1$ from both sides yields $r_1 = r_2$.

Case 3: $r_1 \neq r_2$ We prove that this case cannot happen by contradiction. Collecting the q_i on the left and the r_i on the right

$$b \cdot q_1 + r_1 = b \cdot q_2 + r_2$$

becomes

$$b \cdot (q_1 - q_2) = r_2 - r_1.$$

$$|b \cdot (q_1 - q_2)| = |r_2 - r_1| > 0.$$

Because $r_1 \neq r_2$ and $r_2 < b$ and $r_1 < b$ we have

$$0 < |r_2 - r_1| < b.$$

Because $0 \neq q_1 - q_2$ and $q_1 \in \mathbb{Z}$ and $q_2 \in \mathbb{Z}$ we have

$$|b \cdot (q_1 - q_2)| = b \cdot |q_1 - q_2| \geq b$$

which contradicts

$$|b \cdot (q_1 - q_2)| = |r_2 - r_1| < b.$$

So we must have $r_1 = r_2$ and $q_1 = q_2$ which proves the uniqueness. ■

Listing 5.1.1.6 Naive division algorithm for $a \in \mathbb{N}_0$ and $b \in \mathbb{N}$

```
1  def divmod_naive(a,b):^^I
2    q=0^^I
3    while a>b:
4      a=a-b
5      q=q+1
6    return q,a
```

Theorem 5.1.1.7 Division algorithm for negative integers. *Let a be neagtive integer and d be integers with $d > 0$. There exists unique integers q and r with $0 \leq r < d$ such that $a = dq + r$.*

Proof. The existence of q and r follows from Listing 5.1.1.8.

The uniqueness follows as in the proof of Theorem 5.1.1.5. ■

Listing 5.1.1.8 Naive division algorithm for $a \in \mathbb{Z}$ with $a < 0$ and $\in \mathbb{N}$

```
1  def divmod_negative(a,b):^^I
2    q=0^^I
3    while a <0:
4      a=a+b
5      q=q-1
6    return q,a
```

Together Theorem 5.1.1.5 and Theorem 5.1.1.7 yield:

Theorem 5.1.1.9 Division algorithm. *Let a and d be integers with $d > 0$. There exists unique integers q and r with $0 \leq r < d$ such that $a = dq + r$.*

Remark 5.1.1.10 For $d \in \mathbb{N}$ and $a \in \mathbb{Z}$ we have that $d \mid a$ is equivalent to $a \bmod d = 0$.

Remark 5.1.1.11 Given a and d, we can compute q as the floor and get

$$q = a \operatorname{div} d = \left\lfloor \frac{a}{d} \right\rfloor.$$

After we have q, we can solve for r in the equation $a = dq + r$ and get

$$r = a \bmod d = a - \left\lfloor \frac{a}{d} \right\rfloor \cdot d..$$

Example 5.1.1.12 Division algorithm and floor. Let's see what the division algorithm says in a few examples.

First consider $a = 253$ and $d = 17$. With a calculator we obtain $253/17 \approx 14.88$. Thus $q = \left\lfloor \frac{253}{17} \right\rfloor = 14$. Then $253 = 17 \cdot 14 + r$. Solving for r gives $r = 15$.

Next consider $a = -253$ and $d = 17$. By long division, we see $-253/17 \approx -14.88$. Thus $q = \left\lfloor \frac{-253}{17} \right\rfloor = -15$. Then $-253 = 17(-15) + r$. Solving for r gives $r = 2$. □

Definition 5.1.1.13 Let a and d be integers, with $d > 0$. Let q and r be integers such that $a = dq + r$, with $0 \leq r < d$. We call d the **divisor** and we call a the **dividend**.

We call q the **quotient**, and denote it by $q = a \operatorname{div} d$.

We call r the **remainder**, and denote it $t = a \bmod d$. ◊

Example 5.1.1.14 Quotients and remainders. $17 \operatorname{div} 3 = 5$ and $17 \bmod 3 = 2$ since $17 = 5 \cdot 3 + 2$.

$-17 \operatorname{div} 3 = -6$ and $-17 \bmod 3 = 1$ since $17 = -6 \cdot 3 + 1$. □

Theorem 5.1.1.15 *Every integer is even or odd. No integer is both.*

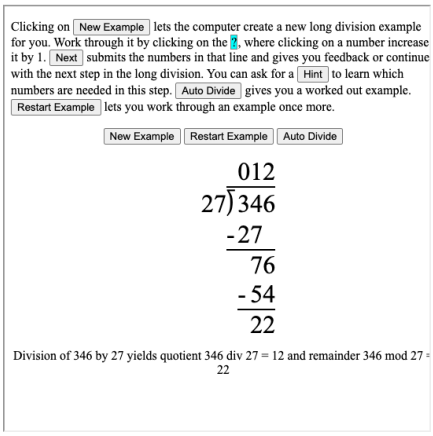
Proof. Let n be an integer.

It follows from and that n is even if and only if $n \bmod 2 = 0$ and from and that n is odd if and only if $n \bmod 2 = 1$.

Because the remainder from the division by 2 is either 0 or 1 but not both by we we have that integer is even or odd but not both. ■

While Listing 5.1.1.6 and Listing 5.1.1.8 are useful for proving Theorem 5.1.1.9 using repeated subtraction or addition to compute quotients and remainders when numbers are a bit larger takes a lot of subtractions or additions. Long division is much more efficient.

Example 5.1.1.16 Long division interactive.



□

In Python, the quotient and remainder can be computed as follows. There is also a function that returns both the quotient and remainder at once.

Listing 5.1.1.17

```
# quotient^^I
q = a // d
# remainder
r = a % d
# or compute them both at the same time
q, r = divmod(a,d)
```

5.1.2 Modular arithmetic

Definition 5.1.2.1 Let a, b, and m be integers, with $m > 0$. We say a is **congruent** to b modulo m, denoted $a \equiv b \pmod{m}$ when $m \mid (a - b)$. ◊

Remark 5.1.2.2 This notion is different (but related) to the mod representing the remainder from the division algorithm. For example, if $a = b \bmod m$, then a is an integer with $0 \leq a < m$. In particular, $b \bmod m$ is an integer. On the other hand, $a \equiv b \pmod{m}$ asserts a relationship between a and b. We never write $b \pmod{m}$ by itself.

Theorem 5.1.2.3 *Let $a, b, m \in \mathbb{Z}$, with $m > 0$. The following are equivalent.*

1. $a \equiv b \pmod{m}$.
2. *There exists $k \in \mathbb{Z}$ such that $a = b + km$.*
3. $a \bmod m = b \bmod m$.

Proof. It is enough to prove $1 \to 2$, $2 \to 3$, and $3 \to 1$.

$1 \to 2$ Suppose $a \equiv b \pmod{m}$. Then $m \mid (a - b)$. Then there exists an integer k such that $a - b = mk$. Solving for a, we have $a = b + km$ as desired.

$2 \to 3$ Suppose there exists $k \in \mathbb{Z}$ such that $a = b + km$. Using the Division algorithm, we write $a = qm + r$ and $b = q'm + r'$, for some integers q, q', r, r' with $0 \leq r, r' < m$. We want to show that $r = r'$. We compute

$$a = b + km = (q'm + r') + km = (q' + k)m + r'.$$

By the uniqueness of the integers in the Division algorithm, we must have $q' + k = q$ and $r' = r$.

$3 \to 1$ Suppose $a \bmod m = b \bmod m$. Let $r = a \bmod m = b \bmod m$. Then by the Division algorithm, there exist integers q and q' such that $a = qm + r$ and $b = q'm + r$. We compute

$$a - b = (qm + r) - (q'm + r) = m(q - q').$$

Since q and q' are integers, we have $m \mid (a - b)$. Thus $a \equiv b \pmod{m}$.

■

Definition 5.1.2.4 The set of integers that are congruent to a modulo m is called the **congruence class** of a modulo m, and is denoted $[a]_m$. In particular,

$$[a]_m = \{a + mk \mid k \in \mathbb{Z}\}.$$

The integer a is called a representative of the congruence class $[a]_m$. ◊

Remark 5.1.2.5 Let $m \in \mathbb{N}$ and $a \in \mathbb{Z}$. By Theorem 5.1.2.3 we have

$$[a]_m = \{a + mk : k \in \mathbb{Z}\} = \{b : b \in \mathbb{Z} \wedge b \equiv a \pmod{m}\} = \{b : b \in \mathbb{Z} \wedge b \bmod m = a \bmod m\}.$$

It follows directly that if $c \in [a]_m$ then $[c]_m = [a]_m$

Example 5.1.2.6 Congruence classes modulo 7. The congruence class of 3 modulo 7 is

$$[3] = \{3 + 7k \mid k \in \mathbb{Z}\} = \{\ldots, -11, -4, 3, 10, 17, \ldots\}$$

This is equal to the congruence class of 10 since

$$[10] = \{10 + 7k \mid k \in \mathbb{Z}\} = \{\ldots, -11, -4, 3, 10, 17, \ldots\}$$

In fact, the congruence class of 3 is equal to the congruence class of any integer that is congruent to 3 modulo 7,

$$\cdots = [-4] = [3] = [10] = [17] = \cdots.$$

□

Example 5.1.2.7 Congruence classes modulo 5. There are five congruence classes modulo 5, namely

$$\begin{aligned}
[0]_5 &= \{0 + 5k \mid k \in \mathbb{Z}\} = \{\ldots, -10, -5, 0, 5, 10, \ldots\} \\
[1]_5 &= \{1 + 5k \mid k \in \mathbb{Z}\} = \{\ldots, -9, -4, 1, 6, 11, \ldots\} \\
[2]_5 &= \{2 + 5k \mid k \in \mathbb{Z}\} = \{\ldots, -8, -3, 2, 7, 12, \ldots\} \\
[3]_5 &= \{3 + 5k \mid k \in \mathbb{Z}\} = \{\ldots, -7, -2, 3, 8, 13, \ldots\} \\
[4]_5 &= \{4 + 5k \mid k \in \mathbb{Z}\} = \{\ldots, -6, -1, 4, 9, 14, \ldots\}.
\end{aligned}$$

and we have

$$[0]_5 \cup [1]_5 \cup [2]_5 \cup [3]_5 \cup [4]_5 = \mathbb{Z}.$$

□

The following result says that we can define arithmetic on congruence classes.

Theorem 5.1.2.8 *Let $a, \hat{a}, b, \hat{b}, m \in \mathbb{Z}$, with $m > 0$. If $a \equiv \hat{a} \pmod{m}$ and $b \equiv \hat{b} \pmod{m}$, then*

1. $a + b \equiv \hat{a} + \hat{b} \pmod{m}$;
2. $ab \equiv \hat{a}\hat{b} \pmod{m}$.

Proof.

1. Suppose $a \equiv \hat{a} \pmod{m}$ and $b \equiv \hat{b} \pmod{m}$. Then there exist integers k and ℓ such that $a = \hat{a} + km$ and $b = \hat{b} + \ell m$. We compute
$$a + b = (\hat{a} + km) + (\hat{b} + \ell m) = (\hat{a} + \hat{b}) + m(k + \ell).$$
Since k and ℓ are integers, $k + \ell$ is an integer. Thus $a + b \equiv \hat{a} + \hat{b} \pmod{m}$ by Theorem 5.1.2.3.
2. Similarly, we compute
$$ab = (\hat{a} + km)(\hat{b} + \ell m) = (\hat{a}\hat{b}) + m(k\hat{b} + \ell\hat{a} + k\ell m).$$
Since $\hat{a}$, $\hat{b}$, k, m, and ℓ are integers, $(k\hat{b} + \ell\hat{a} + k\ell m)$ is an integer. Thus $ab \equiv \hat{a}\hat{b} \pmod{m}$ by Theorem 5.1.2.3.

■

The result says that we can take any element in the congruence class of a and add it to any element in the congruence class of b, and we will get an element of the congruence class of $a+b$, and the analogous result for multiplication. The division algorithm ensures that each congruence class modulo m will contain a unique representative $0 \leq a < m$, so we can transfer the arithmetic on congruence classes to an arithmetic on congruence class representatives.

Example 5.1.2.9 Arithmetic modulo 10. Suppose $a \equiv 5 \pmod{10}$ and $b \equiv 2 \pmod{10}$. Then

$$a^2 - b \equiv 5^2 - 2 \equiv 23 \equiv 3 \pmod{10}.$$

□

Definition 5.1.2.10 Let m be a positive integer. The **integers mod** m, denoted $\mathbb{Z}_m$, as a set is

$$\mathbb{Z}_m = \{0, 1, \ldots, m-1\}.$$

We endow this set with arithmetic, called **modular arithmetic** For a and b in $\mathbb{Z}_m$, we define the sum and product by

$$a \oplus_m b = (a+b) \bmod m \quad \text{and} \quad a \otimes_m b = (ab) \bmod m.$$

◊

Example 5.1.2.11 Addition and multiplication table modulo 6. We compute the addition and multiplication table for $\mathbb{Z}_6$.

$\oplus_6$	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

$\otimes_6$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

□

5.1.3 Modular inverses

Definition 5.1.3.1 Let $m \in \mathbb{N}$ and $a \in \mathbb{Z}_m$ and $b \in \mathbb{Z}_m$. If $a \oplus_m b = 0$ we call b an additive inverse of a modulo m. ◊

Example 5.1.3.2 Additive inverses modulo 6. We list additive inverse modulo 6.

- An additive inverse modulo 6 of 0 is 0, because $0 \oplus_6 0 = (0+0) \bmod 6 = 0 \bmod 6 = 0$.
- An additive inverse modulo 6 of 1 is 5, because $1 \oplus_6 5 = (1+5) \bmod 6 = 6 \bmod 6 = 0$.
- An additive inverse modulo 6 of 2 is 4, because $2 \oplus_6 4 = (2+4) \bmod 6 = 6 \bmod 6 = 0$.
- An additive inverse modulo 6 of 3 is 3, because $3 \oplus_6 3 = (3+3) \bmod 6 = 6 \bmod 6 = 0$.

- An additive inverse modulo 6 of 4 is 2, because $4 \oplus_6 2 = (4+2) \bmod 6 = 6 \bmod 6 = 0$.
- An additive inverse modulo 6 of 5 is 1, because $5 \oplus_6 1 = (5+1) \bmod 6 = 6 \bmod 6 = 0$.

□

Definition 5.1.3.3 Let $m \in \mathbb{N}$ and $a \in \mathbb{Z}_m$ and $b \in \mathbb{Z}_m$. If $a \otimes_m b = 1$ we call b a multiplicative inverse of a modulo m. ◊

Example 5.1.3.4 Multiplicative inverses modulo 6. We list multiplicative inverses modulo 6.

- There is no multiplicative inverse modulo 6 of 0, because $0 \otimes_6 a = (0 \cdot a) \bmod 6 = 0$ for all $a \in \mathbb{Z}_6$.
- A multiplicative inverse modulo 6 of 1 is 1, because $1 \otimes_6 1 = (1 \cdot 1) \bmod 6 = 1 \bmod 6 = 1$.
- There is no multiplicative inverse modulo 6 of 2, because $2 \otimes_6 a = (2 \cdot a) \bmod 6 \neq 1$ for all $a \in \mathbb{Z}_6$.
- There is no multiplicative inverse modulo 6 of 3, because $3 \otimes_6 a = (3 \cdot a) \bmod 6 \neq 1$ for all $a \in \mathbb{Z}_6$.
- There is no multiplicative inverse modulo 6 of 4, because $4 \otimes_6 a = (4 \cdot a) \bmod 6 \neq 1$ for all $a \in \mathbb{Z}_6$.
- A multiplicative inverse modulo 6 of 5 is 5, because $5 \otimes_6 5 = (5 \cdot 5) \bmod 6 = 25 \bmod 6 = 1$.

□

Definition 5.1.3.5 Let $m \in \mathbb{N}$ and $a \in \mathbb{Z}_m$. When there exists $b \in \mathbb{Z}_m - \{0\}$ such that $a \otimes_m b = 0$ then we call a a zero divisor modulo m. ◊

Example 5.1.3.6 Zero divisors modulo 6. We list zero divisors modulo 6.

- 2 is a zero divisor modulo 6, because $2 \otimes_6 3 = (2 \cdot 3) \bmod 6 = 6 \bmod 6 = 0$.
- 3 is a zero divisor modulo 6, because $3 \otimes_6 2 = (3 \cdot 2) \bmod 6 = 6 \bmod 6 = 0$.
- 4 is a zero divisor modulo 6, because $4 \otimes_6 3 = (4 \cdot 3) \bmod 6 = 12 \bmod 6 = 0$.

□

5.1.4 Primitive roots

Example 5.1.4.1 Powers of 3 modulo 5. The powers of 3 modulo 5 are

$$3^0 \bmod 5 = 1 \bmod 5 = 1$$
$$3^1 \bmod 5 = 3 \bmod 5 = 3$$
$$3^2 \bmod 5 = 9 \bmod 5 = 4$$
$$3^3 \bmod 5 = 27 \bmod 5 = 2$$
$$3^4 \bmod 5 = 81 \bmod 5 = 1$$
$$3^5 \bmod 5 = 243 \bmod 5 = 3$$
$$\vdots$$

Thus all elements of $\mathbb{Z}_5 - \{0\} = \{1, 2, 3, 4\}$ are powers of 3 modulo 5. □

Definition 5.1.4.2 A **primitive root** modulo a prime p is an integer r in $\mathbb{Z}_p$ such that every nonzero element of $\mathbb{Z}_p$ is a power of r. ◊

Remark 5.1.4.3 Let $m \in \mathbb{N}$. We have $a \in \mathbb{N}$ is a primitive root modulo m if and only if

$$\{(a^n) \bmod m \mid n \in \mathbb{N} \cup \{0\}\} = \mathbb{Z}_m - \{0\}.$$

Example 5.1.4.4 Is 1 **a primitive root modulo** 5 **?** We determine whether 1 is a primitive root modulo 5. We have $1^n = 1$ for all $n \in \mathbb{N} \cup \{0\}$. Thus there is no $n \in \mathbb{N} \cup \{0\}$ such that $1^n = 2$. But for 1 to be a primitive root modulo 5, all elements of $\mathbb{Z}_5 - \{0\} = \{1, 2, 3, 4\}$ would have to be powers of 1. Thus 1 is not a primitive root modulo 5. □

Example 5.1.4.5 Is 2 **a primitive root modulo** 5 **?** We determine whether 2 is a primitive root modulo 5. We have

$$2^0 \bmod 5 = 1 \bmod 5 = 1$$
$$2^1 \bmod 5 = 2 \bmod 5 = 2$$
$$2^2 \bmod 5 = 4 \bmod 5 = 4$$
$$2^3 \bmod 5 = 8 \bmod 5 = 3$$

So all elements of $\mathbb{Z}_5 - \{0\} = \{1, 2, 3, 4\}$ are powers of 2 modulo 5. Thus 2 is a primitive root modulo 5. □

We have already noticed in Example 5.1.4.1 above that 3 is a primitive root modulo 5.

Example 5.1.4.6 Is 4 **a primitive root modulo** 5 **?** We determine whether 4 is a primitive root modulo 5.\\ We have

$$4^0 \bmod 5 = 1 \bmod 5 = 1$$
$$4^1 \bmod 5 = 4 \bmod 5 = 4$$
$$4^2 \bmod 5 = 16 \bmod 5 = 1$$
$$4^3 \bmod 5 = (4 \cdot 4^2) \bmod 5 = (4 \cdot 1) \bmod 5 = 4$$
$$4^4 \bmod 5 = (4 \cdot 4^3) \bmod 5 = (4 \cdot 4) \bmod 5 = 1$$

We observe that the powers of 4 modulo 5 alternate between 1 and 4.

But for 4 to be a primitive root modulo 5, all elements of $\mathbb{Z}_5 - \{0\} = \{1, 2, 3, 4\}$ would have to be powers of 4. For example, 2 is not a power of 4. So 4 is a not primitive root modulo 5.

Thus we have found that 2 and 3 are primitive roots modulo 5 while 1 and 4 are not primitive roots modulo 5. □

Definition 5.1.4.7 Let $m \in \mathbb{N}$, $b \in \mathbb{Z}_m$, and $a \in \mathbb{Z}_m$.

We call the smallest $n \in \mathbb{N} \cup \{0\}$ such that $b^n \bmod m = a$ the discrete logarithm of a to the base b provided such an n exists.

If $b^n \bmod m = a$ we write $\log_b^{\otimes_m} a = n$. ◊

Problem 5.1.4.8 Find $\log_7^{\otimes} 9$**..** Let $a \otimes b = (a \cdot b) \bmod 11$. Find $\log_7^{\otimes} 9$.

Solution. We need to find $n \in \mathbb{N} \cup \{0\}$ such that $7^{n\otimes} = 9$. We compute

$7^{1\otimes} = 7$ $\quad$ $7^{2\otimes} = (7 \cdot 7) \bmod 11 = 5$
$7^{3\otimes} = 5 \otimes 7 = (5 \cdot 7) \bmod 11 = 2$ $\quad$ $7^{4\otimes} = (2 \otimes 7) = (2 \cdot 7) \bmod 11 = 3$
$7^{5\otimes} = 3 \otimes 7 = (3 \cdot 7) \bmod 11 = 10$ $\quad$ $7^{6\otimes} = (10 \otimes 7) = (10 \cdot 7) \bmod 11 = 4$
$7^{7\otimes} = 4 \otimes 7 = (4 \cdot 7) \bmod 11 = 6$ $\quad$ $7^{8\otimes} = (6 \otimes 7) = (6 \cdot 7) \bmod 11 = 9$

Thus $\log_7^{\otimes} 9 = 8$. □

Remark 5.1.4.9 Let $m \in \mathbb{N} - \{1\}$ and $b \in \mathbb{Z}_m$.

Then $\log_b^{\otimes_m} a$ exists for all $a \in \mathbb{Z}_m - \{0\}$ if and only if b is a primitive root modulo m.

