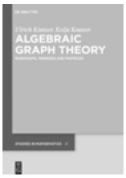
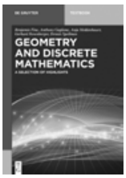
Alexander Kheyfits

A Primer in Combinatorics

Also of Interest



Algebraic Graph Theory
Morphisms, Monoids and Matrices
Ulrich Knauer, Kolja Knauer, 2019
ISBN 978-3-11-061612-5, e-ISBN (PDF) 978-3-11-061736-8,
e-ISBN (EPUB) 978-3-11-061628-6



Geometry and Discrete Mathematics
A Selection of Highlights
Benjamin Fine, Anthony Gaglione, Anja Moldenhauer, Gerhard Rosenberger, Dennis Spellman, 2018
ISBN 978-3-11-052145-0, e-ISBN (PDF) 978-3-11-052150-4,
e-ISBN (EPUB) 978-3-11-052153-5

Algebraic Combinatorics
Eiichi Bannai, Etsuko Bannai, Tatsuro Ito, Rie Tanaka, 2021
ISBN 978-3-11-062763-3, e-ISBN (PDF) 978-3-11-063025-1,
e-ISBN (EPUB) 978-3-11-062773-2

Combinatorics and Finite Fields
Difference Sets, Polynomials, Pseudorandomness and Applications
Edited by Kai-Uwe Schmidt, Arne Winterhof, 2019
ISBN 978-3-11-064179-0, e-ISBN (PDF) 978-3-11-041714-2,
e-ISBN (EPUB) 978-3-11-041719-7

Alexander Kheyfits

A Primer in Combinatorics

2nd edition

DE GRUYTER

Mathematics Subject Classification 2020
Primary: 05-01; Secondary: 97K, 91C

Author
Alexander Kheyfits
USA
alexander.kheyfits@gmail.com

ISBN 978-3-11-075117-8
e-ISBN (PDF) 978-3-11-075118-5
e-ISBN (EPUB) 978-3-11-075124-6

Library of Congress Control Number: 2021940401

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Preface to the second edition

Two new sections were added to this edition. Section 2.6, "Graph coloring", covers that subject in some detail and, in particular, expounds the famous four color problem. Its current solution requires an essential computer time. We avoid any use of computers by limiting to the five colors variant, which is done in detail. Another new section, Chapter 6, "Secondary structures of the RNA", shows how modern applications of graph theory lead to new classes of graphs. We have also refreshed problems and exercises to many sections.

Acknowledgment

This edition was made possible due to many professional efforts of Steven Elliot, Ute Skambraks, Vilma Vaičeliūnienė and the entire staff of De Gruyter, to whom the author is very thankful.

https://doi.org/10.1515/9783110751185-201

Preface to the first edition

Combinatorial analysis or *combinatorics*, for short, deals with enumerative problems where one must answer the question "How many ?" or "In how many ways?" Other problems are concerned with the existence of certain combinatorial objects subject to various constraints. These kinds of problems are considered in this book.

Combinatorial problems, methods and graphical models are abundant in many areas ranging from engineering and financial science to humanitarian disciplines like sociology, psychology, medicine and social sciences, not to mention mathematics and computer science. As parts of discrete mathematics, combinatorics and graph theory have become indispensable parts of introductory and advanced mathematical training for everyone dealing not only with quantitative but also with qualitative data.

Moreover, combinatorics and graph theory have a remarkable and uncommon feature—to begin its study, one needs no background but elementary algebra and common sense. Even simple combinatorial problems often lead to interesting, sometimes difficult questions and allow an instructor to introduce various important mathematical ideas and concepts and to show the nature of mathematical reasoning and proof. These qualities make combinatorics and graph theory an excellent choice for an introductory mathematical class for students of any age, level and major.

This is a text for a one-semester course in combinatorics with elements of graph theory. It can be used in two modes. The first three chapters cover an introductory material and can be (and have actually been) used for an undergraduate class in combinatorics and/or discrete mathematics, as well as for a problem-solving seminar aimed at undergraduate and even motivated high-school students.

Chapters 4 and 5 are of more advanced level and the whole book includes enough material for an entry-level graduate course in combinatorics. For the mathematically inclined reader, the material has been developed systematically and includes all the proofs. After this book, the reader can study more advanced courses, e. g. [1, 9, 10, 22, 51]. At the same time, the reader who is primarily interested in applying combinatorial methods can skip (most of) the proofs and concentrate on problems and methods of their solution.

In Chapter 1 we introduce basic combinatorial concepts, such as the sum and product rules, combinations, permutations, and arrangements with and without repetition. Various particular elementary methods of solving combinatorial problems are also considered throughout the book, such as, for instance, the trajectory method in Section 1.4 or Ferrers diagrams in Section 4.4. In Section 1.6 we apply the methods of Sections 1.1–1.5 to develop the elementary probability theory for random experiments with finite sample spaces. Our goal in this section is not to give a systematic exposition of probability theory, but rather to show some meaningful applications of the combinatorial methods developed earlier.

Chapter 2 contains an introduction to graph theory. After setting up the basic vocabulary in Sections 2.1–2.2, in the next three sections we study properties of trees,

https://doi.org/10.1515/9783110751185-202

Eulerian and planar graphs, and some problems of graph coloring and graphical enumeration. Many other graph theory problems appear in Chapters 3–5. As an application of the methods developed in Chapters 1–2, in Chapter 3 we give an elementary introduction to hierarchical clustering algorithms. This topic has likely never appeared in textbooks before.

Chapter 4 is devoted to more advanced methods of enumerative combinatorics. Sections 4.1–4.2 cover inversion formulas, including the Möbius inversion, and the Principle of Inclusion–Exclusion. The method of generating functions is developed in Section 4.3. Generating functions are introduced as analytical objects, the sums of converging power series. In Section 4.4 we consider several applications of the method of generating functions, in particular partitions and compositions of integer numbers and linear recurrence relations (difference equations) with constant coefficients. The Pólya–Redfield enumeration theory is considered in Section 4.5.

The last chapter of the book is concerned with combinatorial existence problems. The Ramsey theorem and its applications are considered in Section 5.1. The Dirichlet (pigeonhole) principle follows immediately. Section 5.2 treats Hall's theorem on systems of distinct representatives (the marriage problem) and some of its equivalent statements, namely, König's theorem on zero-one matrices and Dilworth's theorem on chains in partially ordered sets. An example of an extremal combinatorial problem (the assignment problem) is also considered here.

Section 5.3 contains an introduction to the theory of balanced block designs. We consider only recursive methods of construction of block designs since deep algebraic results are beyond the scope of this book. Finally, Section 5.4 is devoted to the systems of triples concluding with the proof, due to Hilton [30] of the necessary and sufficient conditions of the existence of Steiner's triple systems.

The author's credo in teaching mathematics involves advancing from examples and model problems to theory and then back to problem solving. This approach works especially well in combinatorics. Every section of the book starts with simple model problems. Discussing and solving these problems, we derive the basic concepts and definitions. Then, we study essential properties of the concepts developed and again solve problems to illustrate the ideas, methods, and their applications. In particular, some parts of proofs are left as problems to be solved by the reader. Studying the solutions of typical problems in the book, the reader can quickly grasp the methods of solving various combinatorial problems and apply these methods to a range of similar problems in any subject. Thus the book can be used as a self-study guide by the reader interested in solving combinatorial problems.

More than 800 problems constitute an integral part of the text. Many problems are drawn from literature, some are folklore, and some are original. Many problems are solved in the text, scores of other problems and exercises are in the end of each section. Additional problems can be found in the books cited in the list of references, specifically, in [11, 13, 29, 38, 39, 53]. Interesting topics for further reading and individual projects can be found in [4]. Solutions, answers or hints to selected problems and exercises are given in the end of the book.

Combinatorial problems often provide natural intuitive motivation and models for important mathematical ideas and concepts, such as operations on sets, various classes of functions, classes of binary relations, and many others. Primary combinatorial concepts, permutations, combinations and alike, can be naturally defined in terms of set theory operations and functions. In the text, we systematically use this approach that can be traced (at least) as far back as C. Berge's monograph [8]. Not to mention its conciseness and theoretical merits, this set-theory based approach is often advantageous in problem solving, and we demonstrate this in the text using many examples. This approach removes the ambiguity that is often present in combinatorial problems, especially when different objects must be identified, and significantly reduces the number of student errors.

It is the author's experience that freshmen usually master this approach with ease and successfully apply it to problem solving. For the reader unfamiliar with the language and basics of set theory, Section 1.1 systematically develops some standard terminology, which is used in the following sections. The reader familiar with naive set theory can skip Section 1.1 and refer back to it as needed.

Very few non-elementary concepts are included in the text. No concept beyond the precalculus level appears before Section 4.3. Two calculus-level concepts, those of derivatives of elementary functions and of converging series, appear in Section 4.3 on generating functions. From this point on the book can be subtitled "Combinatorics through the eyes of an analyst". Even the notion of a converging series can be eliminated and replaced by the finitary concept of generating polynomials, that is, truncated power series, and we solve a few problems to demonstrate the method. This approach makes the method of generating functions accessible to the reader without any calculus background at all, though calculations become more tedious.

It should be noted that these days many college students take at least one calculus class, but afterwards they see no actual application of calculus. Therefore, some nontrivial examples of applications of calculus ideas and methods are appropriate. The same can be said of the few elementary algebraic concepts (groups, rings) appearing in Chapters 4 and 5.

The book is self-contained; all the concepts and definitions used are defined and explained by examples. The Index includes references to important groups of problems and specific methods of their solution, such as "coloring problems" or "method of generating functions". Throughout the text, we use several abbreviations: GF stands for generating function(s), EGF for exponential generating function(s), SDR for system(s) of distinct representatives, and BIBD for balanced incomplete block design(s). Theorems, lemmas, problems, etc., have three-digit numbering, thus, Problem 1.2.3 refers to the third problem in the text of Section 1.2 of Chapter 1, while **1.2.3** means Exercise 1.2.3 in the end of Section 1.2. Figures have two-digit numbering, thus

Fig. 2.3 refers to the third figure in Chapter 2. The symbol □ indicates the ends of the proofs of statements or solutions of problems.

Combinatorial problems and graphical models have been studied by many outstanding scientists for thousands of years. The web site www.degruyter.com of de Gruyter GmbH contains many interesting links describing the history of these developments and lives of the people involved. The coffee cup icon indicates that there is information available at the web site. Any remarks, corrections and suggestions about the book can be sent to akheyfits@gc.cuny.edu.

Acknowledgments

Chapter 3 is a revised version of Module 03-1 in the DIMACS series of educational modules, written when the author participated in Reconnect 1998 and Reconnect 1999 conferences at the DIMACS Center at Rutgers University of New Jersey. The author is grateful to the DIMACS Center, its Director Professor Fred Roberts and Professor Melvin Janowitz for their hospitality and the kind permission to include Module 03-1 in this text, and to Professor Catherine McGeoch for her generous help.

It is finally the author's great pleasure to thank Simon Albroscheit, Robert Plato, Friederike Dittberner and the staff of de Gruyter GmbH for their friendly and highly professional handling of the whole publishing process.

Contents

Preface to the second edition —— V

Preface to the first edition —— VII

Part I: **Introductory combinatorics and graph theory**

1 Basic counting —— 3
1.1 Combinatorics of finite sets —— **3**
1.2 The sum and product rules —— **28**
1.3 Arrangements and permutations —— **36**
1.4 Combinations —— **41**
1.5 Permutations with identified elements —— **66**
1.6 Probability theory on finite sets —— **71**

2 Basic graph theory —— 89
2.1 Vocabulary —— **89**
2.2 Connectivity in graphs —— **97**
2.3 Trees —— **106**
2.4 Eulerian graphs —— **118**
2.5 Planarity —— **121**
2.6 Graph coloring —— **124**

3 Hierarchical clustering and dendrogram graphs —— 131
3.1 Introduction —— **131**
3.2 Model example —— **134**
3.3 Hubert's single-link algorithm —— **147**
3.4 Hubert's complete-link algorithm —— **154**
3.5 Case study —— **168**

Part II: **Combinatorial analysis**

4 Enumerative combinatorics —— 175
4.1 The inclusion–exclusion principle —— **175**
4.2 Inversion formulas —— **187**
4.3 Generating functions I. Introduction —— **192**
4.4 Generating functions II. Applications —— **214**
4.5 Enumeration of equivalence classes —— **236**

5 Existence theorems in combinatorics —— 257
5.1 Ramsey's theorem —— **257**
5.2 Systems of distinct representatives —— **268**
5.3 Block designs —— **283**
5.4 Systems of triples —— **291**

6 Secondary structures of the RNA —— 301
6.1 RNAs, graphs, and the Cauchy–Hadamard formula —— **301**
6.2 Counting the primary structures —— **303**
6.3 Diagrams —— **305**
6.4 Secondary structures —— **307**
6.5 Asymptotic enumeration of the secondary structures. Examples —— **310**

Answers/solutions to selected problems —— 317

Bibliography —— 325

Index —— 327

Part I: **Introductory combinatorics and graph theory**

1 Basic counting

In this chapter we introduce some basic concepts of enumerative combinatorics. In Section 1.1 the language is being prepared—we discuss the axiom of mathematical induction and operations on sets, binary relations, important classes of functions (mappings). This language of sets and mappings is systematically used in Sections 1.2–1.5 to introduce the sum and product rules, arrangements, permutations, and combinations with and without repetition. As an important application of the methods developed, in Section 1.6 we consider some basic notions of the probability theory in the case of finite sample spaces.

1.1 Combinatorics of finite sets

In this introductory section we review a few fundamental set-theory notions and calculate cardinalities of basic set theory objects—the unions, intersections, and Cartesian products of sets. All proofs are based on the Principle of Mathematical Induction.

Coffee-time browsing

- www-history.mcs.st-and.ac.uk/Biographies/Cantor.html (Cantor's biography)
- www.socialresearchmethods.net/kb/dedind.php (Deduction & Induction)
- http://www-history.mcs.st-and.ac.uk/Mathematicians/Nicomachus.html (Nicomachus' biography)
- www.gap-system.org/ history/Biographies/Al-Kashi.html (Al-Kashi's biography)
- scienceworld.wolfram.com/biography/Abel.html (Abel's biography)
- ecee.colorado.edu/~bart/book/stirling.htm (Stirling's approximation for factorials)
- http://mathworld.wolfram.com/Factorial.html (factorials)
- http://en.wikipedia.org/wiki/George_Boole (Boole's biography)
- http://dimacs.rutgers.edu/ (DIMACS center)
- http://www.encyclopedia.com/doc/1E1-Hypsicle.html (Hypsicle's biography)

Throughout we mostly deal with finite sets, thus we accept a naïve point of view, do not introduce axioms of the set theory, do not distinguish sets, classes, etc. Any collection of different elements is called a set and is denoted by braces, $\{x_1, x_2, \ldots\}$, here $x_1, x_2, \ldots$ are the *elements* of this set. A set X can also be introduced by the defining property of its elements, that is, the property P such that every element of X has this property, but no other element possesses it. In this case we write $X = \{x \mid P(x)\}$. If x is an element of a set X, we write $x \in X$, otherwise $x \notin X$.

A set that contains no element is called the *empty* set and is denoted by $\emptyset$. A detailed exposition of naive set theory can be found, for example, in [25].

https://doi.org/10.1515/9783110751185-001

Example 1.1.1. The set $D = \{d \mid d \text{ is a Hindu–Arabic digit}\}$ consists of ten elements $0, 1, 2, \ldots, 9$, thus, $0 \in D, 5 \in D$, but $10 \notin D$. The set $T = \{1, 2, 3\}$ consists of the first three positive integer numbers, $3 \in T$ but $0 \notin T$.

It is important that sets are unordered collections, that is, $\{a, b\} = \{b, a\}$, $\{a, b, c\} = \{b, a, c\} = \{a, c, b\}$, and similar statements hold true for any number of elements. Moreover, a set cannot contain repeating elements, that is, $\{a, b, a\} = \{a, b\}$.

Thus, sets are *primary, undefined* objects. Another major *undefined* object is the set of all natural numbers[1] $\mathbf{N} = \{1, 2, \ldots, n, \ldots\}$. The set of the first n natural numbers is denoted by $\mathbf{N}_n = \{1, 2, \ldots, n\}$ and for brevity is called a *natural segment*, or more specifically, a *natural n-segment*; thus, $\mathbf{N}_3 = \{1, 2, 3\}$ and $\mathbf{N}_1 = \{1\}$. The *whole numbers* $\mathbf{W}$ include all natural numbers and zero, that is, $\mathbf{W} = \{0, 1, 2, \ldots, n, \ldots\}$. The set of all *integer numbers*, positive, negative, and zero, is denoted by $\mathbf{Z}$. Denote also $\mathbf{Z}_p = \{0, 1, \ldots, p-1\}$ for any natural p; in particular, $\mathbf{Z}_2 = \{0, 1\}$. The set of all *real numbers* is denoted by $\mathbf{R}$, and $\mathbf{R}_+$ stands for the set of all nonnegative (including 0) real numbers.

We notice that two words "for all" often appear in mathematical texts. As an abbreviation for this expression, a special symbol $\forall$ is used, called the *universal quantifier*. Thus, a sentence

"A property $P(x)$ holds true for all the elements of a set X"

can be shortened to

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

Likewise, the symbol $\exists$, called the *existential quantifier* serves as an abbreviation for the expression "there exists". For example, the expression

$$(\exists x \in X)P(x)$$

means that there exists at least one element in the set X that possesses the property P. These expressions are often shortened to $\forall x P(x)$, $\exists x \; P(x)$, if it is clear what set X is referred to.

Definition 1.1.1. A set X is called a *subset* of a set Y, if $x \in Y$ whenever $x \in X$, that is,[2] $\forall x(x \in X \Rightarrow x \in Y)$; this is denoted by $X \subset Y$.

Example 1.1.2.
(1) $\mathbf{N}_1 \subset \mathbf{N}_3$ but not vice versa.
(2) $\mathbf{N} \subset \mathbf{W} \subset \mathbf{Z} \subset \mathbf{R}$.

---

1 It is customary in computer science and mathematical logic to treat 0 as a natural number, that is define $\mathbf{N} = \{0, 1, 2, \ldots\}$. For our goals, however, it is more convenient to assume that $0 \notin \mathbf{N}$.

2 In definitions $\Rightarrow$ stands for the implication "if...then".

Basic combinatorial concepts are defined in this text in terms of functions (mappings) and equivalence classes. The concept of a mapping itself is also a *primary, undefined notion*; the following paragraph is not a mathematical definition rather it is an intuitive description of mappings (functions).

Let X and Y be two sets; if to each element $x \in X$ there corresponds the uniquely defined element $y \in Y$, denoted by $y = f(x)$, then it is said that a mapping (or a function or a transformation) f is given with the domain $X = \operatorname{dom}(f)$ and the codomain $Y = \operatorname{codom}(f)$; it is denoted by $f : X \to Y$ or $x \overset{f}{\longmapsto} y$.

Now we can *define* the new concepts in terms of mappings.

Definition 1.1.2. Given a map $f : X \to Y$ and an element $x \in X$, then the element $y = f(x) \in Y$ is called the *image* of the element x with respect to the mapping f; in turn, x is called a *preimage* or an *inverse* image of y. Denote the total preimage of an element $x \in X$, that is, the set of all its preimages, by

$$f^{-1}(y) = \{x \in X \mid f(x) = y\}.$$

The set of all images is called the range of a function (mapping) f and is denoted by $\operatorname{Ran}(f)$ or $f(X)$, thus,

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

in particular, *Ran*$(f) \subset Y$.

Definition 1.1.3. Two mappings, $f : X \to Y$ and $g : X_1 \to Y_1$, are called *equal* if[3] $X = X_1$, $Y = Y_1$, and $f(x) = g(x)$, $\forall x \in X = X_1$.

Example 1.1.3. Consider the mappings

$$f : \mathbf{R} \to \mathbf{R}, \quad g : \mathbf{R}_+ \to \mathbf{R}, \quad h : \mathbf{R} \to \mathbf{R}_+, \quad k : \mathbf{R}_+ \to \mathbf{R}_+,$$

all four given by the same formula $f(x) = x^2, g(x) = x^2, h(x) = x^2, k(x) = x^2$, but with different domains or codomains. These four mappings are pairwise different.

Definition 1.1.4.

(1) A mapping $f : X \to Y$ is called *injective* (or univalent), if no element of Y has more than one preimage.

(2) A mapping $f : X \to Y$ is called *onto* or *surjective*, if each element of Y has at least one preimage.

(3) A mapping $f : X \to Y$ is called *bijective* or a *one-to-one correspondence*, if it is both injective and surjective.

---

3 In definitions “if” always means “if and only if”.

Problem 1.1.1. What mappings in Example 1.1.2 are injective? Surjective? Bijective? Neither?

Definition 1.1.5. A set X is called *finite* if it is the empty set $\emptyset$ or if it can be put in a one-to-one correspondence with a set $\mathbf{N}_k$ with some $k = 1, 2, \ldots$; the quantity $k \in \mathbf{N}$ is called the number of elements or the *cardinality* of X and is denoted by $|X| = k$. Otherwise, the set is called *infinite*. We set $|\emptyset| = 0$ by definition. The set of natural numbers $\mathbf{N}$, as well as any set that can be put in a one-to-one correspondence with $\mathbf{N}$, is called *countable*.

Problem 1.1.2. Are natural segments finite? Explain why the set of natural numbers $\mathbf{N}$ is infinite. Prove that the sets of even positive integers $\{2, 4, 6, \ldots\}$, odd positive integers $\{1, 3, 5, \ldots\}$, prime numbers $\{2, 3, 5, 7, 11, \ldots\}$ are countable.

Problem 1.1.3. Prove that the set of integers $\mathbf{Z}$ is infinite. Is it countable? Explain why the set of real numbers $\mathbf{R}$ is infinite. Is it countable?

To introduce our next topic, the axiom of mathematical induction, we first discuss an example. We want to find an explicit formula for the sum

$$1 + 3 + 5 + \cdots + (2n - 1)$$

of n consecutive odd numbers, which is valid for every $n = 1, 2, 3, \ldots$. To guess the formula, we consider the three sums, $1+3 = 4$, $1+3+5 = 9$, $1+3+5+7 = 16$. We notice that all these sums are squares of integer numbers: if $n = 2$ then $4 = 2^2$, for $n = 3$, $9 = 3^2$, and for $n = 4$, $16 = 4^2$. It is natural now to guess that all such sums are squares. We can check a few more cases, for example, $1+3+5+7+9 = 5^2$, $1+3+5+7+9+11 = 6^2$, $1 + 3 + 5 + 7 + 9 + 11 + 13 = 7^2$; the shortest sum, comprising one addend, $1 = 1^2$, also supports the guess.

Thus, we claim that the equation

$$1 + 3 + \cdots + (2n - 1) = n^2$$

holds true for all natural $n = 1, 2, 3, \ldots$. We have checked this equation for several values of n, however, by no means can we verify infinitely many numerical equations for infinitely many natural numbers. Therefore, we have to develop a new method capable to solve similar problems, that is, the problems involving a parameter, which can take on infinitely many integer values. Thus, this method must reflect some fundamental properties of the infinite set of natural numbers.

This method is called the principle (or the axiom or the postulate) of mathematical induction.

The axiom of mathematical induction

Consider a set of statements or formulas $S_{n_1}, S_{n_1+1}, S_{n_1+2}, \ldots$, numbered by all integer numbers $n \geq n_1$. Usually $n_1 = 1$ or $n_1 = 0$, but it can be any integer number.

(1) Firstly, suppose the statement S_{n_1}, called the *basis step of induction*, is valid. In applications of the method of mathematical induction the verification of S_{n_1} is an independent problem. This step may be sometimes trivial, but it cannot be skipped.
(2) Secondly, suppose that for each natural $n \geq n_1$ we can prove a conditional statement $S_n \Rightarrow S_{n+1}$, that is, we can prove the validity of S_{n+1} for each specified natural $n > n_1$ assuming the validity of S_n, and this conditional statement is valid for all natural $n \geq n_1$. This part of the method is called the *inductive step*. The statement S_n is called the *inductive hypothesis* or *inductive assumption*.
(3) *If we can independently show these two steps, then the principle of mathematical induction claims that all of the statements S_n, for all natural $n \geq n_1$ are valid.*

This method of proof is accepted as an axiom, because nobody can actually verify infinitely many statements $S_n, n \geq n_1$; the method cannot be justified without using some other, maybe even less obvious properties of the set of natural numbers. Mathematicians have been using this principle for centuries and never arrived at a contradiction. Therefore, we accept the method of mathematical induction without a proof, as a postulate, and believe that this principle properly reflects certain fundamental properties of the infinite set **N** of natural numbers. We will apply the method (axiom) of mathematical induction many times in the sequel chapters, the method will be employed in each proof in this chapter, however, sometimes the method presents itself only implicitly, through some known results that have been already proved by using mathematical induction. In the following problem we give a detailed example of an application of the method of mathematical induction.