5.1.5 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) divides

(b) factor

(c) multiple

(d) divisor

(e) dividend

(f) quotient

(g) remainder

(h) two integers are congruent modulo an integer

(i) congruence class of an integer modulo an integer

(j) $\mathbb{Z}_m$

(k) primitive root

2. State precisely the Division algorithm. Be sure to set up any notation that is required.

3. Does 13 divide these numbers? Justify.

(a) 39

(b) -26

(c) 0

(d) 1

(e) 57

4. What is the quotient and remainder for these?

(a) 44 divided by 7

(b) -123 divided by 12

(c) 253 divided by 15

(d) 0 divided by 17

(e) -100 divided by 100

5. Suppose a and b are integers such that $a \equiv 11 \pmod{17}$, and $b \equiv 3 \pmod{17}$. Find the integer c with $0 \leq c \leq 16$ such that satisfies these.

(a) $c \equiv 13a \pmod{17}$

(b) $c \equiv 5a \pmod{17}$

(c) $c \equiv -2a + 3b \pmod{17}$

(d) $c \equiv ab \pmod{17}$

(e) $c \equiv a^3 - b^2 \pmod{17}$

6. Evaluate these quantities.

(a) $-15 \bmod 6$

(b) $-2 \bmod 12$

(c) $124 \bmod 7$

(d) $244 \bmod 24$

(e) $1245 \bmod 3$

7. Find a primitive root modulo 11.

8. What is the congruence class of 5 modulo 7?

9. Compute the addition and multiplication table for $\mathbb{Z}_5$.

10. Let a, b, c be integers, where $a, b \neq 0$. Follow the steps below to prove that if $a \mid b$ and $b \mid c$, then $a \mid c$.

(a) First, we recognize what we want to prove is in the form "if p, then q." What is hypothesis p? What is the conclusion q?

(b) When we want to prove a statement of the form "if p, then q" directly, we assume p is true and try to show q. This is commonly where we set some notation, for example, ***Suppose*** a, b, c ***are integers such that*** $a \mid b$ ***and*** $b \mid c$***.***

(c) Next we need to recall what $x \mid y$ or *divides* means. It should have something to do with the existence of an integer. We say $x \mid y$ or x *divides* y if there exists an integer k such that

(d) Now apply the definition to our situation where we have $a \mid b$ and $b \mid c$. e.g., ***Then, there exist(s) ….***

(e) Check back above to be sure that the two integers whose existence is guaranteed have different names. They need different names because they need not be the same integer.

(f) Now look back to the goal q that we set above. We should have something like "$a \mid c$". Use the definition to re-express this goal. ***We want to show there exists an integer … such that ….***

(g) Now use the two equations from (d) to produce the integer whose existence we want from (f). Use equations or complete sentences. Don't just write sentence fragments of expressions.

(h) Draw the conclusion.

11. Follow the steps below to prove that $n^3 - n$ is divisible by 3 for every integer n.

(a) Remember that divisibility statements can be rewritten as congruence conditions. e.g., a is divisible by b can be written as $a \equiv 0 \pmod{b}$. What is the congruence condition for this problem?

(b) Rewrite the above, alerting the proof reader that our statement is equivalent to the desired statement. Be sure to specify that we will show it for every integer n. e.g., ***It suffices to show ... for every ...***

(c) When we want to prove a universal statement, we can fix a generic one to consider. e.g., ***Let n be an integer.***

(d) Turn this into a finite check by looking at each congruence class separately. e.g., ***There are ... cases to consider.***

(e) For each case, write down the assumption that n is in the congruence class under consideration. Use properties of modular arithmetic from class to prove the result in that case. Repeat for each case.

12. Let a, b, c, d, m be integers, with $m > 1$. Follow the steps below to prove that if $a \equiv b \pmod{m}$ and $c \equiv d \pmod{m}$, then $a + c \equiv b + d \pmod{m}$.

(a) First, recognize what we want to prove is in the form "if p, then q." What is p? What is q?

(b) As before, to prove a statement of the form "if p, then q" directly, we assume p is true and try to show q. This is commonly where we set some notation as well. ***Suppose .... We want to show ....***

(c) Above are some statements about congruences. Use the definition or theorems to state our assumption and our goal in terms of divisibility or existence of integers.

(d) Use arithmetic to get from the assumptions to the goal.

(e) Draw the conclusion.

13. Prove that if n is an odd, positive integer, then $n^2 \equiv 1 \pmod{8}$.

14. Find a counterexample to the following statement about congruences.

> Let a, b, c, and m be integers with $m > 2$. If $ac \equiv bc \pmod{m}$, then $a \equiv b \pmod{m}$.

5.2 Integer representations and applications

Outcomes

1. Convert between representations of integers in different bases, including binary and hexadecimal representations.
2. Apply basis representation to text encoding.
3. Compute powers with fast exponentiation.

5.2.1 Integer representations

We typically use **decimal** (base 10) notation to represent integers. For example,253 means $2 \cdot 10^2 + 5 \cdot 10^1 + 3 \cdot 10^0$. The choice of base 10 is most likely due to the fact that we have 10 fingers.

There is no good mathematical reason to use base 10, and in fact we can use any integer $b > 1$ as a base.

Working with other bases is important. By abstracting to a general base b, we find that we understand the usual decimal operations better. Working in other bases also has applications.

For example, **binary** (base 2) expansions are used to develop fast exponentiation techniques. This is a vital component of modern computer encryption schemes necessary for today's digital world. Without secure means of communication, things such as online shopping would be impossible. In the next section we use base 256 for ASCII encoding, a character encoding standard for electronic communication.

Theorem 5.2.1.1 *Let $b \in \mathbb{N} - \{1\}$ and $a \in \mathbb{N} \cup \{0\}$. Then there are $n \in \mathbb{N} \cup \{0\}$ and $a_i \in \mathbb{Z}_b$ where $0 \leq i \leq n$ such that*

$$a = \sum_{i=0}^{n} a_i \cdot b^i.$$

Proof. We prove the theorem by induction.

For $n \in \mathbb{N}$ let $P(n)$ be "for $a < b^n$ there are $a_0, \ldots, a_{n-1} \in \mathbb{Z}_b$ such that $a = \sum_{i=0}^{n-1} a_i \cdot b^i$".

Let $n = 1$ and $a < b^1 = b$. Then $a \in \mathbb{Z}_b$. Set $a_0 = a$. Then $\sum_{i=0}^{1-1} a_i \cdot b^i = a_0 = a$. Thus $P(1)$ is true.

Fix $k \in \mathbb{N}$. Suppose $P(k)$ is true, that is "for $a < b^k$ there are $a_0, \ldots, a_{k-1} \in \mathbb{Z}_b$ such that $a = \sum_{i=0}^{k-1} a_i \cdot b^i$".

Let $a \in \mathbb{N}$ with $a < b^{k+1}$. By the division algorithm there are integers q and r with $a = bq + r$ and $0 \leq r < b$. Then $b^{k+1} > a = bq + r \geq bq$. Division by b yields $b^k > q$.

By our induction hypothesis that $P(k)$ is true, there are $q_0, \ldots, q_{k-1} \in \mathbb{Z}_b$ such that $q = \sum_{i=0}^{k-1} q_i \cdot b^i$. Thus

$$a = bq + r = bq + r = b\left(\sum_{i=0}^{k-1} q_i \cdot b^i\right) + r = \sum_{i=1}^{k} q_{i-1} \cdot b^i + r.$$

Setting $a_i = q_{i-1}$ and $a_0 = r$ we get

$$a = \sum_{i=1}^{k} q_{i-1} \cdot b^i + r = \sum_{i=0}^{k} a_i \cdot b^i.$$

We have shown that $P(1)$ is true and when $P(k)$ is true then $P(k+1)$ is true. Hence by the principle of mathematical inducion $P(n)$ is true for all $n \in \mathbb{N}$. ■

Such a representation of $a \in \mathbb{N} \cup \{0\}$ can be obtained with Listing 5.2.1.9.

Definition 5.2.1.2 The expression $\sum_{i=0}^{k} a_i b^i$ of the theorem is known as a **base b expansion**, denoted $(a_k a_{k-1} \ldots a_1 a_0)_b$ ♢

Definition 5.2.1.3 We call $a = \sum_{i=0}^{k} a_i b^i$ in Theorem 5.2.1.1 the base b expansion of a. We also write $a = (a_n, a_{n-1}, \ldots, a_2, a_1, a_0)_b$ and call this the base b representation of a. The parenthesis and commas in the base b representation are commonly omitted. ♢

Example 5.2.1.4 Decimal representation. We give numbers in base b and decimal (base 10) representation.

- $10_2 = 1 \cdot 2^1 + 0 \cdot 2^0 = 2$
- $1001_2 = 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 8 + 1 = 9$

- $253_6 = 2 \cdot 6^2 + 5 \cdot 6^1 + 3 \cdot 6^0 = 2 \cdot 36 + 5 \cdot 6 + 3 \cdot 1 = 105$

□

Example 5.2.1.5 Base b representation. Let $b \in \mathbb{N} - \{1\}$.

- $0_b = 0$
- $1_b = 1$
- $10_b = b$
- $100_b = b^2$
- $1000_b = b^3$

□

The positional values of a n digit base b number are

$$b^{n-1},\, b^{n-2}, \ldots,\, b^2,\, b^1,\, b^0.$$

For base b with $2 \leq b \leq 36$ we use the first b characters from

$$0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ A\ B\ C\ D\ E\ F\ G\ \ldots\ Z$$

for the base b digits in the base b representation, where the value of A is 10, the value of B is 11 and so on.

Example 5.2.1.6 Hexadecimal expansion. We give base 16 examples.

- $A_{16} = 10 \cdot 16^0 = 10$
- $A0_{16} = 10 \cdot 16 + 0 \cdot 16^0 = 160$
- $A1_{16} = 10 \cdot 16 + 1 \cdot 16^0 = 161$
- $1A_{16} = 1 \cdot 16 + 10 \cdot 16^0 = 26$
- $D_{16} = 13 \cdot 16^0 = 13$
- $DADA_{16} = 13 \cdot 16^3 + 10 \cdot 16^2 + 13 \cdot 16^1 + 10 \cdot 16^0 = 56026$
- $FF_{16} = 15 \cdot 16^1 + 15 \cdot 16^0 = 255$

□

Example 5.2.1.7 Hexadecimal to decimal representation.

$$\begin{aligned}(1FAD)_{16} &= 1 \cdot 16^3 + F \cdot 16^2 + A \cdot 16^1 + D \cdot 16^0 \\ &= 1 \cdot 16^3 + 15 \cdot 16^2 + 10 \cdot 16^1 + 13 \cdot 16^0 \\ &= 8109.\end{aligned}$$

□

How do we go the other way? Namely, given an integer n in decimal form and a positive base b, how can we compute the base b-expansion of n?

We can compute the base b expansion using div and mod. See the following examples. The first writes out the steps. The second is more condensed, given as a table.

Example 5.2.1.8 Decimal to binary representation. Compute the **binary** (base 2) expansion of 67.

$$67 = 33 \cdot 2 + \boxed{1}$$

$$\begin{aligned} 33 &= 16 \cdot 2 + \boxed{1} \\ 16 &= 8 \cdot 2 + \boxed{0} \\ 8 &= 4 \cdot 2 + \boxed{0} \\ 4 &= 2 \cdot 2 + \boxed{0} \\ 2 &= 1 \cdot 2 + \boxed{0} \\ 1 &= 0 \cdot 2 + \boxed{1}. \end{aligned}$$

The boxed terms give the base 2 expansion of 67 as

$$67 = (1000011)_2.$$

□

In general, to convert $a \in \mathbb{N}$ to base b representation, we proceed as follows. Write:

$$\begin{aligned} a &= q_0 \cdot b + r_0 \text{ where } r_0 \in \mathbb{Z}_b \\ q_0 &= q_1 \cdot b + r_1 \text{ where } r_1 \in \mathbb{Z}_b \\ q_1 &= q_2 \cdot b + r_2 \text{ where } r_2 \in \mathbb{Z}_b \\ &\vdots \\ q_{n-3} &= q_{n-2} \cdot b + r_{n-2} \text{ where } r_{n-2} \in \mathbb{Z}_b \\ q_{n-2} &= q_{n-1} \cdot b + r_{n-1} \text{ where } r_{n-1} \in \mathbb{Z}_b. \\ \text{Until } q_{n-1} &= 0 \end{aligned}$$

Then $a = \sum_{i=0}^{n-1} r_i \cdot b^i$, that is,

$$a = (((\ldots(r_{n-1} \cdot b + r_{n-2}) \cdot b + \ldots r_2) \cdot b + r_1) \cdot b + r_0.$$

The function Listing 5.2.1.9 assumes that $\mathsf{a} \in \mathbb{N} \cup \{0\}$ and $\mathsf{b} \in \mathbb{N} - \{1\}$.

Listing 5.2.1.9

```
def base_b_expansion(b,a):
  """
  The list L of coefficients of the base b
  expansion of a, such that b = L[0]+L[1]*b+L[2]+b^2+...
  """
  L = [ ]
  while a != 0:
    q,r = divmod(a,b)
    L.append(r)
    a = q
  return L
```

Example 5.2.1.10 Decimal to binary representation. We find the base 2 representation $(r_{n-1} \ldots r_1 r_0)_2$ of 253.

$$\begin{aligned} 253 \bmod 2 &= 1 && \text{thus } r_0 = 1 \\ 253 \operatorname{div} 2 &= 126 \\ 126 \bmod 2 &= 0 && \text{thus } r_1 = 0 \\ 126 \operatorname{div} 2 &= 63 \\ 63 \bmod 2 &= 1 && \text{thus } r_2 = 1 \\ 63 \operatorname{div} 2 &= 31 \\ 31 \bmod 2 &= 1 && \text{thus } r_3 = 1 \end{aligned}$$

$$
\begin{aligned}
31 \operatorname{div} 2 &= 15 & \\
15 \bmod 2 &= 1 & \text{thus } r_4 = 1 \\
15 \operatorname{div} 2 &= 7 & \\
7 \bmod 2 &= 1 & \text{thus } r_5 = 1 \\
7 \operatorname{div} 2 &= 3 & \\
3 \bmod 2 &= 1 & \text{thus } r_6 = 1 \\
3 \operatorname{div} 2 &= 1 & \\
1 \bmod 2 &= 1 & \text{thus } r_7 = 1 \\
1 \operatorname{div} 2 &= 0 &
\end{aligned}
$$

Thus $253 = 11111101_2$. □

Example 5.2.1.11 Decimal to hexadecimal. We find the base 16 representation $(r_{n-1} \dots r_1 r_0)_{16}$ of 253.

$$
\begin{aligned}
253 \bmod 16 &= 13 & \text{thus } r_0 = 13 \\
253 \operatorname{div} 16 &= 15 & \\
15 \bmod 16 &= 15 & \text{thus } r_1 = 15 \\
15 \operatorname{div} 16 &= 0 &
\end{aligned}
$$

Thus $253 = 15 \cdot 16^1 + 13 \cdot 16^0 = FD_{16}$. □

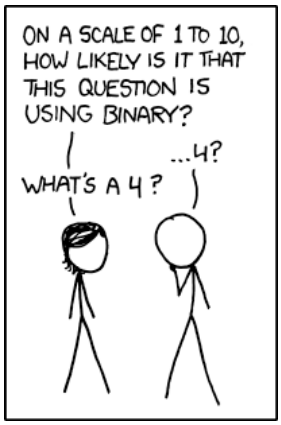


Figure 5.2.1.12 1 to 10. (https://xkcd.com/953/) If you get an 11/100 on a CS test, but you claim it should be counted as a `C', they'll probably decide you deserve the upgrade.

5.2.2 Text encoding

The **American Standard Code for Information Interchange** (ASCII) is a standard way to represent characters as numbers. For example, a space is represented by 32, a comma is 44, and a period is 46. The capital letters are also 2 digit integers, starting with 65 for A and going to 90 for Z. The Python functions `chr` and `ord` to convert the ASCII to characters. e.g., `chr(66)` returns the string A. If we want to go the other way, `ord('A')` returns the integer 65. This is known as *encoding*.

We list the printable ASCII with codes from 32 to 126.

Listing 5.2.2.1 Printables ASCII characters

```
1 for x in range(32,127):
2 print(x,":",chr(x),"␣␣",end="")
```

Example 5.2.2.2 ASCII characters and codes. Listing 5.2.2.1 above prints:

```
 32 :     33 : !   34 : "   35 : #   36 : $   37 : %   38 : &   39 : '
 40 : (   41 : )   42 : *   43 : +   44 : ,   45 : -   46 : .   47 : /
 48 : 0   49 : 1   50 : 2   51 : 3   52 : 4   53 : 5   54 : 6   55 : 7
 56 : 8   57 : 9   58 : :   59 : ;   60 : <   61 : =   62 : >   63 : ?
 64 : @   65 : A   66 : B   67 : C   68 : D   69 : E   70 : F   71 : G
 72 : H   73 : I   74 : J   75 : K   76 : L   77 : M   78 : N   79 : O
 80 : P   81 : Q   82 : R   83 : S   84 : T   85 : U   86 : V   87 : W
 88 : X   89 : Y   90 : Z   91 : [   92 : \   93 : ]   94 : ^   95 : _
 96 : `   97 : a   98 : b   99 : c  100 : d  101 : e  102 : f  103 : g
104 : h  105 : i  106 : j  107 : k  108 : l  109 : m  110 : n  111 : o
112 : p  113 : q  114 : r  115 : s  116 : t  117 : u  118 : v  119 : w
120 : x  121 : y  122 : z  123 : {  124 : |  125 : }  126 : ~
```

□

In order to encode messages longer than one character, we will view each number as a digit in a base 256 expansion of an integer M.

Example 5.2.2.3 Text encoding. Suppose I want to encode the message `Help!` using ASCII. We have $\texttt{ord}(H) = 72$, $\texttt{ord}(e) = 101$, $\texttt{ord}(l) = 108$, $\texttt{ord}(p) = 112$, $\texttt{ord}(!) = 33$. Wo the encoded message is

$$M = (72, 101, 108, 112, 33)_{256}.$$

That means

$$\begin{aligned} M &= 72 \cdot 256^4 + 101 \cdot 256^3 + 108 \cdot 256^2 + 112 \cdot 256^1 + 33 \cdot 256^0 \\ &= 310939250721. \end{aligned}$$

□

Remark 5.2.2.4 ASCII allows use to convert messages written as strings into sequences of integers. We take this a step further interpreting the sequence as the base 256 expansion of a single integer. We will use this later when we look at RSA encryption Subsection 5.5.4.

Example 5.2.2.5 Text decoding. We can **decode** a message by computing the base 256 expansion of the encoded message and decoding each character. For example, suppose the encoded message is $M = 310939249775$. The base 256 expansion of M is

$$M = (72, 101, 108, 108, 111)_{256}.$$

Then $\texttt{chr}(72) = H$, $\texttt{chr}(101) = e$, $\texttt{chr}(108) = l$, $\texttt{chr}(108) = l$, $\texttt{chr}(111) = o$.

So the decoded message is `Hello'. □

Watch the video https://youtu.be/23J2doHaJ6U for additional details about implementation.

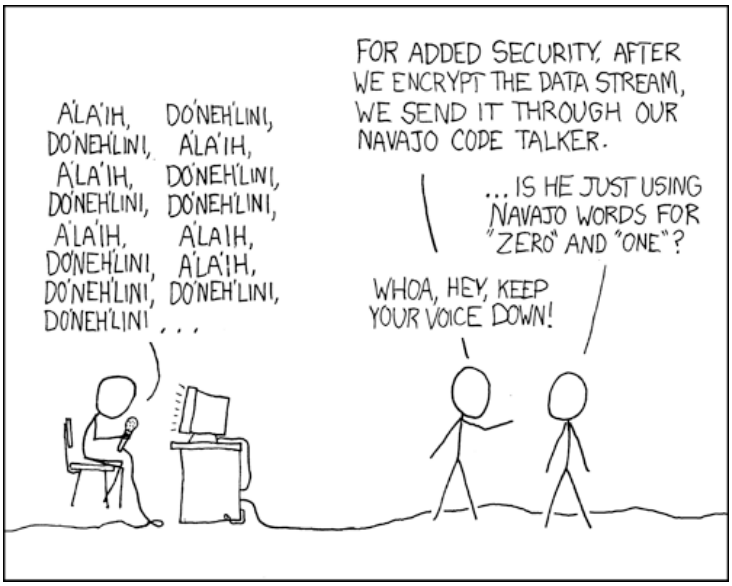


Figure 5.2.2.6 Code Talkers. (https://xkcd.com/257/) As far as I can tell, Navajo doesn't have a common word for `zero'. do-neh-lini means `neutral'.

5.2.3 Repeated Squaring

Let b, n, and m be positive integers. For cryptographic applications, it is important to compute $b^n \bmod m$ efficiently. Note that in theory, the computation is easy. For example, we can compute

$$b^n = \underbrace{b \cdot b \cdots b}_{n \text{ times}}.$$

Then divide the result by m to get the remainder. There are at least two problems with this in practice. The first is that computing b^n in this way requires $n-1$ multiplications. If n is large, this is slow. We can get around this using a method called **fast exponentiation** which we develop in this section and the following section.

Theorem 5.2.3.1 *Let $m \in \mathbb{N}$ and let $a \otimes b = (a \cdot b) \bmod m$. For $b \in \mathbb{Z}_m$ and $n \in \mathbb{N} \cup \{0\}$ we set*

$$b^{n\otimes} = (b^n) \bmod m = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ copies of } b}$$

Let $k \in \mathbb{N}$. We can compute $b^{(2^k)\otimes} = (b^{(2^k)}) \bmod m$ with k operations $\otimes$.

Example 5.2.3.2 Modular powers. Let the operation $\otimes : \mathbb{Z}_{101} \times \mathbb{Z}_{101} \to \mathbb{Z}_{101}$ be given by $a \otimes b = (a \cdot b) \bmod 101$.\\ For $b \in \mathbb{Z}_{101}$ and $n \in \mathbb{N} \cup \{0\}$ we set

$$b^{n\otimes} = (b^n) \bmod 101 = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ copies of } b}$$

We compute $2^{16\otimes} = (2^{16}) \bmod 101$ following this definition of exponentiation.

$$\begin{aligned}
2^{2\otimes} &= 2 \otimes 2 = (2 \cdot 2) \bmod 101 = 4 \\
2^{3\otimes} &= 2 \otimes 4 = (2 \cdot 4) \bmod 101 = 8 \\
2^{4\otimes} &= 2 \otimes 8 = (2 \cdot 8) \bmod 101 = 16 \\
2^{5\otimes} &= 2 \otimes 16 = (2 \cdot 16) \bmod 101 = 32
\end{aligned}$$

$$
\begin{aligned}
2^{6\otimes} &= 2 \otimes 32 = (2 \cdot 32) \bmod 101 = 64 \\
2^{7\otimes} &= 2 \otimes 64 = (2 \cdot 64) \bmod 101 = 128 \bmod 101 = 27 \\
2^{8\otimes} &= 2 \otimes 27 = (2 \cdot 27) \bmod 101 = 54 \\
2^{9\otimes} &= 2 \otimes 54 = (2 \cdot 54) \bmod 101 = 108 \bmod 101 = 7 \\
2^{10\otimes} &= 2 \otimes 7 = (2 \cdot 7) \bmod 101 = 14 \\
2^{11\otimes} &= 2 \otimes 14 = (2 \cdot 14) \bmod 101 = 28 \\
2^{12\otimes} &= 2 \otimes 28 = (2 \cdot 28) \bmod 101 = 56 \\
2^{13\otimes} &= 2 \otimes 56 = (2 \cdot 56) \bmod 101 = 112 \bmod 101 = 11 \\
2^{14\otimes} &= 2 \otimes 11 = (2 \cdot 11) \bmod 101 = 22 \\
2^{15\otimes} &= 2 \otimes 22 = (2 \cdot 22) \bmod 101 = 44 \\
2^{16\otimes} &= 2 \otimes 44 = (2 \cdot 44) \bmod 101 = 88
\end{aligned}
$$

We have computed that $2^{16\otimes} = (2^{16}) \bmod 101 = 88$ with 15 operations $\otimes$. $\square$

Example 5.2.3.3 Repeated squaring. Let the operation $\otimes : \mathbb{Z}_{101} \times \mathbb{Z}_{101} \to \mathbb{Z}_{101}$ be given by $a \otimes b = (a \cdot b) \bmod 101$.

For $b \in \mathbb{Z}_{101}$ and $n \in \mathbb{N} \cup \{0\}$ we set

$$
b^{n\otimes} = (b^n) \bmod 101 = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ copies of } b}
$$

We have

$$
\begin{aligned}
b^{2n\otimes} &= (b^{2n}) \bmod 101 = (b^n)^2 \bmod 101 \\
&= (b^n \cdot b^n) \bmod 101 = (b^{n\otimes}) \otimes (b^{n\otimes}).
\end{aligned}
$$

We compute $2^{16\otimes} = (2^{16}) \bmod 101$ with a method called repeated squaring.

$$
\begin{aligned}
2^{2\otimes} &= 2 \otimes 2 = (2 \cdot 2) \bmod 101 = 4 \\
2^{4\otimes} &= 2^{2\otimes} \otimes 2^{2\otimes} = 4 \otimes 4 = (4 \cdot 4) \bmod 101 = 16 \\
2^{8\otimes} &= 2^{4\otimes} \otimes 2^{4\otimes} = 16 \otimes 16 = (16 \cdot 16) \bmod 101 = 256 \bmod 101 = 54 \\
2^{16\otimes} &= 2^{8\otimes} \otimes 2^{8\otimes} = 54 \otimes 54 = (54 \cdot 54) \bmod 101 = 2916 \bmod 101 = 88
\end{aligned}
$$

We have computed that $2^{16\otimes} = (2^{16}) \bmod 101 = 88$ with 4 operations $\otimes$. $\square$

Theorem 5.2.3.4 *Let $m \in \mathbb{N}$ and let $a \otimes b = (a \cdot b) \bmod m$. For $b \in \mathbb{Z}_m$ and $n \in \mathbb{N} \cup \{0\}$ we set*

$$
b^{n\otimes} = (b^n) \bmod m = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ copies of } b}
$$

Let $k \in \mathbb{N}$. We can compute $b^{(2^k)\otimes} = (b^{(2^k)}) \bmod m$ with k operations $\otimes$.

Proof. For $n \in \mathbb{N}$ we have

$$
b^{2n\otimes} = (b^{2n}) \bmod m = (b^n)^2 \bmod m = (b^n \cdot b^n) \bmod m = (b^{n\otimes}) \otimes b^{n\otimes}
$$

To compute $b^{(2^k)\otimes}$ we compute

$$
\begin{aligned}
b^{2\otimes} &= b \otimes b \\
b^{2^2\otimes} &= b^{2\otimes} \otimes b^{2\otimes} \\
&\vdots
\end{aligned}
$$

$$b^{2^{k-1}\otimes} = b^{2^{k-2}\otimes} \otimes b^{2^{k-2}\otimes}$$
$$b^{2^{k}\otimes} = b^{2^{k-1}\otimes} \otimes b^{2^{k-1}\otimes}$$

which needs k operations $\otimes$. ■

The following function assumes that $\mathtt{m} \in \mathbb{N} \cup \{0\}$ and $\mathtt{b} \in \mathbb{N} - \{1\}$ and $\mathtt{k} \in \mathbb{N}$.

Listing 5.2.3.5

```
1  def repeated_squaring(b,k,m):
2    """
3  ␣␣compute␣b^(2^k)␣mod␣m
4  ␣␣"""
5    s = b # this will be the answer
6    while k != 0:
7      s = (s*s) % m
8      k = k-1
9    return a
```

Remark 5.2.3.6 When we compute $b^{(2^k)\otimes}$ using repeated squaring in the process we also compute

$$b^{(2^0)\otimes}, b^{(2^1)\otimes}, \ldots, b^{(2^{k-1})\otimes}$$

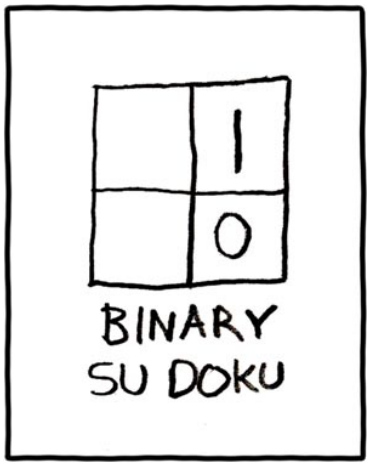


Figure 5.2.3.7 Su Doku. (https://xkcd.com/74/) This one is from the Red Belt collection, of 'medium' difficulty

5.2.4 Fast exponentiation

We will develop the algorithm by first looking at a couple of examples.

Perhaps the most important observation is that we wrote the exponent as a sum of powers of 2 and that these are exactly the exponents of the numbers that we obtain from repeated squaring. In the following example we emphasize this step by first computing the base 2 expansion of the exponent.

Example 5.2.4.1 Fast exponentiation. Compute 3^{11}.

First, we compute the binary expansion of $11 = (1011)_2$. Then

$$3^{11} = 3^{2^3+2^1+1} = 3^8 \cdot 3^2 \cdot 3.$$

Notice that we can compute 3^8 as $(3^4)^2$. Similarly, $3^4 = (3^2)^2$. We have by repeated squaring,

$$3 = 3$$

$$3^2 = 9$$
$$3^4 = 9^2 = 81$$
$$3^8 = 81^2 = 6561.$$

Thus, using the above,

$$3^{11} = 3^{2^3+2^1+1} = 3^8 \cdot 3^2 \cdot 3 = 6561 \cdot 9 \cdot 3 = 177147.$$

In this computation, we cut the number of required multiplications in half. For larger exponents n, the savings are even more extreme. □

If we compute mod after each operation the computation is more efficient. We can save even more time when we only care about the remainder modulo some integer.

Example 5.2.4.2 Modular fast exponentiation. Compute $3^{11} \bmod 7$.

Solution. As in Example 5.2.4.1, we will use the square and multiply technique to get the exponentiation faster. As before, we compute the binary expansion of $11 = (1011)_2$. Then

$$3^{11} \bmod 7 = 3^{2^3+2^1+1} \bmod 7 = (3^8 \cdot 3^2 \cdot 3) \bmod 7.$$

We make additional time and memory savings by keeping the integers small. We achieve this by reducing mod 7 at every stage, since we only want the result modulo 7.

We have by repeated squaring,

$$3 \bmod 7 = 3$$
$$3^2 \bmod 7 = 2$$
$$3^4 \bmod 7 = 2^2 \bmod 7 = 4$$
$$3^8 \bmod 7 = 4^2 \bmod 7 = 2.$$

Thus $3^{11} \bmod 7 = 2 \cdot 2 \cdot 3 \bmod 7 = 5$. □

The **fast exponentiation** algorithm allows us to compute powerefficiently compute a^n.

Given positive integers a and n with $m > 1$, compute a^n by following these steps.

1. Compute the base 2 expansion of the exponent n
2. Make a list of successive squares p_i, where $p_0 = a \bmod m$ and $p_i = p_{i-1}^2 \bmod m$ for $1 < i \leq k$.
3. Compute
$$p_0^{a_0} p_1^{a_1} \dots p_k^{a_k} \bmod m.$$

Note that when $a_i = 0$, the term $p_i^{a_i} = 1$ and so does not contribute to the product. S

When computing modulo some $m \in \mathbb{N}$ with $m \geq 2$ we keep numbers small by reducing modulo m after each multipliction.

This approach is a non-recursive version of the algorithm from Example 2.4.4.4.

Theorem 5.2.4.3 *Let $m \in \mathbb{N}$ and let $a \otimes b = (a \cdot b) \bmod m$. For $b \in \mathbb{Z}_m$ and*

$n \in \mathbb{N} \cup \{0\}$ we set

$$b^{n\otimes} = (b^n) \bmod m = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ copies of } b}$$

We can compute $b^{n\otimes} = (b^n) \bmod m$ with less than $2 \cdot k$ operations $\otimes$ if $k+1$ is the number of digits of the base 2 representation of n.

Proof. We write the base 2 expansion of n as

$$n = \sum_{i=0}^{k} r_i 2^i = r_0 \cdot 2^0 + r_1 \cdot 2^1 + \cdots + r_k \cdot 2^k$$

We have

$$\begin{aligned} b^{n\otimes} &= b^{(r_0 \cdot 2^0 + r_1 \cdot 2^1 + \cdots + r_k \cdot 2^k)\otimes} \\ &= b^{(r_0 \cdot 2^0)\otimes} \otimes b^{(r_1 \cdot 2^1)\otimes} \otimes \cdots \otimes b^{(r_k \cdot 2^k)\otimes} \\ &= \left(b^{(2^0)\otimes}\right)^{r_0} \otimes \end{aligned} \qquad (*)$$

In Theorem 5.2.3.4 we had seen that we can compute all of

$$b^{(2^0)\otimes}, b^{(2^1)\otimes}, \ldots, b^{(2^k)\otimes}$$

with k operations $\otimes$.\\ There are at most k operations $\otimes$ (outside the powers) in $(*)$. So in total we can compute $b^{n\otimes}$ in at most $2 \cdot k$ operations $\otimes$. ■

Example 5.2.4.4 Modular fast exponentiation. Let's compute $3^{67} \bmod 253$.

First we compute the base 2 expansion of $67 = (1000011)_2$ as in Example 5.2.1.8. Note that this is seven digits ($k = 6$), so we need to compute the seven successive squares $p_0, \ldots p_6$.

$$\begin{aligned} p_0 &= 3 \bmod 253 = 3 \\ p_1 &= 3^2 \bmod 253 = 9 \\ p_2 &= 9^2 \bmod 253 = 81 \\ p_3 &= 81^2 \bmod 253 = 236 \\ p_4 &= 236^2 \bmod 253 = 36 \\ p_5 &= 36^2 \bmod 253 = 31 \\ p_6 &= 31^2 \bmod 253 = 202. \end{aligned}$$

We summarize the computations so far.

i	a_i	p_i
0	1	$3 \bmod 253 = \boxed{3}$
1	1	$3^2 \bmod 253 = \boxed{9}$
2	0	$9^2 \bmod 253 = 81$
3	0	$81^2 \bmod 253 = 236$
4	0	$236^2 \bmod 253 = 36$
5	0	$36^2 \bmod 253 = 31$
6	1	$31^2 \bmod 253 = \boxed{202}$

The terms corresponding to a 1 in the binary expansion of 67 have been boxed.

It is these terms that we multiply together. It follows that

$$3^{67} \bmod 253 = 3 \cdot 9 \cdot 202 \bmod 253 = 141.$$

□

Example 5.2.4.5 Modular fast exponentiation. Let the operation $\otimes : \mathbb{Z}_{101} \times \mathbb{Z}_{101} \to \mathbb{Z}_{101}$ be given by $a \otimes b = (a \cdot b) \bmod 101$. For $b \in \mathbb{Z}_{101}$ and $n \in \mathbb{N} \cup \{0\}$ we write

$$b^{n\otimes} = \underbrace{b \otimes b \otimes \cdots \otimes b}_{n \text{ copies of } b}$$

We compute $2^{18\otimes} = (2^{18}) \bmod 101$. In a previous example we had computed:

$$\begin{aligned}
2^{2\otimes} &= 2 \otimes 2 = 4 \\
2^{4\otimes} &= 2^{2\otimes} \otimes 2^{2\otimes} = 4 \otimes 4 = 16 \\
2^{8\otimes} &= 2^{4\otimes} \otimes 2^{4\otimes} = 16 \otimes 16 = 256 \bmod 101 = 54 \\
2^{16\otimes} &= 2^{8\otimes} \otimes 2^{8\otimes} = 54 \otimes 54 = 2916 \bmod 101 = 88
\end{aligned}$$

We have computed these values in 4 operations $\otimes$. Using this we get:

$$\begin{aligned}
2^{18\otimes} &= 2^{(2+16)\otimes} \\
&= 2^{2\otimes} \otimes 2^{16\otimes} \\
&= 4 \otimes 88 \\
&= (4 \cdot 88) \bmod 101 \\
&= 352 \bmod 101 = 49
\end{aligned}$$

This took another 2 operations $\otimes$. So we computed $2^{18\otimes}$ in $4+2 = 6$ operations $\otimes$.

Computing $2^{18\otimes}$ following the definition

$$2^{18\otimes} = \underbrace{2 \otimes 2 \otimes \cdots \otimes 2}_{18 \text{ copies of } 2}$$

needs 17 operations $\otimes$. □

In our Python implementation of fast exponentiation 5.2.4.6 we go through the base 2 digits of the exponent in the same loop that we use to compute the powers $s = b^{2^n}$ of the base b.

Since, in a computer, numbers are represented in base 2 representation looping over the digits of the base 2 representation, actually means going through the digits that are stored in the computer already.

The function in Listing 5.2.4.6 assumes that $\mathtt{m} \in \mathbb{N} \cup \{0\}$ and $\mathtt{b} \in \mathbb{N} - \{1\}$ and $\mathtt{b} \in \mathbb{N} \cup \{0\}$.

Listing 5.2.4.6 Fast Exponentiation

```
def power_mod_fast(b,n,m):
  """
␣␣compute␣b^n␣mod␣m
␣␣"""
  a = 1 # this will be the answer
  s = b # b^(2^i)
  while n != 0:
    q,r = divmod(n,2)
    if r==1:
      a = a*s % m
    n = q
    s = s*s % m
   return a
```

5.2.5 Casting out nines

Theorem 5.2.5.1 *A positive integer is congruent to the sum of its decimal digits modulo* 9.

Proof. Consider the decimal (base 10) expansion of a:

$$a = \sum_{i=0}^{n} a_i 10^i$$

Then the base 10 representation of a is $a = (a_n \dots a_1 a_0)$. We take both sides modulo 9 and obtain.

$$\begin{aligned} a \bmod 9 &= \left(\sum_{i=0}^{n} a_i 10^i \right) \bmod 9 \\ &= \left(\sum_{i=0}^{n} a_i (10 \bmod 9)^i \right) \bmod 9 \end{aligned}$$

Because $10 \bmod 9 = 1$ we have

$$\begin{aligned} a \bmod 9 &= \left(\sum_{i=0}^{n} a_i 1^i \right) \bmod 9 \\ &= \left(\sum_{i=0}^{n} a_i \right) \bmod 9. \end{aligned}$$

Which means that $a \equiv \left(\sum_{i=0}^{n} a_i \right) \pmod 9$. ■

Remark 5.2.5.2 This result is the reason the **casting out nines** technique for checking arithmetic works. See the video Casting Out Nines[1] by Numberphile for additional details.

Since an integer is divisible by 9 if and only if it is congruent to 0 modulo 9, Theorem 5.2.5.1 gives the following divisibility criterion.

Corollary 5.2.5.3 Divisibility by nine. *A natural number is divisible by* 9 *if and only if the sum of its decimal digits is divisible by* 9.

[1] youtu.be/FlndIiQa20o

Proof. Let $a \in \mathbb{N}$. By Theorem 5.2.5.1 we know a is congruent to the sum of its decimal digits modulo 9. Thus a is congruent to 0 modulo 9 if and only if the sum of its decimal is congruent to 0 modulo 9. That is, a is divisible by 9 if and only if the sum of its decimal digits is divisible by 9. ■

A similar argument gives divisibility rules for 3 and 11.

5.2.6 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) base b expansion

(b) binary

(c) decimal

(d) hexadecimal

2. Compute the base 5 expansion of 253.

3. Watch the following videos

- How to count to 1000 on two hands[1] posted by 3Blue1Brown
- How high can you count on your fingers?[2] by James Tanton posted by TED-Ed ? (Spoiler: much higher than 10)

and learn to count in binary on your hands.

4. Compute the integer representing the ASCII encoding of these messages.

(a) apple

(b) Radiohead

(c) discrete math

(d) secret

5. These integers represent messages encoded in ASCII as described in Subsection 5.2.2. Decode these messages.

(a) 6582119

(b) 280991720293

(c) 311107740793

(d) 221077791181978131135567265 61

(e) The following is an integer that is too long to fit in a line.

$$19714861788808749212047758235827272137541227451278547143875494253988 30$$

6. Compute $12^{321} \bmod 456$. Follow the steps below if you get stuck.

(a) Compute the base 2 expansion of 321.

(b) Check that you got $321 = (101000001)_2$.

[1] youtu.be/1SMmc9gQmHQ
[2] youtu.be/UixU1oRW64Q

(c) Compute the successive powers p_i, remembering to reduce mod 456. Did you notice anything that helps this computation go faster?

(d) Multiply the correct terms together to compute 12^{321} mod 456.

7. Compute 7^{447} mod 645 using fast exponentiation. Show your work.

8. Compute 3^{447} mod 645 using fast exponentiation. Show your work.

9. Convert the decimal expansion of each of these integers to a binary expansion. that is , Convert base 10 to base 2.

(a) 1

(b) 156

(c) 765

(d) 23

(e) 245

10. Convert the decimal expansion of each of these integers to a hexadecimal expansion. that is , Convert base 10 to base 16.

(a) 123

(b) 45326

(c) 12

(d) 157

(e) 149987

11. Convert the binary expansion of each of these integers to a decimal expansion. that is , Convert base 2 to base 10.

(a) $(10011)_2$

(b) $(111)_2$

(c) $(101010)_2$

(d) $(111000)_2$

(e) $(1011011)_2$

12. Convert the hexadecimal expansion of each of these integers to a decimal expansion. that is , Convert base 16 to base 10.

(a) $(A123B)_{16}$

(b) $(81C)_{16}$

(c) $(ABBA)_{16}$

(d) $(DA3)_{16}$

(e) $(253)_{16}$

13. Prove that a positive integer is divisible by 5 if and only if the last digit is divisible by 5.

14. Prove that a positive integer is divisible by 3 if and only if the sum of its decimal digits is divisible by 3.

15. Prove that a positive integer is divisible by 11 if and only if the alternating sum of its decimal digits is divisible by 11.

5.3 Primes and greatest common divisors

Objectives

To introduce some fundamental concepts from number theory, including primality, prime factorization, and greatest common divisors. To introduce some important conjectures about primes.

5.3.1 Primes

Definition 5.3.1.1 An integer $p > 1$ is **prime** if the only positive factors of p are 1 and p. A positive integer greater than 1 that is not prime is called **composite** ◊

Remark 5.3.1.2 In later courses, we extend the notion of prime to all integers. For simplicity, in this course we restrict the prime versus composite distinction to integers greater than 1. Even in the extended notion, 1 is *not* prime.

Example 5.3.1.3 First primes. The first few primes are

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, \ldots.$$

□

As a stronger version of Example 3.2.1.2 we have:

Theorem 5.3.1.4 Fundamental Theorem of Arithmetic. *Every integer that is greater than* 1 *can be written uniquely as a prime or a product of two or more primes, where the primes are written in order of non-decreasing size.*

Proof. For a proof see, for example, *An inductive proof of fundamental theorem of arithmetic*[1] by Kevin Buzzard. ■

Example 5.3.1.5 Factorization.

$$\begin{aligned} 100 &= 2 \cdot 2 \cdot 5 \cdot 5 = 2^2 5^2 \\ 253 &= 11 \cdot 13 \\ 7007 &= 7 \cdot 7 \cdot 11 \cdot 13 = 7^2 \cdot 11 \cdot 13 \\ 23498357349 &= 3 \cdot 53 \cdot 397 \cdot 372263 \end{aligned}$$

□

The following Theorem says than if an integer is composite, it must have a "small" prime factor, where ``small'' means less than or equal to the square root of the number.