Problem 1.1.4. Show that, for every natural $n, n = 1, 2, 3, \ldots$,

$$1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

Solution. Here $n_1 = 1$ and S_n stands for the equation above, thus, S_1 denotes the equation $1^2 = \frac{1}{6}1(1+1)(2\cdot 1+1)$, which is certainly true, S_2 denotes $1^2+2^2 = \frac{1}{6}2(2+1)(2\cdot 2+1)$, which is true as well; S_3 is also a valid statement $1^2 + 2^2 + 3^2 = \frac{1}{6}3(3+1)(2\cdot 3+1)$. Therefore, we have the basis of induction (of course, it was enough to verify only one statement S_1) and we have to validate the inductive step.

To do that, we have to prove S_{n+1}, *assuming* that S_n is valid for some unspecified but fixed natural $n = n'$. In this problem we must prove the statement (the equation) $S_{n'+1}$, which reads

$$S_{n'+1} : 1^2 + 2^2 + \cdots + (n')^2 + (n'+1)^2 = \frac{1}{6}(n'+1)(n'+2)(2n'+3)$$

assuming that $S_{n'}$ is valid, that is, using the equation

$$S_{n'} : 1^2 + 2^2 + \cdots + (n')^2 = \frac{1}{6}n'(n'+1)(2n'+1)$$

as if it were correct. Its validity in general is unknown yet, however, in the procedure we suppose it to be true. It is worth repeating that our reasoning must be valid for any natural number n, that is, the reasoning can use only properties common to all natural numbers. For instance, we cannot assume that n is an odd number.

To simplify notation, in the sequel we drop the apostrophe and write in all formulas n assuming it to be fixed. We observe that the left-hand side of

$$S_{n+1} : 1^2 + 2^2 + \cdots + n^2 + (n+1)^2 = \frac{1}{6}(n+1)(n+2)(2n+3)$$

contains the left-hand side of S_n, and the latter in our inductive reasoning is considered to be known. This observation gives us the idea of the proof. Since we assume that the equation

$$S_n : 1^2 + 2^2 + \cdots + (n)^2 = \frac{1}{6}n(n+1)(2n+1)$$

holds true, we employ S_n to transform the left-hand side of S_{n+1} as follows,

$$\begin{aligned} S_{n+1} &: [1^2 + 2^2 + \cdots + n^2] + (n+1)^2 \\ &= \left[\frac{1}{6}n(n+1)(2n+1)\right] + (n+1)^2 = \frac{1}{6}(n+1)(n+2)(2n+3). \end{aligned}$$

Thus we have derived the statement S_{n+1} from S_n for an arbitrary fixed natural n. Since we completed both steps of the principle of mathematical induction, we claim that S_n is valid for all natural n. □

Next we introduce a useful notation. In the preceding problem we had to deal with sums with variable limits. To simplify many formulas, it is convenient to use the *summation or sigma notation*. The sum $a_1 + a_2 + \cdots + a_n$ is denoted by $\sum_{k=1}^{k=n} a_k$. Here k is called the *summation index*, 1 and n are the *lower* and *upper limits of summation*. Usually we simplify the upper index and write the sums as $\sum_{k=1}^{n} a_k$. For example, $\sum_{k=1}^{n} \frac{1}{k}$ means

$$\sum_{k=1}^{n} \frac{1}{k} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n}.$$

If $n = 1$, this is just

$$\sum_{k=1}^{1} \frac{1}{k} = \frac{1}{1},$$

if $n = 2$, it becomes

$$\sum_{k=1}^{2} \frac{1}{k} = \frac{1}{1} + \frac{1}{2},$$

if $n = 3$, it means

$$\sum_{k=1}^{3} \frac{1}{k} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3}.$$

Using the sigma-notation, Problem 1.1.4 can be stated as

$$\sum_{k=1}^{n} k^2 = \frac{1}{6} n(n+1)(2n+1), \quad n = 1, 2, 3, \ldots.$$

The summation index is often called *dummy* index, for it can be replaced by any character, which collides with no other indeterminate in the formula. For example, we can write

$$\sum_{k=1}^{n} \frac{1}{k} = \sum_{l=1}^{n} \frac{1}{l} = \sum_{i=1}^{n} \frac{1}{i},$$

but it is ambiguous to write $\sum_{n=1}^{n} \frac{1}{n}$.

Similarly to the summation notation, we can abbreviate any other operation with several operands. For instance, the product $a_1 \cdot a_2 \cdots a_n$ can be written as

$$\prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdots a_n;$$

for example, $\prod_{k=1}^{4} k^2 = 576$.

The following problem shows some useful properties of the sigma notation. In the end of this section the reader finds other problems concerning this symbol.

Problem 1.1.5. Prove that

(1) $\sum_{k=m}^{n} (-a_k) = -\sum_{k=m}^{n} a_k$;

(2) $\sum_{k=m}^{n} (a_k + b_k) = \sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k$;

(3) $\sum_{k=m}^{n} (b a_k) = b \sum_{k=m}^{n} a_k$ for any constant b;

(4) $\sum_{k=m}^{n} a_{k+l} = \sum_{k=m+l}^{n+l} a_k$.

Problem 1.1.6. Prove by mathematical induction that, for every natural n,

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

or using the summation notation, $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

As another example, we consider an ancient Greek problem.

Problem 1.1.7. (Nicomachus) Partition all odd numbers into groups consisting of $1, 2, 3, \ldots, n, \ldots$ consecutive odd numbers, namely,

$$\mathbf{N} = \{1\} \cup \{3, 5\} \cup \{7, 9, 11\} \cup \{13, 15, 17, 19\} \cup \ldots.$$

If we add up the numbers within each group, we discover (cf. the discussion after Problem 1.1.3) the equations $3 + 5 = 8 = 2^3$, $7 + 9 + 11 = 27 = 3^3$, $13 + 15 + 17 + 19 = 4^3$, etc.; certainly $1 = 1^3$. Show that this is a general pattern, that is, demonstrate that the sum of odd numbers in the nth group is n^3 for any natural n.

Solution. It is convenient here to denote odd numbers by $2k - 1, k = 1, 2, 3, \ldots$. Since the nth group contains n numbers, the preceding $n - 1$ groups altogether contain, by Problem 1.1.6, $1+2+\cdots+(n-1) = (1/2)(n-1)n$ odd numbers. Thus the problem reduces to proving the equation

$$\sum_{k=k_1}^{k_2} (2k - 1) = n^3$$

where we must determine the indices k_1 and k_2 so that $2k_1-1$ is the smallest odd number in the nth group and $2k_2 - 1$ is the largest one.

We notice that in the equation $m = 2k-1$ the number k means the "serial number" of the odd number m in the series of all odd numbers. Indeed,

if $1 = 2k - 1$, then $k = 1$, that is, 1 is the first odd number;
if $3 = 2k - 1$, then $k = 2$, which means 3 is the second odd number;
if $5 = 2k - 1$, then $k = 3$, and 5 is the third odd number, etc.

Therefore, since the first $n-1$ groups contain first $(n-1)n/2$ odd numbers, the first odd number in the nth group is the $(\frac{(n-1)n}{2} + 1)$st odd number, which implies $k_1 = \frac{(n-1)n}{2} + 1$. By the same token, $k_2 = \frac{n(n+1)}{2}$, thus to solve the problem we have to prove that

$$\sum_{k=\frac{(n-1)n}{2}+1}^{\frac{n(n+1)}{2}} (2k - 1) = n^3.$$

It is not hard to prove this by the straightforward mathematical induction, but it is simpler to use the properties mentioned in Problem 1.1.5 and to transform the sum on the left side as

$$2 \sum_{k=\frac{(n-1)n}{2}+1}^{\frac{n(n+1)}{2}} k - \sum_{k=\frac{(n-1)n}{2}+1}^{\frac{n(n+1)}{2}} 1$$
$$= 2\left\{ \sum_{k=1}^{\frac{n(n+1)}{2}} k - \sum_{k=1}^{\frac{(n-1)n}{2}} k \right\} - \left(\frac{n(n+1)}{2} - \frac{(n-1)n}{2} \right),$$

apply twice Problem 1.1.6 to the sums in braces and simplify the resulting expression. □

Problem 1.1.8. Where in the proof did we use the mathematical induction?

The mathematical induction is a very powerful method of proof; however, sometimes we can find an easier approach.

Problem 1.1.9. Show that for every natural n

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{n\cdot(n+1)}=\frac{n}{n+1}.$$

Give two proofs, by mathematical induction and by making use of telescoping sums—which method is simpler in this problem?

Solution. We do only the second proof. A sum $\sum_{k=1}^{2n} a_k$ is said to be *telescoping* if $a_3 = -a_2, a_5 = -a_4, \ldots, a_{2n-1} = -a_{2n-2}$, thus all the addends but the first and the last one, cancel out and the sum is $a_1 + a_{2n}$. We remark that

$$\frac{1}{k(k+1)}=\frac{1}{k}-\frac{1}{k+1},$$

thus the sum in the problem is telescoping and we get

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{n\cdot(n+1)}=1-\frac{1}{n+1}=\frac{n}{n+1},$$

thus proving the claim. □

Remark 1.1.1. Did we really avoid mathematical induction?

It is essential that neither the first nor the second step in an inductive proof can be omitted. For instance, consider a polynomial $P(x) = x^2 + x + 41$. Computing $P(1) = 43$, $P(2) = 47$, $P(3) = 53$, we observe that all these values are prime numbers; $P(0) = 41$ is also prime. A reasonable hypothesis springs up that the value $P(n)$ is prime for any whole n. Such a conclusion is called *incomplete induction*, since it is based on a finite set of observations and has not been confirmed by the inductive step. Without this validation the incomplete induction can lead to false conclusions. Indeed, if we continue the numerical experiment with the polynomial above, we discover that all numbers $P(0), \ldots, P(39)$ are prime, but $P(40) = 1681 = 41^2$ is a composite number, thus invalidating our guess.

The following result may look simple, although it is fundamental in solving combinatorial problems. It is this property that underlines, for instance, the following trivial fact: if 25 students attend a class, and there are only 24 chairs in the classroom, then either two students will have to share a chair or one student will have to stand. The latter is obvious, but there are many non-obvious problems where this result is useful. Even if this theorem is not mentioned explicitly, it is present in any enumerative problem. We prove it only for finite sets.

Theorem 1.1.1. *For any finite sets X and Y,*
(1) $|X| \le |Y|$ *if and only if there exists an injective mapping* $f : X \to Y$;
(2) $|X| \ge |Y|$ *if and only if there exists a surjective mapping* $f : X \to Y$;
(3) $|X| = |Y|$ *if and only if there exists a bijective mapping* $f : X \to Y$.

Proof. Denote $X = \{x_1, \dots, x_n\}$ and $Y = \{y_1, \dots, y_m\}$. If there exists an injective mapping $f : X \to Y$ such that $f(x_i) = y_{j_i}, 1 \le i \le n$, then all images $y_{j_i}, 1 \le i \le n$, must be different for f is injective, thus, there are at least as many y_js as x_is, that is, $n \le m$. On the other hand, if $n = |X| \le m = |Y|$, we can straightforwardly construct a required injective mapping $f : X \to Y$, for instance as $f(x_i) = y_i, 1 \le i \le n$, which proves part (1) of the problem. Part (2) can be proved likewise and part (3) follows from parts (1) and (2). □

Problem 1.1.10. To facilitate memorization of telephone numbers, they can be expressed as certain combinations of digits and relevant words; for example, it is easier to remember 1-800-333-TOLL, than 1-800-333-8655. To this end, the dialing keys on telephone handsets are marked by both digits and letters. What relationship between the set of digits $\{0, 1, \dots, 9\}$ and the English alphabet allows us to use this approach?

Next we introduce operations on sets. Working on any problem involving sets, we always assume, explicitly or implicitly, that all sets under consideration are subsets of a certain ambient totality, called the *universal set* U. This is our universe and nothing exists in the problem beyond U. This remark is important when we compute the complement of a set.

Definition 1.1.6.

(1) If $\{X_l\}_{l=1}^{\omega}$ is a family of sets X_l, then the collection of all elements x, belonging to at least one of the sets $X_l, l = 1, 2, \dots$, is called the *union* of the sets X_l and is denoted by

$$\bigcup_{l=1}^{\omega} X_l = \{x \mid \exists l \ge 1, x \in X_l\}.$$

(2) The collection of all elements x belonging to *each one* of the sets $X_l, l = 1, 2, \dots, \omega$, is called the *intersection* of the sets X_l and is denoted by

$$\bigcap_{l=1}^{\omega} X_l = \{x \mid x \in X_l, \forall l \ge 1\}.$$

If $X \cap Y = \emptyset$, the sets X and Y are called disjoint.

(3) The difference of sets X and Y, denoted by $X \setminus Y$, is the set of all those elements of X, which do not belong to Y and is denoted by

$$X \setminus Y = \{x \mid x \in X \text{ and } x \notin Y\}.$$

(4) The complement of a set X, denoted by $\overline{X}$ or X^c, is the set of all the elements of the universal set U that do not belong to X,

$$\overline{X} = \{x \in U \mid x \notin X\};$$

it is obvious that $\overline{X} = U \setminus X$.

Problem 1.1.11. Find $\mathbf{N}_1 \cup \mathbf{N}_3, \mathbf{N}_1 \cap \mathbf{N}_3, \mathbf{N}_1 \setminus \mathbf{N}_3, \mathbf{N}_3 \setminus N_1$—the natural segments $\mathbf{N}_l$ were introduced at the beginning of this section.

The following problem lists important properties of these operations, some of them are similar to the well-familiar properties of the addition and multiplication of numbers. These properties are also valid not only for two, but for any finite collection of sets.

Problem 1.1.12.

(1) The union and intersection of sets are commutative,

$$X \cup Y = Y \cup X,$$
$$X \cap Y = Y \cap X,$$

and associative operations,

$$X \cup (Y \cup Z) = (X \cup Y) \cup Z = X \cup Y \cup Z,$$
$$X \cap (Y \cap Z) = (X \cap Y) \cap Z = X \cap Y \cap Z,$$

they satisfy two distributive laws,

$$X \cup (Y_1 \cap Y_2) = (X \cup Y_1) \cap (X \cup Y_2),$$
$$X \cap (Y_1 \cup Y_2) = (X \cap Y_1) \cup (X \cap Y_2).$$

(2) The complement is connected with the union and intersection by de Morgan laws,

$$\overline{X \cap Y} = \overline{X} \cup \overline{Y},$$
$$\overline{X \cup Y} = \overline{X} \cap \overline{Y}.$$

The properties we have already considered, are useful in many enumerative problems. For instance, the definition of the union of two sets directly implies the following statement.

Lemma 1.1.1. *If X and Y are finite disjoint sets, that is, $|X| < \infty$, $|Y| < \infty$, and $X \cap Y = \emptyset$, then $|X \cup Y| = |X| + |Y|$.*

By the axiom of mathematical induction, this lemma immediately extends to any finite collection of sets.

Lemma 1.1.2. *If $|X_i| < \infty, i = 1, 2, \ldots, m$, and $X_i \cap X_j = \emptyset$ for all $1 \leq i, j \leq m, i \neq j$, then $|X_1 \cup X_2 \cup \cdots \cup X_m| = |X_1| + |X_2| + \cdots + |X_m|$.*

We will often use the following notion.

Definition 1.1.7. It is said that non-empty and mutually disjoint sets

$$X_\alpha, X_\beta, X_\gamma, \ldots$$

make a *partition* of a set X, if $X = X_\alpha \cup X_\beta \cup X_\gamma \cup \cdots$, where the order of sets is immaterial. The number of all partitions of an n-element set is called the *Bell number* B_n; see Problem 1.1.19.

Example 1.1.4. Thus, the set $\mathbf{Z}_e = \{\dots, -4, -2, 0, 2, \dots\}$ of all even numbers including zero, and the set $\mathbf{Z}_o = \{\dots, -3, -1, 1, 3, \dots\}$ of all odd numbers form a partition of the set of integers, $\mathbf{Z} = \mathbf{Z}_e \cup \mathbf{Z}_o$ and $\mathbf{Z}_e \cap \mathbf{Z}_o = \emptyset$, while $\mathbf{Z}'{}_e = \{\dots, -4, -2, 2, \dots\}$ and $\mathbf{Z}_o = \{\dots, -3, -1, 1, 3, \dots\}$ do not.

Problem 1.1.13. Find a set $\mathbf{Z}''$ such that $\{\mathbf{Z}'_e, \mathbf{Z}_0, \mathbf{Z}''\}$ is a partition of $\mathbf{Z}$; the set $\mathbf{Z}'_e$ was introduced in Example 1.1.4.

Problem 1.1.14. Prove that the total preimages of all the elements in the range of any mapping make a partition of the domain of this mapping.

The result of this problem implies immediately

Lemma 1.1.3. *If X and Y are finite sets and $f : X \to Y$ is a surjective mapping, then*

$$|X| = \sum_{y \in Y} |f^{-1}(y)|.$$

In particular, if all total preimages have the same cardinality n_0, then

$$|X| = n_0 |Y|. \qquad \square$$

Example 1.1.5. Let a mapping $f : X \to Y$, where $X = \{-3, -2, -1, 0, 1, 2, 3\}$ and $Y = \{0, 1, 4, 9\}$, be given by $f(x) = x^2$. Then $f^{-1}(\{0\}) = \{0\}$, $f^{-1}(\{1\}) = \{-1, 1\}$, $f^{-1}(\{4\}) = \{-2, 2\}$, $f^{-1}(\{9\}) = \{-3, 3\}$. Here

$$|X| = \sum_{y \in Y} |f^{-1}(y)| = 7.$$

If $f_1 : X_1 \to Y_1, f_1(x) = x^2$, where $X_1 = \{-3, -2, -1, 1, 2, 3\}$ and $Y_1 = \{1, 4, 9\}$, then $|Y_1| = 3, n_0 = 2$ and $|X_1| = 6 = 2 \cdot 3$.

In many problems we have to distinguish ordered and unordered totalities. The latter are sets and as such, are denoted by braces, $\{a\}$, $\{a, b\} = \{b, a\}, \dots$. However, unlike the two-element set $\{a, b\}$, *ordered pairs*, denoted by parentheses, (a, b), are characterized by the profound property

$$(a, b) = (b, a) \quad \text{if and only if} \quad a = b,$$

and a definition must preserve this property. Such a definition can be given in terms of mappings.

Definition 1.1.8. An *ordered pair* with the first element a and the second element b is a mapping $f : \{1, 2\} \to \{a, b\}$, where $a = f(1)$ and $b = f(2)$. This pair is denoted by (a, b).

The next definition introduces a useful mathematical model dealing with ordered totalities.

Definition 1.1.9. Given two sets X and Y, the set of all ordered pairs (x, y) with $x \in X, y \in Y$ is called the *Cartesian* or *direct product* of these sets in this order and is denoted by

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

Problem 1.1.15. Compute $\mathbf{N}_1 \times \mathbf{N}_3, \mathbf{N}_3 \times \mathbf{N}_1, \mathbf{N}_2 \times \mathbf{N}_3, \mathbf{N}_3 \times \mathbf{N}_2$, and find the cardinal numbers of these sets.

An ordered totality of n elements $a_1, a_2, \dots, a_n$ is called an n-tuple or n-vector and is denoted by $(a_1, a_2, \dots, a_n)$; thus, 2-tuples are ordered pairs. To avoid confusion with unordered sets, ordered totalities are denoted by *parentheses*.

Problem 1.1.16. Define n-tuples in terms of mappings. Give a definition of the Cartesian product of three or more sets.

In many problems it is necessary to consider not the entire Cartesian product but only its subsets.

Definition 1.1.10. Given two sets X and Y, any subset ϱ of their Cartesian product $X \times Y$ is called a *binary relation* between elements of X and Y. If $Y = X$, that is, $\varrho \subset X \times X$, ϱ is called a (binary) relation on the set X.

Example 1.1.6. For instance, if $X \times Y = \mathbf{N} \times \mathbf{N}$, we can consider $\varrho_0 = \emptyset$, or $\varrho_1 = \{(1, 1), (1, 2)\}$, or $\varrho_2 = \{(1, 3)\}$, or $\varrho_3 = \{(3, 1)\}$; it is worth repeating that $\varrho_2 \neq \varrho_3$. We say that $1 \in X = \mathbf{N}$ is in the relation ϱ_2 with $3 \in Y = \mathbf{N}$ *but not vice versa*, that is, $3 \in Y = \mathbf{N}$ is not in the relation ϱ_2 with $1 \in X = \mathbf{N}$.

Problem 1.1.17. How many binary relations do exist between the natural segments $\mathbf{N}_1$ and $\mathbf{N}_3$?

Definition 1.1.10 of binary relations is very general. In applications we are usually interested in more specific classes of binary relations. In the following definitions we consider only relations *on* a set X.

Definition 1.1.11.

(1) A binary relation $\varrho \subset X \times X$ is called *reflexive* if each element of X is in this relation with itself, that is,

$$(\forall x \in X)\big((x, x) \in \varrho\big).$$

(2) A binary relation $\varrho \subset X \times X$ is called *symmetric* if for all $x, y \in X$, the element y is in the relation ϱ with x whenever the element x is in the relation ϱ with y, that is,

$$(\forall x, y \in X)\big((x, y) \in \varrho \Rightarrow (y, x) \in \varrho\big).$$

(3) A binary relation $\varrho \subset X \times X$ is called *transitive* if for all $x, y, z \in X$, the element x is in the relation ϱ with z whenever the element x is in the relation ϱ with y and y is in this relation with z, that is,

$$(\forall x, y, z \in X)(((x, y) \in \varrho \,\&\, (y, z) \in \varrho) \Rightarrow ((x, z) \in \varrho)).$$

(4) A binary relation $\varrho \subset X \times X$ is called *antisymmetric* if for all $x, y \in X$, the elements x and y cannot simultaneously be in the relation ϱ with one another unless $x = y$, that is,

$$(\forall x, y \in X)(((x, y) \in \varrho \,\&\, (y, x) \in \varrho) \Rightarrow (y = x)).$$

An important class of binary relations is introduced in the following definition.

Definition 1.1.12. A reflexive, symmetric, and transitive binary relation $\varrho \subset X \times X$ is called an *equivalence relation* on the set X. If $\varrho \subset X \times X$ is an equivalence relation on X and an ordered pair $(x, y) \in \varrho$, then the elements x and y are called equivalent (with respect to ϱ); this equivalence is denoted by $x \overset{\varrho}{\sim} y$ or simply $x \sim y$.

If ϱ is an equivalence relation on X, then a subset $X_0 \subset X$ consisting of all pairwise equivalent elements of X, is called an *equivalence class*.

The family of all equivalence classes with respect to an equivalence relation ϱ on a set X is called the *factor set* of X with respect to this equivalence relation ϱ and is denoted by X/ϱ or $X/\sim$. Examples of equivalence relations are considered in the end of this section.

Problem 1.1.18. Prove that any equivalence class is non-empty, any two different equivalence classes are disjoint, and the union of all the equivalence classes with respect to an equivalence relation on a set X is equal to X. Thus, the equivalence classes make up a partition of X.

The converse assertion is also true.

Problem 1.1.19. Prove that any partition of a set generates an equivalence relation on this set such that the factor set of this equivalence relation is precisely the family of all the parts of the partition. Therefore, there is a one-to-one correspondence between the partitions of a set and the equivalence relations on the set and the number of the equivalence relations on an n-set X is equal to the Bell number B_n; see Definition 1.1.7.

The following class of binary relations also often occurs in applications.

Definition 1.1.13. A reflexive, antisymmetric, and transitive binary relation $\varrho \subset X \times X$ is called a relation of *partial order* or just a partial order on the set X. If ϱ is a partial order and $(x, y) \in \varrho$, then we write $x \prec y$. If $x \prec y$ or $y \prec x$, the elements x and y are called *comparable* (with respect to the order ϱ). A set with a relation of partial order on it is called a *partially ordered set (poset)*. If any two elements of a poset are comparable, that is, either $x \prec y$ or $y \prec x$ for all $x, y \in X$, then the set is called a *chain* or a *linearly* (sometimes *totally*) ordered set.

In the following statements we create our "combinatorial toolkit"—we compute cardinal numbers of major set-theory constructions. The next statement directly follows from Lemma 1.1.4 and Problem 1.1.18.

Lemma 1.1.4. *Let an equivalence relation be given on a finite set X such that all the equivalence classes have the same cardinality k. Then the cardinality of the factor set, that is, the number of the equivalence classes is*

$$\frac{|X|}{k}. \tag{1.1.1}$$

□

Next we calculate the cardinality of the union of finite sets. We will need the following properties, whose proofs are left to the reader.

Problem 1.1.20. For any (not necessarily finite) sets X and Y,
(1)

$$X \cup Y = X \cup (Y \setminus X), \tag{1.1.2}$$

(2)

$$Y = (Y \cap X) \cup (Y \setminus X), \tag{1.1.3}$$

where the sets on the right in both (1.1.2) and (1.1.3) are disjoint, that is, $X \cap (Y \setminus X) = \emptyset$ and $(Y \cap X) \cap (Y \setminus X) = \emptyset$.

Theorem 1.1.2. *If* $|X| < \infty$ *and* $|Y| < \infty$, *then*

$$|X \cup Y| = |X| + |Y| - |X \cap Y|.$$

Proof. It is sufficient to apply Lemma 1.1.1 to identities (1.1.2)–(1.1.3). □

We extend this statement to any finite family of sets.

Theorem 1.1.3. *If* $|X_i| < \infty, 1 \leq i \leq k$, *then*

$$\begin{aligned} &|X_1 \cup X_2 \cup \cdots \cup X_k| \\ &\quad = |X_1| + |X_2| + \cdots + |X_k| - |X_1 \cap X_2| - \cdots - |X_{k-1} \cap X_k| \\ &\qquad + |X_1 \cap X_2 \cap X_3| + \cdots + (-1)^{k-1}|X_1 \cap X_2 \cap \cdots \cap X_k|. \end{aligned} \tag{1.1.4}$$

Proof. To prove (1.1.4) for any k, we use the mathematical induction on the number k of sets. If $k = 1$, then formula (1.1.4) is obvious, which already makes the basis of induction. Moreover, if $k = 2$, (1.1.4) reduces to Theorem 1.1.2. Suppose that the statement is valid for any union of k sets, and consider a union of $k+1$ sets $X_1 \cup X_2 \cup \cdots \cup X_k \cup X_{k+1}$. Now, Theorem 1.1.2 with $X = X_1 \cup X_2 \cup \cdots \cup X_k$ and $Y = X_{k+1}$ implies the equation

$$|X_1 \cup \cdots \cup X_k \cup X_{k+1}| = |X_1 \cup X_2 \cup \cdots \cup X_k| + |X_{k+1}| - |(X_1 \cup X_2 \cup \cdots \cup X_k) \cap X_{k+1}|.$$

By the distributive law (Problem 1.1.12)

$$(X_1 \cup X_2 \cup \cdots \cup X_k) \cap X_{k+1} = (X_1 \cap X_{k+1}) \cup \cdots \cup (X_k \cap X_{k+1}).$$

Applying the inductive hypothesis to the unions $X_1 \cup X_2 \cup \cdots \cup X_k$ and $(X_1 \cap X_{k+1}) \cup \cdots \cup (X_k \cap X_{k+1})$, we get the result. □

Consider now Cartesian products. A proof of the following proposition is left as an exercise to the reader.

Lemma 1.1.5. *For any, not necessarily finite sets X, Y_1, Y_2,*

$$X \times (Y_1 \cup Y_2) = (X \times Y_1) \cup (X \times Y_2).$$

Moreover, if $Y_1 \cap Y_2 = \emptyset$, then $(X \times Y_1) \cap (X \times Y_2) = \emptyset$. □

Theorem 1.1.4. *If $|X| < \infty$ and $|Y| < \infty$, then*

$$|X \times Y| = |X| \cdot |Y|. \tag{1.1.5}$$

Proof. It is worth mentioning that the symbol $\times$ in (1.1.5) on the left means the set-theory operation—the Cartesian product of two sets, while the symbol $\cdot$ on the right indicates the usual arithmetic multiplication of whole numbers. We customarily omit the symbol $\cdot$ and write $|X||Y|$.

To prove the assertion, we carry the mathematical induction on the cardinal number $k = |Y|$. If $k = 1$, then Y is a 1-element set, $Y = \{y\}$. Denoting $X = \{x_1, x_2, \dots, x_n\}$, we have $X \times Y = \{(x_1, y), (x_2, y), \dots, (x_n, y)\}$, hence, $|X \times Y| = |X| = |X| \cdot |Y|$ and the basis of induction is valid.

To make the inductive step, we fix a set X with $|X| = n$, assume that (1.1.5) holds for all k-element sets with a fixed natural k, and consider an arbitrary $(k+1)$-element set Y. Choose an element $y \in Y$ and consider two subsets of Y, $Y_1 = \{y\}$ and $Y_2 = Y \setminus Y_1$; it is clear, that $|Y_1| = 1$, $|Y_2| = k$, $Y = Y_1 \cup Y_2$, and $Y_1 \cap Y_2 = \emptyset$. Due to the inductive assumption, $|X \times Y_2| = |X||Y_2|$, moreover, we have seen at the basis step that $|X \times Y_1| = |X|$. By making use of Lemmas 1.1.5 and 1.1.1 we derive the equation

$$|X \times Y| = |X \times Y_1| + |X \times Y_2| = |X|(1 + |Y_2|) = |X||Y|,$$

which completes the inductive step of the proof. The statement follows by the axiom of mathematical induction. □

Theorem 1.1.4 and the axiom of mathematical induction imply immediately the following.

Theorem 1.1.5. *If* $|X_i| < \infty, 1 \leq i \leq k$, *then*

$$|X_1 \times X_2 \times \cdots \times X_k| = |X_1| \cdot |X_2| \cdots |X_k|.$$

Definition 1.1.14.

(1) The class of all mappings with the domain X and codomain Y is called the *power set*[4] and is denoted by Y^X.
(2) The class of all injective mappings with the domain X and codomain Y is denoted by $\mathrm{Inj}(Y^X)$.
(3) The class of all surjective mappings with the domain X and codomain Y is denoted by $\mathrm{Surj}(Y^X)$.
(4) The class of all bijective mappings with the domain X and codomain Y is denoted by $\mathrm{Bij}(Y^X)$.
(5) The set of all subsets of a set X, including the empty set $\emptyset$ and the set X itself, is called the set of subsets or the Boolean of X and is denoted by 2^X. The set of all k-element subsets of X is denoted by 2^X_k.

Let us stipulate that, if $X = \emptyset$, then there is only one "empty" mapping, belonging to $Y^\emptyset$, that is, $|Y^\emptyset| = 1$. Also, it is obvious that $2^X_k = 0$ whenever $k < 0$ or $k > |X|$.

Example 1.1.7. Let $X = \{a, b, c\}$ and $Y = \{1, 2\}$. Then $Y^X = \{f_1, f_2, \dots, f_8\}$, where the mappings $f_i, 1 \leq i \leq 8$, are given by the following charts:

$$f_1 : \left\{ \begin{array}{c} f_1(a) = 1 \\ f_1(b) = 1 \\ f_1(c) = 1 \end{array} \right\}, \quad f_2 : \left\{ \begin{array}{c} f_2(a) = 1 \\ f_2(b) = 1 \\ f_2(c) = 2 \end{array} \right\}, \quad f_3 : \left\{ \begin{array}{c} f_3(a) = 1 \\ f_3(b) = 2 \\ f_3(c) = 1 \end{array} \right\},$$

$$f_4 : \left\{ \begin{array}{c} f_4(a) = 2 \\ f_4(b) = 1 \\ f_4(c) = 1 \end{array} \right\}, \quad f_5 : \left\{ \begin{array}{c} f_5(a) = 1 \\ f_5(b) = 2 \\ f_5(c) = 2 \end{array} \right\}, \quad f_6 : \left\{ \begin{array}{c} f_6(a) = 2 \\ f_6(b) = 1 \\ f_6(c) = 2 \end{array} \right\},$$

$$f_7 : \left\{ \begin{array}{c} f_7(a) = 2 \\ f_7(b) = 2 \\ f_7(c) = 1 \end{array} \right\}, \quad f_8 : \left\{ \begin{array}{c} f_8(a) = 2 \\ f_8(b) = 2 \\ f_8(c) = 2 \end{array} \right\}.$$

Notice that in this example there are $8 = 2^3$ different mappings, that is, $|Y^X| = |Y|^{|X|}$. This is a particular case of the subsequent Theorem 1.1.6. First we introduce a convenient notation and prove a lemma.

Definition 1.1.15. Given a mapping $f : X \to Y$ and a subset $Z \subset X$, the mapping $f|_Z : Z \to Y$ such that $f|_Z(x) = f(x), \forall x \in Z$, is called the *restriction* of f onto Z.

4 The set of all subsets of X is sometimes also called the power set.

Lemma 1.1.6. *If X_1, X_2, and Y are finite sets and $X_1 \cap X_2 = \emptyset$, then*

$$|Y^{X_1 \cup X_2}| = |Y^{X_1}| \cdot |Y^{X_2}|.$$

Proof. We establish a one-to-one correspondence between the power set $Y^{X_1 \cup X_2}$ and the Cartesian product $Y^{X_1} \times Y^{X_2}$. Consider a mapping $f \in Y^{X_1 \cup X_2}$ and denote its restrictions $f|_{X_i}$ onto X_i by f_i, $i = 1, 2$. Introduce a mapping

$$H : Y^{X_1 \cup X_2} \to Y^{X_1} \times Y^{X_2}$$

by the rule $H(f) = (f_1, f_2)$; here on the right we have an ordered pair of two restrictions of the mapping f. We prove that H is a bijection, that is, H is a one-to-one correspondence we are looking for.

We have to prove that H is both injective and onto. To prove the former, we consider two different mappings $f, g \in Y^{X_1 \cup X_2}$. Since $f \neq g$, there exists an element $x_0 \in X = X_1 \cup X_2$ such that $f(x_0) \neq g(x_0)$. If $x_0 \in X_1$, then, by the definition of a restriction, $f_1(x_0) \neq g_1(x_0)$ at the point x_0, thus, the restrictions are different maps, $f_1 \neq g_1$. If $x_0 \in X_2$, then $f_2 \neq g_2$ on the same basis. In both cases $H(f) = (f_1, f_2) \neq (g_1, g_2) = H(g)$, which proves that H is injective.

To prove that H is surjective, we pick an arbitrary ordered pair

$$(f_1^0, f_2^0) \in Y^{X_1} \times Y^{X_2}$$

and find its preimage with respect to H. In order for the mapping H to be *onto*, there must exist a mapping $f^0 \in Y^{X_1 \cup X_2}$ such that $H(f^0) = (f_1^0, f_2^0)$. We define this mapping f^0 explicitly

$$f^0(x) = \begin{cases} f_1^0(x) & \text{if } x \in X_1, \\ f_2^0(x) & \text{if } x \in X_2. \end{cases}$$

The mapping f^0 is well-defined since X_1 and X_2 are disjoint sets by the assumption. Obviously, $H(f^0) = (f_1^0, f_2^0)$, thus, H is a surjective mapping. Since all sets here are finite, by Theorems 1.1.1 and 1.1.4 we get the equation

$$\left|Y^{X_1 \cup X_2}\right| = \left|Y^{X_1} \times Y^{X_2}\right| = \left|Y^{X_1}\right|\left|Y^{X_2}\right|. \qquad \square$$

The next two statements explain the choice of notation for the power set Y^X and for the Boolean 2^X.

Theorem 1.1.6. *If X and Y are finite non-empty sets, then*

$$\left|Y^X\right| = |Y|^{|X|}.$$

Proof. The conclusion follows immediately if we set $X_2 = \emptyset$ in Lemma 1.1.6. However, it is useful to give here another proof by mathematical induction on the cardinality of X. If $|X| = 1$, say $X = \{x\}$, then the statement is clear, for $Y^{\{x\}}$ contains exactly as many mappings as there are elements in Y. Indeed, an image for the unique element $x \in X$ can be chosen in $|Y|$ ways, and each choice generates exactly one mapping from X to Y, so that $|Y^X| = |Y| = |Y|^{|X|}$.

Suppose now that the statement is valid for all k-element sets, and consider a set X with $k+1$ elements. Select an element $x_1 \in X$ and introduce two sets, $X_1 = \{x_1\}$ and $X_2 = X \setminus X_1$. Since $1 + |X_2| = |X|$, we have by Lemma 1.1.6 and the inductive assumption

$$|Y^X| = |Y^{X_1}| \cdot |Y^{X_2}| = |Y| \cdot |Y|^{|X_2|} = |Y|^{|X|},$$

which proves the theorem. □

Definition 1.1.16. Let A be an arbitrary subset of a set X, $A \subset X$, thus $0 \le |A| \le |X|$. A function $f_A \in Y^X$ given by

$$f_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A, \end{cases}$$

is called the *characteristic function* of a subset $A \subset X$.

Theorem 1.1.7. *If X is a finite set, then* $|2^X| = 2^{|X|}$.

Proof. We reduce the statement to Theorem 1.1.6. Consider the 2-element set $Y = \{0, 1\}$ and the power set Y^X. Since $|Y| = 2$, to prove the theorem it is sufficient to set up a one-to-one correspondence between the two sets Y^X and 2^X. As in Lemma 1.1.6, we will prove that the mapping

$$H : 2^X \to Y^X, \quad H(A) = f_A, \quad \forall A \subset X,$$

is bijective. To prove that H is injective, we select two different subsets $A, B \subset X$, $A \ne B$. Thus, there exists an element $x_0 \in (A \setminus B) \cup (B \setminus A)$. If $x_0 \in A \setminus B$, then $f_A(x_0) = 1 \ne 0 = f_B(x_0)$ and mappings f_A and f_B are different. The same conclusion, $f_A \ne f_B$, follows if $x_0 \in B \setminus A$. Thus, $H(A) = f_A \ne H(B) = f_B$, and H is an injective mapping.

To prove that H is onto, we consider a mapping $f^0 \in Y^X$ and the subset $A_0 = (f^0)^{-1}(\{1\}) \subset X$. We immediately see that $H(A_0) = f^0$, which proves that H is surjective and, together with the preceding part, proves that H is bijective. Now Theorem 1.1.7 follows straightforwardly from Theorem 1.1.6. □

Example 1.1.8. Let $X = \{a, b, c\}$, $|X| = 3$. Then

$$\begin{aligned}
2^X_0 &= \{\emptyset\}, \quad |2^X_0| = 1, \\
2^X_1 &= \{\{a\}, \{b\}, \{c\}\}, \quad |2^X_1| = 3, \\
2^X_2 &= \{\{a, b\}, \{a, c\}, \{b, c\}\}, \quad |2^X_2| = 3,
\end{aligned}$$

$$2_3^X = \{X\}, \quad |2_3^X| = 1,$$

and $|2^X| = 1 + 3 + 3 + 1 = 8 = 2^3$.

Definition 1.1.17. We recall that the *n-factorial*, denoted by $n!$, is the function defined for all natural numbers $n \in \mathbf{N}$ as the product of the first n natural numbers,

$$n! = 1 \cdot 2 \cdots (n-1) \cdot n.$$

We also define $0! = 1$.

Example 1.1.9. $1! = 1, 2! = 1 \cdot 2 = 2, 3! = 1 \cdot 2 \cdot 3 = 6, 4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$.

Problem 1.1.21.
(1) Compute 11!.
(2) Compute $\frac{201!}{199!}$.
(3) Simplify $n \cdot (n-1)!$, $\frac{(n+2)!}{(n+2)(n+1)}$.

Remark 1.1.2. Using some calculus, we can prove the Stirling asymptotic formula

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \quad n \to \infty. \tag{1.1.6}$$

Here $e \approx 2.7182818$ is the base of natural logarithms; "*asymptotic*" means that the ratio of the left-hand-side and the right-hand-side of (1.1.6) tends to 1 as $n \to \infty$. For example, when $n = 7$, formula (1.1.6) computes 7! with a relative error slightly more than 1 %.

We recall that 2_k^X is a set, not the cardinality of this set.

Theorem 1.1.8. *If* $|X| = n < \infty$, *then for* $0 \le k \le n$

$$|2_k^X| = \frac{n!}{(n-k)!k!}. \tag{1.1.7}$$

Proof. We will carry mathematical induction on $n = |X|$. Since the claim contains two natural parameters, n and k, we reformulate the statement of the theorem by binding one of them. □

Theorem 1.1.9. *Equation* (1.1.7) *holds true for every nonnegative integer n and for all integers k, $0 \le k \le n$.*

Proof. It is convenient in this proof to use $n = 0$ as the basis of induction. Since $0 \le k \le n$, for $n = 0$ there is the only value of k, $k = 0$. Hence $X = \emptyset$, $2_0^X = \{\emptyset\}$, $|2_0^X| = 1 = 2^0$, and (1.1.7) in the case $n = 0$ follows.

To make the inductive step, we choose an $n \ge 1$ and assume that equation (1.1.7) is valid for any n-element set. Consider a set X such that $|X| = n + 1$. If $k = n + 1$, then

$2_k^X = \{X\}, |2_k^X| = 1$, and (1.1.7) is valid. To verify (1.1.7) when $k \leq n$, we pick an element $x_0 \in X$ and split 2_k^X in two subsets, $2_k^X = A \cup B$, where A consists of all k-element subsets of X containing x_0 and $B = 2_k^X \setminus A$, thus, subsets in B do not contain x_0. Therefore, these subsets, which are elements of B, can be considered as k-element subsets of the set $X \setminus \{x_0\}$. Since $|X \setminus \{x_0\}| = n$, the inductive assumption is applicable to B, and we have $|B| = n!/((n-k)!k!)$.

On the other hand, if any $\alpha \in A$, that is, α is a k-element subset of X, then $\alpha \in 2_k^X, |\alpha| = k$, and by definition of A, $\alpha \ni x_0$. Therefore, $\alpha \setminus \{x_0\}$ is a $(k-1)$-element subset of the set $X \setminus \{x_0\}$. Hence the elements of A can be put in a one-to-one correspondence with $(k-1)$-element subsets of $X \setminus \{x_0\}$, thus $|A| = |2_{k-1}^{X\setminus\{x_0\}}|$. By the inductive assumption, $|A| = n!/((n-k+1)!(k-1)!)$. Since $A \cap B = \emptyset$, Lemma 1.1.1 implies

$$\begin{aligned} |2_k^X| &= \frac{n!}{(n-k+1)!(k-1)!} + \frac{n!}{(n-k)!k!} \\ &= \frac{n!}{(n-k+1)!k!}(k+n-k+1) = \frac{(n+1)!}{(n-k+1)!k!}. \end{aligned}$$

The proof of Theorem 1.1.9 and that of Theorem 1.1.8 are complete. □

Corollary 1.1.1. *Consider an n-element set X. Applying Theorems* 1.1.7, 1.1.8, *and Lemma* 1.1.2 *to the Boolean 2^X, we deduce the equation*

$$1 + n + \frac{n(n-1)}{2} + \frac{n(n-1)(n-2)}{3!} + \cdots + \frac{n!}{(n-k)!k!} + \cdots + n + 1 = 2^n. \qquad \square$$

Next we calculate the number of injective mappings, $\operatorname{Inj}(Y^X)$, for finite sets X and Y.

Theorem 1.1.10. *Let X and Y be two finite sets and $0 < n = |X| \leq m = |Y|$. Then*

$$|\operatorname{Inj}(Y^X)| = \frac{m!}{(m-n)!}. \tag{1.1.8}$$

Remark 1.1.3. If $n = 0$, then in agreement with (1.1.8) we define

$$|\operatorname{Inj}(Y^X)| = 1,$$

assuming that there exists the unique "empty mapping" with the empty domain.

To prove (1.1.8), we first consider a special case $m = n$.

Lemma 1.1.7. *If $|X| = |Y| = n, 0 < n < \infty$, then $|\operatorname{Inj}(Y^X)| = n!$.*

Proof. We again use the mathematical induction. The conclusion is clear if $n = 1$, because in this case there is the unique mapping from $X = \{x\}$ to $Y = \{y\} : x \mapsto y$, and this mapping is certainly injective (as well as surjective and hence bijective).

Now we assume the statement to be valid for all n-element sets, and select two $(n+1)$-element sets X and $Y = \{y_1, y_2, \ldots, y_{n+1}\}$. Pick an element $x_0 \in X$. The set $|\operatorname{Inj}(Y^X)|$

breaks down into the union of $n+1$ disjoint sets $A_1, A_2, \ldots, A_{n+1}$ such that an injective mapping $f : X \longmapsto Y$ belongs to the set $A_i, 1 \le i \le n+1$, if and only if $f(x_0) = y_i$. For a fixed image $f(x_0) = \hat{y} \in Y$, the set $X \setminus \{x_0\}$ can be injectively mapped into the set $Y \setminus \{f(x_0)\}$ in $n!$ ways due to the inductive assumption. Altogether, we have $|\operatorname{Inj}(Y^X)| = (n+1) \cdot n! = (n+1)!$. □

Define now the following equivalence relation on the set $\operatorname{Inj}(Y^X)$.

Two mappings $f, g \in \operatorname{Inj}(Y^X)$ *are equivalent if and only if they have the same range, that is,* $f(X) = g(X)$.

Problem 1.1.22. Verify that this is an equivalence relation in the sense of Definition 1.1.12 such that $|f(X)| = |X|$.

End of proof of Theorem 1.1.10. All mappings in any equivalence class have the same range $f(X)$. This range is an n-element subset of Y. Hence, there exists a one-to-one correspondence between the factor set and the set of all n-element subsets of Y, which is denoted by 2^Y_n. Lemma 1.1.7 implies that the cardinality of each equivalence class is $n!$. Now by Lemma 1.1.4, $|2^Y_n| = |\operatorname{Inj}(Y^X)|/n!$, and Theorem 1.1.8 yields equation (1.1.8), $|\operatorname{Inj}(Y^X)| = m!/(m-n)!$. □

Remark 1.1.4. Thus, for any finite sets X and Y we have found the numbers of injective, bijective and arbitrary mappings from X and Y. There is no such a simple formula for the number of surjective mappings. We will find that number in Section 4.1.

Several statements have already been proved by making use of a simple and powerful method—by establishing a one-to-one correspondence between the set in question and another set, whose cardinality can be found easier than the former, and we will use this approach again and again—see, for instance, the solution of Problem 1.4.16.

We end this section with a notation, which is convenient in many instances. Let the symbol $b \pmod p$ denote the remainder after dividing b over p.

Definition 1.1.18. For integer numbers a, b and a natural p, we write $a \equiv b \pmod p$ if p divides the difference $a - b$; in other words, the difference $a - b = kp$ with an integer k, or p divides both a and b with the same remainder. In this case the numbers a and b are called *congruent modulo* p.

For example, $7 \pmod 3 = 7 \pmod 2 = 1$, while $7 \pmod 4 = 3$; 5 and 11 are congruent modulo 2, $5 \equiv 11 \pmod 2$, but 5 and 4 are not.

Problem 1.1.23. Prove that the congruence is an equivalence relation on the set $\mathbf{Z}$ of integer numbers and describe its factor sets. Does this statement remain true on the set of natural numbers $\mathbf{N}$? The same question regarding the set of whole numbers $\{0, 1, 2, \ldots\}$.

Exercises 1.1.

Exercise 1.1.1. Compute the sums

(1) $\sum_{k=0}^{5} \frac{k}{k+2}$,

(2) $\sum_{k=1}^{5} \frac{1}{k}$,

(3) $\sum_{k=5}^{1} \frac{1}{k}$—here the summation index is decreasing,

(4) $\sum_{m,n=1}^{\infty} \frac{1}{(n+1)^{m+1}}$.

(5) The following transformation, called the *Abel transformation* or discrete summation by parts, is useful in many problems involving sums. Consider two finite or infinite sequences $\{a_k\}$ and $\{b_k\}, k = 1, 2, \ldots$, and the sequence of their pairwise products $\{a_k \cdot b_k\}, k = 1, 2, \ldots$. Introduce the partial sums of these sequences $B_n = \sum_{k=1}^{n} b_k$ and $S_n = \sum_{k=1}^{n} a_k \cdot b_k, k = 1, 2, \ldots$. Prove that for all $n \geq 2$

$$S_n = \sum_{k=1}^{n-1} (a_k - a_{k+1}) B_k + a_n B_n. \tag{1.1.9}$$

Use (1.1.9) to find the sums

(6) $\sum_{k=1}^{n} q^k, q$ is a constant,

(7) $\sum_{k=1}^{n} kq^k$,

(8) $\sum_{k=1}^{n} k \cos(kx)$ for a fixed number x.

Exercise 1.1.2. Prove the following statements by mathematical induction.

(1) $1^3 + 2^3 + \cdots + n^3 = [\frac{1}{2} n(n+1)]^2, \forall n \in \mathbf{N}$,

(2) (Al-Kashi) $1^4 + 2^4 + \cdots + n^4 = \frac{1}{30}(6n^5 + 15n^4 + 10n^3 - n), \forall n \in \mathbf{N}$,

(3) $2^n < n!$ for any natural $n \geq 4$.

Exercise 1.1.3. Find $\sum_{k=1}^{n} (2k-1)^3$.

Exercise 1.1.4. Prove by mathematical induction that for any natural n

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \frac{1}{i_1 i_2 \cdots i_k} = n,$$

where the sum runs over all k-tuples of natural numbers $i_1 < i_2 < \cdots < i_k$ for each $k = 1, 2, \ldots, n$.

Exercise 1.1.5. A sequence $\{a_1, a_2, \ldots, a_n, \ldots\}$ is called an *arithmetic progression* or an *arithmetic sequence*, if $a_{j+1} = a_j + d$ for each $j \geq 1$, where a constant d is called the *common difference* of the progression and a_1 is its first term. Find by mathematical induction an explicit formula for the general term a_n of an arithmetic progression and for the sum $\sum_{n=k}^{l} a_n$ of its $l+1$ consecutive terms. In particular, find the sum of the first l terms of an arithmetic sequence.