Theorem 5.3.1.6 *If n is composite, then n has a prime factor less than or equal to $\sqrt{n}$.*

Proof. We prove the theorem by contradiction. Let n be composite. Then there are $a \in \mathbb{N} - \{1\}$ and $b \in \mathbb{N} - \{1\}$ such that $n = a \cdot b$. Suppose $a > \sqrt{n}$ and $b > \sqrt{n}$. Then

$$a \cdot b > \sqrt{n} \cdot \sqrt{n}$$

[1] wwwf.imperial.ac.uk/~buzzard/maths/research/notes/inductive_proof_of_fundamental_theorem_of_arithmetic.pdf

$$= \left(\sqrt{n}\right)^2$$
$$= n$$

which contradicts $a \cdot b = n$. So one of a and b must be less than or equal to $\sqrt{n}$. As the prime factors of a and b are less than or equal to a respectively b, we have that n must have a prime factor that is less than or equal to $\sqrt{n}$. ■

Theorem 5.3.1.6 can be used to prove primality of certain integers. The theorem says that if an integer is composite, then it must have a small prime factor. The contrapositive says that if an integer does not have a small prime factor, then it must be prime.

Example 5.3.1.7 Primality proof. Prove that 523 is prime.

Answer. Theorem 5.3.1.6 says that if 523 is composite, it will have a prime factor less than $\sqrt{523}$. Since $\sqrt{523} \approx 22.9$, it suffices to show that no prime less than 22 divides 523. That means we need to check $2, 3, 5, 7, 11, 13, 17, 19$. A quick computation shows that

$$523 \bmod 2 = 1$$
$$523 \bmod 3 = 1$$
$$523 \bmod 5 = 3$$
$$523 \bmod 7 = 5$$
$$523 \bmod 11 = 6$$
$$523 \bmod 13 = 3$$
$$523 \bmod 17 = 13$$
$$523 \bmod 19 = 10.$$

Thus 523 is prime. □

We can use the **Sieve of Eratosthenes** to find the primes up to some bound. This is a simple, ancient algorithm for finding all prime numbers up to any given limit.

The sieve works by iteratively marking as composite (that is , not prime) the multiples of each prime, starting with the first prime number, 2.

Explore the sieving process in Figure 5.3.1.8. Click on a number to have all its multiples marked by changing the field color to red and crossing them out. Numbers that you have clicked appear on green background. When there are no white fields left, the numbers in green fields are all the prime numbers up to 192.

First click on 2, then click on the first number that remained white, which is 3 and so on. To illustrate the sieving better multiples are marked with a delay and different shades of red are used for multiples of different numbers.

Figure 5.3.1.8 Interactive Sieve of Eratosthenes up to 192. Click on a number to sieve out its multiples. Auto Sieve goes through the complete sieving process. Reset Sieve clears the sieve.

Through sieving, we could enumerate the primes up to a bound N. This number is denoted $\pi(N)$. Only 5.0848% of the positive integers less than 1000000000 are prime. See Table 5.3.1.9 for additional data. It looks like the frequency with which primes occur is diminishing. It might be reasonable to guess that eventually prime numbers become so rare that beyond a certain bound, they no longer occur. This is not the case, as was known to the Greeks over 2000 years ago. The first known proof is due to Euclid (c.~300 BC). See Theorem 5.3.1.10. It turns out, the proportion $\frac{\pi(N)}{N}$ of integers up to N that

are prime does decrease, but at an ever decreasing rate. If you are interested, you can read more about the **Prime Number Theorem** and its fascinating history on Wolfram Mathworld[2]

Table 5.3.1.9

bound (N)	number of primes $\pi(N)$	percentage ($\frac{\pi(N)}{N}\%$)
10	4	40.0000 %
100	25	25.0000%
1000	168	16.8000%
10000	1229	12.2900%
100000	9592	9.5920%
1000000	78498	7.8498%
10000000	664579	6.6458%
100000000	5761455	5.7615%
1000000000	50847534	5.0848%

Theorem 5.3.1.10 Infinitude of primes. *There are infinitely many primes.*

Proof. We use proof by contradiction. Suppose there are finitely many primes. Label them as $p_1, p_2, \ldots, p_n$. Consider the integer $q = p_1 p_2 \ldots p_n + 1$. By Example 3.2.1.2, we have that q is prime or can be expressed as a product of two or more primes. Since $q \bmod p_i = 1$ for $i = 1, 2, \ldots, n$, we have that q is not divisible by any prime. Thus q is prime. This gives the desired contradiction since q is a prime that is not on our list. Therefore, there are infinitely many primes. ■

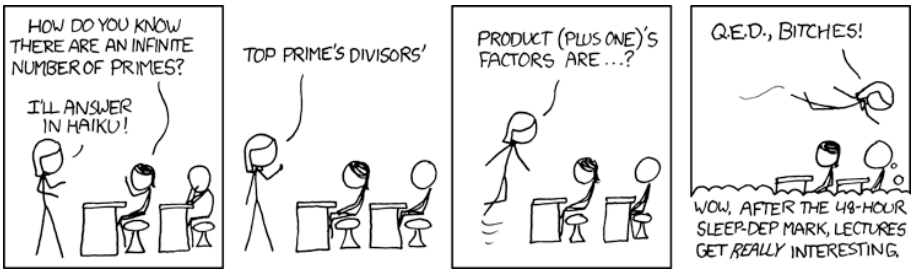


Figure 5.3.1.11 Haiku Proof. (https://xkcd.com/622/) After somewhere around 40 hours, there's no academic reason to go to the class. Only go for the hallucinations.

5.3.2 Two conjectures about primes

Though we were able to prove there are infinitely many primes, there are several open questions about primes.

Definition 5.3.2.1 *Twin primes* are pairs of primes that differ by 2. ◊

Example 5.3.2.2 Some twin primes. These are twin prime pairs:
3 and 5, 5 and 7, 11 and 13, 17 and 19 □

It is conjectured that there are infinitely many twin primes.

Conjecture 5.3.2.3 Twin prime conjecture. *There are infinitely many primes p such that $p + 2$ is also prime.*

While the twin prime conjecture is still open, there have been several amazing recent advances in this direction. See the video Gaps between Primes[1] posted by Numberphile for some of this story.

[2] mathworld.wolfram.com/PrimeNumberTheorem.html
[1] www.youtube.com/watch?v=vkMXdShDdtY

Another famous open problem in number theory concerns decomposing integers into sums of primes. In a letter to Leonhard Euler in 1742, Christian Goldbach conjectured that every odd integer n, $n > 5$, is the sum of three primes. Euler replied that this conjecture is equivalent to the conjecture that every even integer n, $n > 2$, is the sum of two primes. For example, $4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $10 = 3 + 7$, and so on.

Although no proof of the Goldbach conjecture has been found, the conjecture has been computationally checked to hold up to $4 \cdot 10^{18}$, see [4].

Conjecture 5.3.2.4 Goldbach conjecture. *Every even integer n, $n > 2$, is the sum of two primes.*

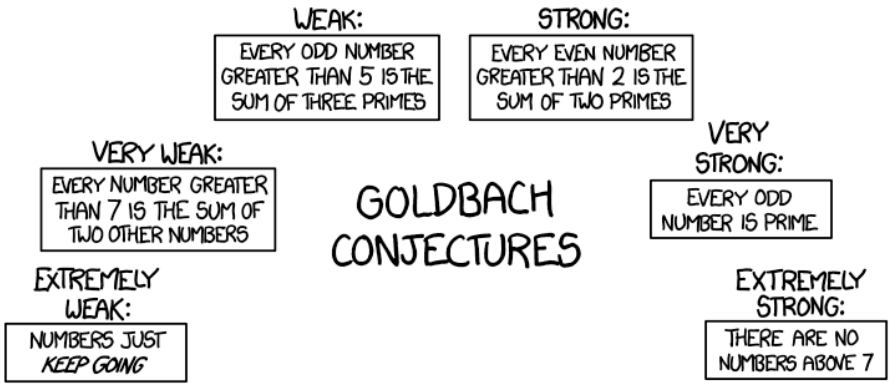


Figure 5.3.2.5 Goldbach Conjectures ({https://xkcd.com/1310/\}) The weak twin primes conjecture states that there are infinitely many pairs of primes. The strong twin primes conjecture states that every prime p has a twin prime $(p + 2)$, although $(p + 2)$ may not look prime at first. The tautological prime conjecture states that the tautological prime conjecture is true.

5.3.3 Greatest common divisor and Euclidean algorithm

Definition 5.3.3.1 Let a and b be integers, not both 0. The **greatest common divisor** of a and b, denoted $\gcd(a, b)$, is the largest integer d such that $d \mid a$ and $d \mid b$. We say a is **relatively prime** or **coprime** to b if $\gcd(a, b) = 1$. ◊

Remark 5.3.3.2 It is easy to see from the definition that $\gcd(a, b) = \gcd(b, a)$.

Example 5.3.3.3 Greatest common divisor. The greatest common divisor of 10 and 15 is 5 since 5 divides both 10 and 15, and it is the largest integer to do so. We write $\gcd(10, 15) = 5$.

The integers 15 and 22 are relatively prime since $\gcd(22, 15) = 1$. □

Example 5.3.3.4 Special cases of greatest common divisors. Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Then

1. $\gcd(a, b) = \gcd(b, a)$,
2. $\gcd(a, a) = a$, as the largest natural number that divides a is a.
3. $\gcd(a, 0) = a$, as 0 is divisible by all natural numbers a.
4. $\gcd(a, 1) = 1$, as the largest divisor of 1 is 1 and all natural numbers a are divisible by 1.

□

Definition 5.3.3.5 For a positive integer n, the **Euler phi function** or **Euler totient function** evaluated at n, denoted $\phi(n)$, is the number of non-negative integers less than n that are relatively prime to n. ◊

Example 5.3.3.6 Relatively prime to 10. There are four positive integers less than 10 that are relatively prime to 10: $\{1, 3, 7, 9\}$. Thus $\phi(10) = 4$. □

Example 5.3.3.7 Relatively prime to 7. There are six positive integers less than 7 that are relatively prime to 7: $\{1, 2, 3, 4, 5, 6\}$. Thus $\phi(7) = 6$. □

Remark 5.3.3.8 For p prime, it is easy to show that $\phi(p) = p - 1$. One can also show that for p and q distinct primes, $\phi(pq) = (p - 1)(q - 1)$.

Definition 5.3.3.9 Let a and b be positive integers. The **least common multiple** of a and b, denoted $\text{lcm}(a, b)$, is the smallest positive integer ℓ such that $a \mid \ell$ and $b \mid \ell$. ◊

For integers we can factor, or for small integers where we can use trial division, it is straightforward to compute greatest common divisors and least common multiples.

Example 5.3.3.10 Greatest common divisor and least common multiple. Let's compute the greatest common divisor and least common multiple of 24 and 30. We have that $24 = 2^3 \cdot 3$, and $30 = 2 \cdot 3 \cdot 5$. They have $2 \cdot 3$ in common, so greatest common divisor is

$$\gcd(30, 24) = 2 \cdot 3 = 6.$$

For the least common multiple, we need to include all the prime factors that arise and keep the larger exponent, so the least common multiple is

$$\text{lcm}(30, 24) = 2^3 \cdot 3 \cdot 5 = 120.$$

Note that $24 \cdot 30 = 720 = 6 \cdot 120$ so that

$$24 \cdot 30 = \gcd(30, 24)\text{lcm}(30, 24).$$

This turns out to be true in general. See Theorem 5.3.3.11. □

Theorem 5.3.3.11 *Let a and b be positive integers. Then*

$$ab = \gcd(a, b)\text{lcm}(a, b)$$

Proof. Try this at home. Hint: Use the prime factorizations of a and b guaranteed from the Fundamental Theorem of Arithmetic. Express the gcd and lcm of a and b in terms of the factorizations. Compare the product of these with the prime factorization of the product of a and b. ■

Factoring or trial division works for computing greatest common divisors, but for large a and b, it is horribly inefficient. There is a better method known as the **Euclidean algorithm** . As an additional bonus, the theorem above shows that fast computation of gcd leads to fast computation of lcm. The Euclidean algorithm works because of the following results.

Theorem 5.3.3.12 *Let g be a natural number and let a, b, and q be integers.*

1. *If g divides a and g divides b then g also divides $a - (b \cdot q)$.*
2. *If g divides $a - (b \cdot q)$ and g divides b then g also divides a.*

Proof.

1. As g divides a there exists an integer s such that $a = g \cdot s.\backslash\backslash$ As g divides

b there exists an integer t such that $b = g \cdot t$.\\ We now have

$$a - (b \cdot q) = (g \cdot s) - ((g \cdot t) \cdot q) = g \cdot (s - (t \cdot q)).$$

Thus $a - (b \cdot q)$ is a multiple of g which means that g divides $a - (b \cdot q)$.

2. As g divides $a - (b \cdot q)$ there exists an integer s such that $a - (b \cdot q) = g \cdot s$.

   As g divides b, there exists an integer s such that $b = g \cdot t$.

   We now have

   $$a = (a - (b \cdot q)) + (b \cdot q) = (g \cdot s) + ((g \cdot t) \cdot q) = g \cdot (s + (t \cdot q))$$

   Thus a is a multiple of g which means that g divides a.

■

Theorem 5.3.3.13 *Let g be a natural number and let a, b, and q be integers.*

1. $\gcd(a - (b \cdot q), b) = \gcd(a, b)$
2. $\gcd(a \bmod b, b) = \gcd(a, b)$.

Proof.

1. By the theorem above natural numbers g that divide a and b also divide $a - (b \cdot q)$.

   All natural numbers g that divide $a - (b \cdot q)$ and b also divide a.

   Thus the common divisors of a and b and the common divisors of $a - (b \cdot q)$ are the same.

   So, in particular, the greatest common divisor of a and b is equal to the greatest common divisor of $a - (b \cdot q)$ and b.

2. For $q = a \operatorname{div} q$ we have $a \bmod b = a - (b \cdot q)$.

   We get

   $$\gcd(a \bmod b, b) = \gcd(a - (b \cdot q), b) = \gcd(a, b)$$

■

It may look like these Theorems do not help, as it just turns one gcd computation into another. The real power here comes from two facts:

- $\gcd(a, b) = \gcd(b, a)$, so we can arrange that b is
- $a \bmod b$ is strictly less than b.

That means that Theorem 5.3.3.13 allows us to compute $\gcd(a, b)$ by computing $\gcd(A, B)$, where A and B are smaller than a and b. Nothing prevents us from repeatedly using this result, so we can keep using the result until we are computing $\gcd(d, 0)$, which is equal to d.

Example 5.3.3.14 Euclidean algorithm. Let's compute $\gcd(252, 198)$ using Theorem 5.3.3.13.

Since $252 \bmod 198 = 54$,

$$\gcd(252, 198) = \gcd(198, 54).$$

Since $198 \bmod 54 = 36$, we have

$$\gcd(198, 54) = \gcd(54, 36).$$

Since $54 \bmod 36 = 18$, we have

$$\gcd(54, 36) = \gcd(36, 18).$$

Since $36 \bmod 18 = 0$, we have

$$\gcd(36, 18) = \gcd(18, 0) = 18.$$

Thus the greatest common divisor of 252 and 198 is 18.

$$\gcd(252, 198) = 18.$$

□

It is quite natural to implement the Euclidean Algorithm in Python using recursion.

Listing 5.3.3.15 Recursive Euclidean algorithm

```
1  def euclid_recursive(a,b):
2    """
3  ␣␣a␣and␣b␣natural␣numbers␣with␣a>␣b
4  ␣␣compute␣gcd(a,b)
5  ␣␣"""
6    if b==0:
7      return a
8    return euclid_recursive(b,a%b)
```

For larger inputs a version with a while loop is more efficient.

Listing 5.3.3.16 Euclidean algorithm

```
1  def euclid(a,b):
2    """
3  ␣␣a␣and␣b␣natural␣numbers␣with␣a>␣b
4  ␣␣compute␣gcd(a,b)
5  ␣␣"""
6    while b!=0:
7      r = a%b
8      print(r)
9      a = b
10     b = r
11   return a
```

Theorem 5.3.3.17 Bézout's theorem. *If a and b are positive integers, there exist integers s and t such that*

$$\gcd(a, b) = as + bt.$$

By keeping track of some extra information in the Euclidean algorithm, we get the solution to the Bézout equation as well. The extended version is called **Extended Euclidean Algorithm** .

Example 5.3.3.14 shows the computation of $\gcd(252, 198)$ using Theorem 5.3.3.13. In the following example, we go through the additional bookkeeping required to solve the Bézout equation.

Example 5.3.3.18 Find integers s and t such that

$$\gcd(252, 198) = 252s + 198t.$$

Solution. We construct a table to help with the bookkeeping. We have columns q, r, s, and t. The q column keeps track of the quotients. The r column keeps track of the remainder. The s and t columns show integers such

that

$$r = 252s + 198t$$

on every row. We proceed using the Euclidean algorithm to reduce r to $\gcd(252, 198)$.

Step 1 Initialize the table.

q	r	s	t
	252	1	0
	198	0	1

Step 2 Compute the quotient

$$252 \operatorname{div} 198 = \left\lfloor \frac{252}{198} \right\rfloor = 1,$$

and enter it in the box in the q column as shown.

q	r	s	t
	252	1	0
$\boxed{1}$	198	0	1

This step may be easier to understand in a more general setting. The general pattern is repeated throughout the computation, so we go through it in more detail here. We have the table filled out as below. We want to compute the quotient q shown in a box below.

q	r	s	t
	r_1	s_1	t_1
$\boxed{q}$	r_2	s_2	t_2

To do so, we compute the quotient

$$q = r_1 \operatorname{div} r_2 = \left\lfloor \frac{r_1}{r_2} \right\rfloor,$$

and enter it in the box in the q column as shown.

Step 3 Use the value of q shown in ***bold*** to compute the next row values for r, s, and t. Compute the next r value: $252 - \boxed{1} \cdot 198 = 54$. Compute the next s value: $1 - \boxed{1} \cdot 0 = 1$. Compute the next t value: $0 - \boxed{1} \cdot 1 = -1$. Enter these three values in their respective boxes as shown.

q	r	s	t
	252	1	0
$\boxed{1}$	198	0	1
	$\boxed{54}$	$\boxed{1}$	$\boxed{-1}$

This step may be easier to understand in a more general setting. The general pattern is repeated throughout the computation, so we go through it in more detail here. We have the table filled out as below. We just computed the quotient q shown in ***bold*** below. We want to compute the boxed quantities r_3, s_3 and t_3.

q	r	s	t
	r_1	s_1	t_1
$\boxed{q}$	r_2	s_2	t_2
	$\boxed{r_3}$	$\boxed{s_3}$	$\boxed{t_3}$

To do so, we compute

$$\begin{aligned} r_3 &= r_1 - \boxed{q} r_2 \\ s_3 &= s_1 - \boxed{q} s_2 \\ t_3 &= t_1 - \boxed{q} t_2. \end{aligned}$$

You can also think of it as

$$(r_3, s_3, t_3) = (r_1, s_1, t_1) - \boxed{q}(r_2, s_2, t_2).$$

Enter these values in the boxed spots in the table as shown.

Step 4 Now the bottom two rows look like the table at the end of Step 1, so we proceed to compute the next quotient as in Step 2

$$198 \operatorname{div} 54 = \left\lfloor \frac{198}{54} \right\rfloor = 3,$$

and enter the value in the box in the q column as shown.

q	r	s	t
	252	1	0
1	198	0	1
$\boxed{3}$	54	1	-1

Step 5 Now the bottom two rows look like the table at end of Step 2, so we proceed to compute the next values of r, s, and t as in Step 3

$$(198, 0, 1) - 3(54, 1, -1) = (36, -3, 4),$$

and enter the values in their respective boxes as shown.

q	r	s	t
	252	1	0
1	198	0	1
3	54	1	-1
	$\boxed{36}$	$\boxed{-3}$	$\boxed{4}$

Step 6 Compute the next quotient

$$54 \operatorname{div} 36 = \left\lfloor \frac{54}{36} \right\rfloor = 1.$$

q	r	s	t
	252	1	0
1	198	0	1
3	54	1	-1
$\boxed{1}$	36	-3	4

Step 7 Compute the next row of r, s, and t

$$(54, 1, -1) - 1(36, -3, 4) = (18, 4, -5).$$

q	r	s	t
	252	1	0
1	198	0	1
3	54	1	-1
1	36	-3	4
	$\boxed{18}$	$\boxed{4}$	$\boxed{-5}$

Step 8 Compute the next quotient

$$36 \operatorname{div} 18 = \left\lfloor \frac{36}{18} \right\rfloor = 2.$$

q	r	s	t
	252	1	0
1	198	0	1
3	54	1	-1
1	36	-3	4
$\boxed{2}$	18	4	-5

Step 9 Compute the next row of r, s, and t. We can get by with less work in this step because in the r column, $36-2\cdot 18 = 0$. That signals us that the computation is done, and the row above is the one we want. We can signify this in the table by entering 0 in the r column and putting $-$ in the s and t columns, since those values do not matter.

q	r	s	t
	252	1	0
1	198	0	1
3	54	1	-1
1	36	-3	4
2	18	4	-5
	$\boxed{0}$	$\boxed{-}$	$\boxed{-}$

Step 10 Draw the conclusion. We have the following table computed. How do we interpret it?

q	r	s	t
	252	1	0
1	198	0	1
3	54	1	-1
1	36	-3	4
2	$\boxed{18}$	$\boxed{4}$	$\boxed{-5}$
	0	$-$	$-$

We look in the row *above* the row where we have a 0 in the r column. The r value is $\gcd(252, 198)$, and the values of s and t in that row satisfy

$$\gcd(252, 198) = 252s + 198t.$$

Thus we have shown that $\gcd(252, 198) = \boxed{18}$, and that

$$\boxed{18} = 252 \cdot \boxed{4} + 198 \cdot \boxed{-5}.$$

□

Example 5.3.3.19 We use the Extended Euclidean Algorithm to compute $\gcd(1184339, 137632)$ and to find integers s and t such that

$$\gcd(1184339, 137632) = 1184339s + 137632t.$$

We compute the table, following the steps as in Example 5.3.3.18.

q	r	s	t
	1184339	1	0
8	137632	0	1
1	83283	1	-8
1	54349	-1	9
1	28934	2	-17
1	25415	-3	26
7	3519	5	-43
4	782	-38	327
2	$\boxed{391}$	$\boxed{157}$	$\boxed{-1351}$
	0	$-$	$-$

The computation below shows that $\gcd(1184339, 137632) = \boxed{391}$ and

$$\boxed{391} = 1184339 \cdot \boxed{157} + 137632 \cdot \boxed{-1351}.$$

□

Here is Python code that will do Extended Euclidean Algorithm to find a solution to the Bézout equation.