Exercise 1.1.6. Prove that a sequence $\{a_n\}, n \geq 1$, is an arithmetic progression if and only if $a_{n+1} + a_{n-1} = 2a_n, \forall n \geq 2$.

Exercise 1.1.7. (Hypsicle from Alexandria) Let $\{a_1, \ldots, a_n, a_{n+1}, \ldots, a_{2n}\}$ be an arithmetic progression with an even number of terms. Prove that $\sum_{k=n+1}^{2n} a_k - \sum_{k=1}^{n} a_k = bn^2$, where b is an integer number.

Exercise 1.1.8. A sequence $\{a_1, a_2, \ldots, a_n, \ldots\}$ is called a *geometric progression* or a *geometric sequence* if $a_{j+1} = q \cdot a_j$ for each $j \geq 1$, where q is called the *common ratio* of the progression and a_1 is its first term. Find an explicit formula for the general term a_n of the geometric progression and for the sum $\sum_{n=k}^{l} a_n$ of its $l - k + 1$ consecutive terms.

Exercise 1.1.9. Prove that a sequence $\{a_n\}, n \geq 1$, is a geometric progression if and only if $a_{n+1} \cdot a_{n-1} = a_n^2, \forall n \geq 2$.

Exercise 1.1.10. Find a closed-form expression for the sum $\sum_{k=1}^{n}(k^2 + k)$.

Exercise 1.1.11. How many zeros are in the end of the number 5!? 53!? 100!?

Exercise 1.1.12. Which is bigger, 300! or 100^{300}?

Exercise 1.1.13. Use mathematical induction to prove the fundamental theorem of arithmetic:

Any natural number $n > 1$ can be uniquely, up to the ordering of factors, written as a product of prime numbers. If n is prime, then the product contains only one factor.

Exercise 1.1.14. Find a flaw in the following "inductive proof" of the claim that all girls have sky-blue eyes:

The reader definitely knows at least one such a girl, which establishes the basis of induction. Suppose now that in any group of n girls all the girls have sky-blue eyes and deduce that, if so, then any group G of $n + 1$ girls possesses the same property. Indeed, let g be any girl in G. Consider an n-element group $G_1 = G \setminus \{g\}$ consisting of n girls. By the inductive assumption, all girls in G_1 have sky-blue eyes. Choose a girl g_1 in G_1; it is obvious that g and g_1 are two different girls. Next we remove g_1 from G_1 and replace her with g, that is, consider a set $G_2 = (G_1 \setminus \{g_1\}) \cup \{g\}$. The set G_2 also consists of n elements, hence by the inductive assumption, all girls in G_2 have sky-blue eyes. In particular, $g_1 \in G_2$, therefore, she also has sky-blue eyes, which in turn means that all girls in $G = G_1 \cup \{g\}$ have sky-blue eyes. Now the principle of mathematical induction implies the claim.

Exercise 1.1.15. Compare the sequences $a_n = 2^n$ and $b_n = n^2, n = 1, 2, \ldots$. We immediately verify that $a_1 = 2 > b_1 = 1$, while $a_2 = b_2$, $a_3 < b_3$, and $a_6 > b_6$. Determine which inequality, $a_n \geq b_n$ or $a_n \leq b_n$, is valid for all $n \geq n_0$, that is, for all but finitely many subscripts n. Find the smallest such n_0 and prove the correct inequality, $a_n \geq b_n$ or $a_n \leq b_n$, for all $n \geq n_0$.

Exercise 1.1.16. Prove the following modification of the axiom of mathematical induction: *If the statements S_1 and S_2 are valid and statements S_n and S_{n+1} together imply S_{n+2} for all natural n, then all the statements S_n, $n = 1, 2, \dots$ are valid.*

Exercise 1.1.17. Does the pair of sets $\mathbf{Z}_e = \{\dots, -4, -2, 0, 2, \dots\}$ and $\mathbf{Z}'_o = \{\dots, -3, -1, 0, 1, 3, \dots\}$ make up a partition of $\mathbf{Z}$?

Exercise 1.1.18. Prove that

$$\frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\cdots\left(\frac{1}{2}-n+1\right) = \frac{(-1)^{n-1}(2n-2)!}{2^{2n-1}(n-1)!}. \tag{1.1.10}$$

Exercise 1.1.19. Let $|X| = |Y| < \infty$. Prove that in this case $f \in Y^X$ is injective if and only if it is surjective, thus, in the case $|X| = |Y| < \infty$ the three properties (to be injective, to be surjective, and to be bijective) are equivalent.

Exercise 1.1.20. The binary relations below are given as sets of ordered pairs on appropriate sets. Are they reflexive, symmetric, antisymmetric, transitive, or neither? For the equivalence relations, describe their factor sets.

(A) The relations on the set $\{a, b, c, d\}$:
  1. $\varrho_1 = \{(a, a), (b, b), (c, c), (d, d)\}$,
  2. $\varrho_2 = \{(a, a)\}$,
  3. $\varrho_3 = \{(b, b), (c, c), (d, d)\}$,
  4. $\varrho_4 = \{(a, b), (b, a), (d, d)\}$,
  5. $\varrho_5 = \{(a, b), (b, c), (a, c), (d, d)\}$,
  6. $\varrho_6 = \{(a, a), (a, b), (b, b), (b, c), (c, c), (c, d), (d, d)\}$,
  7. $\varrho_7 = \{(a, a), (b, b), (c, c), (d, d), (a, b), (a, c), (a, d), (b, c), (b, d), (c, d)\}$.

(B) The relations on the set of real numbers $\mathbf{R}$:
  1. $\varrho_8 = \{x, y \in \mathbf{R} \mid x + y = 0\}$,
  2. $\varrho_9 = \{x, y \in \mathbf{R} \mid x + y = 0 \text{ or } x - y = 0\}$.

(C) The relations on the set of integer numbers $\mathbf{Z}$:
  1. $\varrho_{10} = \{m, n \in \mathbf{Z} \mid m = 2n\}$,
  2. $\varrho_p = \{m, n \in \mathbf{Z} \mid p \text{ divides } m - n, \text{ where } p \text{ is a given prime number}\}$.

Exercise 1.1.21. Find the flaw in the following "proof" of the claim: *A symmetric and transitive binary relation ϱ on a set X is an equivalence relation.*

Let $a, b \in X$ and $(a, b) \in \varrho$. Due to the symmetry, $(b, a) \in \varrho$, and by virtue of the transitivity, $(a, b) \in \varrho$ and $(b, a) \in \varrho$ together imply $(a, a) \in \varrho$. Thus, ϱ is reflexive.

Find a counter-example to the claim, that is, construct a symmetric and transitive but not reflexive binary relation.

Exercise 1.1.22. Let P^α stand either for the property P or for its negation. Prove that three properties, reflexivity (R), symmetry (S), and transitivity (T) are independent in totality, that is, for any triple of properties $(R^{\alpha_1}, S^{\alpha_2}, T^{\alpha_3})$ there exists a binary relation possessing this set of properties. By Theorem 1.1.7, to prove the claim it is enough to provide $2^3 = 8$ examples of binary relations.

Exercise 1.1.23. Prove that three properties, reflexivity (R), antisymmetry (AS), and transitivity (T) are independent in totality.

Exercise 1.1.24. How many binary relations are there on the set $\{1, 2, 3, 4, 5\}$? How many among them are reflexive? Symmetric? Antisymmetric? Transitive? How many possess any two or any three of these properties?

Exercise 1.1.25. By the definition, binary relations are sets, therefore, one can form their unions, intersections, etc.
(1) Let ρ and σ be two reflexive binary relations. Is any of the relations $\rho \cap \sigma$ or $\rho \cup \sigma$ reflexive?
(2) Let ρ and σ be two symmetric relations. Is any of the relations $\rho \cap \sigma$ or $\rho \cup \sigma$ symmetric?
(3) Let ρ and σ be two transitive relations. Is any of the relations $\rho \cap \sigma$ or $\rho \cup \sigma$ transitive?

Exercise 1.1.26. Prove Lemma 1.1.3.

Exercise 1.1.27. Is it true that $7 \equiv -8 \pmod 4$?

Exercise 1.1.28. Suppose that the binary relation of acquaintanceship on a set of people is symmetric. Prove that in a party of $n \geq 2$, at least two people have an equal number of acquaintances.

Exercise 1.1.29. Give an example of a binary relation ϱ in a Cartesian product $X \times Y$, which is not a mapping from X to Y. What restrictions should be imposed on ϱ to make it a mapping?

Give a definition a mapping $f : X \to Y$ as a binary relation ϱ in the Cartesian product $X \times Y$.

Exercise 1.1.30. Let $\mathbf{Z}_2 = \{0, 1\}$ and $\mathbf{Z}_2^n$ be the Cartesian product of n copies of $\mathbf{Z}_2$. A *Boolean function* of n variables is a mapping $f : \mathbf{Z}_2^n \to \mathbf{Z}_2$. How many different Boolean functions of n variables are there?

Exercise 1.1.31. Prove the equation

$$n! = \int_0^\infty e^{-t} t^n dt, \quad n = 0, 1, 2, \ldots .$$

1.2 The sum and product rules

In this section we study two important results called the sum rule and the product rule, which demonstrate themselves in many combinatorial problems. We will see that they are nothing but the formulas for calculating the cardinalities of the union and the Cartesian product of finite sets. We introduce these rules by considering simple model problems.

Coffee-time browsing

- www.math.csusb.edu/~history/Mathematicians/Descartes.html (Descartes' biography)
- www.saintjoe.edu/~karend/m122/CountingSlides.ppt (Sum and Product rules)

Problem 1.2.1. In a group of students, each person studies one and only one of three foreign languages: six people take French, eight take German, and nine students take Spanish. How many students are there in the group?

Solution. Denote the set of all students in the group by X, the subset of students studying French by X_F, the subset of students studying German by X_G, and the subset of students studying Spanish by X_S. Since each student studies at least one language, we can represent X as the union,

$$X = X_F \cup X_G \cup X_S.$$

Moreover, these subsets are pairwise disjoint,

$$X_F \cap X_G = X_F \cap X_S = X_G \cap X_S = \emptyset,$$

for none student studies two languages. Thus, by Lemma 1.1.2 with $m = 3$, $|X| = 6 + 8 + 9 = 23$. □

There are many similar problems where the set in question is the union of several disjoint subsets, or this set can be put in a one-to-one correspondence with a union of disjoint sets. Consequently, the cardinality of the set can be calculated by making use of Lemmas 1.1.1 or 1.1.2. It is said in such situation that the solution was derived by the sum rule; both these lemmas are referred to as the sum rule as well. Without using the set theory terminology the rule can be stated as follows.

If one task can be performed in k ways and another task in l ways, and these tasks cannot be done simultaneously, then one of the two tasks can be done in $k + l$ ways. It is clear after our analysis of Problem 1.2.1 that the latter statement is just a descriptive formulation of Lemma 1.1.1, where $|X| = k$ and $|Y| = l$.

The sum rule can also be stated in other terms.

The Sum Rule. If the finite sets $X_1, X_2, \dots, X_m$ form a partition of a set X, then

$$|X| = |X_1| + |X_2| + \cdots + |X_m|. \tag{1.2.1}$$

Evidently, (1.2.1) is equivalent to Lemma 1.1.2.

If there is a sum rule, then there likely is a product rule. To introduce it, we again analyze a model problem.

Problem 1.2.2. Identification cards on Small Planet contain two characters, one capital Latin letter and one Hindu–Arabic digit, for example, "$S-8$". How many various cards are there, if we can use all 26 letters and 10 digits?

Solution. First of all, we have to state unequivocally what cards must be considered identical, and what cards are different. Since we consider a mathematical problem, we do not take into consideration size, color, font, etc. Two cards are considered as different, if they have *different pairs* of symbols, that is, if at least one symbol on either card is distinct from the corresponding symbol on another card. In other words, to say that two cards are identical is just to say that they have both the same letter and the same digit. *We reiterate here this statement, because clear qualification of what objects are distinct in a combinatorial problem and which ones are the same (are identical) is a crucial step in solving the problem; otherwise, two people can read the same words but solve two different problems.*

Another important issue is the *ordering* of characters. In this problem, should we count the cards "$S-8$" and "$8-S$" as different or identical?

As the matter of fact, these are two different problems. Combinatorics itself does not know whether or not the order of elements is substantial, combinatorics only provides necessary means for solving both problems. This is the solver's task to clarify the problem and choose the right approach. The distinction between problems where order of elements is or is not essential, will be discussed in more detail later on in this chapter.

In Problem 1.2.2 we assume that the first character on the card is always a letter and the second one is a digit. Thus from our standpoint, each card is an *ordered pair* of symbols (λ, δ), where λ may be any of 26 letters and δ any of ten digits. Having said the key words "ordered pair", we immediately recognize that these objects make up the Cartesian products of sets and we can use the latter as a mathematical model in our problem. Denote the set of all characters of the English alphabet by $\Lambda = \{A, B, C, \dots, Y, Z\}, |\Lambda| = 26$, and the set of digits by $\Delta = \{0, 1, \dots, 8, 9\}, |\Delta| = 10$. Our discussion implies that there is a one-to-one correspondence between the set of various identification cards we sought for and the Cartesian product $\Lambda \times \Delta$. Thus, by Theorem 1.1.4 the number of different cards is equal to $|\Lambda \times \Delta| = |\Lambda| \cdot |\Delta| = 260$. □

Henceforth, we say that a solution was derived by making use of the product rule.

The Product Rule. The product rule means that we have established a one-to-one correspondence between a set under consideration and some Cartesian product, and computed the cardinality of this product by making use of Theorems 1.1.4–1.1.5. The latter theorems are also referred to as the product rule.

The subsequent problems illustrate the sum and product rules.

Problem 1.2.3. Find the number of car license plates containing four letters and three digits (as in the previous problem, there are 26 letters and 10 digits available).

Solution. Using the same notation and reasoning as in the preceding problem and applying the product rule, we get the answer: there are

$$|\Lambda \times \Lambda \times \Lambda \times \Lambda \times \Delta \times \Delta \times \Delta| = |\Lambda|^4 \cdot |\Delta|^3 = 26^4 \cdot 10^3 = 456\,976\,000$$

license plates. □

Problem 1.2.4. Find the number of license plates containing four letters and either one, or two, or three digits.

Solution. The desired set of license plates Π comprises objects of three types—with one, or with two, or with three digits. Thus, we can set up the equation

$$\Pi = \Pi_1 \cup \Pi_2 \cup \Pi_3,$$

where Π_i denotes the set of plates containing i digits, $i = 1, 2, 3$. Moreover, a claim that a plate contains *one* digit, clearly distinguishes such a plate from those with two or three digits, whence the subsets $\Pi_i, i = 1, 2, 3$, are pairwise disjoint. Hence we can apply equation (1.2.1) and conclude that $|\Pi| = |\Pi_1| + |\Pi_2| + |\Pi_3|$.

The cardinal number $|\Pi_3| = 456\,976\,000$ was found in Problem 1.2.3. In the same way we calculate the cardinal numbers $|\Pi_1|$ and $|\Pi_2|$,

$$|\Pi_1| = |\Lambda \times \Lambda \times \Lambda \times \Lambda \times \Delta| = |\Lambda|^4 \cdot |\Delta| = 26^4 \cdot 10 = 4\,569\,760,$$
$$|\Pi_2| = |\Lambda \times \Lambda \times \Lambda \times \Lambda \times \Delta \times \Delta| = |\Lambda|^4 \cdot |\Delta|^2 = 26^4 \cdot 10^2 = 45\,697\,600,$$

and the total number of plates is

$$|\Pi| = |\Pi_1| + |\Pi_2| + |\Pi_3| = 4\,569\,760 + 45\,697\,600 + 456\,976\,000 = 507\,243\,360.$$ □

Problem 1.2.5. Three polyhedron-shape beads having, respectfully, six faces (cube), eight faces (octahedron), and ten faces (decahedron) are rolled simultaneously. Their faces are numbered, respectively, from 1 through 6, from 1 through 8, and from 1 through 10. After each roll we write down the numbers on the face they landed.
(1) In how many different ways can these beads land?
(2) In how many different ways can these beads land, if at least two of them fall on the faces marked with a 1?

Solution. (1) The result of each roll can be written as an ordered (since, for instance, a 7 cannot occur in the first position) triple (a, b, c), where $1 \le a \le 6$, $1 \le b \le 8$ and $1 \le c \le 10$. Thus, we can directly use the product rule, implying that there are $6 \times 8 \times 10 = 480$ variants of landing these beads.

(2) Let P_{cu} be the set of all possible results of the landing such that the octahedron and the decahedron read a 1, and the cube shows any face; obviously (or by the product rule again), $|P_{cu}| = 6 \times 1 \times 1 = 6$. The sets P_{oc} and P_{de} are defined in a similar way, $|P_{oc}| = 8$ and $|P_{de}| = 10$. After that we are compelled to apply the sum rule and to compute the "answer": $6 + 8 + 10 = 24$.

However, the sum rule does not apply here and this "answer" is wrong, since the three sets P_{cu}, P_{oc} and P_{de} *are not mutually exclusive*, they have a non-empty intersection, containing one element, namely the triple $(1, 1, 1)$. To take this intersection into account, it is convenient to introduce three other sets, $\widehat{P}_{cu}, \widehat{P}_{oc}$ and $\widehat{P}_{de}$, where $\widehat{P}_{cu}$ stands for the set of all possible results of landing of the beads such that the octahedron and the decahedron read a 1, but the cube shows any face but a 1; $\widehat{P}_{oc}$ and $\widehat{P}_{de}$ are defined similarly. It is clear now that $|\widehat{P}_{cu}| = 5$, since one of the six faces of a cube is now excluded, and $|\widehat{P}_{oc}| = 7$, $|\widehat{P}_{de}| = 9$.

The three "hatted" sets are disjoint, but there appears now another obstacle: these sets do not exhaust all the ordered triples in the problem. Introduce the set $P_1 = \{(1, 1, 1)\}$ corresponding to the case when all three beads land on a 1; thus, $|P_1| = 1$. These four sets, $\widehat{P}_{cu}, \widehat{P}_{oc}, \widehat{P}_{de}$, and P_1, partition the set of all the possible outcomes in the problem, and by the sum rule we have the answer, $5 + 7 + 9 + 1 = 22$. □

Problem 1.2.6. In how many ways can one choose two movies in different genres out of five different comedies, seven different thrillers, and ten different dramas?

Solution. Combining the Sum and product rules, we arrive at the answer: $5 \times 7 + 5 \times 10 + 7 \times 10 = 155$ variants. □

Problem 1.2.7. A Combi Club has 18 members. In how many ways can the members elect the President and the Treasurer of the Club?

Solution. Let S be the set of the Club members, $|S| = 18$. If a student s_1 was elected the Club President, then there are only 17 candidates the Treasurer can be chosen from. Thus, there are 17 ways to elect the President and the Treasurer given that the student s_1 is to be the President. If the student s_2 is to be the President, we also have 17 possibilities, etc. Since these 18 options do not intersect, we can apply the sum rule and get $\underbrace{17 + 17 + \cdots + 17}_{\text{18 addends}} = 17 \times 18 = 306$ different results of the elections. □

Remark 1.2.1. We notice in this problem another issue, important in many combinatorial problems. In our solution we *implicitly* assume that one student cannot serve simultaneously as the President and the Treasurer. In Problem 1.2.7 such repeating choices were not allowed, but it may be another way elsewhere. Actually in the business world we often see a person who simultaneously is the President and the CEO of a company. To distinguish these two kinds of problems, we say that a problem allows or does not allow *repetition*. It is worth emphasizing that, like the order of elements, the assumption of (non)repetition depends on a particular problem, combinatorics only provides the means for solving both kinds of problems.

The end of the solution of this problem is similar to applying the product rule. However, the number 17 here is not a cardinality of a specific set—the sets of candidates to elect the Treasurer for different Presidents elected are different, even though they all have the same cardinality. It may be convenient to model this and similar problems by a special drawing, *tree of alternatives* similar to one in Fig. 1.1. We study such drawings in more detail in Chapter 2.

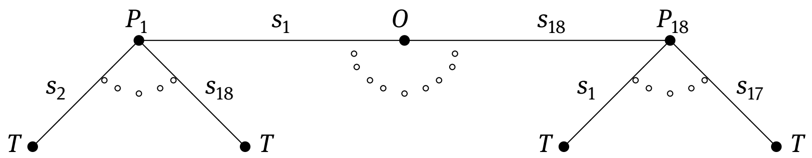


Figure 1.1: The tree of alternatives in Problem 1.2.7.

This tree represents all possible outcomes of the voting. Since any one of the 18 students $s_1, s_2, \dots, s_{18}$ can be the Club President, 18 first-level branches, incident to the root O and labeled by $s_1, s_2, \dots, s_{18}$, represent 18 possible results of the President election; Fig. 1.1 displays only few branches corresponding to s_1 and s_{18}. If the President has been elected, there are only 17 candidates for the Treasurer, however, the sets of candidates are all different. Indeed, if the student s_1 has been elected as the President (this case is depicted by the subtree at the vertex P_1 in Fig. 1.1) then only the students $s_2, s_3, \dots, s_{18}$ may run for the Treasurer, hence there are 17 second-level branches incident to the left vertex P and labeled by $s_2, \dots, s_{18}$. If the student s_{18} has been elected as the President (this case is depicted by the subtree at the vertex P_{18}) then only the students $s_1, s_2, \dots, s_{17}$ may run for the Treasurer, hence there are 17 second-level branches incident to the right vertex P and labeled by $s_1, \dots, s_{17}$; likewise, the tree has 16 intermediate branches between s_1 and s_{18}. The tree has $18 \times 17 = 306$ pendant vertices representing all possible results of the voting.

The tree in Fig. 1.1 is *regular*, that is, every vertex except for the pendant ones, has the same number of incident second-level branches. In other problems these quantities can be different. To solve such problems, the sum rule may be of use.

Problem 1.2.8. Four people—A, B, C, and D took part in a car race. A student has only partial information on the results of the race. It is known that B lost to A, C was not the last one, and there were no ties. How many different results are possible in the race?

Solution. The tree of alternatives for the race is drawn in Fig. 1.2. Since B finished after A, A can finish either first, or second, or third. The first level of the tree, above the broken line $\alpha - -\alpha$, has three branches, representing these alternatives. Next, if A is first, then B can be either second, or third or fourth; if A is second, B can be either third or fourth; and if A is third, B can be only the last one. The second level of the tree,

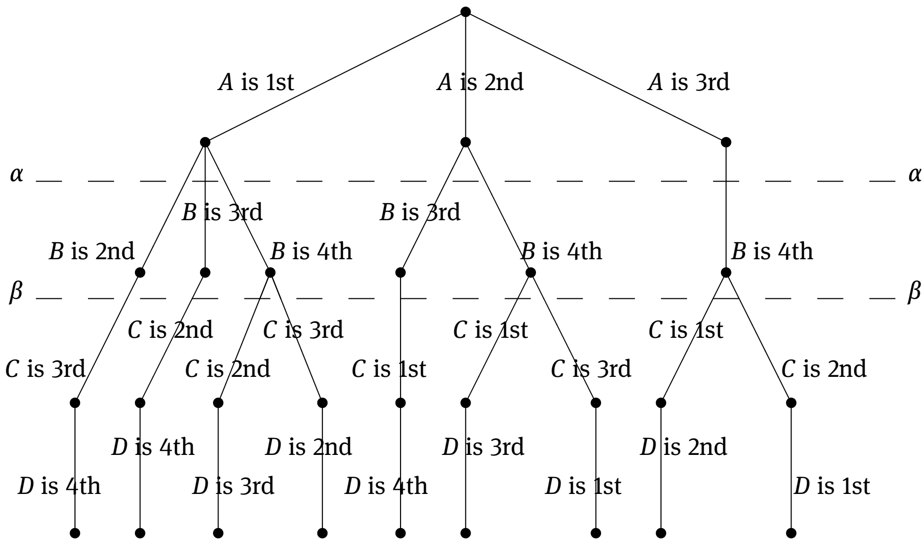


Figure 1.2: The tree of alternatives in Problem 1.2.8.

above the broken line $\beta - -\beta$, represents these alternatives. The entire tree (Fig. 1.2) has nine pendant vertices corresponding to nine possible results of the race. □

Exercises and Problems 1.2.

Exercise 1.2.1. Bob participates in two sweepstakes simultaneously. In the first one he can win one out of four books, in the other—one out of five tapes. How many different pairs of prizes can he bring home?

Exercise 1.2.2. Betty takes part in two book raffles. In the first one she can win one of four different books. In the second raffle she can win one of five different books, but one of them is the same as in the first drawing. If one wins the same book twice, she may change one of them for another book distinct from any book

Exercise 1.2.3. A dog and a cat can peacefully sit and dine side by side, but if two dogs or two cats are sitting alongside, they start fighting. In how many ways can n dogs and n cats be peacefully seated at a round table?

Exercise 1.2.4. How many a license plates consisting of one letter and two digits are there if at least one of these digits is to be 9?

Exercise 1.2.5. How many license plates consisting of three letters and four digits are there if at least one of these digits is 9?

Exercise 1.2.6. Four cars take part in a race. How many ways are there to finish the race if ties between the second and the next places are allowed but the winner cannot make a tie?

Exercise 1.2.7. Among the following nine sets, what combinations of them make up partitions of the set of natural numbers $\mathbf{N}$?
(1) $\mathbf{N}_1 = \{1\}$,
(2) $\mathbf{N}_2 = \{1, 2\}$,
(3) $\mathbf{T}_0 = \{n = 3k \mid k \in \mathbf{N}\}$,
(4) $\mathbf{T}_1 = \{n = 3k + 1 \mid k \in \mathbf{N}\}$,
(5) $\mathbf{T}_2 = \{n = 3k + 2 \mid k \in \mathbf{N}\}$,
(6) $\mathbf{P}$—the set of all prime numbers,
(7) $\mathbf{P}^{\mathbf{c}}$—the set of all composite numbers greater than 1,
(8) $\mathbf{N}$,
(9) $\emptyset$.

Exercise 1.2.8. A gentleman has eight shirts and five ties. In how many ways can he choose a shirt and a tie to go out, if he cannot combine a shirt S_1 with ties T_1 and T_2, and also a shirt S_2 with ties T_1 and T_3?

Exercise 1.2.9. A family consisting of mother, father, four daughters, and two sons participates in a mixed doubles badminton tournament, where each team consists of a female and a male player. How many various family teams are possible if the youngest daughter does not want to be on a team with her elder brother?

Exercise 1.2.10. Prove the following combination of the sum and product rules, when subsets may have different cardinalities. Let ϱ be a subset of a direct product $X \times Y$ of finite sets $X = \{a_1, \dots, a_m\}$ and $Y = \{b_1, \dots, b_n\}$. Then

$$|\varrho| = \sum_{k=1}^{m} \operatorname{card}(a_k, \cdot) = \sum_{l=1}^{n} \operatorname{card}(\cdot, b_l), \tag{1.2.2}$$

where $\operatorname{card}(a_k, \cdot)$ or $\operatorname{card}(\cdot, b_l)$ is, respectively, the number of ordered pairs in ϱ with the first element a_k or with the second element b_l.

Thus, ϱ is a binary relation between X and Y and equation (1.2.2) allows us to compute the cardinality of an arbitrary binary relation.

Exercise 1.2.11. How many divisors does the number $2^3 3^4 5^5 7^6$ have? Find the sum of all divisors.

Exercise 1.2.12. A student put 5 sheets of paper in a shredder. The shredder cut some of these sheets into 5 parts, then cut some of these pieces into 5 parts, and so on. When the shredder stopped, the student found 2006 small pieces of paper in the shredder. Is this count correct?

Exercise 1.2.13. In how many ways is it possible to place three rooks on the 8×8 chessboard so that no two of them can attack one another?

Exercise 1.2.14. In the year of 2006 there were 2 006 meetings of student clubs in a Big Club College, each meeting attended by 40 students. For any two meetings, exactly one student attended both of them. Prove that there was a student who attended all 2006 meetings.

Exercise 1.2.15. How many four-digit natural numbers are multiple of 7?

Exercise 1.2.16. The vertices of a triangle belong to the set of the vertices of a given convex n-gon, but no side of the triangle is an entire side of the n-gon. How many such triangles are there?

Exercise 1.2.17. All integer numbers from 1 through 2 222 222 are written in a row. How many times each of the digits 0, 1, 2 appears in this series of digits?

1.3 Arrangements and permutations

In this section we deal with ordered totalities of objects, called here arrangements. To introduce them, we consider a model problem.

Coffee-time browsing

– www.usna.edu/Users/math/wdj/book/node156.html (Listing permutations)

Problem 1.3.1 (Problem 1.2.7 revisited). Combi Club has 18 members. In how many ways can the members elect the President and the Treasurer of the Club?

Solution. The following solution is similar to the solution of this problem in Section 1.2, but we put it in different terms. Suppose that the Election Board reports the results of the voting, using the form

$P =$	$T =$

and fills in two blank spaces with the names of the students elected. To convert this form to standard mathematical notation, we introduce two sets, the 2-element set $C = \{P, T\}$ symbolizing the positions to be filled in and the set $S = \{s_1, \dots, s_{18}\}$ of the Club members. Denote the result of an election by $(s(P), s(T))$, where $s(P)$ signifies the student having been elected to preside, and $s(T)$ stands for the student having been elected to count money.

We see that the result of every voting can be described as a mapping v with the domain C and the codomain S. Choosing the President of the Club, we associate with the element $P \in C$ an element $s' = s(P) \in S$; choosing the Treasurer, we associate an element $s'' = s(T) \in S$ with the element $T \in C$.

Suppose that a student cannot simultaneously serve as the President and the Treasurer, that is, different persons have to be elected for these two positions; in our notation it must be $s' = s(P) \neq s'' = s(T)$. Thus the mapping v is to be injective, $v \in \mathrm{Inj}(S^C)$. Vice versa, each injective mapping $v \in \mathrm{Inj}(S^C)$ can be interpreted as the result of some voting in this Club. Hence, we see for ourselves that there is a one-to-one correspondence between the set of all possible results of the election and the set $\mathrm{Inj}(S^C)$ of all injective mappings from C into S. Now Theorem 1.1.10 with $n = 2$ and $m = 18$ implies that there are $18!/16! = 18 \cdot 17 = 306$ different outcomes of the election, as we have already found in Problem 1.2.7.

Suppose now that one person may be elected for both positions. In this case we have to take into account not only injective but all mappings from C to S, that is the entire power set S^C. By Theorem 1.1.6, we get $|S^C| = 18^2 = 324$ different results of this voting. □

Remark 1.3.1. The difference $324 - 306 = 18$ gives the number of possible outcomes of the voting, when one student is elected for both offices.

Considering this problem as a model, we give the following definitions. It is clear that the answer does not depend on particular sets, like C and S in Problem 1.3.1, it only depends upon their cardinalities, therefore, for the domains of mappings in these definitions we always use natural segments $\mathbf{N}_n$ with various n. For instance, in Problem 1.3.1 $n = 2$.

Definition 1.3.1. Let A be a finite set, $|A| = m \in \mathbf{N}$. An arbitrary mapping $f : \mathbf{N}_n \to A$ is called an *n-arrangement* **with repetition** of the elements of the set A, or more precisely, arrangement of m elements taken n at a time.

Let the element $a_i \in A$ be the image of the element $i \in \mathbf{N}_n$ under the mapping $f, a_i = f(i)$. Since arrangements are *ordered* totalities, we denote the arrangements with repetition by $(a_1, a_2, \ldots, a_n)$, using the same notation as for n-tuples. If $|A| = m$, then the number of n-arrangements with repetition is denoted by $A_{\text{rep}}(m, n)$.

Theorem 1.3.1. *By Theorem* 1.1.6, *the number of n-arrangements with repetition is*

$$A_{\text{rep}}(m, n) = |A^{\mathbf{N}_n}| = |A|^n = m^n. \tag{1.3.1}$$ □

This number certainly depends upon the cardinality m of the set A, but not on the specific nature of its elements.

Definition 1.3.2. Let A be a finite set, $|A| = m \in \mathbf{N}$. Any **injective** mapping $f : \mathbf{N}_n \to A$ is called an *n-arrangement* **without repetition** of the elements of A, or more precisely, arrangement of m elements without repetition, taken n at a time.

We often omit the specification "without repetition", assuming that an "n-arrangement" always means an arrangement without repetition, but "with repetition" must be specified. Arrangements with and without repetition are denoted by the same symbol $(a_1, a_2, \ldots, a_n)$. If $|A| = m$, the number of n-arrangements without repetition is denoted[5] by $A(m, n)$.

---

5 Sometimes the notations $P(m, n)$ and ${}_mP_n$ are used, and these arrangements are called n-permutations.

Theorem 1.3.2. *By Theorem* 1.1.10*,*

$$A(m,n) = \left|\mathrm{Inj}(A^{\mathbf{N}_n})\right| = \frac{m!}{(m-n)!}, \quad 1 \le n \le m. \tag{1.3.2}$$

We also set by definition

$$A(m,0) = 1 \quad \text{and} \quad A(m,n) = 0 \quad \text{if } n < 0 \text{ or } n > m. \tag{1.3.3}$$

□

Remark 1.3.2. In other words, arrangements without repetition of the elements of a set A can be considered as *ordered* n-subsets of A. Thus, to introduce the arrangements in a proper way, we have to either accept *ordered sets* as a primary, undefined concept, or define them through another notion. At the same time the arrangements with repetition can contain several copies of the same element, although no set can contain repeating elements. Therefore, the arrangements cannot be defined as sets. To unify definitions, it is convenient to introduce arrangements both with and without repetition as mappings, as it has been done above.

Definition 1.3.3. In the case $m = n$ the arrangements (without repetition) are called *permutations* (of n elements) or n-permutations; their number is denoted by $P(n)$.

Theorem 1.3.3. *By Lemma* 1.1.7*, the number of n-permutations is*

$$P(n) = A(n,n) = n!, \quad n > 0. \tag{1.3.4}$$

□

Remark 1.3.3. Therefore, the permutations of a set A are bijective mappings. If the elements of A are ordered, for instance, they are numbered by natural numbers, like $A = \{a_1, a_2, \dots, a_m\}$, then any permutation gives a reordering of A, for example, $(a_{i_1}, a_{i_2}, \dots, a_{i_m})$. This sequence of elements, that is, the ordered image under the original bijection, is also often called a permutation of the set A. We return to permutations in Section 4.5.

Example 1.3.1. Thus, if a Board consists of seven members, they can be seated in a row in $P(7) = 7! = 5040$ ways.

Problem 1.3.2. Prove a recurrence relation

$$P(m) = A(m,m) = A(m,n) \times A(m-n, m-n), \quad 0 \le n \le m.$$

Remark 1.3.4. If formula (1.3.4) has been proven independently, say by mathematical induction, then we can deduce (1.3.2) from (1.3.4) and Problem 1.3.2.

Problem 1.3.3. A bus route has nine stops, excluding the departure stop; there are 23 passengers in the bus. In how many ways can they get off the bus?

Solution. First of all we have to clarify which runs of the bus we treat as different. We consider two runs to be different, if there is at least one stop such that the sets (not the quantity!) of passengers, leaving the bus at this stop in the first run and in the second run, differ. Denote the set of all passengers by $P, |P| = 23$, and the set of the stops by $S, |S| = 9$. Now we can associate a mapping $r : P \to S$ with every run of the bus. Namely, if a passenger p gets off the bus at a stop s, then $r(p) = s$. Next we notice that there is a one-to-one correspondence between the bus runs and these mappings. By Theorem 1.3.1, there are 9^{23} different runs of the bus. □

Exercises and Problems 1.3.

Exercise 1.3.1. Compute $A(m, n)$ for all $m, n, 0 \le n \le m \le 5$, and $P(n)$ for all $n, 0 \le n \le 10$.

Exercise 1.3.2 (Problem 1.2.7 revisited again). Suppose that, for certain personal reasons, Ann and Alex cannot serve as officers together and Bob cannot be the treasurer. In how many ways can the officers of the Combi Club be elected?

Exercise 1.3.3. Given 6 different balls and 4 different urns, in how many ways can we place 4 balls in 4 urns, one ball in an urn?

Exercise 1.3.4. We have 7 tasks to do. In how many ways can we choose 5 of them to perform one task a day during 5 consecutive weekdays?

Exercise 1.3.5. How many 9-digit natural numbers are there containing every digit $1, 2, \ldots, 9$ once?

Exercise 1.3.6. How many 10-digit natural numbers are there containing each digit $0, 1, 2, \ldots, 9$ once?

Exercise 1.3.7. How many 10-digit numbers are there with the sum of digits equal 4?

Exercise 1.3.8. Find the sum of all integer numbers containing digits $1, 2, 3, 4$, such that any digit occurs in each number once.

Exercise 1.3.9. Find the sum of all 4-digit integers containing digits $1, 2, \ldots, 9$, such that any digit occurs in each number no more than once.

Exercise 1.3.10. Town Infiniburg occupies the entire plane. It has s straight parallel streets. In addition, it has t more straight streets such that none among them is parallel to any one among the other $s + t - 1$ streets. Moreover, no three streets have a common intersection. Into how many blocks have the streets split the town?

Exercise 1.3.11.
(1) Consider the first 1 000 000 natural numbers. What numbers make the majority among them: those whose decimal representation contains a 1, or those without a 1?
(2) Solve the same problem for the first 10 000 000 natural numbers.
(3) How many among the first 1 000 000 natural numbers contain exactly one of the digits $2, 3$, and 4?

(4) How many among the first 1 000 000 natural numbers contain exactly one digit 2 and two digits 3?

Exercise 1.3.12. How many of each of the digits $0, 1, 2, \dots, 9$ must be used to represent all integer numbers from 1 through 9 999 inclusive? From 1 through $10^k - 1$?

Exercise 1.3.13. How many 6-digit odd numbers without repeating digits are there? How many such numbers begin with a 1?

Exercise 1.3.14. How many permutations of the 10 digits $0, 1, \dots, 9$ contain either the sequence 246 or the sequence 578, but not both?

Exercise 1.3.15. How many permutations of the 10 digits $0, 1, \dots, 9$ contain the sequence 246 or the sequence 680, but not both?

Exercise 1.3.16. How many permutations of the 10 digits contain either the sequence 246, or the sequence 680, or both?

Exercise 1.3.17. How many different 10-digit natural numbers are there consisting only of digits $1, 2$, and 3, if a 3 appears precisely two times?

Exercise 1.3.18. A combination lock has 5 disks with 12 different symbols on each. Only one combination opens the lock. Assuming that it takes 10 seconds to change a combination, what is the maximum time necessary to open the lock at random?

Exercise 1.3.19. How many different pairs of disjoint subsets does an n-element set have?

Exercise 1.3.20. Solve the equations for integer n and k:
(1) $A(n, 2) = 20$,
(2) $P(n) = 5P(k)$.

Exercise 1.3.21. There are n traffic lights in Lighttown, each with three standard colors—green, yellow, and red.
(1) How many different combinations of signals can they show?
(2) Answer the same question, if the lights TL_1 and TL_2 can only be either both yellow or in opposite 'green-red' state, that is, if one of them is green, then another must be red and vice versa.

Exercise 1.3.22. How many ways are there to assign 12 players to 5 coaches for practice?

Exercise 1.3.23. How many 10-digit phone numbers are there such that 0 and 9 do not appear among the first four digits?

Exercise 1.3.24. How many 3-digit natural multiples of 3 are there which contain a digit 9 in their decimal representation? We recall that an integer number is divisible by 3 if and only if 3 divides the sum of all its digits.

Exercise 1.3.25. How many 6-digit natural numbers divisible by 9 are there such that their last digit is 9? We recall that an integer number is divisible by 9 if and only if 9 divides the sum of all its digits.

Exercise 1.3.26. Consider all 10^5 whole 5-digit numbers attaching, if necessary, a few zeros in front of such a number, like 00236. How many of them contain exactly one digit 0, one 1, one 2, and one 3?

Exercise 1.3.27. Show that the elements of an n-element set can be ordered in $n!$ ways.

Exercise 1.3.28. How many 4-arrangements of the letters a, b, c, d, e, f are there if they
(1) begin with an a?
(2) contain the letter a?
(3) contain two letters a, b?
(4) contain the letters a, b in this order?

Exercise 1.3.29. Find the number of arrangements of n different objects taken r at a time, if each arrangement must contain p specified objects from the given n. When do such arrangements exist?

Exercise 1.3.30. Find the number of arrangements of n different objects taken r at a time, if each arrangement must contain p specified objects from the given n, but cannot contain any of the other q specified objects (assuming $p + q \leq n$).

Exercise 1.3.31.
(1) A college prepares three-student teams for a tournament. How many such teams can be made, if the students can be distinguished only by their standing—freshmen, sophomores, juniors, seniors?
(2) To get the Mass Award, the college must have at least 25 teams. Is it possible to get this award, if this year the school has no seniors? If the school has only freshmen and sophomores?

1.4 Combinations

This section deals with unordered totalities of objects. The binomial coefficients and Catalan numbers inevitably make their presence felt here. We also consider the trajectory method.

Coffee-time browsing

– mathforum.org/dr.math/faq/faq.pascal.triangle.html (Pascal triangle)
– www-history.mcs.st-and.ac.uk/Biographies/Pascal.html (Pascal's biography)
– www.gap-system.org/~history/Mathematicians/Catalan.html (Catalan's biography)
– http://en.wikipedia.org/wiki/Walther_von_Dyck (von Dyck's biography)
– http://en.wikipedia.org/wiki/Dyck_language (Dyck language)

- www.answers.com/topic/leopold-kronecker (Kronecker's biography)
- www.gap-system.org/~history/Biographies/Vandermonde.html (Vandermonde's biography)
- www.gap-system.org/~history/Biographies/Kaplansky.html (Kaplansky's biography)

Problem 1.4.1. The same Combi Club with 18 members (see Problems 1.2.7 and 1.3.1) has to send two of its members to a meeting. How many ways are there to select these two delegates assuming that both have the same rights and responsibilities?

Solution. As in Problem 1.2.7, we have to choose two different people. However, unlike Problem 1.3.1, this problem emphasizes that the order of the members chosen makes no difference, only the two selected names matter. We immediately recall that these are sets, where the order of elements does not count. Thus, any 2-member delegation can be viewed as a 2-element subset of the same set S of the Club members, $|S| = 18$, and we have to compute the number of 2-element subsets in an 18-element set. Theorem 1.1.8 with $n = 18$ and $k = 2$ yields $|2^S_2| = 18!/(2! \cdot 16!) = 153$ delegations. □

Considering this analysis, we give the following definition.

Definition 1.4.1. Given a set X, any k-element subset of X is called a combination (a *k-combination without repetition*) of the elements of X taken k at a time. The number of k-combinations of the elements of an n-element set X is hereafter denoted by $C(n, k)$; sometimes the symbols ${}_nC_k$ and C^k_n are also used. These quantities are also called *binomial coefficients* and denoted by $\binom{n}{k}$. For integer $n \geq 0$ we use both symbols $C(n, k) = \binom{n}{k}$ interchangeably, for other n we will write only $\binom{n}{k}$.

In Section 1.1, the set of all such subsets, that is, the set of k-combinations was denoted by 2^X_k. By Theorem 1.1.8, if $|X| = n$, then, for any $0 \leq k \leq n$,

$$|2^X_k| = \frac{n!}{(n-k)! \cdot k!}.$$

Clearly, this number does not depend on a particular set X, so long as its cardinal number is $|X| = n$.

Theorem 1.4.1. *We immediately deduce from the latter formula the number of combinations (binomial coefficients),*

$$C(n, k) = \frac{n!}{(n-k)! \cdot k!}, \quad 0 \leq k \leq n. \tag{1.4.1}$$

An n-element set cannot contain k-element subsets with $k > n$. Thus, for $k > n$ and $k < 0$ we set $C(n, k) = 0$. □

Corollary 1.4.1. *Now Corollary* 1.1.1 *can be stated as*

$$C(n, 0) + C(n, 1) + \cdots + C(n, k) + \cdots + C(n, n) = 2^n. \qquad □$$

The binomial coefficients with a negative upper index $-n < 0$, that is, $n > 0$, are defined as

$$\begin{aligned}\binom{-n}{k} &= (-1)^k \frac{n(n+1)\cdots(n+k-1)}{k!} = \frac{(-n)(-n-1)\cdots(-n-k+1)}{k!} \\ &= (-1)^k \binom{n+k-1}{k} = (-1)^k C(n+k-1, k).\end{aligned}$$

Some important properties of the binomial coefficients are discussed in the sequel problems, including problems in the end of this section. Many more problems are scattered in the literature, in addition to the references mentioned above; see, for example, [35, Sect. 1.2.6].

Problem 1.4.2. Show that, for $0 \le k \le n$,

$$C(n,k) = C(n-1,k-1) + C(n-1,k). \tag{1.4.2}$$

Solution. The equation easily follows from (1.4.1), but we shall prove it using specifically combinatorial reasoning useful in many instances (cf. the proof of Theorem 1.1.9). Choose any n-element set X and an element $a \in X$. A k-element subset $Y \subset X$ either contains this a, or does not. If $a \in Y$, then $Y \setminus \{a\}$ is a $(k-1)$-subset of the $(n-1)$-element set $X \setminus \{a\}$, otherwise, Y itself is a k-element subset of $X \setminus \{a\}$. Since the sets of subsets $2^{X\setminus\{a\}}_{k-1}$ and $2^{X\setminus\{a\}}_{k}$ are disjoint—no set can consist of k elements and $k-1$ elements simultaneously, thus by definition of combinations and the sum rule we get (1.4.2). □

In the following chart, called Pascal's triangle, every number, except for the unities at the boundary, is equal to the sum of its two upper neighbors, $2 = 1 + 1, 3 = 1 + 2 = 2 + 1, \dots$

$$\begin{array}{c} 1 \\ 1 \quad 1 \\ 1 \quad 2 \quad 1 \\ 1 \quad 3 \quad 3 \quad 1 \\ 1 \quad 4 \quad 6 \quad 4 \quad 1 \\ 1 \quad 5 \quad 10 \quad 10 \quad 5 \quad 1 \\ \dots \end{array}$$

Since $C(n,0) = 1$, Problem 1.4.2 implies that all the entries in this numerical triangle are consecutive binomial coefficients. Indeed, the upper-most $1 = C(0,0)$, let us call this row the zero row. The next, first row contains $1 = C(1,0)$ and $1 = C(1,1)$, after that we have $1 = C(2,0)$, $2 = C(2,1)$, $1 = C(2,2)$, the third row starts with $1 = C(3,0)$, followed by $3 = C(3,1)$, and so on. The sum of entries in the nth row is 2^n by Corollary 1.1.1. Pascal's triangle often appears in various problems.

The following properties of the binomial coefficients are often helpful. The solutions of the next two problems are left to the reader.

Problem 1.4.3. Prove that

$$C(n,k) = C(n,n-k). \tag{1.4.3}$$

Problem 1.4.4. Use the combinatorial interpretation of binomial coefficients to prove the binomial formula or binomial theorem

$$(a+b)^n = a^n + na^{n-1}b + C(n,2)a^{n-2}b^2 + \cdots + C(n,k)a^{n-k}b^k + \cdots + nab^{n-1} + b^n. \tag{1.4.4}$$

Evidently, the coefficients of a^n and b^n here can be written as $C(n,0) = 1$ and those of $a^{n-1}b$ and ab^{n-1} as $C(n,1) = n$.

Problem 1.4.5. No three diagonals of a convex decagon[6] intersect at one point. In how many segments are the diagonals split by the intersection points?

Solution. First, we find the number of the points of intersection. Any such point comes from two intersecting diagonals connecting four vertices of the decagon. So that, each 4-element subset of the set of vertices generates exactly one intersection point, and we obtain $C(10,4) = 10!/(4!(10-4)!) = 210$ intersection points. Some of these points are incident to the four segments we sought. However, there are segments that are incident to only one intersection point and a vertex of the decagon. Therefore, if we multiply 210 by 4 (because an interior intersection point connects 4 segments) we count the former segments once, but the latter segments twice. To overcome this discrepancy, we notice that each vertex has 7 incident diagonals, therefore, there are $7 \times 10 = 70$ segments incident to all the vertices of decagon. Finally, the total number of segments is $\frac{1}{2}(4 \times 210 + 70) = 455$. The factor $\frac{1}{2}$ occurs here because a segment has two end points and the expression $4 \times 210 + 70$ counts them separately. □

Problem 1.4.6. Where in the solution did we use the condition that three diagonals cannot intersect at a point?

Problem 1.4.7. Prove that in the case of an n-gon there are

$$2C(n,4) + \frac{n(n-3)}{2}$$

such segments.

Problem 1.4.8. In how many ways can one choose three different numbers in the set $\mathbf{N}_{300} = \{1, 2, \ldots, 299, 300\}$, so that 3 divides their sum?

Solution. Since the three numbers chosen must be different (we will do without this assumption in Problem 1.4.12) and their order makes no difference, each triple is a 3-element subset of the given set. But we cannot immediately apply 3-combinations,

6 A polygon with 10 sides and 10 vertices.

for not every ordered triple verifies the problem. We notice that when we divide an integer number by 3, there are exactly three possible remainders, $0, 1$, and 2, and to satisfy the condition, the remainders for each triple either must be the same or must be pairwise different. If the three remainders are different, that is, they are $0, 1, 2$, then the numbers themselves are also different, and we have to select one number out of 100 numbers $\{1, 4, 7, \ldots, 295, 298\}$, another number from the set $\{2, 5, 8, \ldots, 296, 299\}$, and the third number from the set $\{3, 6, 9, \ldots, 297, 300\}$; hence, we have 100^3 such triples. The reader can put this result in the formal framework of the 3-arrangements with repetition.