Listing 5.3.3.20 Extended Euclidean algorithm

```
def XGCD(a,b):
  """
␣␣Return␣[d,s,t],␣where␣d␣=␣(a,b)␣and␣s,␣t␣are␣integers␣
    such
␣␣that␣d␣=␣as␣+␣bt.␣␣␣Uses␣Extended␣Euclidean␣Algorithm.
␣␣"""
  # set up first 2 rows
  r1, s1, t1, r2, s2, t2 = a, 1, 0, b, 0, 1
  r = r2
  while r!= 0:   # while remainder is not 0
    q, r = divmod(r1,r2)  # compute quotient and remainder
    s = s1 - q*s2
    t = t1 - q*t2
    # now shift everything
    r1, s1, t1, r2, s2, t2 = r2, s2, t2, r, s, t
    # want data just before remainder 0
  return r1, s1, t1
```

Lemma 5.3.3.21 $a, b, c \gcd(a, b) = 1 a \mid bca \mid c$

Proof. By Bézout's theorem, there exist integers s and t such that

$$1 = sa + tb.$$

Multiply both sides by c to get

$$c = cas + cbt.$$

It is clear that $a \mid csa$, and $a \mid bc$ by assumption, so $a \mid cbt$. Thus a divides the sum c as desired. ■

Although we cannot divide both sides of a congruence by an integer, the following result says that we can if the integer is relatively prime to the modulus.

Theorem 5.3.3.22 *Let m be a positive integer. Let a, b, and c be integers. If $ac \equiv bc \pmod{m}$, and $\gcd(c, m) = 1$, then $a \equiv b \pmod{m}$.*

Proof. Suppose $ac \equiv bc \pmod{m}$ and $\gcd(c, m) = 1$. Then $m \mid (ac - bc)$, so $m \mid c(a - b)$. Since $\gcd(c, m) = 1$, Lemma 5.3.3.21 implies $m \mid (a - b)$. Then $a \equiv b \pmod{m}$ as desired. ■

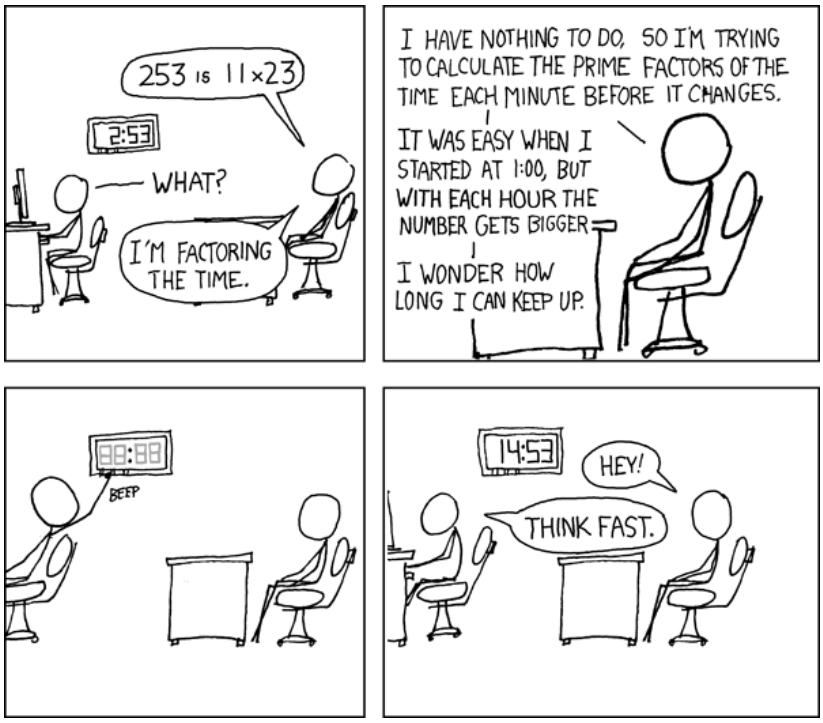


Figure 5.3.3.23 Factoring the Time.}{(https://xkcd.com/247/) I occasionally do this with mile markers on the highway.}

5.3.4 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) prime

(b) composite

(c) twin primes

(d) greatest common divisor of integers

(e) relatively prime or coprime integers

(f) least common multiple

(g) Euler phi function

2. State precisely the Fundamental Theorem of Arithmetic. Be sure to set up any notation that is required.

3. State precisely Bézout's theorem. Be sure to set up any notation that is required.

4. Compute $\phi(15)$.

5. Compute $\phi(24)$.

6. Which positive integers less than 12 are relatively prime to 12?

7. Which positive integers less than 25 are relatively prime to 25?

8. Determine whether these integers are prime.

(a) 21

(b) 100

(c) 101

(d) 253

(e) 91

9. Compute the greatest common divisor and least common multiple of the following integers.

(a) 131 and 19

(b) 260 and 77

(c) 46 and 34

(d) 132 and 192

(e) 293 and 37

(f) 15 and 87

(g) 57 and 93

10. With the Euclidean algorithm compute the greatest common divisor and least common multiple of the following integers.

(a) 36503 and 5017

(b) 8479 and 12017

(c) 15089 and 16999

(d) 11371 and 10541

(e) 14453 and 26671

11. With the Euclidean algroithm compute the greatest common divisor and least common multiple of the following integers.

(a) 3599 and 5917

(b) 9701 and 8633

(c) 23707 and 5809

(d) 8413 and 12709

(e) 19303 and 5917

12. Show that $\gcd(75, 53) = 1$. Find integers s and t such that

$$75s + 53t = 1.$$

Use these to find an integral solution to

$$75x + 53y = 13.$$

Check your solutions by plugging back in.

13. Show that $\gcd(75, 10) = 5$. Use this to show that

$$75x + 10y = 13$$

has no integral solutions.

14. Explain why there are no integers s and t such that $25s + 30t = 1$.

15. Find integers s and t such that $2018s + 253t = 1$.

16. Use the Extended Euclidean Algorithm to find integers s and t such that

$$4321s + 12367t = 149.$$

Check your solution by plugging back in.

17. Use the Extended Euclidean Algorithm to find integers s and t such that

$$5293s + 8509t = 67.$$

Check your solution by plugging back in.

18. Use the Extended Euclidean Algorithm to find integers s and t such that

$$27263s + 44377t = 199.$$

Check your solution by plugging back in. Use the Sieve of Eratosthenes to find primes up to 100. (You should find 25 primes less than 100.)

5.4 Solving congruences

Outcomes

1. Solve linear congruences
2. Solve simultaneous systems of linear congruences.

5.4.1 Linear congruences

Definition 5.4.1.1 A **linear congruence** is a congruence of the form

$$ax \equiv b \pmod{m},$$

where m is a positive integer, a and b are integers, and x is a variable. ◊

We want to solve linear congruences. As motivation, note that if we wanted to solve $ax = b$, and a has a multiplicative inverse a^{-1}, then we would just multiply both sides by a^{-1} to solve, and get $x = a^{-1}b$. e.g., The solution to $7x = 3$ is $x = \frac{3}{7}$. The same idea works for linear congruences if there exists an integer $\bar{a}$ such that $a\bar{a} \equiv 1 \pmod{m}$. Such an integer is called an *inverse* of a modulo m.

Definition 5.4.1.2 Let a and m be integers, with $m > 1$. A (multiplicative) **inverse** of a modulo m is an integer $\bar{a}$ such that $a\bar{a} \equiv 1 \pmod{m}$. ◊

Example 5.4.1.3 Modular inverses. 5 is an inverse of 7 modulo 17, since $5 \cdot 7 \equiv 1 \pmod{17}$. Note that an inverse is not unique. 22 is another inverse of 7 modulo 17, since $22 \cdot 7 \equiv 1 \pmod{17}$. Inverses can be negative. For example, -12 is an inverse of 7 modulo 17. In fact, there are infinitely many inverses of 7 modulo 17. Every integer in the congruence class of 5 modulo 17 is an inverse of 7 modulo 17

$$[5] = \{5 + 17k \mid k \in \mathbb{Z}\}.$$

□

In the example above, we see that the inverses of 7 modulo 17 form a congruence class. If an integer a has an inverse modulo m, this is what happens in general.

Theorem 5.4.1.4 *If* $\gcd(a, m) = 1$*, then the inverse of* a *modulo* m *exists. Furthermore, it is unique modulo* m*.*

Proof. Suppose $\gcd(a, m) = 1$. Then by Bézout, there exists integers s and t such that $as + mt = 1$. It follows that $mt = 1 - as$, so $as \equiv 1 \pmod{m}$. Thus s is an inverse of a modulo m.

Next, we show uniqueness. Suppose s and s' are inverses of a modulo m. We need to show $s \equiv s' \pmod{m}$. Since s and s' are inverses of a modulo m, there exists integers k and k' such that $as = 1 + km$ and $as' = 1 + k'm$. Then

$$a(s - s') = as - as' = (1 + km) - (1 + k'm) = m(k - k').$$

Then m divides $a(s - s')$ and $\gcd(a, m) = 1$, so Lemma 5.3.3.21 implies m divides $s - s'$. It follows that $s \equiv s' \pmod{m}$, as desired. ■

Using modular arithmetic, we can see a few things that will simplify computations. Suppose $\gcd(a, m) = 1$ so that a has an inverse modulo m. An inverse of a will be an inverse for every integer in the congruence class of a modulo m. We can use this to simplify our problem when $|a| > m$. Furthermore, once an inverse is found, every integer in the congruence class of the inverse is also an inverse. This is helpful when we want to find an inverse with additional properties. For example, we may want to find an inverse that is positive and "small" since we are doing computations by hand.

Example 5.4.1.5 Modular inverse of 253 modulo 8. Let's find a positive integer that is an inverse of 253 modulo 8.

Since $253 \bmod 8 = 5$, this reduces to finding an inverse of 5 modulo 8. Since $3 \cdot 5 \equiv 15 \equiv -1 \pmod{8}$, we have that -3 is an inverse of 5 modulo 8. Every integer in the congruence class $[-3]$ is also an inverse. We want a positive inverse, so we can take $-3 + 8 = 5$ as an inverse. In fact, since the inverse is unique modulo 8, we just proved that this is the smallest positive integer that is an inverse of 253 modulo 8. □

We can use the modular inverse to solve linear congruences $ax \equiv b \pmod{m}$ when $\gcd(a, m) = 1$. Namely, $x \equiv \bar{a}b \pmod{m}$. When $\gcd(a, m) \neq 1$, the solution is a bit more subtle.

For small m, the inverse of a is easy to compute by inspection. For larger m, use Extended Euclidean algorithm. Specifically, if $\gcd(a, m) = 1$, there exists integers s and t such that

$$1 = as + tm.$$

Then

$$1 \equiv as \pmod{m}.$$

5.4.2 Chinese Remainder Theorem

Theorem 5.4.2.1 Chinese Remainder Theorem. *Let* $m = m_1 m_2 \cdots m_r$

with $\gcd(m_i, m_j) = 1$ *for all* $i \neq j$. *Then the system*

$$\left\{ \begin{array}{ll} x \equiv a_1 & (\text{mod } m_1) \\ x \equiv a_2 & (\text{mod } m_2) \\ \quad \vdots & \\ x \equiv a_r & (\text{mod } m_r) \end{array} \right\}$$

has a unique solution modulo m, *given by*

$$x = a_1e_1 + a_2e_2 + \cdots + a_re_r, \quad \textit{where } e_i = w_i \cdot t_i,$$

with

$$w_i = \frac{m}{m_i} \quad \textit{and} \quad t_iw_i \equiv 1 \pmod{m_i}.$$

In other words, the solution set is exactly the congruence class of x *modulo* m,

$$[x] = \{x + km \mid k \in \mathbb{Z}\}.$$

Remark 5.4.2.2 Roughly speaking, the e_i is 1 in the mod m_i direction and 0 in the mod m_j direction for $i \neq j$.

Steps to solve Chinese Remainder Theorem> 5.4.2.1 problem:

1. Identify a_i and m_i.
2. Compute m and w_i.
3. Compute t_i, the inverse of w_i modulo m_i.
4. There is a unique solution modulo m

Example 5.4.2.3 System of two congruences. Use the Chinese Remainder Theorem> 5.4.2.1 to find all of the solutions to following the system of congruences.

$$\left\{ \begin{array}{ll} x \equiv 7 & (\text{mod } 12) \\ x \equiv 3 & (\text{mod } 5) \end{array} \right\}$$

Solution. Following the notation above, we have $a_1 = 7$, $m_1 = 12$, $a_2 = 3$, and $m_2 = 5$. Since $\gcd(12, 5) = 1$, we can use Chinese Remainder Theorem> 5.4.2.1. We are guaranteed a unique solution modulo $m = 12 \cdot 5 = 60$. It has the form

$$x \equiv 7 \cdot e_1 + 3e_2 \pmod{60}.$$

Recall that e_1 is an integer such that $e_1 \equiv 1 \pmod{12}$ and $e_1 = 0 \pmod 5$. We can compute e_1 as $e_1 = w_1t_1$, where $w_1 = m/m_1$, and t_1 is the inverse of w_1 modulo m_1. Then $w_1 = 5$. By inspection, we can take $t_1 = 5$, since $5 \cdot w_1 \equiv 1 \pmod{12}$. Then $e_1 = 5 \cdot 5 = 25$.

Similarly, $e_2 = w_2t_2$, where $w_2 = m/m_2 = 12$. We compute the inverse of 12 modulo 5, but this is the same as the inverse of 2 modulo 5, since $12 \equiv 2 \pmod 5$. By inspection, we can take $t_2 = 3$, so $e_2 = 12 \cdot 3 = 36$.

Then

$$x \equiv 7 \cdot 25 + 3 \cdot 36 \equiv 283 \equiv 43 \pmod{60}.$$

Equivalently,

$$x = 43 + 60k, \quad \text{for } k \in \mathbb{Z}.$$

□

Example 5.4.2.4 System of four congruences. Find all the solutions to

$$\left\{\begin{array}{ll} x \equiv 1 & (\text{mod } 2) \\ x \equiv 2 & (\text{mod } 3) \\ x \equiv 3 & (\text{mod } 5) \\ x \equiv 4 & (\text{mod } 11) \end{array}\right\}.$$

Solution. Since $2, 3, 5, 11$ are pairwise coprime, we can use Chinese Remainder Theorem> 5.4.2.1. There is a unique solution modulo $m = 2 \cdot 3 \cdot 5 \cdot 11 = 330$. Let $w_i = \frac{m}{m_i}$. Then

$$\begin{aligned} w_1 &= 165 \\ w_2 &= 110 \\ w_3 &= 66 \\ w_4 &= 30. \end{aligned}$$

We compute t_i such that $t_i w_i \equiv 1 \pmod{m_i}$. Then t_1 is the inverse of 165 modulo 2. We take $t_1 = 1$. Similarly, t_2 is inverse of 110 modulo 3, which is the same as the inverse of 2 modulo 3. We pick $t_2 = 2$. Similarly, we can take $t_3 = 1$ and $t_4 = 7$. Then

$$\begin{aligned} e_1 &= t_1 w_1 = 1 \cdot 165 = 165 \\ e_2 &= t_2 w_2 = 2 \cdot 110 = 220 \\ e_3 &= t_3 w_3 = 1 \cdot 66 = 66 \\ e_4 &= t_4 w_4 = 7 \cdot 30 = 210. \end{aligned}$$

Then

$$\begin{aligned} x &\equiv 1 \cdot e_1 + 2e_2 + 3e_3 + 4e_4 \pmod{330} \\ &\equiv 1 \cdot 165 + 2 \cdot 220 + 3 \cdot 66 + 4 \cdot 210 \end{aligned}$$

Thus $x = 323 + 330k$, for $k \in \mathbb{Z}$. Equivalently, $x \equiv 323 \pmod{330}$. □

Example 5.4.2.5 Action figures. My son had 500 action figures before we moved to Greensboro. He has not bought any new ones, but he lost a few in the process of the move, and he wants to know how many he has now. He can only count accurately to 10, but he knows that you are a number theorist, and he has faith in you. He reports that there is an odd number left. When you tell him that is not enough information, he reports that there is 1 left over if he lines them up 5 at a time, 2 left over if he lines the up 7 at a time, and 3 left over if he lines them up 9 at a time. How many action figures does he have?

Solution. Let x be the number of action figures my son has. Then

$$\begin{array}{lll} & x \equiv 1 \pmod{2} & \text{since } x \text{ is odd} \\ x \equiv 1 \pmod{5} & & \text{since there is 1 left over in rows of 5} \\ x \equiv 2 \pmod{7} & & \text{since there is 2 left over in rows of 7} \\ x \equiv 3 \pmod{9} & & \text{since there is 3 left over in rows of 9} \end{array}$$

Note that we can solve for x using Chinese Remainder Theorem> 5.4.2.1 since

$$\gcd(2,5) = \gcd(2,7) = \gcd(2,9) = \gcd(5,7) = \gcd(5,9) = \gcd(7,9) = 1.$$

Let's follow the steps to solve a Chinese Remainder Theorem> 5.4.2.1 problem:

1. We identify a_i and m_i. We have
2. We compute
3.
4. We compute

In other words, $x = 471 + 630k$ for some integer k. Since my son has less than 500 action figures, he must have 471 action figures. □

5.4.3 Fermat's little theorem and Euler's generalization

Theorem 5.4.3.1 Fermat's little theorem. *Let p be a prime number and a an integer. If $p \nmid a$, then $a^{p-1} \equiv 1 \pmod{p}$.*

Proof. We prove that if $1 \leq a < p$ then $a^{p-1} \equiv 1 \pmod{p}$.

First we proof that if $1 \leq x \leq p-1$ and $1 \leq y \leq p-1$ with $x \neq y$ then $(a \cdot x) \bmod p \neq (a \cdot y) \bmod p$ by contradiction.

Suppose $(a \cdot x) \bmod p = (a \cdot y) \bmod p$. Then $a \cdot (x-y) \bmod p = 0$. Because p is prime and $p \nmid a$ we have that $p \mid a(x-y)$ implies $p \mid (x-y)$. Because $1 \leq x \leq p-1$ and $1 \leq y \leq p-1$ and $x \neq y$ we have $-(p-1) < x - y < p-1$. Thus $x - y = 0$, which contradicts $x \neq y$.

Because of the above we have

$$\{1, \ldots, p-1\} = \big\{(a \cdot 1) \bmod p, \ldots, \big(a \cdot (p-1)\big) \mod p\big\}.$$

Thus

$$\left(\prod_{b=1}^{p-1} b\right) \bmod p = \left(\prod_{b=1}^{p-1} a \cdot b\right) \mod p = \left(a^{p-1} \cdot \prod_{b=1}^{p-1} b\right) \mod p.$$

Therefore $a^{p-1} \bmod p = 1$. ■

Fermat's Little Theorem 5.4.3.1 is easily generalized to all integers. For a prime number p and an integer a we obtain $a^p \bmod p = a$

Example 5.4.3.2 Small example for Little Fermat. Let $p = 5$ and $x = 2$. Then $2^{5-1} \bmod 5 = 16 \bmod 5 = 1$.

Let $p = 5$ and $x = 3$. Then $3^{5-1} \bmod 5 = 81 \bmod 5 = 1$. □

With Euler's ϕ function and the Chinese Remainder Theorem 5.4.2.1 we can generalize Fermat's Little Theorem 5.4.3.1 to composite moduli

Theorem 5.4.3.3 Euler's generalization. *Let m be an integer greater than one and let a be an integer. If $\gcd(a, m) = 1$, then $a^{\phi(m)} \equiv 1 \pmod{m}$.*

Example 5.4.3.4 Little Fermat's power. Compute $8^{222} \bmod 11$.

Solution. Note that 11 is prime and $\gcd(11, 8) = 1$. Since $11 - 1 = 10$, Fermat's little theorem implies $8^{10} \equiv 1 \pmod{11}$. We rewrite the exponent 222 as

$$222 = 22 \cdot 10 + 2$$

and use properties of exponents to get

$$8^{222} \equiv 8^{22 \cdot 10 + 2} \equiv 1^{22} \cdot 8^2 \pmod{11} \equiv 9 \pmod{11}.$$

Thus $8^{222} \bmod 11 = 9$. □

5.4.4 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) linear congruence

(b) inverse of an integer modulo an integer

2. State precisely the Chinese Remainder Theorem. Be sure to set up any notation that is required.

3. State precisely the Fermat's little theorem and Euler's generalization. Be sure to set up any notation that is required.

4. Show that 25 is an inverse of 13 modulo 36.

5. Find an inverse of 5 modulo 7 by inspection.

6. Use the Extended Euclidean Algorithm to find an inverse of 68 modulo 253.

7. Use Chinese Remainder Theorem> 5.4.2.1 to find the smallest positive solution to the system of congruences

$$\left\{ \begin{array}{ll} x \equiv 2 & (\text{mod } 3) \\ x \equiv 1 & (\text{mod } 4) \\ x \equiv 3 & (\text{mod } 5) \end{array} \right\}.$$

8. Use Fermat's Little Theorem 5.4.3.1 to compute 7^{222} mod 11.

9. Use Fermat's little theorem to compute 3^{302} mod 5.

10. Find all the solutions to

$$\left\{ \begin{array}{ll} x \equiv 2 & (\text{mod } 13) \\ x \equiv 6 & (\text{mod } 15) \end{array} \right\}.$$

What is the smallest positive solution?

11. Find all the solutions to

$$\left\{ \begin{array}{ll} x \equiv 2 & (\text{mod } 7) \\ x \equiv 3 & (\text{mod } 8) \\ x \equiv 4 & (\text{mod } 5) \end{array} \right\}.$$

What is the smallest positive solution?