Next, if each of the three remainders is 0, then there are $C(100, 3)$ such triples. The cases when the remainder is 1 or 2, give in addition $2C(100, 3)$ choices. Altogether, we get by the sum rule $100^3 + 3C(100, 3) = 1\,485\,100$ triples. □

To introduce combinations with repetition, we analyze a sweet model problem.

Problem 1.4.9. A college cafeteria sells four kinds of pastries: biscuits (**B**), doughnuts (**D**), muffins (**M**), and napoleons (**N**). In how many ways can a student buy seven pastries?

Solution. A crucial point in this problem is to clarify in what way two purchases of pastries can be distinct from one another. Certainly, they can contain different quantities of similar items. For instance, one student bought three **B**s and four **M**s while another student bought four **B**s and three **M**s; of course, we consider these two purchases as different. Now, what if each of these two students bought, say, seven **B**s? As physical objects, all these pastries are different, but once again we do not consider physical entities, rather corresponding mathematical symbols. If we think this way, both purchases of seven biscuits have the same notation (**B**, **B**, **B**, **B**, **B**, **B**, **B**). Therefore, in this problem any two symbols **B** are indistinguishable, and we must *identify* them. The same applies to symbols **D**, **M**, and **N**.

Moreover, suppose a student bought seven pastries, put them on a tray and then shuffled them up on the tray. It is natural not to consider this new ordering of the same seven pastries as a new buy, this is exactly the same purchase. Thus, since ordering does not count, we cannot consider strings (**B**, **B**, **B**, **B**, **B**, **B**, **B**), (**B**, **B**, **B**, **B**, **D**, **D**, **D**), etc., as subsets of some set, for no set can contain the same element twice. This is a typical problem about *combinations with repetition*, where one has to count the number of families of the same cardinality, containing elements of different types, provided that two families are considered to be different if and only if there is at least one type of elements, which in these two families is represented by different quantities of the elements. At the same time neither order of the elements, nor what elements of any type are included, matters.

This heuristic description is actually an informal definition of the combinations with repetition, and the reader can skip the following formal definition, which translates the description in the formal set-theory language. Before deriving the formula for the number of combinations with repetition in Theorem 1.4.2, to illustrate the proof, we apply the method to solve Problem 1.4.9. □

Solution of Problem 1.4.9 *(continued).* Since the order of pastries (objects) is immaterial, we fix any order; let it be, say, **B**, **D**, **M**, **N**. Suppose we bought 3 biscuits, 2 doughnuts, 1 muffin and 1 napoleon. If we write **B**, **B**, **B**, **D**, **D**, **M**, **N**, this string represents the buy but does not help us and we want to develop better way to represent the outcomes. If we write just seven zeros, $0, 0, 0, 0, 0, 0, 0$, this is much simpler but does not represent the buy, since we do not know which zeros represent biscuits, etc. But since we know that the left-most zeros represent biscuits, we can insert a separator, say 1, which separates the zeros representing biscuits from the zeros representing doughnuts, etc., therefore the string of 10 zeros and ones, $0, 0, 0, 1, 0, 0, 1, 0, 1, 0$, represents the buy above in unique way. For instance, the string $0, 0, 0, 0, 0, 1, 1, 1, 0, 0$ means that the student bought 5 biscuits and 2 napoleons. We immediately see that there is a one-to-one correspondence between our buys and the set of all strings containing 7 zeros and 3 units in any order. The latter can be easily found to be $C(10, 3) = C(10, 7) = 120$. □

Now we give a formal definition of combinations with repetition.

Definition 1.4.2. Consider a set X, any its n-partition

$$X = X_1 \cup X_2 \cup \cdots \cup X_n$$

and a natural number r. On the set 2^X_r of all r-element subsets of X we introduce an equivalence relation (see Problem 1.4.10) as follows:

Two subsets $A, B \in 2^X_r$ are said to be *equivalent*, $A \sim B$, if

$$\begin{aligned} |A \cap X_1| &= |B \cap X_1|, \\ |A \cap X_2| &= |B \cap X_2|, \\ &\vdots \\ |A \cap X_n| &= |B \cap X_n|, \end{aligned}$$

that is, the sets A and B are equivalent if and only if they contain an equal number of elements of the subset X_1, and an equal number of elements of the subset $X_2, \ldots,$ and an equal number of elements of the subset X_n. The equivalence relation partitions the set 2^X_r into disjoint equivalence classes. These equivalence classes, that is, the elements of the factor set $2^X_r / \sim$, are called r-combinations with repetition or with identified elements from elements of n types, or more precisely, combinations with identified elements of the subsets $X_i, 1 \le i \le n$, taken r at a time.

The number of r-combinations with repetition depends on n and r, but not upon a specific set X, so that we denote this quantity by $C_{\text{rep}}(n, r)$.

Problem 1.4.10. Verify that the binary relation in Definition 1.4.2 is an equivalence relation in the sense of Definition 1.1.12.

Theorem 1.4.2. *If*

$$1 \le r \le \min_{1 \le i \le n} |X_i|, \tag{1.4.5}$$

then

$$C_{\text{rep}}(n, r) = C(n + r - 1, r) = C(n + r - 1, n - 1) = \frac{(n + r - 1)!}{(n - 1)!r!}. \tag{1.4.6}$$

Proof. Consider the equivalence relation in Definition 1.4.2 and choose an element-representative in every equivalence class. These representatives make up an r-combination with repetition of the elements of n types. Associate with this r-combination a sequence of r symbols 0 and $n - 1$ symbols 1 as follows. First, write down as many 0s as there are elements of the first type, that is, the elements of the subset X_1 in this r-combination; if there is no element of the first type, we do not write a 0. After that write a 1, which separates two groups of 0s corresponding to different types of elements. Then write as many 0s as there are elements of the second type (from the subset X_2) in this r-combination and again write a separator 1, and so on; but we do not write a 1 after the very last, nth group of 0s.

In this way, we have constructed a one-to-one correspondence between all r-combinations with repetition from elements of n types and the sequences of r 0s and $n - 1$ 1s. This one-to-one correspondence is useful, because we can easily find the number of the latter sequences. Indeed, this number is equal to the number of ways to choose, without ordering, r places for 0s among the given $n + r - 1$ places and fill out the remaining $(n + r - 1) - r = n - 1$ places with 1s; or, which is the same, to select $n - r$ places for 1s. Now formula (1.4.6) follows immediately. □

Problem 1.4.11. Where in the proof was the condition (1.4.5) used?

Second solution of Problem 1.4.9. We apply (1.4.6) with $n = 4$ and $r = 7$ and as before, we compute $C_{\text{rep}}(4, 7) = C(10, 3) = C(10, 7) = 120$ ways to buy seven pastries. □

Problem 1.4.12. We solve again Problem 1.4.8, allowing now the triples with two or all three equal numbers.

Solution. This provision does not change the number of triples whose elements have different remainders after dividing by 3, there are still 100^3 such triples. Consider now the numbers with the remainder 1, that is, the elements of the set $\{1, 4, \ldots, 298\}$. The cases of numbers with the remainders 2 or 3 are similar. Since the ordering is immaterial, triples $\{1, 4, 7\}$ and $\{1, 7, 4\}$ must be identified; however, now we should count also triples with repeating elements, like $\{1, 4, 4\}$ or $\{4, 4, 4\}$. This is again a typical problem concerning the combinations with repetition. Actually, we have in the problem not 100 different elements, but 100 various types of elements and we have to select three elements of these types, which can be done in $C_{\text{rep}}(100,3) = C(102,3)$ ways. All in all, there are $100^3 + 3C_{\text{rep}}(100,3) = 1\,515\,100$ such triples. □

If the restriction (1.4.5), which guarantees that the entire combination can consist of identical elements, fails, the scheme is not immediately applicable. Nonetheless, problems with $r > \min_{1\le i\le n} |X_i|$ can be solved using Theorem 1.4.2 and the sum rule. Consider the following modification of Problem 1.4.9.

Problem 1.4.13. A college cafeteria sells the same four kinds of pastries: biscuits (**B**), doughnuts (**D**), muffins (**M**), and napoleons (**N**); however, only three muffins remain in stock. In how many ways can a student buy seven pastries?

Solution. Since now there are only three objects of the **M** type and $3 < 7$, the condition (1.4.5) fails and we cannot immediately apply formula (1.4.6). Nevertheless, we can use it if we partition the set of all possible purchases in four disjoint subsets.
(0) No muffin was bought.
(1) One muffin was bought.
(2) Two muffins were bought.
(3) Three muffins were bought.

By making use of Theorem 1.4.2, in case (0) we have $C_{\text{rep}}(3,7)$ purchases, since we have to buy seven items of three types.

In case (1), we have $C_{\text{rep}}(3,6)$ purchases, since now in addition to one muffin bought, we have to buy six more pastries of three other types. In case (2), there are $C_{\text{rep}}(3,5)$ purchases, because we buy two muffins and five pastries of the other three types. Finally, in case (3) there are $C_{\text{rep}}(3,4)$ purchases.

By the sum rule, we have

$$\begin{aligned} &C_{\text{rep}}(3,7) + C_{\text{rep}}(3,6) + C_{\text{rep}}(3,5) + C_{\text{rep}}(3,4) \\ &\quad = C(9,7) + C(8,6) + C(7,5) + C(6,4) = 100 \text{ purchases.} \end{aligned}$$ □

Problem 1.4.14. Show that, if instead of (1.4.5) we have

$$r_1 = |X_1| < r \le \min_{2\le i\le n} |X_i|, \tag{1.4.7}$$

then the number of r-combinations with repetition of elements of n types is

$$C_{\text{rep}}(n,r) = C(n+r-1, n-1) - C(n+r-r_1-2, n-1). \tag{1.4.8}$$

Solution. Arguing as in Problem 1.4.13, we represent the number sought as

$$\begin{aligned} &C_{\text{rep}}(n-1,r) + C_{\text{rep}}(n-1,r-1) + \cdots + C_{\text{rep}}(n-1,r-r_1) \\ &\quad = C(n+r-2,n-2) + \cdots + C(n+r-r_1-2,n-2). \end{aligned}$$

To obtain (1.4.8), we rewrite each addend here by formula (1.4.2) as

$$C(m, l) = C(m + 1, l + 1) - C(m, l + 1)$$

and then combine like terms. □

In the rest of this section we solve various enumerative problems.

Problem 1.4.15. How many whole-number solutions (that is, consisting of nonnegative integer numbers) does the equation

$$x_1 + x_2 + \cdots + x_k = n \tag{1.4.9}$$

have?

Solution. Introducing new unknowns $y_i = x_i + 1, 1 \leq i \leq k$, we will look for positive integer solutions of the equivalent equation $y_1 + y_2 + \cdots + y_k = n + k$. If we represent the number $n + k$ on the right-hand side of the latter as the sum of $n + k$ unities, we immediately realize that solving the problem is equivalent to splitting $n + k$ identical items (in our case, 1s) into k non-empty groups such that the ith group contains $y_i \geq 1$ 1s. To this end, we arrange these $n + k$ 1s in a row and observe that there are $n + k - 1$ spaces (gaps) between these 1s. To split the 1s into k groups, we choose $k - 1$ places among these $n + k - 1$ gaps and insert some separators; for example, we can insert 0s in these gaps. This insertion can be done in $C(n + k - 1, k - 1) = C_{\text{rep}}(k, n)$ ways, which is the number of solutions of equation (1.4.9). □

It is worth repeating that we have established a one-to-one correspondence between the set of solutions of (1.4.9) and a set with the known cardinality, namely, the set of all n-combinations with repetition of the elements of k types. In the following problem we systematically exploit the same approach of the reduction of the set at question to a set with a simpler structure, whose cardinality is known or can be found easier.

Problem 1.4.16. Compute the sum of all natural numbers whose digits go either in increasing order or in decreasing order.

Solution. Let us denote a k-digit natural number a with digits (from left to right) $a_1, a_2, \ldots, a_k$ by overline, $a = \overline{a_1 a_2 \ldots a_k}$. The set of all natural numbers with strictly increasing digits is denoted by INC and the set of numbers with strictly decreasing digits is denoted, respectively, by DEC; the sum of all numbers in a set X is denoted by SUM(X).

Denote by DEC_0 the set of all numbers with decreasing digits, whose last digit is zero and let $\text{DEC}_1 = \text{DEC} \setminus \text{DEC}_0$; we have $\text{DEC}_1 \bigcap \text{DEC}_0 = \emptyset$ and so that $\text{SUM(DEC)} = \text{SUM}(\text{DEC}_0) + \text{SUM}(\text{DEC}_1)$. We immediately observe that there is a one-to-one correspondence between DEC_0 and DEC_1, given by

$$b \in \mathrm{DEC}_1 \Leftrightarrow 10b \in \mathrm{DEC}_0\,.$$

Thus, $\mathrm{SUM}(\mathrm{DEC}_0) = 10\,\mathrm{SUM}(\mathrm{DEC}_1)$ and $\mathrm{SUM}(\mathrm{DEC}) = 11\,\mathrm{SUM}(\mathrm{DEC}_1)$.

Pick a number a', whose digits go in increasing order, say

$$a' = \overline{a_1a_2\dots a_k} \in \mathrm{INC}, \quad a_1 < a_2 < \cdots < a_k,$$

and consider the number

$$b' = \overline{b_1b_2\dots b_k}, \tag{1.4.10}$$

where $b_j = 10 - a_j$. The left-most digit of a' cannot be 0, $a_1 \neq 0$, while all other digits must be bigger than a_1, thus, $1 \le a_j \le 9$ for $1 \le j \le 9$. In turn, this implies

$$1 \le b_j = 10 - a_j \le 9, \quad 1 \le j \le 9,$$

therefore, $b' \in \mathrm{DEC}_1$. For example, if $k = 3$ and $a' = 139$, then $b' = 971$; we observe that $a' + b' = 1\,110 = (10/9)(10^3 - 1)$. We generalize this observation in the following problem.

Problem 1.4.17. Prove that this observation is not a coincidence, that is, if a' is a k-digit number with digits going in increasing order and b' is defined by (1.4.10), then $a' + b' = (10/9)(10^k - 1)$.

Next we notice that the pairing $a' \Leftrightarrow b'$ establishes a one-to-one correspondence between the sets INC and DEC_1. For each $k = 1, 2, \dots, 9$, the set INC contains $C(9,k)$ k-digit numbers, and every number $a \in \mathrm{INC}$ can be derived by removal of certain digits from the string 123456789. Thus,

$$\begin{aligned}\mathrm{SUM}(\mathrm{INC}) + \mathrm{SUM}(\mathrm{DEC}_1) &= \sum_{k=1}^{9} C(9,k)(10/9)(10^k - 1)\\ &= (10/9)((1+10)^9 - (1+1)^9) = (10/9)(11^9 - 2^9).\end{aligned}$$

Denote the latter number by $x = (10/9)(11^9 - 2^9)$ and let ${}_9\,\mathrm{DEC}$ be a subset of the set DEC consisting of numbers, whose left-most digit is a 9. The sets ${}_9\,\mathrm{DEC}$ and $\mathrm{DEC}' = \mathrm{DEC} \setminus {}_9\,\mathrm{DEC}$ make a partition of DEC,

$$\mathrm{DEC} = {}_9\,\mathrm{DEC} \cup \mathrm{DEC}'\,.$$

We also notice that DEC' consists of all numbers with decreasing digits, including 0, whose first digit is not 9 and

$$\mathrm{SUM}(\mathrm{DEC}) = \mathrm{SUM}({}_9\,\mathrm{DEC}) + \mathrm{SUM}(\mathrm{DEC}').$$

Another one-to-one correspondence, now between the sets INC and DEC′, is established by

$$a'' = \overline{a_1 a_2 \dots a_k} \in \text{INC} \Leftrightarrow b'' = \overline{(9-a_1)(9-a_2)\dots(9-a_k)} \in \text{DEC}';$$

we immediately see that $a'' + b'' = 10^k - 1$. Thus, denoting $11^9 - 2^9 = y$, we have

$$\text{SUM(INC)} + \text{SUM}(\text{DEC}') = \sum_{k=1}^{9} C(9,k)(10^k - 1) = 11^9 - 2^9 = y.$$

Between the sets $\text{INC} \bigcup \{0\}$ and ${}_9\,\text{DEC}$ there also exists a one-to-one correspondence by virtue of the pairing

$$a''' = \overline{(9-b_1)(9-b_2)\dots(9-b_k)} \in \text{INC} \Leftrightarrow b''' = \overline{9b_1 b_2 \dots b_k} \in_9 \text{DEC}$$

for $k \geq 1$; if $k = 0$, we set $9 \Leftrightarrow 0$, therefore $a''' + b''' = 10^{k+1} - 1$. Denoting $10 \cdot 11^9 - 2^9 = z$, we derive from this that

$$\text{SUM(INC)} + \text{SUM}({}_9\text{DEC}) = \sum_{k=0}^{9} C(9,k)(10^{k+1} - 1) = 10 \cdot 11^9 - 2^9 = z$$

and

$$\begin{aligned} y + z &= \text{SUM(INC)} + \text{SUM}(\text{DEC}') + \text{SUM(INC)} + \text{SUM}({}_9\text{DEC}) \\ &= 2\,\text{SUM(INC)} + \text{SUM(DEC)}. \end{aligned}$$

Combining these linear equations for SUM(INC) and SUM(DEC), we find

$$\text{SUM(INC)} = (1/9)(11x - y - z) \quad \text{and} \quad \text{SUM(DEC)} = (11/9)(y + z - 2x).$$

However, the sets INC and DEC are not disjoint, their intersection consists of 9 one-digit numbers with the total sum of 45. Thus, the sum we look for is

$$\begin{aligned} &\text{SUM(INC)} + \text{SUM(DEC)} - 45 \\ &\quad = (80/81)11^{10} - (35/81)2^{10} - 45 = 25\,617\,208\,995. \end{aligned}$$ □

Definition 1.4.3. For a real number x, let $[x]$ denote its *integer part*, that is, the largest integer number not exceeding x; it is also called the *floor function* and is denoted by $\lfloor x \rfloor$. For example, $[3.14] = 3$, $[-3.14] = -4$, $[3] = 3$.

Problem 1.4.18. Find the number of n-arrangements with repetition from the set $A = \{0, 1\}$, containing an even number of 0s.

Solution. Since we have defined an arrangement as a mapping, to specify such an arrangement (that is, a mapping) we have to choose preimages for 0s, and the number of preimages must be an even number $2k$, $0 \le 2k \le n$. We suppose that the arrangement $(1, 1, \dots, 1)$ without 0s satisfies the condition; this corresponds to the case $k = 0$. Hence, by the sum rule there are

$$S = C(n, 0) + C(n, 2) + \cdots + C(n, 2[n/2])$$

such arrangements. Setting $a = 1$ and $b = 1$ in the binomial expansion (1.4.4) yields

$$2^n = (1 + 1)^n = C(n, 0) + C(n, 1) + C(n, 2) + \cdots + C(n, n),$$

and setting $a = 1$ and $b = -1$ in (1.4.4) yields

$$0 = (1 - 1)^n = C(n, 0) - C(n, 1) + C(n, 2) - \cdots + (-1)^n C(n, n).$$

Adding these two equations gives $2^n = 2S$, thus $S = 2^{n-1}$. □

Remark 1.4.1. We know from (1.3.1), that without any parity restriction there are 2^n n-arrangements from a 2-element set $A = \{0, 1\}$. Hence, among them there are 2^{n-1} arrangements with an even number of 0s and $2^n - 2^{n-1} = 2^{n-1}$ arrangements with an odd number of 0s.

Another solution of Problem 1.4.18 *is of interest.* Let us denote the number of arrangements we sought for by S_n. All these arrangements fall into two disjoint classes: those beginning with a 1, $(1, a_2, \dots, a_n)$, and those beginning with a 0, $(0, a_2, \dots, a_n)$. In the first case an $(n - 1)$-arrangement $(a_2, \dots, a_n)$ contains an even number of 0s, therefore there are S_{n-1} such arrangements. In the second case $a_1 = 0$, thus the $(n - 1)$-arrangement $(a_2, \dots, a_n)$ contains an odd number of 0s, that is, S_{n-1} less than the total number of $(n - 1)$-arrangements. By the sum rule,

$$S_n = S_{n-1} + (2^{n-1} - S_{n-1}) = 2^{n-1}.$$ □

Problem 1.4.19. Find the number of n-arrangements with repetition from the set $A = \{0, 1, 2\}$, containing an even number of 0s.

Solution. Hereafter we refer to the solution of Problem 1.4.18. If $2k$ preimages of 0 have been chosen, then by virtue of (1.3.1) the images for the remaining $n - 2k$ preimages can be assigned in 2^{n-2k} ways, and these images are either 1 or 2. Using the sum and product rules, as in Problem 1.4.18, we get the formula

$$2^n C(n, 0) + 2^{n-2} C(n, 2) + \cdots + 2^{n-q} C(n, q), \quad \text{where } q = 2[n/2].$$

To compute this sum explicitly, we add the equations (Cf. Problem 1.4.18)

$$3^n = (2 + 1)^n = 2^n C(n, 0) + 2^{n-1} C(n, 1) + \cdots + 2^0 C(n, n)$$

and

$$1 = (2-1)^n = 2^n C(n,0) - 2^{n-1}C(n,1) + \cdots + (-1)^n 2^0 C(n,n),$$

which gives $(1/2)(3^n + 1)$. □

Problem 1.4.20. Find the number of n-arrangements with repetition from the set $A = \{0,1,2,3\}$, containing an even number of 0s and an even number of 1s.

Solution. Problem 1.4.19 readily implies that there are $(1/2)(3^n+1)$ arrangements without 0s. If an arrangement contains two 0s, then their preimages can be chosen in $C(n,2)$ ways. For the remaining $n-2$ preimages, their $n-2$ images, containing an even number of 1s and any numbers of 2s and 3s, can be chosen in $(1/2)(3^{n-2}+1)$ ways—we again use here the result of Problem 1.4.19, with $n-2$ instead of n. Continuing in the same way and using the sum and product rules, we derive

$$\begin{aligned}&\frac{1}{2}(3^n+1)C(n,0) + \frac{1}{2}(3^{n-2}+1)C(n,2) + \cdots + \frac{1}{2}(3^{n-q}+1)C(n,q)\\ &= \frac{1}{2}(C(n,0) + \cdots + C(n,q)) + \frac{1}{2}(3^n C(n,0) + \cdots + 3^{n-q}C(n,q)),\end{aligned}$$

where $q = 2[n/2]$. The first sum on the right-hand side of this equation was found in Problem 1.4.19. To find the second sum, we proceed similarly, using the expansions $(3 \pm 1)^n$. Finally, we get the answer, $4^{n-1} + 2^{n-1}$. □

We solve these problems again in Section 4.3 (Problem 4.3.17) using the method of GF.

Problem 1.4.21. Consider 10^n n-digit nonnegative integer numbers. Two numbers are said to be equivalent, if one can be derived from another by permuting some digits. For example, four-digit numbers 3213 and 3231 are equivalent. If after permuting the left-most digit is 0, we still consider the number as having n digits.

(1) How many classes of equivalence, that is, pairwise nonequivalent numbers are there?
(2) The same question if a number cannot contain more than one digit 0 and more than one digit 9.

Solution. (1) If all digits in any number are different, then every equivalence class contains $n!$ numbers—obviously, in this case $n \le 10$ and there are $C(10,n)$ equivalence classes. But digits may repeat, and we have to use combinations with repetition—two numbers are equivalent, if there is at least one digit occurring a different number of times in these two numbers. Therefore, we have n objects of 10 types, that is, there are $C_{\text{rep}}(10,n)$ equivalence classes.

(2) In this case the factor set splits into four disjoint subsets:
(a) Numbers containing neither 0 nor 9.
(b) Numbers containing one 9 and no 0.
(c) Numbers containing one 0 and no 9.
(d) Numbers containing one 0 and one 9.

In case (a) we have $C_{\text{rep}}(8,n)$ equivalence classes, in cases (b) and (c) there are $C_{\text{rep}}(8,n-1)$ classes, in case (d) there are $C_{\text{rep}}(8,n-2)$ classes. Altogether we have

$$C_{\text{rep}}(8,n) + 2C_{\text{rep}}(8,n-1) + C_{\text{rep}}(8,n-2) = \frac{2(n+5)!(2n^2+12n+21)}{7!n!}$$

nonequivalent numbers. □

Problem 1.4.22. Let $(a_1, a_2, \ldots, a_{n+p})$ denote $(n+p)$-arrangements with repetition from the elements of the set $A = \{-1, 1\}$ containing n numbers -1 and p numbers 1. Denote $f(k) = \sum_{l=1}^{k} a_l$. Find the number of these arrangements such that $f(k) \geq 0$ for each $k = 1, 2, \ldots, n+p$.

Solution. In this problem we use the *trajectory method* ([18, Chap. 3], see also [22, p. 127, No. 2.7.13]), which sometimes gives an easy and very transparent solution. Introduce an orthogonal coordinate system in the plane and consider points

$$Z_0 = (0,0), \quad Z_k = \big(k, f(k)\big), \quad 1 \leq k \leq n+p.$$

A broken line consisting of $n+p$ segments consecutively connecting the points Z_0 and Z_1, Z_1 and Z_2, Z_2 and Z_3,..., Z_{n+p-1} and Z_{n+p}, is called the *trajectory* or *Dyck path* corresponding to the arrangement $(a_1, a_2, \ldots, a_{n+p})$. Among these $n+p$ segments, p are directed upward and have the slope $+1$, and n are directed downward and have the slope -1, hence it is easy to find the coordinates of the point Z_{n+p}, namely, $Z_{n+p} = (n+p, p-n)$. To determine a particular trajectory, it suffices to choose p places for the upward segments among the given $n+p$ places or, which is the same, n places for the downward segments. Therefore, the total number of the trajectories is $C(n+p, p) = C(n+p, n)$. We have to find how many of them do not drop below the X-axis, but it is easier to compute the number of trajectories that do drop below it, that is, which have common points with the horizontal line $y = -1$. Let T be such a trajectory, and k_0 be the left-most common point of T and the line $y = -1$.

Consider another trajectory $\overline{T}$ that coincides with T from 0 to k_0, and is the mirror reflection of T at the line $y = -1$ to the right of k_0. This procedure sets a one-to-one correspondence between the set of all trajectories joining the points 0 and Z_{n+p} and crossing the line $y = -1$, on the one hand, and the set of trajectories joining the points 0 and $\overline{Z}_{n+p} = (n+p, n-p-2)$, on the other hand.

If a trajectory, connecting 0 and $\overline{Z}_{n+p}$, has u upward and d downward segments, then

$$\begin{cases} u + d = n + p, \\ u - d = n - p - 2. \end{cases}$$

Solving this system of linear equations we find $d = p + 1$. Hence the number of trajectories crossing the line $y = -1$ is $C(n + p, p + 1)$, and the number of trajectories in question is

$$C(n + p, p) - C(n + p, p + 1) = \frac{p + 1 - n}{p + 1} C(n + p, p). \tag{1.4.11}$$

This implies in particular that the trajectories we looked for exist only if $p \geq n$, though this is clear from the problem without calculations. □

Remark 1.4.2. If $p = n$, then (1.4.11) becomes $\frac{1}{n+1} C(2n, n)$; these numbers, called the Catalan numbers, occur in many combinatorial and other problems, see, e. g. [2, 51] and Problem 4.4.10; we denote them Cat_n.

Exercises 1.4.

Exercise 1.4.1.
(1) An urn contains 12 different balls. In how many ways is it possible to draw 8 of them without return?
(2) With return?
(3) An urn contains 12 identical balls. In how many ways is it possible to draw 8 of them without ordering and without return?
(4) Without ordering but with return?

Exercise 1.4.2. Calculate the binomial coefficients $C(m, n)$ for all $-2 \leq n \leq m \leq 6$.

Exercise 1.4.3. Prove that $k!$ divides the product of any natural number n and its $k - 1$ successors. For example, for $k = 5$ and $n = 3$, $5! = 120$ divides the product $3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 = 2\,520$.

Exercise 1.4.4. Prove the following properties of the binomial coefficients for any natural n and appropriate values of all other parameters.
(1) $\sum_{k=m}^{n} C(k, m) = C(n + 1, m + 1)$,
(2) $\sum_{k=1}^{n} kC(n, k) = n2^{n-1}$,
(3) $\sum_{k=2}^{n} k(k - 1)C(n, k) = n(n - 1)2^{n-2}, n \geq 2$,
(4) Extend the two preceding equations, so that the left-hand side reads $\sum_{n_0}^{n}$ with $1 \leq n_0 \leq n$,
(5) $\sum_{k=0}^{n} (2k + 1)C(n, k) = (n + 1)2^{n}$,
(6) $\sum_{k=0}^{n} \frac{1}{k+1} C(n, k) = \frac{1}{n+1}(2^{n+1} - 1)$,

(7) $\sum_{k=0}^{n} \frac{(-1)^k}{k+1} C(n,k) = \frac{1}{n+1}$,
(8) $\sum_{k=1}^{n} (n+1-k)k^2 = \frac{1}{3} C(n+1,2)C(n+2,2)$,
(9) (*Vandermonde's identity*)

$$\sum_{k=0}^{n} C(m,k)C(l,n-k) = C(m+l,n), \quad n \le \min\{m; l\},$$

(10) Use Vandermonde's identity above to prove the formula

$$\sum_{k+l=m} \frac{l}{m} \binom{s-i}{k} \binom{i}{l} = \frac{i}{s} \binom{s}{m},$$

(11) $\sum_{k=l}^{m} (-1)^k C(m,k)C(k,l) = (-1)^m \delta_{ml}$, where the Kronecker symbol δ_{ml} (Kronecker's delta) is defined for all non-equal integers m and l by $\delta_{ml} = 0$ and for $m = l$ by $\delta_{ll} = 1$,
(12) for natural m and n, prove the identity

$$(-1)^n C(-n, m-1) = (-1)^m C(-m, n-1),$$

(13) for $0 \le k \le n$, find the maximum value of the binomial coefficients $C(n,k)$, and determine, for each n, how many binomial coefficients $C(n,k), k = 0, 1, \dots, n$, are equal to this maximum value,
(14) prove that[7] $\sum_{k=1}^{m} \frac{(-1)^{k-1}}{k} C(m,k) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}$,
(15) prove the identity for the harmonic numbers H_m,

$$H_{[n/2]} - H_n = \sum_{k=1}^{n} \frac{(-1)^k}{k}.$$

Exercise 1.4.5. The binary, ternary, ..., decimal, ... numerical systems represent any natural number by making use of a fixed number of digits—for instance, the two digits, 0 and 1, in the binary system, and the ten digits, $0, 1, \dots, 8, 9$, in decimal system. Another representation of the integer numbers, called the *combinatorial representation*, uses binomial coefficients to write down any natural number as a sum of a fixed number of addends.

(1) Prove that, given an integer number $k \ge 1$, any natural number n can be written as

$$n = C(d_1, 1) + C(d_2, 2) + \cdots + C(d_k, k),$$

and this representation is unique if we require, in addition, that $0 \le d_1 < d_2 < \cdots < d_k$.

7 The numbers $H_m = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}$, $m = 1, 2, \dots$, are called *harmonic numbers*.

(2) Find the combinatorial representations of $n = 1\,000$ and $n = 1\,000\,000$ with $k = 5, 8, 10$.
(3) Given n and k, estimate d_k in the combinatorial representation of n.

Exercise 1.4.6. Yet another useful representation of integer numbers, called *factorial representation*, uses factorials instead of powers or binomial coefficients.
(1) Prove that any whole number n can be written as

$$n = f_1 \cdot 1! + f_2 \cdot 2! + f_3 \cdot 3! + \cdots,$$

and this representation is unique if we also assume that $0 \le f_i \le i$.
(2) Find the factorial representations of $n = 1\,000$ and $n = 1\,000\,000$.
(3) Given n, estimate the number of addends in the factorial representation of n.

Exercise 1.4.7. Consider families with 5 children, without twins. If we assume that the family composition depends on the order the kids were born, then among these families there is one family with all 5 girls, 5 families with one boy and 4 girls, etc. List all families with 5 kids. The answer is $32 = 2^5$. Explain this answer, using a combinatorial argument.

Exercise 1.4.8. Assuming that boys and girls have equal chances to be born, what part of all families with 6 children have 4 girls and 2 boys?

Exercise 1.4.9. 2^n people depart from the upper point of Pascal's triangle, which was defined after Problem 1.4.2. At each point, including the upper-most one, half of them move to the left and another half to the right. How many people arrive at each point of the nth row?

Exercise 1.4.10. How many functions

$$f : \{1, 2, \ldots, 2\,006\} \to \{2\,005, 2\,006, 2\,007\}$$

are there such that the number $f(1) + f(2) + \cdots + f(2\,006)$ is even? Is odd?

Exercise 1.4.11. The following two identities connect the Catalan numbers Cat_n, defined in Remark 1.4.2, and the binomial coefficients $C(m, n)$.
(1) $\sum_{k=0}^{n} (C(n,k))^2 = (n+1)C_n$,
(2) $C_n = C(2n, n) - C(2n, n-1)$.

Exercise 1.4.12. Find the coefficients of x^{19} and x^{21} after expanding the polynomial $(x^8 + x^5 + 1)^{20}$ by the binomial formula (1.4.4) and combining like terms.

Exercise 1.4.13. Find the number of $2n$-dimensional vectors $(\alpha_1, \alpha_2, \ldots, \alpha_{2n})$ such that $\alpha_i = \pm 1, 1 \le i \le 2n$, $\sum_{i=1}^{k} \alpha_i \ge 0$ for $k = 1, 2, \ldots, 2n-1$, and $\sum_{i=1}^{2n} \alpha_i = 0$.

Exercise 1.4.14. Given n points in the plane, how many different lines, connecting them pairwise, can be drawn, if no three among the points are collinear, that is, lie on the same line.

Exercise 1.4.15. Among k points in the plane, l lie on the same line, while no three points among the others lie on the same line.
(1) How many lines are necessary in order to connect all these points pairwise?
(2) How many triangles are there with vertices at these points?

Exercise 1.4.16. Find the largest number of parts that a plane can be divided by
(1) 7 lines,
(2) l lines,
(3) 3 circumferences,
(4) m circumferences.

Exercise 1.4.17. Thirteen resorts are located by the shore of a convex lake. For every 2, and every $3, \ldots$, and for all 13 resorts there is a route connecting them. Each route is a convex polygon (or a line segment for two ports) with vertices at the resorts and is served by a separate ferry. How many ferries are necessary for all these routs?

Exercise 1.4.18. Among given 15 points in a plane, 6 lie on a line, however, no other 3 points are collinear. How many lines containing at least two given points are there?

Exercise 1.4.19. Given 15 points in a plane, 6 among them lie on a circumference, however, no other 4 points belong to a circumference. How many circumferences containing at least three given points are there?

Exercise 1.4.20. At a meeting of the Combi Club, if two attending students know each other, they have no more mutual acquaintances. At the same time, if two participants do not know each other, then they have exactly two common acquaintances at the meeting. Prove that every participant is familiar with the same number of attendees.

Exercise 1.4.21. There are 10 mutually intersecting lines, such that no three of them have a common point of intersection. How many circumferences tangent to any three lines among the given 10 are there?

Exercise 1.4.22. Three points are said to be collinear if they lie on a line. It is known that for any three non-collinear points in three-dimensional space there exists the unique plane containing these three points. Suppose that among k points in space, l are coplanar, that is, lie in the same plane, while no four points among the others are coplanar. How many different planes do exist, such that each plane contains a triple of given points?

Exercise 1.4.23. Consider three non-collinear points in a plane and draw p lines through the first point, q lines through the second one, and r lines through the third one. No three among these $p+q+r$ lines intersect at a common point and no two are parallel. How many triangles are made up by the intersections of these lines?

Exercise 1.4.24. A family of l parallel lines is crossed by another family of k parallel lines making several parallelograms. How many different parallelograms are there in this figure?

Exercise 1.4.25. A beetle, moving in horizontal or vertical direction, can visit only points with integer coordinates. It starts at the origin and must return back to the origin after traveling $2m$ units. How many different routes does the beetle have?

Exercise 1.4.26. P parents and S students attend a school meeting. In how many ways can they be seated in a row, if at least one parent must sit between any two students? The same question if they sit by a round table.

Exercise 1.4.27. Prove that, for any natural p,

$$\sum_{r\geq 1}\frac{p}{r(r+p)}=\sum_{r\geq 1}\frac{(-1)^{r-1}}{r}C(p,r).$$

Exercise 1.4.28. Find n and m such that

$$C(n,m):C(n,m+1):C(n,m+2)=22:20:15,$$

where $a:b$ stands for the ratio a to b.

Exercise 1.4.29 (Compare with Problem 1.4.18). Consider a polynomial

$$(1+x+x^2)^n=a_0+a_1x+a_2x^2+\cdots+a_{2n}x^{2n}.$$

Prove that

$$a_0+a_3+a_6+\cdots=a_1+a_4+a_7+\cdots=a_2+a_5+a_8+\cdots=3^{n-1}.$$

Exercise 1.4.30. The numbers $T_n\equiv C(n+1,n)=\frac{n(n+1)}{2}$ (Fig. 1.3) are called *triangular numbers*. Prove by mathematical induction that

$$(-1)^{n+1}T_n=\sum_{k=1}^{n}(-1)^{k+1}k^2.$$

Exercise 1.4.31. There are n identical black balls and n identical white balls. In how many ways is it possible to choose n balls containing at least one ball of each color? Extend the problem for 3, 4 and more colors.

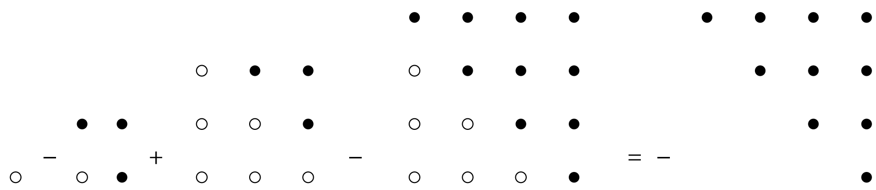


Figure 1.3: Triangular numbers $T_1 - T_4$.

Exercise 1.4.32. In how many ways can $m + n$ balls be chosen among $2m$ identical white and $3n$ identical black balls?

Exercise 1.4.33. No two students at the Even College have the same performance, that is, every two students get different grades at least at one test. Moreover, no student performs better than any other one, that is, for every two students s_1 and s_2, s_1 performs better than s_2 at some test but worse at some other test. This semester, every student takes $2n$ classes. Prove that there are at most $C(2n, n)$ students at the school.

Exercise 1.4.34. Prove that, for a prime p, $C(p, n) \equiv 0 \pmod p$ for $1 \le n \le p - 1$, and $C(p-1, n) \equiv (-1)^n \pmod p$ for $0 \le n \le p - 1$.

Exercise 1.4.35. In how many ways is it possible to choose six different numbers from the set $\mathbf{N}_{49} = \{1, 2, \ldots, 49\}$, so that the difference of two of them is 1? Such a pair of numbers does not have to be unique.

Exercise 1.4.36. How many divisors does the number $2^5 3^4 5^3 7^2 11$ have?

Exercise 1.4.37. For how many integers from 1 to 9 999 is the sum of their digits equal to 9?

Exercise 1.4.38. In how many ways can one choose 4 colors from given seven colors? Assuming that one of the given seven colors is red, what is the answer if red must enter the chosen combination? What is the answer to the latter question if red is not among the given colors?

Exercise 1.4.39. A high school offers classes in English, French, German, Italian, and Spanish. How many bilingual dictionaries must the school library buy for the students?

Exercise 1.4.40.
(1) In how many ways is it possible to split a 20-element set into ten 2-element sets?
(2) In how many ways is it possible to split a 21-element set into ten 2-element sets and one 1-element set?
(3) In how many ways is it possible to split a 21-element set into seven 3-element sets?

Exercise 1.4.41. At a grocery store, there are five identical bottles of apple juice, seven bottles of orange juice, and eight bottles of grape juice. In how many ways can a student buy three bottles of juice for a party?

Exercise 1.4.42. The following statement is called the Kaplansky lemma:

n different books are ordered on a shelf. Prove that, for $k \le n/2$, there are $\frac{n}{n-k}C(n-k,k)$ ways to choose k books, so that no two neighboring books are chosen.

Exercise 1.4.43. A department store has 12 kinds of shoes in Kate's size. In how many ways can she buy 4 pairs of different shoes? What if the shoes bought can repeat?

Exercise 1.4.44. The city of Oldnewburg has the shape of a rectangle, and all its streets are parallel to the sides of the given rectangle. The City Hall is located in the South-West corner of the city. 2^x sheriffs leave the City Hall, half of them due East and another half due North. Officers who reach any street crossing, do the same: half of them goes to East and another half is due North. Eventually m sheriffs arrived at the crossing of the kth and lth streets. Is there any relation between the numbers k, l, m, and x? Compute x in terms of k, l, and m.

Exercise 1.4.45. 16 scouts are searching for their friend who got lost in the woods. Among them there are only 4 boys, who know the area. In how many ways can they make two equal groups for the search, if each group must have two guides knowing the area?

Exercise 1.4.46. Sixteen friends reserved 8 identical double cabins for a cruise. In how many ways can they occupy the cabins?

Exercise 1.4.47.
(1) How many pairwise products can be made from the numbers $1, 2, \ldots, 100$?
(2) How many among them are a multiple of 3?

Exercise 1.4.48. How many 7-digit phone numbers are there with the same last four digits?

Exercise 1.4.49. There are 8 banks of lights in a school hall controlled by 8 different switches. Students decided that at a graduation dance no more than two banks of lights are to be on. In how many ways is it possible to set these 8 switches?

Exercise 1.4.50. Draw k lines through each of the 3 given points in the plane.
(1) At how many points do these $3k$ lines intersect if no two of the lines are parallel and no three intersect at a point (the intersections at the given 3 points do not count)?
(2) Answer the same question if there are four points in the plane.

Exercise 1.4.51. There are l lines and p points on each of these lines, such that no three points on different lines are collinear. How many triangles with vertices at these points are there?

Exercise 1.4.52. A number of n rays in a plane have a common vertex. How many angles do they make?

Exercise 1.4.53. The cafeteria in John's school has a very stable menu, every day they offer the same 13 tasty meals. During a day John can consume any number, from 0 to 13, of these dishes. For how many days can he buy meals at the cafeteria without repetition, that is, no two days have the same selection of dishes? How many dishes will he eat during this time?

Exercise 1.4.54. John's friends Nancy and Kate also decided to have a new menu every day, but Nancy would eat an even number of dishes every day, while Kate would eat an odd number. Who will have to repeat her menu sooner, Nancy or Kate?

Exercise 1.4.55. Generalize Exercises 1.4.53–1.4.54 if the cafeteria has n meals instead of 13.

Exercise 1.4.56. Solve again Problem 1.4.9 under the additional assumption that, among the seven pastries, the student must buy at least four donuts. Compare the answer with the answer to Problem 1.4.13.

Exercise 1.4.57. Let $n = p_1^{k_1} \times \cdots \times p_l^{k_l}$ be the prime factorization of a natural number $n > 1$, that is, $1 < p_1 < \cdots < p_l$ are distinct primes and $k_1, \ldots, k_l$ are arbitrary natural numbers. Find the number and the sum of all natural divisors of n. First solve the problem for $l = 1, 2$ and 3.

Exercise 1.4.58.
(1) What is the smallest natural number with exactly 6 divisors?
(2) With no more than 6 divisors?

Exercise 1.4.59. Given a finite set X, $|X| = k$, find the number of pairs of subsets A, B of X, such that $A \cup B = X$.

Exercise 1.4.60. A standard deck of playing cards consists of 52 cards of 4 suits; spades and clubs are black, diamonds and hearts are red. Each suit contains cards of 13 denominations: 9 numbered cards $2, 3, 4, \ldots, 9, 10$, and 4 face cards: J (a Jack), Q (a Queen), K (a King), and A (an Ace). Find in how many ways it is possible to draw five cards from a standard deck, so that among these five cards there are
(1) 10, J, Q, K, and A of the same suit (a royal flush).
(2) Five adjacent cards of the same suit not starting at 10 (a straight flush).
(3) Five (not necessarily adjacent) cards of the same suit (a flush).
(4) Four cards of the same denomination (four of a kind).

(5) Three cards of the same denomination and two cards of two other different values (three of a kind).
(6) Three cards of the same denomination and two cards of another denomination (full house).
(7) Four cards of four different denominations.

Exercise 1.4.61. In how many ways can k cards, comprising cards of all 4 suits, be dealt from a standard deck of cards if
(1) $k = 4$?
(2) $k = 5$?
(3) $k = 6$?

Exercise 1.4.62. Solve the previous problem if a deck consists of $4n$ cards of four different suits and cards are numbered consecutively from 1 through n.

Exercise 1.4.63. How many are (that is, with the coefficient of 1) monomials are there of degree k in l variables?

Exercise 1.4.64. In how many ways is it possible to distribute 25 identical coins among 4 students?

Exercise 1.4.65. How many solutions in whole numbers does the equation $x_1 + x_2 + x_3 + x_4 = 15$ have?

Exercise 1.4.66. How many solutions in positive integer numbers does the equation $x_1 + x_2 + x_3 + x_4 = 15$ have if $x_2 \geq 2$ and $x_3 \geq 5$?

Exercise 1.4.67. How many solutions in natural numbers does the equation $x_1 + x_2 + x_3 + x_4 = 15$ have if, in addition, $x_2 \geq 2$ and $1 \leq x_3 \leq 5$?

Exercise 1.4.68. How many solutions in whole numbers does the inequality $x_1 + x_2 + x_3 + x_4 \leq 15$ have?

Exercise 1.4.69. How many solutions in integer numbers does the equation $x_1 + x_2 + \cdots + x_k = n$ have subject to restrictions $x_1 > n_1, x_2 > n_2, \ldots, x_k > n_k$?

Exercise 1.4.70. How many solutions in integer numbers does the inequality $|x_1| + |x_2| \leq 100$ have?

Exercise 1.4.71. How many natural numbers not exceeding 10 000 000 are there with the sum of their digits equal to 9?

Exercise 1.4.72. In how many ways can six coins be chosen from an ample supply of pennies, nickels, dimes, and quarters?

Exercise 1.4.73. A raffle ticket at the Combi Club party costs \$5. In a line to the counter, each member has either a \$5 or a \$10 bill. Therefore, the line would stop if the next student has a \$10 bill but the treasurer has no change. To avoid such halts, the treasurer prepared t \$5 bills to give change. If p students have \$10 bills and q have \$5 bills, in how many ways can the students make up a line to buy the tickets without interruptions?

Exercise 1.4.74. How many diagonals does a convex 30-gon have?

Exercise 1.4.75. Find n if a convex n-gon has 35 diagonals.

Exercise 1.4.76. No three diagonals of a convex n-gon have a point in common. In how many regions is the n-gon divided by its diagonals?

Exercise 1.4.77. This problem refers to the binomial formula (1.4.4).
(1) Compute $(x \pm y)^n$ for $n = 1, \ldots, 5$. Determine the largest coefficient(s) in these expansions.
(2) Find the coefficient of $x^8 y^5$ in $(x - 2y)^{10}$.

Exercise 1.4.78. In how many ways can you read the word **MAGIC** in the following diagram?

$$\begin{array}{ccccc} & & M & & \\ & A & & A & \\ G & & G & & G \\ & I & & I & \\ & & C & & \end{array}$$

Exercise 1.4.79. A binary string is a 0-1-sequence, for instance, 01100011 is a binary string of length 8. A binary code, that is, a set of binary strings is designed to represent a set of 35 objects, each object is coded by a string. Every string of the code contains k zeros and l unities, and $k + l = n$. Find k, l, n such that n has the smallest possible value.

Exercise 1.4.80. Let us call two real numbers equivalent if they have the same integer part.
(1) Prove that this is an equivalence relation.
(2) What is the cardinality of the factor set of this equivalence relation considered on the set of all positive real numbers less than 10?

Boolean functions were defined in Exercise 1.1.30 where the reader has computed that there are 2^{2^n} Boolean functions with n variables. However, some of these functions actually depend on less than n arguments in the following sense.

Definition 1.4.4. Given a Boolean function $f(z_1, \ldots, z_n)$, a *variable* z_i, $1 \le i \le n$, is called *essential* if there are values

$$z_1^0, \ldots, z_{i-1}^0, z_{i+1}^0, \ldots, z_n^0 \in Z_2 = \{0, 1\},$$

such that

$$f(z_1^0,\dots,z_{i-1}^0,0,z_{i+1}^0,\dots,z_n^0) \neq f(z_1^0,\dots,z_{i-1}^0,1,z_{i+1}^0,\dots,z_n^0).$$

Otherwise a variable is called *unessential* or *fictitious*. For example, for the Boolean function $f(z_1,z_2) = z_1 \wedge (z_2 \vee \overline{z_2})$, where $\bar{z}$ denote the negation of z, z_1 is an essential variable, while z_2 is a fictitious one.

Denote the number of Boolean functions with precisely n essential variables by $\mathrm{Bess}(n)$.

Exercise 1.4.81. How many n-digit integers whose digits go in non-decreasing order are there?

Exercise 1.4.82.
(1) Verify that $\mathrm{Bess}(0) = \mathrm{Bess}(1) = 2, \mathrm{Bess}(2) = 10$ and compare these numbers with $2^{2^n}, n = 0,1,2$. Find all Boolean functions with no more than 2 essential variables.
(2) Prove that

$$\begin{aligned}\mathrm{Bess}(n) = 2^{2^n} &- C(n,n-1)\,\mathrm{Bess}(n-1)\\ &- C(n,n-1)\,\mathrm{Bess}(n-1) - \cdots - C(n,1)\,\mathrm{Bess}(1) - \mathrm{Bess}(0).\end{aligned}$$

(3) (G. Krylov) Prove that $\mathrm{Bess}(n) = \sum_{k=0}^{n}(-1)^k C(n,k)2^{2^{n-k}}$.