12. The problems below are related. Namely, a computation in one portion may be used in more than one part.

(a) Compute $\gcd(2468, 357)$.

(b) Use the Extended Euclidean Algorithm to solve the Bézout equation for 2468 and 357. Namely, find integers s and t such that

$$\gcd(2468, 357) = 2468s + 357t.$$

(c) Compute an inverse of 357 modulo 2468, or explain why one does not exist.

13. Find an inverse of 19 modulo 241.

14. Show that 937 is an inverse of 13 modulo 2436.

15. Solve the linear congruence $357x \equiv 123 \pmod{2468}$.

16. I have an unknown number of Easter candies. When I arrange them in groups of 20, there are 3 left over. When I arrange them in groups of 41, there are 26 left over. What is the minimum number of candies that I could have?

17. Solve the congruence $2x \equiv 3 \pmod{13}$ by inspection.

18. Solve the linear congruence $19x \equiv 11 \pmod{141}$.

19. Solve the congruence $201x \equiv 5 \pmod{1357}$ using modular inverses.

20. Split up the positive integers less than 13, except 1 and 12 into pairs of integers such that each pair consists of integers that are inverses of each other modulo 13. Use this to show $12! \equiv -1 \pmod{13}$.

21. Suppose you collected shells on the beach with your daughter. When you arrange them in piles of 15, there are 13 left over. When you arrange them in piles of 19, there are 6 leftover. Use the Chinese Remainder Theorem to figure out how many shells you collected. Give the smallest positive solution.

5.5 Cryptography

Outcomes

1. Reproduce cryptographic protocols.
2. Apply both classical and modern encryption methods.

5.5.1 Shift ciphers

We want a method to encrypt, or make secret, a plain text message. One of the earliest know approaches was by Julius Caesar. The idea was to basically shift the alphabet by 3 spots.

Definition 5.5.1.1 The **Caesar cipher** is the shift cipher, where we shift forward by three. ◊

Here are the steps to encrypt using the Caesar cipher.

1. Replace each letter in the message by an element of $\mathbb{Z}_{26}$ equal to 1 less than its position in the alphabet. For example, A is replaced by 0, B is replaced by 1, \dots, Z is replaced by 25. In this step, we are **encoding**
2. Replace each number p by $(p+3)$ mod 26. Equivalently, apply the function $f\colon \mathbb{Z}_{26} \to \mathbb{Z}_{26}$ defined by $f(x) = x+3 \bmod 26$ to each number. This step does the **encrypting**.
3. Translate the numbers back to letters.

Example 5.5.1.2 Let's encrypt CAT using the Caesar cipher.

First, we change CAT to numbers, we get 2 0 19. Apply f to get 5 3 22. Translate back to letters to get FDW.

$$\text{C} \xrightarrow{\text{encode}} 2 \xrightarrow{\text{encrypt}} 5 \xrightarrow{\text{decode}} \text{F}$$
$$\text{A} \xrightarrow{\text{encode}} 0 \xrightarrow{\text{encrypt}} 3 \xrightarrow{\text{decode}} \text{D}$$
$$\text{T} \xrightarrow{\text{encode}} 19 \xrightarrow{\text{encrypt}} 22 \xrightarrow{\text{decode}} \text{W}$$

If we were going to encrypt longer messages, it would be faster to precompute a lookup table as shown in Table 5.5.1.3. □

Table 5.5.1.3

A	B	C	D	E	F	G	H	I	J	K	L	M
D	E	F	G	H	I	J	K	L	M	N	O	P
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Remark 5.5.1.4 Instead of using f, which represents a shift of 3, we can shift by any integer amount

$$f(x) = (x + k) \bmod 26$$

to yield a cipher known as a **shift cipher** . To decode, we use a function

$$g(x) = f^{-1}(x) = x - k \bmod 26.$$

More generally, the encryption function f can be any invertible function $f : \mathbb{Z}_{26} \to \mathbb{Z}_{26}$.

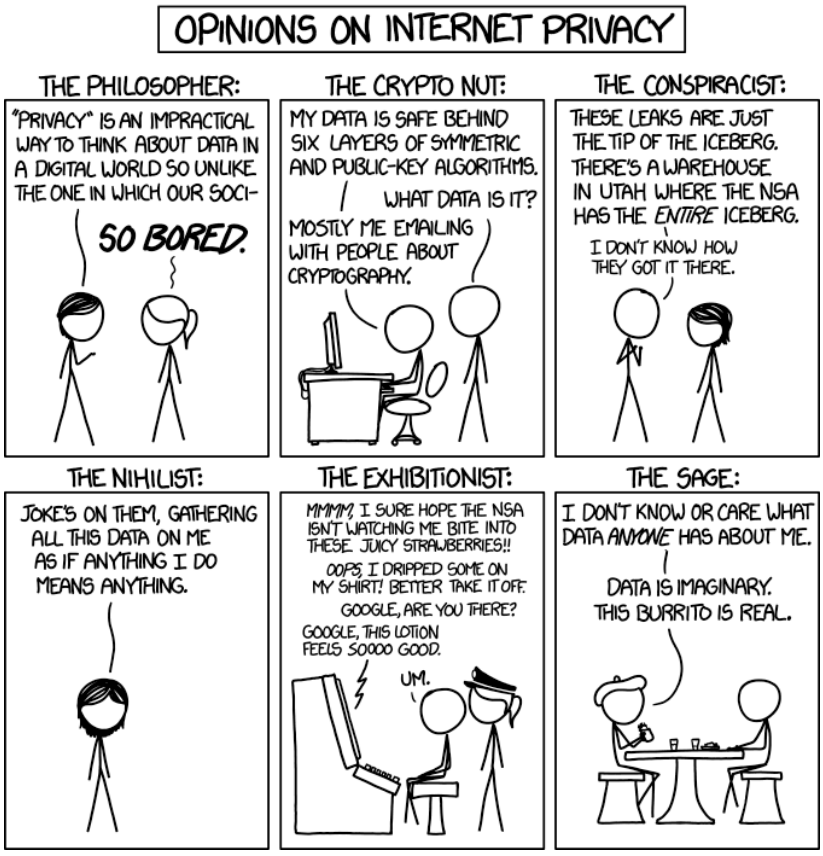


Figure 5.5.1.5 Privacy Opinions. (https://xkcd.com/1269/) I'm the Philosopher until someone hands me a burrito.

5.5.2 Diffie-Hellman

I. INTRODUCTION

WE STAND TODAY on the brink of a revolution in cryptography. The development of cheap digital hardware has freed it from the design limitations of mechanical computing and brought the cost of high grade cryptographic devices down to where they can be used in such commercial applications as remote cash dispensers and computer terminals. In turn, such applications create a need for new types of cryptographic systems which minimize the necessity of secure key distribution channels and supply the equivalent of a written signature. At the same time, theoretical developments in information theory and computer science show promise of providing provably secure cryptosystems, changing this ancient art into a science.

Figure 5.5.2.1 The first paragraph of the article *New Directions in Cryptography* by Whitfield Diffie and Martin Hellman (1976)

Let $\mathbb{Z}_p^{\otimes} = \{1, 2, 3, \ldots, p-1\}$ and let

$$\otimes : \mathbb{Z}_p^{\otimes} \times \mathbb{Z}_p^{\otimes} \to \mathbb{Z}_p^{\otimes}, \quad a \otimes b = (a \cdot b) \bmod p.$$

We have $a^{n\otimes} := a^n \bmod p$.

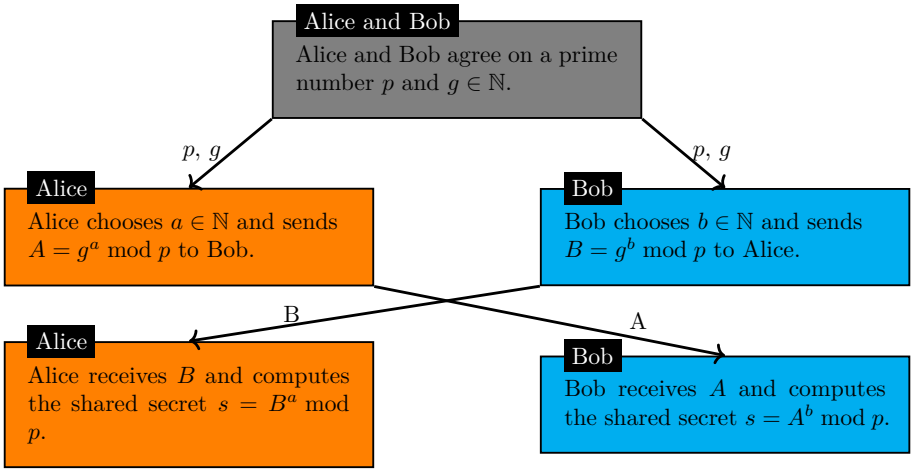


Figure 5.5.2.2 The Diffie-Hellman Key Exchange.

The agreement on p and g takes place over an insecure channel. Alice and Bob generate a shared secret s without sending the secret.

Alice computes

$$s_A = B \otimes a = \left(g^{\otimes b}\right)^{\otimes a} = g^{\otimes (a \cdot b)}.$$

Bob computes

$$s_B = A^{\otimes b} = \left(g^{\otimes a}\right)^{\otimes b} = g^{\otimes (b \cdot a)}.$$

Since $a \cdot b = b \cdot a$ Alice and Bob share the secret $s_A = s_B$.\\

Assume Eve has eavesdropped on the communication between Alice and Bob and now knows the p, g, A, and B. To obtain the shared secret s, Eve needs either Alice's secret a or Bob's secret b, which can only be obtained by finding

- the discrete logarithm of A to base g in $(\mathbb{Z}_p^{\otimes}, \otimes)$ or
- the discrete logarithm of B to base g in $(\mathbb{Z}_p^{\otimes}, \otimes)$.

So, the security of the Diffie-Hellman key exchange depends on the difficulty of computing discrete logarithms in $(\mathbb{Z}_p^{\otimes}, \otimes)$.

Remark 5.5.2.3 Instead of in $\mathbb{Z}_p^{\otimes}$ any finite cyclic group can be used. For example the group of points on a elliptic curve.

To give an idea how large the prime p should be for the Diffie-Hellman key exchange to be secure, we present an example for a discrete logarithm that was computed in 2014.

Example 5.5.2.4 Let p be the 180 decimal digit (596 digits in base 2) prime

$$p = 191147927718986609689229466631454649812986246276667354864188$$
$$503638807260703436799058776201365135161278134258296128109200$$
$$046702912984568752800330221777752773957404540495707852046983.$$

The group $(\mathbb{Z}_p^{\otimes}, \otimes)$ is generated by $g = 5$, that is $\mathbb{Z}_p^{\otimes} = \{5^{n\otimes} \mid n \in \mathbb{N}\}$. In 2014 Cyril Bouvier, Pierrick Gaudry, Laurent Imbert, Hamza Jeljeli, and Emmanuel Thom\'e announced that they computed the discrete logarithm to base 5 of

$$a = 681880801095823308798688613309985061517748546004037006257 97$$
$$299927558995162740321112260973638619757922646242302104885437$$
$$536745080299248852065080008358309735875192480724496530325927.$$

It took them under 130 core years (that is, on a single core computer it would take that long) on a parallel computing cluster to find the solution

$$n = 138670566126823584879625861326333326312363943825621039220215$$
$$583346153783336272559955521970357301302912046310782908659450$$
$$758549108092918331352215751346054755216673005939933186397777$$

such that

$$5^n \bmod p = a.$$

They applied the data computed for finding this discrete logarithm in the computation of further discrete logarithms which only required a few hours each. □

The example illustrates that a 180 decimal digits (596 digits in base 2) prime is too small for cryptographic purposes, as the discrete logarithm problem can be solved in a (relatively) short amount of time, provided enough computation

power is available. The commonly recommended size of the prime for the Diffie-Hellman key exchange is 3072 base 2 digits.

Example 5.5.2.5 Let p be the 240 decimal digit (795 digits in base 2) prime

$$p = 124620366781718784065835044608106590434820374651678805754818$$
$$78888328966680118821085503603957027250874750986476843845862 1$$
$$05486553797025393057189121768431828636284694840530161441643 0$$
$$46806687569941524699318570418303051254959437137215902928530 3$$

In 2020 Fabrice Boudot, Pierrick Gaudry, Aurore Guillevic, Nadia Heninger, Emmanuel Thom\'e, and Paul Zimmermann reported that they had computed the discrete logarithm to base 5 of

$$a = 774356626343973985966622216006087686926705588649958206166 31$$
$$71477224217061017234703519702385387550490934249 97$$

to be

$$n = 926031359281441953630949553317328555029610991914376116167294$$
$$20475898744562365366788100548099072093487548258752802923326 4$$
$$47367244150096121629264809207598195062213366889859186681126 9$$
$$28982506005127728321426751244111412371767375547225045851716$$

such that

$$5^n \bmod p = a.$$

□

Figure 5.5.2.6 Protocol. (https://xkcd.com/1323/) Changing the names would be easier, but if you're not comfortable lying, try only making friends with people named Alice, Bob, Carol, etc.

5.5.3 ElGamal

The *ElGamal encryption system* is a public key encryption algorithm created by Taher Elgamal [3] in 1985. It is based on the Diffie-Hellman key exchange.

Let $\mathbb{Z}_p^{\otimes} = \{1, 2, 3, \ldots, p-1\}$ and let

$$\otimes : \mathbb{Z}_p^{\otimes} \times \mathbb{Z}_p^{\otimes} \to \mathbb{Z}_p^{\otimes}, \quad a \otimes b = (a \cdot b) \bmod p$$

As before we have $a^{n\otimes} = a^n \bmod p$. For $s \in \mathbb{Z}_p^{\otimes}$ we write $s^{-1\otimes}$ for the inverse of s with respect to $\otimes$ such that

$$(s \otimes s^{-1\otimes}) \bmod p = 1$$

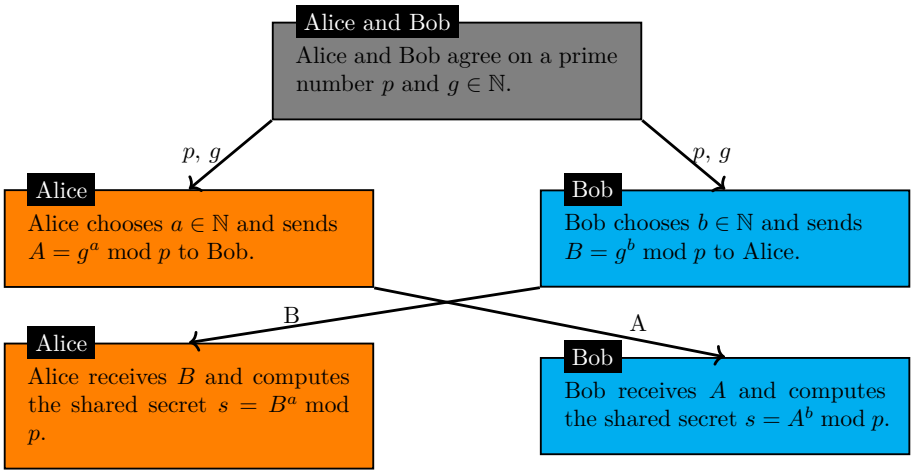


Figure 5.5.3.1 The El Gamal Cryptosystem

We now show that the message M that Bob obtains when decrypting the message is equal to Alice's plain text message m. We have

$$M = X \otimes s^{-1\otimes} = (m \otimes s) \otimes s^{-1\otimes} = m \otimes (s \otimes s^{-1\otimes}) = m \otimes 1 = m.$$

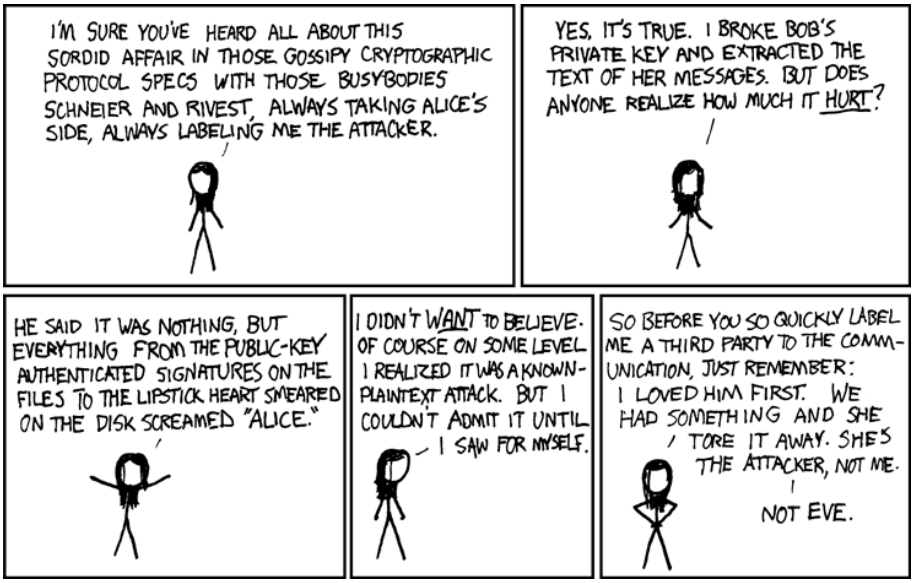


Figure 5.5.3.2 Alice and Bob. (https://xkcd.com/177/) Yet one more reason I'm barred from speaking at crypto conferences.

5.5.4 RSA

Figure 5.5.4.1 Ron Rivest, Adi Shamir, and Leonard Adleman (left to right), inventors of the RSA encryption scheme.

First, we give a bit of history. Ron Rivest, Adi Shamir, and Leonard Adleman (shown in Figure 5.5.4.1) first publicly described this algorithm for public key encryption in 1978 [9]. [1]

They posted one of the first public-key encryption messages using a 129 digit number which later became known as RSA-129 [10].

$$\begin{aligned}\text{RSA-129} =& 114381625757888867669235779976146612010218296721242362562561842935706935245733897830597123563958705058980751475992900268795435413490529510847650949147849619903898133417764638493387843990820577\\ &\cdot 327691329932667095499619881908344614131776429679929425397982885333.\end{aligned}$$

They offered a $100 prize and remarked that using technology and factoring techniques available at that time, it would take 40 quadrillion years to crack. Advances in factoring techniques and computers cracked the code in April 1994 [5] to find that the secret message was:

> The Magic Words are Squeamish Ossifrage [2]

Suppose Alice wants to send Bob an encrypted message. Bob lets her know his public key, a pair of integers (N, e).

Definition 5.5.4.2 The *RSA public encryption key* consists of a pair of integers (N, e), where N is the product of two distinct primes p and q, and the **encryption exponent** e is relatively prime to $\phi(N) = (p-1)(q-1)$. ◊

The **message space** is set of integers $\{1, 2, \cdots, N\}$. To encrypt a message M from the message space, Alice computes the integer C in the message space that satisfies

$$C \equiv M^e \pmod{N}.$$

Notice that with fast exponentiation, this is fast.

Remark 5.5.4.3 RSA encryption works when the messages are integers between 1 and N. As we saw in Subsection 5.2.2, ASCII encoding allows us to encode a message string into an integer. After the message is encoded as an integer, we can encrypt it with RSA encryption.

[1] Clifford Cocks described an equivalent system in 1973, but it was classified by the UK intelligence agency GCHQ until 1997}

[2] According to Wikipedia, Ossifrage is an older name for the lammergeier, a scavenging vulture that is famous for dropping animal bones and live tortoises onto rocks to crack them open. It might perhaps be considered among the least squeamish of creatures.

If an eavesdropper Eve captures C while it is being transmitted, Eve will have a hard time computing the original message M, since taking e th roots is a hard problem known as the **discrete log problem** .

How is it any easier for Bob? The trick is that Bob has a bit of extra information. When constructing the key, Bob chooses N to be a product of two distinct primes p and q, so that $N = pq$. The exponent e is chosen so that the greatest common divisor $\gcd(e, \phi(N)) = 1$. Since Bob knows the factorization of N, he can easily compute the Euler *phi* function $\phi(N) = (p-1)(q-1)$. Then using the Euclidean algorithm, Bob can compute an inverse to e modulo $\phi(N)$, an integer d such that $ed \equiv 1 \pmod{\phi(N)}$. By Bézout's Theorem, there is an integer k so that $ed = 1 + k\phi(N)$. Now Euler's generalization to Fermat's little theorem says that if $\gcd(C, N) = 1$, we have

$$\begin{aligned} C^d &\equiv (M^e)^d \pmod{N} \\ &\equiv M^{1+k\phi(N)} \pmod{N} \\ &\equiv M \cdot (M^{\phi(N)})^k \pmod{N} \\ &\equiv M \pmod{N}. \end{aligned}$$

In fact, this happens even if $\gcd(C, N) \neq 1$. We show this using Chinese Remainder Theorem> 5.4.2.1 in Theorem 5.5.4.6.

Thus, to decrypt the message, Bob does not need to take an e th root of C modulo N. Instead, he can raise C to the d th power and achieve the same result. Thank you, Euler! Again, with fast exponentiation, this is fast.

Note that if Eve can factor N, then she can also decrypt the message. For that reason, Bob must choose p and q "large" enough.

Definition 5.5.4.4 A positive integer d is a **decryption exponent** for RSA public key (N, e), d is an inverse of e modulo $\phi(N)$. ◊

Before proving that RSA works in general, let's look at a small example.

Example 5.5.4.5 Suppose Bob has public key $(N, e) = (55, 3)$. For the purposes of this example, pretend 55 is so large that Eve cannot factor it. Then Alice can encrypt any number from 1 to 55. Suppose Alice wants to send $M = 18$. She computes

$$C = M^e \bmod N = 18^3 \bmod 55 = 2.$$

How does Bob decrypt C? He knows that $\phi(55) = 40$, since he created the key. The decryption exponent d is the inverse if e modulo $\phi(N)$, so it satisfies $ed \equiv 1 \pmod{\phi(N)}$. Bob can compute $d \equiv -13 \pmod{40}$ by Extended Euclidean Algorithm

q	r	s	t
	40	1	0
13	3	0	1
3	1	1	−13
	0	−3	40

or by observing that $3 \cdot 13 = 39 \equiv -1 \pmod{40}$. Thus $d = -13 \bmod 40 = 27$.

To decrypt $C = 2$, Bob computes

$$M = C^d \bmod N = 2^{27} \bmod 55 = 18.$$

What are some of the problems with this example? that is , What kind of attacks should Eve try to decrypt the message?

First, the value $N = 55$ is too easy to factor. Once Eve knows $55 = 5 \cdot 11$, Chinese Remainder Theorem> 5.4.2.1 tells her that

$$\phi(55) = \phi(5)\phi(11) = (5-1)(11-1) = 40.$$

Then she can compute d using Euclidean algorithm just as Bob did to decrypt any intercepted message.

Next, the message space is too small. Notice that if Eve could solve $x^3 \equiv 2 \pmod{55}$, she can find Alice's message. The message space $\{1, 2, \ldots, 55\}$ is small enough that Eve could just compute $x^3 \bmod 55$ for several values of x and quickly find an answer. □

Theorem 5.5.4.6 *Let p and q be prime numbers with $\gcd(p, q) = 1$.*

Then for all integers C and for all $a \in \{0, \ldots, pq - 1\}$ we have $a^{1+C(p-1)(q-1)} \bmod pq = a$.

Proof. We distinguish three cases.

Case I: $p \mid a$ and $\gcd(a, q) = 1$

By Fermat's Little Theorem 5.4.3.1 $a^{(q-1)} \bmod q = 1$. Hence

$$a^{1+C(p-1)(q-1)} \bmod q = a \cdot \left(a^{q-1}\right)^{C(p-1)} = a.$$

Furthermore it follows from $a \bmod p = 0$ that $a^{1+C(p-1)(q-1)} \bmod p = 0$.

By the Chinese Remainder Theorem> 5.4.2.1 $a^{1+C(p-1)(q-1)} \bmod pq$ is determined uniquely. Therefore $a^{1+C(p-1)(q-1)} \bmod pq = a$.

Case II: $q \mid a$ and $\gcd(a, p) = 1$

Swap p and q in Case I.

Case III: $\gcd(a, pq) = 1$

By Fermat's Little Theorem 5.4.3.1 we have $a^{1+C(p-1)(q-1)} \bmod p = a$ and $a^{1+C(p-1)(q-1)} \bmod q = a$. The statement follows with the Chinese Remainder Theorem> 5.4.2.1.

■

We use Theorem 5.5.4.6 above to encrypt a message represented as an integer M as an integer X and decrypt X as shown in Figure 5.5.4.7.

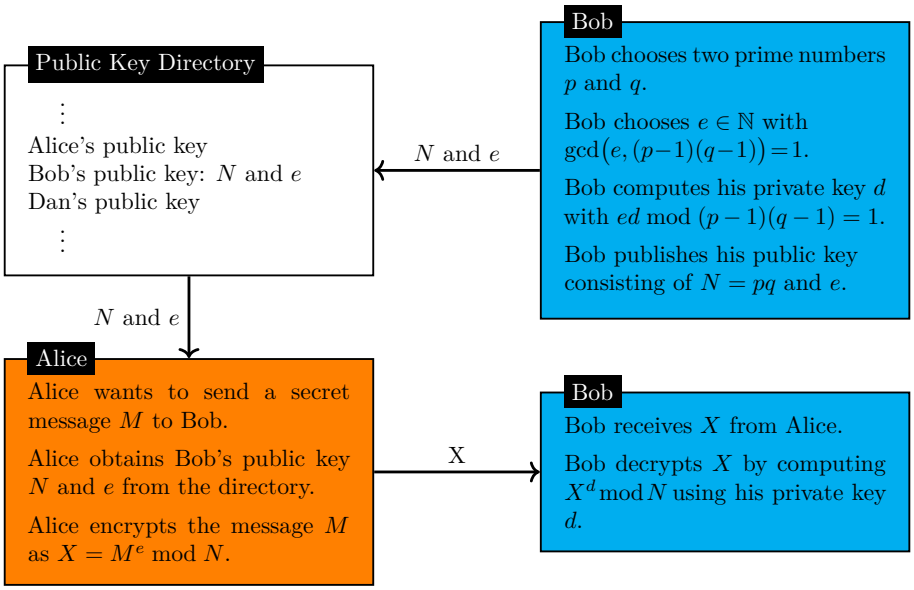


Figure 5.5.4.7 The RSA Cryptosystem. Bob creates a key pair consisting of a private and a public key (top right) and the makes the public key available, for example in a public key directory (top left). To send Bob a secret message Alice obtains Bob's public key and uses it to encrypts a message (bottom left). Bob decrypts the message using his private key (bottom right).

Remark 5.5.4.8 Bob's decription $X^d \bmod N$ in Figure 5.5.4.7 works because $(e \cdot d) \bmod \big((p-1)(q-1)\big) = 1$ and $0 \leq M < N$ and because by Theorem 5.5.4.6 we have

$$\begin{aligned} X^d \bmod n &= (m^e)^d \bmod N \\ &= M^{ed} \bmod N \\ &= M^{1+k(p-1)(q-1)} \bmod N = M \end{aligned}$$

for some non negative integer k

Why is RSA secure? What numbers should I pick? The encryption exponent e is typically chosen to be $e = 2^{16} + 1 = 65537$. This is not for security, but for speed. Because of our fast powering algorithm, this choice of e allows encryption to be done even more quickly.

If p and q are chosen to be *large* primes (for today's technology 100-200 digit primes are large enough), then the claim is that this encryption is secure. To decrypt a message, we either have to be able to take e th roots (that is , solve $x^e \equiv C \pmod N$), or compute d. The first problem is hard, and one of the main ways to try to solve it is via finding d. In order to find d, we must know $\phi(N)$. As long as Bob keeps the factors p and q secret then computing $\phi(N)$ is hard.

In the days of early commercial cryptography, many companies offered "challenges" to measure the state of progress in practical cryptanalysis. RSA used a *Factoring Challenge*. More information can be found at https://en.wikipedia.org/wiki/RSA_Factoring_Challenge.

They posted a set of eight challenge numbers, ranging in size from 576 bits (174 decimal digits) to 2048 bits (617 decimal digits) that made up the challenge [3]. Each number is the product of two large primes, similar to the

[3]The RSA numbers were generated on a computer with no network connection of any kind. The computer's hard drive was subsequently destroyed so that no record would exist,

modulus of an RSA key pair. The first person to submit a correct factorization for any of the challenge numbers was eligible for a cash prize. To date, only four of the eight challenge numbers have been factored.

A reasonable RSA key to use is 1024 bits. Their RSA-1024 has 309 decimal digits.

$$\begin{aligned} N =& 135066410865995223349603216278805969938881475605667027524485 \\ & 143851526510604859533833940287150571909441798207282164471551 \\ & 373680419703964191743046496589274256239341020864383202110372 \\ & 958725762358509643110564073501508187510676594629205563685529 \\ & 475213500852879416377328533906109750544334999811150056977236 \\ & 890927563 \end{aligned}$$

Just how large is N? There is a nice video by Vsauce explaining to see how large 52! is. If you don't have 20 minutes, skip to around the 15:56 minute mark. Even then, we are not done, since 52! is MUCH smaller than N. In fact,

$$N > (52!)^4,$$

so if we repeat the process 52! times, then repeat all of that 52! times, then repeat all of that 52! times, we are still not done. https://www.youtube.com/watch?v=ObiqJzfyACM

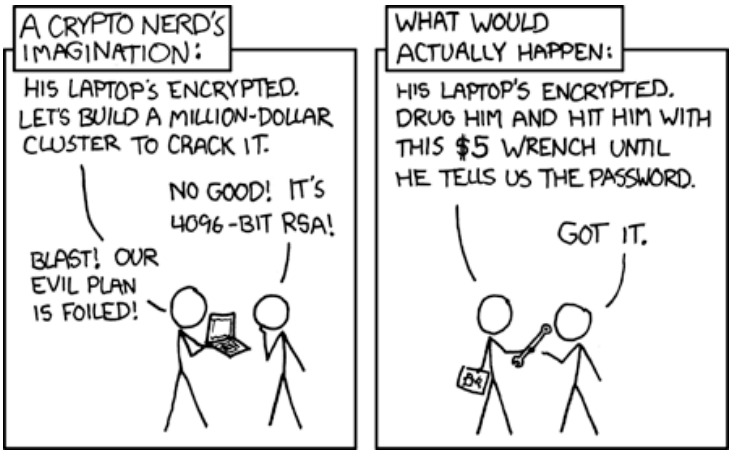


Figure 5.5.4.9 Security. (https://xkcd.com/538/) Actual actual reality: nobody cares about his secrets. (Also, I would be hard-pressed to find that wrench for \$5.)

5.5.5 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required

(a) Caesar cipher

(b) RSA public encryption key

(c) RSA encryption exponent

(d) RSA decryption exponent

anywhere, of the solution to the factoring challenge. Not even the people at RSA knew the factorizations.

2. Encrypt the message UNCG SPARTANS by translating the letters into numbers, applying the given encryption function, then translating the numbers back into letters.

(a) $f(p) = p + 3 \bmod 26$

(b) $f(p) = p + 22 \bmod 26$

(c) $f(p) = -3p \bmod 26$

(d) $f(p) = 5p + 7 \bmod 26$

3. Encrypt the message MIDNIGHT by translating the letters into numbers, applying the given encryption function, then translating the numbers back into letters.

(a) $f(p) = p + 17 \bmod 26$

(b) $f(p) = p - 4 \bmod 26$

(c) $f(p) = -7p \bmod 26$

(d) $f(p) = 3p + 12 \bmod 26$

4. Decrypt these messages that were encrypted using the Caesar cipher.

(a) KHOS

(b) VSDUWDQV

(c) VHFUHW

(d) DEVTXDWXODWH

(e) VXUUHSWLWLRXV

(f) DOLFH ORYHV ERE

5. Decrypt these messages that were encrypted using the encryption function $f(p) = 5p - 3 \bmod 26$.

(a) OGERR

(b) LKYLJLCAR

(c) XKPKNFPTJ

(d) CPC APYRJ RYR

6. Decrypt these messages that were encrypted using the encryption function $f(p) = p + 10 \bmod 26$.

(a) LVEO

(b) LKXKXK

(c) CYVSNKBSDI

(d) OFO VYFOC KVSMO

7. Suppose the cipher text HVS ZONM RCU XIADG CJSF HVS EIWQY PFCKB TCL was produced by encrypting a plain text message using a shift cipher. What is the original plain text message?

8. My RSA public key is $(N, e) = (85, 3)$. Encrypt the number $M = 23$.

9. Encrypt the message $M = 253$ using the RSA encryption scheme with public key $(N, e) = (391, 17)$.

10. Encrypt the current year using the RSA encryption scheme with public key $(N, e) = (343751, 23)$.

11. Show that the RSA public key $(N, e) = (527, 13)$ is too small by computing the decryption exponent.

12. Alice encodes her birthday as an 8-digit number `yyyymmdd`. Suppose she encrypts it using the RSA encryption scheme with public key

$$(N, e) = (25736197, 29),$$

resulting in the cipher text $C = 8141408$. What is her birthday?

Chapter 6

Relations

Phineas: … you believe in us !

Ferb: And we believe in you.

Phineas: And therefore, through the transitive property of belief, you do believe in yourself !

—Phineas and Ferb Season 2 (2010)

Relationships between sets occur in many different contexts. For example, we deal with relationships between people all the time. e.g., Alice is dating Bob. Bob is friends with Eve. Eve and Alice are sisters. We also deal with relationships between other sets. e.g., Four quarters are worth \$1. Ten dimes are worth \$1.

Relationships between elements of sets are described mathematically using a structure called a *relation*, which is just a subset of the Cartesian product of the sets. For example, in the friendship example above, friendship is the relation F and (Bob, Eve) is an element in F.

6.1 Relations and their properties

Outcomes

1. Reproduce the definition of relation and its propoerties.
2. Determine whether a relation is reflexive, symmetric. antisymmetric, or transitive.

6.1.1 Binary relations

Definition 6.1.1.1 Let A and B be sets. A **binary relation** or **relation** R from A to B is a subset of $A \times B$, $R \subseteq A \times B$. We use $a\,R\,b$ to denote the fact that (a, b) is in R, and say a is **related** to b. $\Diamond$

One can think of a relation R as defining a relationship between elements from A to B.

Example 6.1.1.2 People and foods. Let P denote the set of people on earth, and let F be the set of foods. Define a relation L from P to F by $(p, f) \in L$ if and only if p likes f. For example, $(\text{Dan}, \text{potato chips})$ is an

element of L. We can also write that as Dan L potato chips. □

Example 6.1.1.3 Cities and States. Let C be the set of names of cities in the US. Let S be the set of names of states in the US. Let R be the relation from C to S

$$R = \{(a, b) \in C \times S : a \text{ is a city in } b\}.$$

Then (Greensboro, North Carolina) and (Dallas, Texas) are both in R. □

Example 6.1.1.4 Grandfather relation. Let P denote the set of all people (alive or dead). Let G be the relation on P by xGy if and only if y is the grandfather of x. People have two grandfathers (a paternal grandfather and a maternal grandfather), so G does not define a function from P to P. Functions are not allowed to have this ambiguity. □

Example 6.1.1.5 States and cities. Relations are more general than graphs of functions. Let C be the set of names of cities in the US. Let S be the set of names of states in the US. Let R be the relation from C to S

$$R = \{(a, b) \in C \times S : a \text{ is a city in } b\}.$$

Then (Greenville, North Carolina) and (Greenville, South Carolina) are in R. This cannot happen for the graph of a function, since functions have the property that there is a unique element in the range related to each element in the domain. □

Example 6.1.1.6 Graph of a function is a relation. A function $f : A \to B$ defines a subset of $A \times B$ called the graph of the function f, namely

$$\Gamma = \{(a, b) \in A \times B : b = f(a)\}.$$

Clearly Γ is a relation from A to B. □

An example for the graph of a function can be found in Figure 2.3.1.11.

In Figure 2.3.2.6 and Figure 2.3.2.7 we give graphical representations of relations that are not the graphs of functions.

Definition 6.1.1.7 A **relation** on a set A is a relation from A to A. ◊

Example 6.1.1.8 Equality as a relation. Equality defines a relation on $\mathbb{R}$, $E \subseteq \mathbb{R} \times \mathbb{R}$.

$$E = \{(a, b) \in \mathbb{R} \times \mathbb{R} : a = b\}.$$

Since $2 = 2$, we have $2E2$. Similarly, $\pi E \pi$, $\sqrt{2} E \sqrt{2}$, etc. □

Example 6.1.1.9 Three relations on the integers. The following define relations on $\mathbb{Z}$.

1. $R = \{(a, b) \in \mathbb{Z}^2 : a \leq b\}$.
2. $S = \{(a, b) \in \mathbb{Z}^2 : a = b - 3\}$.
3. $T = \{(a, b) \in \mathbb{Z}^2 : a = (b - 12)^2\}$.

We have $4R57$ and $4S7$ and $4T10$ and $4T14$. □

6.1.2 Properties of binary relations

Definition 6.1.2.1 A relation R on a set A is **reflexive** if, for every a in A, we must have aRa ◊

Proof Technique 6.1.2.2 Show R is reflexive. *Suppose R is a relation on a set A, and we want to prove R is reflexive.*

1. *Fix a generic element a in A.*
2. *Deduce aRa.*
3. *Conclude R is reflexive.*

Definition 6.1.2.3 A relation R on a set A is **symmetric** if, for every a and b in A, whenever bRa , we must have $a\,R\,b$. ◊

Proof Technique 6.1.2.4 Show R is symmetric. *Suppose R is a relation on a set A, and we want to prove R is symmetric.*

1. *Fix generic elements a and b in A such that*
2. *Deduce bRa.*
3. *Conclude R is symmetric.*

Definition 6.1.2.5 A relation R on a set A is **antisymmetric** if, for any a and b in A, whenever $a\,R\,b$ and bRa, we must have $a = b$. ◊

Definition 6.1.2.6 A relation R on a set A is **transitive** of $a\,R\,b$ if, for every a, b, and c in A, whenever $a\,R\,b$ and bRc, we must have aRc. ◊

Proof Technique 6.1.2.7 Show R is transitive. *Suppose R is a relation on a set A, and we want to prove R is transitive.*

1. *Fix generic elements a, b, and c in A such that*
2. *Deduce aRc.*
3. *Conclude R is transitive.*

Example 6.1.2.8 Less than or equal to. Let R be the relation on $\mathbb{Z}$ defined by $\leq$, so that $a\,R\,b$ if and only if $a \leq b$.

1. Since $a \leq a$ for every integer a, we have aRa for
2. Since $1 \leq 2$, but $2 \not\leq 1$, we have $1R2$ but not
3. Whenever $a \leq b$ and $b \leq a$, we must have $a = b$.
4. Whenever $a \leq b$ and $b \leq c$, we must have $a \leq c$.

□

Example 6.1.2.9 Non-reflexive, non-antisymmetric, but transitive. Let R be the relation on $\{1, 2, 3, 4\}$ that is given by

$$R = \{(2, 2), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4)\}.$$

Then R is not reflexive since $(1, 1)$ is not in R. It is not symmetric since $(2, 4)$ is in R but $(4, 2)$ is not in R. It is not antisymmetric since $(2, 3)$ is in R and $(3, 2)$ is in R, but $2 \neq 3$. It is transitive since if (a, b) is in R and (b, c) is in R, then (a, c) is in R. □

Example 6.1.2.10 Non-reflexive, non-symmetric, non-transitive. Let R be the relation on $\{1, 2, 3, 4\}$ that is given by

$$R = \{(1, 3), (1, 4), (2, 3), (2, 4), (3, 1), (3, 4)\}.$$

Then R is not reflexive since $(1, 1)$ is not in R. It is not symmetric since $(1, 4)$ is in R, but $(4, 1)$ is not in R. It is not antisymmetric since $(1, 3)$ is in R and $(3, 1)$ is in R, but $1 \neq 3$. It is not transitive since $(1, 3)$ is in R and $(3, 1)$ is in R, but $(1, 1)$ is not in R. □

Example 6.1.2.11 Reflexive, symmetric, transitive, non-antisymmetric. Let R be the relation on $\{1, 2, 3, 4\}$ that is given by

$$R = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}.$$

Then R is reflexive since $(1, 1)$, $(2, 2)$, $(3, 3)$, and $(4, 4)$ are all in R. It is symmetric because whenever (a, b) is in R, we have (b, a) is in R. It is transitive since if (a, b) is in R and (b, c) is in R, then (a, c) is in R. It is not antisymmetric since $(1, 2)$ is in R and $(2, 1)$ is in R, but $1 \neq 2$. □

6.1.3 Combining binary relations

Since relations are sets, we can take unions, intersections, differences, and complements of relations.

Example 6.1.3.1 Students and courses. Let S be the set of students at UNCG, and let C be the set of courses at UNCG. Let T and N be the relations from S to C,

$$T = \{(s, c) \in S \times C : s \text{ has taken }\};$$
$$N = \{(s, c) \in S \times C : s \text{ needs } c \text{ to graduate}\}.$$

Then

$$T \cap N = \{(s, c) \in S \times C : s \text{ has taken } c \text{ and needs } c \text{ to graduate}\};$$

$$T - N = \{(s, c) \in S \times C : s \text{ has taken } c \text{ but does not need } c \text{ to graduate}\}.$$

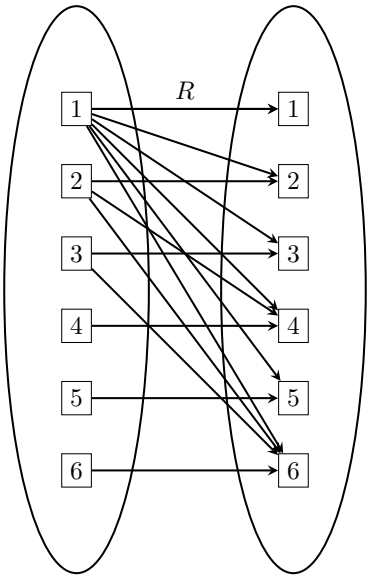
□

Like functions, we can represent relations in many ways.

Example 6.1.3.2 Relation divides. Let R be the relation on $\{1, 2, 3, 4, 5, 6\}$ given by $a\,R\,b$ if and only if a divides b. Then

$$R = \big\{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 2),$$
$$(2, 4), (2, 6), (3, 3), (3, 6), (4, 4), (5, 5), (6, 6)\big\}.$$

We can represent it graphically as



We can represent R in a table by putting an $\times$ in the a th row and b th column if and only if a divides b.

R	1	2	3	4	5	6
1	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
2		$\times$		$\times$		$\times$
3			$\times$			$\times$
4				$\times$		
5					$\times$	
6						$\times$

□

Like functions, we can also compose relations.

Definition 6.1.3.3 Let A, B, and C be sets. Suppose R is a relation from A to B, and S is a relation from B to C. The **composite** or **composition** of R and S, denoted $S \circ R$, is the relation from A to C

$$S \circ R = \{(a, c) \in A \times C : \exists b \in B \text{ such that } a\, R\, b \text{ and } b S c\}.$$

◊

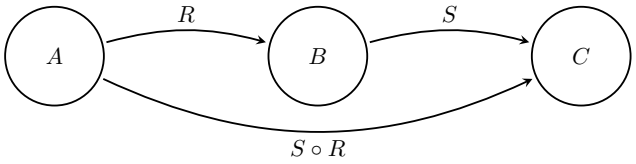


Figure 6.1.3.4 Composition of the two relations $S \subseteq B \times C$ and $R \subseteq A \times B$.

Example 6.1.3.5 Composition of relations. Let $A = \{1, 2, 3\}$, $B = \{1, 2, 3, 4\}$, and $C = \{0, 1, 2\}$. Let R be the relation from A to B

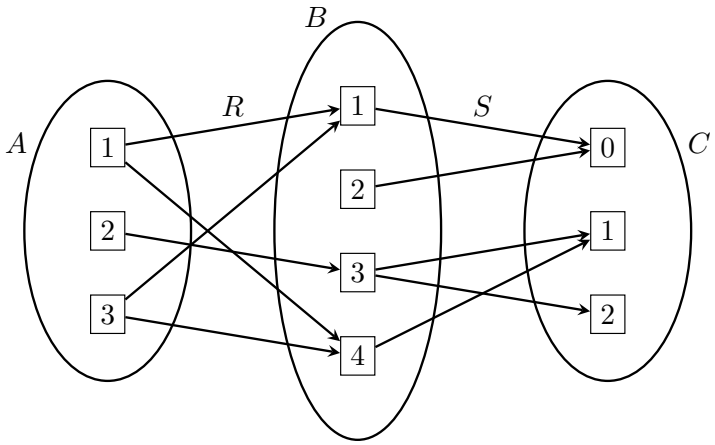
$$R = \{(1, 1), (1, 4), (2, 3), (3, 1), (3, 4)\},$$

and let S be the relation from B to C

$$S = \{(1, 0), (2, 0), (3, 1), (3, 2), (4, 1)\}.$$

Compute the composition $S \circ R$.

One approach is to view this graphically.



Then

$$S \circ R = \{(1, 0), (2, 1), (2, 2), (1, 1), (3, 0), (3, 1)\}.$$

□

6.1.4 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) binary relation from one set to another

(b) relation on a set

(c) reflexive relation

(d) symmetric relation

(e) antisymmetric relation

(f) transitive relation

(g) composition of relations

2. Consider the relation $R = \{(a, b) : a \text{ divides } b\}$ on the set $\{1, 2, 3, 4\}$.

(a) List all of the ordered pairs in R.

(b) Is R reflexive ? Explain.

(c) Is R symmetric ? Explain.

(d) Is R antisymmetric ? Explain.

(e) Is R transitive ? Explain.

3. Which of these relations on the set of all people are reflexive ? Which are symmetric ? Which are antisymmetric ? Which are transitive ? Justify.

(a) $\{(a, b) : a \text{ and } b \text{ are the same height}\}$

(b) $\{(a, b) : a \text{ and } b \text{ live in the same state}\}$

(c) $\{(a, b) : a \text{ and } b \text{ have met}\}$

(d) $\{(a, b) : a \text{ and } b \text{ are blood relatives }\}$

(e) $\{(a, b) : a \text{ and } b \text{ speak a common language}\}$

4. Which of these relations on $\{1, 2, 3, 4\}$ are reflexive ? Which are symmetric ? Which are antisymmetric ? Which are transitive ? Justify.

(a) $\{(1, 1), (2, 2), (3, 3), (4, 4)\}$

(b) $\{(1, 2), (2, 1), (1, 3), (3, 1), (1, 4), (4, 1)\}$

(c) $\{(1, 1), (2, 2), (1, 2), (2, 1), (3, 3), (4, 4)\}$

(d) $\{(1, 1), (2, 2), (1, 2), (2, 1), (1, 3), (3, 1), (3, 3), (4, 4)\}$

5. Consider the following relations R on the set of real numbers. For each of these relations, determine whether or not it is reflexive, symmetric, antisymmetric, and/or transitive. Justify.

(a) $(x, y) \in R$ if and only if $x + y = 0$

(b) $(x, y) \in R$ if and only if $x - y \in \mathbb{Z}$

[1] A person who is related to another through a common ancestor, and not by marriage or adoption.

(c) $(x, y) \in R$ if and only if $x = 253$

(d) $(x, y) \in R$ if and only if $xy = 0$

(e) $(x, y) \in R$ if and only if $x = 1$ or $y = 1$

6. Let $A = \{0, 1, 2, 3, 4\}$, and let $B = \{0, 1, 2, 3\}$. Consider the following relations from A to B. For each of these relations, list the ordered pairs in the relation.

(a) $(a, b) \in R$ if and only if $a = b$

(b) $(a, b) \in R$ if and only if $a > b$

(c) $(a, b) \in R$ if and only if $\gcd(a, b) = 1$

(d) $(a, b) \in R$ if and only if $\text{lcm}(a, b) = 2$

(e) $(a, b) \in R$ if and only if $a \mid b$

(f) $(a, b) \in R$ if and only if $a + b = 4$

7. Let R and S be the relations

$$R = \{(1, 2), (2, 3), (3, 4)\}$$
$$S = \{(1, 1), (1, 2), (2, 1), (2, 3), (3, 1), (3, 4)\}$$

Compute the following relations.

(a) $R \cup S$

(b) $R \cap S$

(c) $R - S$

(d) $S - R$

8. Let R and S be relations on $\{1, 2, 3, 4\}$ defined by

$$R = \{(2, 2), (2, 3), (3, 4), (4, 4)\}$$
$$S = \{(1, 2), (2, 1), (2, 4), (3, 1), (3, 4)\}.$$

Compute the following relations.

(a) $R \circ R$

(b) $R \circ S$

(c) $R \circ R$

(d) $S \circ S$

9. Let P be the relation on the set of people consisting of pairs (a, b), where a is a parent of b. Let S be the relation consisting of pairs (a, b), where a and b are siblings (brothers or sisters). Describe the composition relations $P \circ S$ and $S \circ P$.

10. Let B be the relation on the set of states in the US consisting of pairs (a, b) where a shares a land border with b.

(a) Give three examples of elements in B.

(b) Give three examples of elements not in B.

(c) Is B symmetric ?

(d) Is B transitive ?

11. Give an example of a relation on a set that is

(a) reflexive, but not symmetric.

(b) symmetric, but not reflexive.

(c) reflexive and symmetric, but not transitive.

6.2 Equivalence relations

Outcomes

1. Determine whether a relation is a equivalence relation
2. Determine the equivalence classes of equivalence relations

6.2.1 A special class of binary relation

Definition 6.2.1.1 A relation R on a set A is called an **equivalence relation** if R is reflexive, symmetric, and transitive. ◊

Equivalence relations are important throughout mathematics and computer science. It makes the notion of equivalent objects precise.

Definition 6.2.1.2 Two elements a and b are **equivalent** if they are related by an equivalence relation. ◊

This is a notion that makes precise when two thing are "the same" up to differences that we are willing to ignore. We have seen this before.

For example, we regularly think of a \$1 note as "the same" as four quarters, though they are different. According to the United States Mint[1], a quarter is 5.670 grams. That means \$1000 in quarters weighs about 50 pounds. On the other hand, a \$1 note weighs about 1 gram, so \$1000 in \$1 notes weighs about 2.2 pounds. Would you rather run a race carrying \$1000 in \$1 notes or quarters ?

For a more mathematical example, we think of $\frac{2}{6}$ as equal to $\frac{1}{3}$. In what sense are they the same ? In this case, two objects are equal as rational numbers, but differ in representation.

Remark 6.2.1.3 We often use $\sim$ instead for equivalence relations, so if a and b are equivalent, we might write $a \sim b$.

\ How do we prove a relation R on a set A is an equivalence relation ?

We need to verify that R is reflexive, symmetric, and transitive. Recall that these are characterized by universal conditions. To show something is true "for all x in A", we fix a generic element a in A, and show that the conditions are satisfied for the generic element a. Then since the condition is satisfied for the generic element, it is satisfied by all the elements. How do we do this in practice ?

Proof Technique 6.2.1.4 Show R is an equivalence relation. *Suppose R is a relation on a set A, and we want to prove R is an equivalence relation. The proof has three parts.*

1. *Use Proof Technique 6.1.2.2 to show R isreflexive.*
2. *Use Proof Technique 6.1.2.4 to show R is symmetric.*

[1] www.usmint.gov/learn/coin-and-medal-programs/coin-specifications

3. *Use Proof Technique 6.1.2.7 to show R is transitive.*

4. *Conclude R is an equivalence relation.*

Example 6.2.1.5 Same absolute value relation. Let R be the relation on $\mathbb{Z}$ defined by $a\,R\,b$ if and only if $|a| = |b|$.

Show that R is an equivalence relation.

Solution. We check the three defining characteristics.

Reflexive Let $a \in \mathbb{Z}$. Then $|a| = |a|$, so aRa. Thus R is reflexive.

Symmetric Let $a, b \in \mathbb{Z}$. Suppose $a\,R\,b$ so that $|a| = |b|$. Then $|b| = |a|$, so bRa. Thus R is symmetric.

Transitive Let $a, b, c \in \mathbb{Z}$. Suppose $a\,R\,b$ and bRc so that $|a| = |b|$ and $|b| = |c|$. Then $|a| = |c|$, so aRc. Thus R is transitive.

Therefore R is an equivalence relation. □

Example 6.2.1.6 Differerence is an integer relation. Define a relation R on $\mathbb{R}$ by $a\,R\,b$ if and only of $a - b \in \mathbb{Z}$.

Is R an equivalence relation ?

Solution. We check the three defining characteristics.

Reflexive Let $a \in \mathbb{R}$. Then $a - a = 0 \in \mathbb{Z}$, so aRa. Thus R is reflexive.

Symmetric Let $a, b \in \mathbb{R}$. Suppose $a\,R\,b$ so that $a - b = k \in \mathbb{Z}$. Then $b - a = -k \in \mathbb{Z}$, so bRa. Thus R is symmetric.

Transitive Let $a, b, c \in \mathbb{R}$. Suppose $a\,R\,b$ and bRc so that $a - b = k \in \mathbb{Z}$ and $b - c = \ell \in \mathbb{Z}$. Then $a - c = (a - b) + (b - c) = k + \ell \in \mathbb{Z}$ and because the sum of integers is an integer, we have aRc. Thus R is transitive.

Therefore R is an equivalence relation. □

Theorem 6.2.1.7 *Let $m > 1$ be an integer. Let R be the relation on $\mathbb{Z}$*

$$R = \{(a, b) \in \mathbb{Z}^2 : a \equiv b \pmod{m}\}.$$

Then R is an equivalence relation.

Proof. We check the three defining conditions.

Reflexive Let $a \in \mathbb{Z}$. Then $m \mid (a - a)$, so $a \equiv a \pmod{m}$. Then aRa. Thus R is reflexive.

Symmetric Let $a, b \in \mathbb{Z}$. Suppose $a\,R\,b$ so that $a \equiv b \pmod{m}$. Then $a - b = km$ for some integer k. Then $b - a = (-k)m$, so $b \equiv a \pmod{m}$. Then bRa. Thus R is symmetric.

Transitive Let $a, b, c \in \mathbb{Z}$. Suppose $a\,R\,b$ and bRc so that $a \equiv b \pmod{m}$ and $b \equiv c \pmod{m}$. Then $a = b + km$ for some integer k, and $b = c + \ell m$ for some integer ℓ. Then

$$a = b + km = (c + \ell m) + km = c + (\ell + k)m.$$

Since k and ℓ are integers, $k + \ell$ is an integer. Then $a \equiv c \pmod{m}$, and so aRc. Thus R is transitive.

Therefore R is an equivalence relation. ■

Figure 6.2.1.8 Soda Sugar Comparisons. (https-//xkcd.com/1793/) The key is portion control, which is why I've switched to eating smaller cans of frosting instead of full bottles.

6.2.2 Equivalence classes and partitions

Definition 6.2.2.1 Let R be an equivalence relation on A, and let a be an element of A. The **equivalence class** of a, denoted $[a]_R$ is the set of elements that are related to a. In other words,

$$[a] = \{b \in A : a \, R \, b\}.$$

◊

Remark 6.2.2.2 Since equivalence relations are symmetric, this is the same as

$$[a] = \{b \in A : bRa\}.$$

Example 6.2.2.3 Congruence classes. Let R denote the congruent modulo 7 equivalence relation. That is, $a \, R \, b$ when $a \equiv b \pmod 7$. Then

$$\begin{aligned}[0] &= \{\ldots, -14, -7, 0, 7, 14, \ldots\} \\ [14] &= \{\ldots, -14, -7, 0, 7, 14, \ldots\} \\ [3] &= \{\ldots, -11, -4, 3, 10, 17, \ldots\}\end{aligned}$$

We $[a] = [a']$ precisely when aRa'.

Equivalence classes of this kind are called congruence classes 5.1.2.4. □

Definition 6.2.2.4 A **partition** of a set A is a collection of non-empty subsets $A_i \subseteq A$, $i \in I$, such that

1. $A_i \cap A_j = \emptyset$ for all $i \neq j$;
2. $A = \bigcup_{i \in I} A_i$.

$\Diamond$

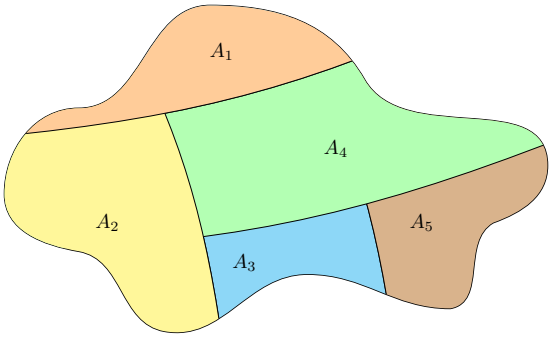


Figure 6.2.2.5 Partition of a set into the sets A_1, A_2, A_3, A_4, and A_5.

Theorem 6.2.2.6 *Let $\sim$ be an equivalence relation on A. Then for all $a, b \in A$, the following are equivalent.*

1. $a \sim b$
2. $[a] = [b]$
3. $[a] \cap [b] \neq \emptyset$

Proof. We prove $1 \to 2 \to 3 \to 1$.

$1 \to 2$	Suppose $a \sim b$. Let $c \in [a]$. Then $c \sim a$. By transitivity, we have $c \sim b$. Thus $c \in [b]$. Therefore $[a] \subseteq [b]$. Now let $c \in [b]$. Then $b \sim c$. By transitivity, we have $a \sim c$. Thus $c \in [a]$. Therefore $[b] \subseteq [a]$. Then $[a] = [b]$, as desired.
$2 \to 3$	Suppose $[a] = [b]$. Since $\sim$ is reflexive, $a \in [a]$, so $[a] \neq \emptyset$. Then $[a] \cap [b] = [a]$.
$3 \to 1$	Suppose $[a] \cap [b] \neq \emptyset$. Let $c \in [a] \cap [b]$. Then $a \sim c$ and $c \sim b$.By transitivity, we have $a \sim b$.

■

Theorem 6.2.2.7 *Let $\sim$ be an equivalence relation on a set A. Then the equivalence classes form a partition of A. Conversely, given a partition $\{A_i : i \in I\}$ of a set A, there is an equivalence relation $\sim$ that has the sets A_i as equivalence classes.*

Example 6.2.2.8 Same birthday. Consider the equivalence relation R on the set of people in the world defined by $a\,R\,b$ if and only if a has the same birthday (month and date).

Then [Thom Yorke] is the set of all people with a birthday on October 7. The equivalence relation partitions the set of all people into equivalence classes, where each equivalence class contains all the people with a given birthday. Specifically, there are 366 equivalence classes (don't forget February 29!). □

Given a partition of a set A, we compute the corresponding equivalence relation R on A using the subsets in the partition as equivalence classes.

Example 6.2.2.9 Partitions. Suppose $A = \{1, 2, 3, 4, 5, 6\}$. Given the partition $A_1 = \{1, 2\}$, $A_2 = \{3, 5, 6\}$, and $A_3 = \{4\}$, the ordered pairs in the equivalence relation R produced by this partition is

$$R = \{(1,1), (1,2), (2,2), (2,1), (3,3), (3,5), (3,6), (5,3), (5,5), (6,3), (5,6), (6,5), (6,6), (4,4)\}.$$

□

Example 6.2.2.10 Number of equivalence relations. How many different equivalence relations are there on $A = \{1, 2, 3\}$?

Solution. By Theorem 6.2.2.7, it is enough to enumerate partitions of A.

1. $A = \{1, 2, 3\}$
2. $A = \{1, 2\} \cup \{3\}$
3. $A = \{1\} \cup \{2, 3\}$
4. $A = \{1, 3\} \cup \{2\}$
5. $A = \{1\} \cup \{2\} \cup \{3\}$

Thus there are five equivalence relations on A. □

6.2.3 Exercises

1. Give the definition for these terms. Be sure to set up any notation that is required.

(a) equivalence relation

(b) equivalent elements

(c) equivalence class

(d) partition of a set

2. Define an equivalence relation on the set of restaurants in Greensboro.

3. Define an equivalence relation on the set of classes offered at UNCG.

4. Let $\sim$ be an equivalence relation on a set A, and let a and b be elements of A. Prove that if a is equivalent to b, then the equivalence class of a is equal to the equivalence class of b. In other words, show if $a \sim b$, then $[a] = [b]$.

5. Consider the relation $R = \{(a, b) : a \text{ divides } b\}$ on the set $\{1, 2, 3, 4\}$. Is R an equivalence relation ? Justify.

6. Let R be the relation on the set of all people, defined by

$$R = \{(a, b) : a \text{ and } b \text{ were born in the same month}\}.$$

Prove that R is an equivalence relation.

7. Suppose $\sim$ is an equivalence relation on a set A, and a and b are elements of A. Prove that if $[a] \cap [b] \neq \emptyset$, then $a \sim b$.

8. Let R be the relation on $\mathbb{Z}^2$ such that $(a, b)R(c, d)$ if and only if $a+d = b+c$. Is R an equivalence relation ? Justify your answer.

9. Let R be the relation on $\mathbb{Z}^2$ such tha $(a, b)R(c, d)$ if and only if $ad = bc$. Is R an equivalence relation ? Justify your answer.

10. Which of these relations on the set of all people are equivalence relations ? Justify your answers.

(a) $\{(a, b) : a \text{ and } b \text{ are the same height}\}$

(b) $\{(a, b) : a \text{ and } b \text{ live in the same state}\}$

(c) $\{(a, b) : a \text{ and } b \text{ have met}\}$

(d) $\{(a, b) : a \text{ and } b \text{ are blood relatives}\}$ [1]

(e) $\{(a, b) : a \text{ and } b \text{ speak a common language}\}$

11. Which of these relations on $\{1, 2, 3, 4\}$ are equivalence relations ? Justify your answers.

(a) $\{(1, 1), (2, 2), (3, 3), (4, 4)\}$

(b) $\{(1, 2), (2, 1), (1, 3), (3, 1), (1, 4), (4, 1)\}$

(c) $\{(1, 1), (2, 2), (1, 2), (2, 1), (3, 3), (4, 4)\}$

(d) $\{(1, 1), (2, 2), (1, 2), (2, 1), (1, 3), (3, 1), (3, 3), (4, 4)\}$

12. Let A and B be sets, and let $f : A \to B$ be a function. Let R be the relation on A defined by

$$R = \{(a_1, a_2) \in A^2 : f(a_1) = f(a_2)\}.$$

(a) Prove that R is an equivalence relation.

(b) Describe the equivalence classes of R.

[1] A person who is related to another through a common ancestor, and not by marriage or adoption.

Bibliography

[1] Caldwell, Chris K. and Xiong, Yeng, *What is the Smallest Prime?*, Journal of Integer Sequences, (2012), https://cs.uwaterloo.ca/journals/JIS/VOL15/Caldwell1/cald5.pdf

[2] Diffie, Whitfield and Hellman, Martin E., *New directions in cryptography*, Institute of Electrical and Electronics Engineers. Transactions on Information Theory, **IT-22**, (1976), no. 6, 644-654, https://ee.stanford.edu/~hellman/publications/24.pdf

[3] ElGamal, Taher, *A public key cryptosystem and a signature scheme based on discrete logarithms*, Institute of Electrical and Electronics Engineers. Transactions on Information Theory, **31**, (1985), no. 4, 469-472, http://dx.doi.org/10.1109/TIT.1985.1057074

[4] Oliveira e Silva, Tomàs and Herzog, Siegfried and Pardi, Silvio *Empirical verification of the even Goldbach conjecture and computation of prime gaps up to* $4 \cdot 10^{18}$ Mathematics of Computation, **83** (2014) no. 288 2033--2060 https://doi.org/10.1090/S0025-5718-2013-02787-1

[5] Atkins, Derek and Graff, Michael and Lenstra, Arjen K. and Leyland, Paul C. *The magic words are squeamish ossifrage (extended abstract)* in Advances in cryptology---{ASIACRYPT} '94 ({W}ollongong, 1994) Lecture Notes in Comput. Sci. **917** 263--277 Springer, Berlin (1995)

[6] Saidak, Filip, *A new proof of Euclid's theorem*, The American Mathematical Monthly, **113**, (2006), no. 10, 937-938, http://dx.doi.org/10.2307/27642094

[7] Euclid, *The thirteen books of Euclid's Elements*, Translated with introduction and commentary by Thomas L. Heath, 2nd ed, Dover Publications, Inc., New York, (1956)

[8] Gauss, Carl Friedrich, *Disquisitiones arithmeticae* Translated and with a preface by Arthur A. Clarke, Revised by William C. Waterhouse, Cornelius Greither and A. W. Grootendorst and with a preface by Waterhouse, Springer-Verlag, New York, (1986)

[9] Rivest, R. L. and Shamir, A. and Adleman, L., *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the Association for Computing Machinery, **21**, (1978), no. 2, 120-126, http://dx.doi.org/10.1145/359340.359342

[10] Gardner, M., *Mathematical games: A new kind of cipher that would take millions of years to break*, Scientific American, (August 1977), 120-124

[11] Weber, H., *Leopold Kronecker*, Jahresbericht der Deutschen Mathematiker Vereinigung, **2**, (1891), 5-31

[**12**] Merkle, Ralph C, *Secure Communications Over Insecure Channels*, Communications of the ACM, **21**, (1978), no. 4, 294-299

[**13**] Rosen, Kenneth H., *Discrete Mathematics and Applications*, (2012), 978-0-07-338309-5 McGraw-Hill Higher Education

[**14**] *Bounded gaps between primes* Annals of Mathematics **179** (2014) 1121-1174 https://www.jstor.org/stable/24522787

Index

Colophon

This book was authored in PreTeXt.