Exercise 1.4.83. How many Boolean functions of three variables satisfy the equation $f(\overline{z_1},\overline{z_2},\overline{z_3}) = f(z_1,z_2,z_3)$?

Exercise 1.4.84. How many integers between 0 and 10^n are there which do not contain the same two digits doing together?

Exercise 1.4.85. In how many ways can you select 6 cards out of a standard deck of 52 cards such that they contain cards of every suit?

Exercise 1.4.86. Consider a k-gon spanned by k vertices of a convex n-gon, $k \leq n$. How many such k-gons do exist, if at least s vertices of the n-gon lie between every two vertices of a k-gon?

Exercise 1.4.87. Ms. Matrix and Mr. Radical ran for the President of the Combi Club. After each ballot vote was cast, Matrix has never been behind Radical. Prove that, if each candidate received exactly n votes, then there are Cat_n ways to count the votes, where Cat_n is the nth Catalan number.

Exercise 1.4.88. A section of 41 students passed the session of the three tests, and nobody failed any exam, i. e., all the grades were “Excellent”, “Good”, and “Satisfactory”. Prove that at least five students passed the session with the same grades.

Exercise 1.4.89. Among the first 999,999 positive integers, how many include the digit 3 in their decimal representation?

1.5 Permutations with identified elements

Objects, considered in this section, resemble the combinations with repetition–they involve indistinguishable elements, which must be identified; however unlike the combinations, ordering of the elements is also important. We again begin with a model problem.

Coffee-time browsing

- www.absoluteastronomy.com/topics/Multinomial_theorem
- http://www.gap-system.org/~history/Mathematicians/Bose.html
- www.gap-system.org/~history/Mathematicians/Dirac.html
- nobelprize.org/nobel_prizes/physics/.../einstein-bio.html
- http://en.wikipedia.org/wiki/Enrico_Fermi#Biography
- http://en.wikipedia.org/wiki/James_Clerk_Maxwell
- http://en.wikipedia.org/wiki/Ludwig_Boltzmann

Problem 1.5.1. In how many ways is it possible to order the letters of the word **DAD**? The same question about the words **ARMADA** and **LETTER**?

Solution. In this and similar problems "words" like **DDA**, which we cannot find in a dictionary, are also acceptable sequences of characters called *strings*. The difficulty of this problem is due to the presence of two identical characters **D**, for transpositions of these symbols do not generate a new string. Moreover, since a set cannot contain two repeating elements, the three characters **D**, **A**, and **D** of a given word do not constitute a set. To overcome this obstacle, we make the two repeating letters *distinguishable* by supplying subscripts and introducing the set $X = \{A, D_1, D_2\}$. Now we can consider all $3! = 6$ permutations of the elements of this new set,

$$\begin{matrix} (A, D_1, D_2) & (D_1, A, D_2) & (D_1, D_2, A) \\ (A, D_2, D_1) & (D_2, A, D_1) & (D_2, D_1, A) \end{matrix} .$$

If we remove here all the subscripts, then two permutations in each of the three columns become indistinguishable and have to be identified. Thus, we break down $P(3) = 6$ permutations of the elements of set X, taken all three at a time, in three disjoint subsets of pairs of permutations. Each subset consists of two permutations, because two elements "D_1" and "D_2" can be transposed in $P(2) = 2! = 2$ ways, hence the set of 6 permutations is split in three pairs. These three pairs of permutations generate three different strings **ADD**, **DAD** and **DDA**. Therefore, in the problem there are $3!/2! = 3$ essentially different permutations.

Similarly, the letters of the word **ARMADA** can be rearranged in $6!/3! = 120$ ways, therefore, there are 120 strings from the letters of the word **ARMADA**. Now, the characters of the word **LETTER** can be reordered in $6!/(2! \cdot 2!) = 180$ ways; here we have to make *two* independent identifications in the set of all permutations of the elements of the set $\{E_1, E_2, L, R, T_1, T_2\}$—we have to identify permutations that can be derived from one another by transposing the symbols E_1 and E_2, and also to identify permutations, that can be derived from one another by transposing the symbols T_1 and T_2. □

It should be noticed that in the solution we implicitly used Lemma 1.1.4. The method of solution presented above is quite transparent and sufficient in the most of applications. However, it is also necessary to have a formal definition of the permutations with identified elements.

Definition 1.5.1. Given a k-partition of a set $X = X_1 \cup X_2 \cup \cdots \cup X_k$, consider the following equivalence relation on the set of all permutations of the elements of X:

Two permutations are called *equivalent* if one of them can be derived from the other by transposing only the elements of the subset X_1, or only the elements of $X_2, \ldots$, or only the elements of X_k. The elements of the factor set derived are called *permutations* of the elements of the set X *with identified elements* of the subsets $X_1, X_2, \ldots, X_k$; we call them *permutations with repetition* if it is clear what partition of the set X generates them.

Let $|X_i| = n_i, 1 \le i \le k$, and $|X| = n = \sum_{i=1}^{k} n_i$. The number of permutations with repetition is denoted by $C(n; n_1, \ldots, n_k)$; these numbers are also called *multinomial coefficients*.

Theorem 1.5.1. *The following equation holds for the multinomial coefficients,*

$$C(n; n_1, \ldots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdots n_k!}. \tag{1.5.1}$$

Proof. The result follows immediately from (1.3.4) after a k-fold application of Lemma 1.1.4. □

Problem 1.5.2. Show that for any natural l the number $(l!)! \cdot (l!)^{-(l-1)!}$ is integer.

Solution. The following is a pure combinatorial proof. If one considers objects of $(l-1)!$ types, l items of each type, that is, $l \cdot (l-1)! = l!$ items in total, then the expression in the problem is exactly the number of permutations with repetition of this set, given by equation (1.5.1) with $k = (l-1)!, n = l!$, and $n_1 = n_2 = \cdots = n_k = l!$. □

Problem 1.5.3.

(1) Small College runs four mathematical courses for the Liberal Arts students—five sections of *The History of Mathematics*, four sections of *Mathematics in the Arts*, three sections of *Introductory Statistics*, and two sections of *Elementary Combinatorics*. Each section of these courses has, respectively, 28, 25, 15, and 12 seats. All 14 sections are taught by 14 different professors, and $5 \cdot 28 + 4 \cdot 25 + 3 \cdot 15 + 2 \cdot 12 = 309$ students satisfy prerequisites and want to take one class each. In how many ways can these students register for the classes?

(2) Solve the same problem if one professor teaches all sections of *The History of Mathematics*, another professor teaches all sections of *Mathematics in the Arts*,

yet another one teaches all sections of *Introductory Statistics*, and another professor teaches both sections of *Elementary Combinatorics*.

Solution. (1) Since all the professors are different, we have 309 objects of 14 types and by (1.5.1) there are

$$\frac{309!}{(28!)^5(25!)^4(15!)^3(12!)^2}$$

ways these students can register for the classes.

(2) However, if one professor teaches all sections of *The History of Mathematics*, it makes no difference for the student what particular section of a class to register for, and the answer is now

$$\frac{309!}{(28!)^5(25!)^4(15!)^3(12!)^2 5!4!3!2!}. \qquad \square$$

These problems and many others can be conveniently stated by using a general model of objects, say balls, placed in urns. Both urns and balls can be either distinguishable or identical. Therefore, there are in general four possible cases. Two of these cases are treated in the following theorem; we omit its proof, which is similar to the reasoning in the solution of Problem 1.5.3. Two other cases will be considered later.

Theorem 1.5.2. *Given k distinguishable groups of urns, p_1 urns of one type, p_2 urns of another type, etc., then there are*

$$\frac{n!}{(n_1!)^{p_1}\cdots(n_k!)^{p_k}}$$

ways to place $n = n_1p_1 + \cdots + n_kp_k$ different objects into these $p_1 + \cdots + p_k$ urns if all the urns are different, and

$$\frac{n!}{(n_1!)^{p_1}\cdots(n_k!)^{p_k}(p_1)!\cdots(p_k)!}$$

ways to place $n = n_1p_1 + \cdots + n_kp_k$ different objects into $p_1 + \cdots + p_k$ urns if urns within each group are indistinguishable. $\square$

Exercises 1.5.

Exercise 1.5.1. In how many ways is it possible to place 6 identical balls into 4 different urns, so that

(1) No urn is empty?
(2) Exactly 2 urns are empty?
(3) At most 3 urns are empty?
(4) At least 3 urns are empty?

Exercise 1.5.2. A bus with 35 passengers makes 7 stops. In how many ways can the passengers leave the bus, so that exactly 5 of them get off at each stop?

Exercise 1.5.3. Prove that $((2n))! \cdot 2^{-n}$ and $((3n))! \cdot 6^{-n}$ are integer numbers.

Exercise 1.5.4. Prove that the fraction $\frac{n!}{n_1!n_2!\cdots n_k!}$ is an integer number whenever $n_1+n_2+\cdots+n_k \leq n$.

Exercise 1.5.5. A student is preparing to a Spelling Bee contest. She looks for an 11-character word containing four letters **s**, four letters **i**, two letters **p**, and one more consonant. How many dictionary entries should she browse at most?

Exercise 1.5.6. How many four-digit integers can be composed from the digits of number 12 553 322?

Exercise 1.5.7. In how many ways can the letters of the word **ARMADA** be rearranged so that the letters **R** and **M** remain together
(1) in the same (**RM**) order?
(2) in any order?

Exercise 1.5.8. In how many ways can the letters of the word **MISSISSIPPI** be rearranged so that the first occurrence of the letter **I** precedes the first letter **S**?

Exercise 1.5.9. How many natural numbers less than one million contain only digits 7 and 8?

Exercise 1.5.10. How many 4-arrangements of 4 red, 1 green, 1 blue, 1 black, and 1 white balls are there?

Exercise 1.5.11. In how many ways can 30 boy scouts be split in 10 equal groups of 3? In 3 equal groups of 10?

Exercise 1.5.12. In how many ways can the letters a, e, i, o, u, z be arranged so that a and z are adjacent?

Exercise 1.5.13. In how many ways can 13 balls be placed into 6 urns, so that urn 1 contains 3 balls, urn 2 also contains 3 balls, urn 3 contains 1 ball, urn 4 contains 2 balls, urn 5 contains 4 balls, and urn 6 is empty?

Exercise 1.5.14. How many 27-digit natural numbers are there containing the digits $1, 2, \ldots, 9$ if each digit appears three times?

Exercise 1.5.15. In how many ways can we partition a k-element set X in l parts if the first part contains k_1 elements, the second part contains k_2 elements, ..., the lth part contains k_l elements, thus $k_1 + \cdots + k_l = k$?

Exercise 1.5.16. How many r-combinations with repetition from k letters A, l letters B, and m other different characters are there, if each combination contains all symbols A and B (and maybe some other symbols)?

Exercise 1.5.17. There are 15 students in the Combi Club who play ice hockey. In how many ways can their coach make up three sets of five field players? Consider two different cases—when the ordering of the selected five players in mini-teams of 5 makes or does not make difference.

Exercise 1.5.18. There are 18 students in the Combi Club. In how many ways can their ice hockey coach assign three goalies and make up three sets of five field players?

Exercise 1.5.19. Prove the multinomial theorem,

$$(t_1 + t_2 + \cdots + t_k)^n = \sum C(n; n_1, n_1, \ldots, n_k) t_1^{n_1} \cdots t_k^{n_k},$$

where $C(n; n_1, n_1, \ldots, n_k)$ are multinomial coefficients (1.5.1); the sum is taken over all sets of whole numbers n_i such that $n_1 + n_2 + \cdots + n_k = n$.

Exercise 1.5.20. Use the multinomial theorem to find the expansion of $(x_1 + x_2 + x_3)^4$.

Exercise 1.5.21.
(1) How many terms does the expansion $(x_1 + x_2 + x_3)^8$ have?
(2) Use the multinomial theorem to find the coefficient of $x_1^2 x_2 x_3^5$ in $(x_1 + x_2 + x_3)^8$.
(3) What is the constant term (not containing x) in the expansion $(x + \frac{1}{x} - 3)^7$ in powers of x?

Exercise 1.5.22. How many four-digit multiples of 4, composed of the digits 1, 2, 3, 4, and 5 are there?

Exercise 1.5.23. How many permutations with repetition of b identical balls and c identical cubes are there?

Exercise 1.5.24 (Cf. Theorem 1.5.2). In how many ways can n balls be distributed in k different urns if
(1) All balls are different and any urn can contain any number of balls (*Maxwell–Boltzmann*  *statistics*)?
(2) The balls are indistinguishable and any urn can contain any number of balls (*Bose–Einstein statistics*)?
(3) The balls are indistinguishable and any urn can contain no more than one ball (*Fermi–Dirac statistics*)?

Exercise 1.5.25. A number Q is equal to the product of q different prime factors. Prove that there are $S(q, l)$ ways to represent Q as the product of l factors.

1.6 Probability theory on finite sets

In this section we consider probabilistic problems with finite sample spaces. If we in addition assume the hypothesis of equally likely outcomes, then these problems can be straightforwardly translated into combinatorial ones and vice versa. Therefore, these probabilistic problems provide an ample field for applications of the methods we have developed in preceding sections. In particular, we consider applications of these results to calculating the outcomes of lotteries and other games of chance.

Coffee-time browsing

- en.wikipedia.org/wiki/Abacus (Abacus)
- interactive-genetics.hayden-mcneil.com/IG_topics_ma.htm (Genetics and Probability Theory)
- www-history.mcs.st-and.ac.uk/Biographies/Bayes.html (Bayes' biography)
- http://en.wikipedia.org/wiki/Bernoulli_family (Bernoulli family)
- http://www.cut-the-knot.org/Probability/ChevalierDeMere.shtml (De Mere's paradox)
- http://www.sexratio.com/facts.htm (Gender's ratio)

Our world is random, often unpredictable, meaning that the results, the outcomes, of many of our actions cannot be predicted in advance. When a girl starts study at elementary school, her parents have certain expectations, but they cannot predict for sure her college GPA.[8] Another simple and popular example of randomness is tossing a coin. The probability theory studies (some of) such random events by mathematical methods. First we introduce some terminology.

Any operation, procedure, experiment with results that cannot be predicted in advance, like tossing a coin, or rolling a die, or drawing a card from a deck, is referred to as a *random experiment*. This is not a definition, here we just introduce a *primary notion* like the concepts of a set and a function introduced in Section 1.1. When an experiment has the only possible result, the outcome is certainly known in advance and this experiment is not random.

Definition 1.6.1. All the possible results of a random experiment are called its *outcomes*. The totality of all possible outcomes is called the *sample space* S of the experiment. Points of the sample space, that is, outcomes of a random experiment, are also called *elementary events*. Any set E of outcomes, that is, a subset of the sample space $E \subset S$, is called an *event*. The *empty event* $E = \emptyset$ is also called impossible or improbable, the *universal event* $E = S$ is called certain.

8 Grade Point Average.

The outcomes, belonging to a given event, are sometimes called *favorable* outcomes to this event. The sample space depends upon the problem. For instance, when we roll a coin, then in addition to two typical outcomes, heads and tails, a coin might rest on the edge, even though this phenomenon is not easy to observe, or it can roll away and disappear, but the latter two possibilities are practically improbable, negligible. Thus, discussing experiments with flipping a coin, we always consider the sample space consisting of only two points, a head H and a tail T, in symbols $S = \{H, T\}$. If we roll a die (a six-faced cube) with faces marked by the digits $1, 2, 3, 4, 5, 6$, or by dots, then the sample space of this random experiment is $S = \{1, 2, \ldots, 6\}$. However, if a die is marked by $\{1, 2, 3, 4, 5, 5\}$, then the sample space of the experiment is $S = \{1, 2, 3, 4, 5\}$.

As another example, we consider a lottery with prize levels \$1, \$5, \$100, and \$10 000. The drawing is a random experiment and if we have only one ticket, the sample space consists of five points, $S = \{\$0, \$1, \$5, \$100, \$10\,000\}$. However, in some cases we may only be interested in the very fact of winning (W) or losing (L) the game and can choose another sample space $S_1 = \{W, L\}$. Depending on the issue we are interested in, there are also other possible choices for the sample space in this problem.

To correctly solve a problem in the probability theory, we must explicitly specify the sample space of the problem, otherwise different people can read the same words in different ways and arrive at different conclusions. Hereafter we consider only random experiments with *finite* sample spaces.

Problem 1.6.1. Define the sample space in the last example if
(1) You have two tickets.
(2) You have one ticket that costs \$1 and are interested in the net income.

In many problems it is necessary to consider composite events, consisting of simple ones, and combine simple sample spaces in more complex spaces. For example, if we toss two distinguishable coins, the sample space consists of ordered pairs of the symbols H and T; by the product rule, the new sample space contains 4 points, namely, $S = \{(H, H), (H, T), (T, H), (T, T)\}$. If we roll simultaneously 3 different dice, then the sample space consists of $6^3 = 216$ ordered triples, $S = \{(1, 1, 1), (1, 1, 2), \ldots, (6, 6, 6)\}$.

Up to this point we have discussed only sample spaces. The probability theory originates when a certain specific number $p(s)$, called the probability of the outcome s,[9] is assigned to each point s of the sample space S. The set of these values is called a *probability distribution* on the sample space S, because we distribute a certain given "supply" of probability among the points of S. These values cannot be

9 One can often hear in everyday talk, "It's probable" or "That's unlikely." Based on such an individual judgment, some people play lotteries while the others do not, because the latter do not believe that there are reasonable chances to win. Any discussion of such subjective probabilities is beyond the scope of this book.

assigned arbitrarily, they must satisfy certain assumptions, axioms of the probability theory; for more on that see, for example, [18]. We consider the following system of axioms.
(PA1) $p(s) \geq 0$ for any point $s \in S$,
(PA2) if $E = \{s_1, s_2, \ldots, s_k\} \subset S$, then $p(E) = p(s_1) + p(s_2) + \cdots + p(s_k)$,
(PA3) $p(S) = 1$.

Therefore, we have assumed that probability values are nonnegative, the probability of any event E is the sum of the probabilities of elementary events composing E, and the total probability is 1. These axioms immediately imply that, if $E_1, \ldots, E_k$ are any pairwise disjoint events, that is, $E_1, \ldots, E_k \subset S$ and $E_i \cap E_j = \emptyset, 1 \leq i, j \leq k$, then $p(E_1 \cup \cdots \cup E_k) = p(E_1) + \cdots + p(E_k)$, that is, the probability is *finitely additive*. Moreover, for any event E we have $p(E) = p(E \cup \emptyset) = p(E) + p(\emptyset)$, thus, $p(\emptyset) = 0$, the empty event must have zero probability.

In some cases we can conduct a random experiment in reality, for instance, we can toss a coin many times and record the numbers of heads, $n(H)$, and tails, $n(T)$, occurred. If the experiment was repeated n times and the favorable outcomes to an event E were observed $k(E)$ times among the n outcomes, then the frequency ratio $f(E) = \frac{k(E)}{n}$ is called the *experimental* or *frequency probability* of the event E. Clearly, the frequency $f(E)$ depends, among other things, on the length n of the experiment. If with n increasing, $f(E)$ is stabilizing to a number $p(E)$, we can use $f(E)$ as an estimation of the probability $p(E)$ of the event E, but this is only a plausible approximation. For example, there is nothing unusual to get two heads in a row, thus in this series $n = 2$, $p(H) = 1/1 = 1$, and $p(T) = 0$. However, if we use this very short series to estimate the probability of getting a tail, we have $p(T) = f(T)/2 = 0$, which obviously makes no sense. More advanced courses in the probability theory treat in more detail this issue–what is the appropriate length of an experiment.

Any collection of numbers, satisfying axioms (PA1)–(PA3), can be used as a probability distribution. For example, experimenting with a coin and choosing the sample space $S = \{H, T\}$, we can assign $p(H) = 1/3$ and $p(T) = 2/3$. However, unless we have a specifically tailored (very biased) coin, the results of our physical experiments will likely be essentially different from the results predicted by the mathematical model. Thus, to assign a probability distribution, we use either some previous experience (the results of real experiments) or a theory, if it exists.

Probably, it is physically impossible to make a perfect coin, however, real experiments have confirmed that, if a coin was chosen at random, then as the first approximation it is quite realistic to assign the probabilities $p(H) = p(T) = 1/2$. On the other hand, the same experiments show that no real coin satisfies this probability distribution precisely, but exhibits slight deviations from the theoretical probability 1/2. Nevertheless, it is customary in theoretical studies to accept the hypothesis of *equally likely* *probabilities* or equal chances,[10] that is, to assign equal probabilities to each point of the sample space.

Definition 1.6.2. It is said that the assumption of *equally likely probabilities* is valid for a given problem with the sample space $S = \{s_1, s_2, \dots, s_n\}$ if the probability distribution on S is given by

$$p(s_1) = p(s_2) = \cdots = p(s_n) = \frac{1}{n}.$$

Whether or not this assumption holds true in any particular case, should be verified by comparing our calculations with experiments. The following well-known example is illuminative. Our intuition might tell us that the number of girls born must on average be the same as the number of boys, and many computations using the equal probabilities 1/2 as the first approximation, give good results. However, many-year observations have shown that in reality the probability for a new-born baby to be a boy is slightly bigger, namely 0.51, versus 0.49[11] for a girl.

From now on we always suppose the hypothesis of equally likely probabilities to be valid, unless the opposite is explicitly stated.

The goal of this section is to show applications of the developed combinatorial methods and results to the probability theory. First we translate a few basic set-theory notions to probabilistic language. Recall that an event is just a subset of the basic (universal) set, the latter is called here the *sample space*. All the events under consideration are subsets of some fixed sample space S. Therefore, we can define the following operations with events through their set-theory counterparts.

Definition 1.6.3.
(1) The event $\overline{E} = S \setminus E$ is called *complementary* to an event E.
(2) Two events are called *disjoint* or mutually exclusive if their set-theory intersection is empty, that is, if they have no common favorable outcomes. Thus, if E_1 and E_2 are disjoint events, then $p(E_1 \cap E_2) = 0$, and by (PA2) $p(E_1 \cup E_2) = p(E_1) + p(E_2)$.
(3) A system of events $\{E_1, \dots, E_k\}$ is called *exhaustive* if $\bigcup_{i=1}^{k} E_i = S$.

Example 1.6.1. Let us toss a coin and choose the sample space $\{H, T\}$. Then the events "To get an H" and "To get a T" are disjoint, mutually complementary, and together exhaust the sample space. The events "To get an odd number" and "To get a number less than 3" in one rolling of a die are not mutually exclusive. The complementary event to "To get a number less than 3" is "To get a number greater than or equal to 3". Any event and its complement make up an exhaustive system.

10 The terms *probability* or *chance* should not be confused with the term *odds*. The expression "odds in favor of an event E" means the ratio $\frac{p(E)}{p(\overline{E})}$, while "odds against an event E" means the reciprocal ratio $\frac{p(\overline{E})}{p(E)}$.

11 There are data indicating that this gap maybe is shrinking.

The following properties are immediate consequences of the definitions, axioms (PA1)–(PA3), and the results of Section 1.1. It is critical that any probability distribution is finitely additive.

Theorem 1.6.1.
(1) *For any event E,*

$$p(\overline{E}) = 1 - p(E). \tag{1.6.1}$$

(2) *For any events E_1 and E_2,*

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2). \tag{1.6.2}$$

In particular, if E_1 and E_2 are disjoint, that is, $E_1 \cap E_2 = \emptyset$, then $p(E_1 \cap E_2) = 0$, and $p(E_1 \cup E_2) = p(E_1) + p(E_2)$.

Problem 1.6.2. Two fair right tetrahedrons, a green one and a blue one, with faces marked 1 through 4, were tossed. We record the numbers on the faces they landed.
(1) What is the probability that the sum of these numbers is 7?
(2) What is the probability that the sum of these numbers is greater than or equal to 7?
(3) What is the probability that the sum of these numbers is greater than 7?

Solution. In this and similar problems "fair" means that we accept the hypothesis of equally likely outcomes. Since the tetrahedrons are different, the sample space consists of $4 \times 4 = 16$ ordered pairs, $S = \{(1.1), (1,2), \dots, (4,4)\}$, where the pairs $(1,2)$ and $(2,1)$ are different. Hence, the probability of any outcome is 1/16. The sum of 7 can occur as either $3+4$ or $4+3$, and these outcomes are disjoint, since at the same throwing of a tetrahedron we cannot observe both a 3 and a 4. Hence, the answer to part (1) is $\frac{1}{16} + \frac{1}{16} = 2 \times \frac{1}{16} = \frac{1}{8}$.

(2) Since the largest possible outcome in this part is a 4, the sum of 7 or more means either 7 or 8, therefore, comparing with Part (1) of the problem, there is one more favorable outcome, the pair $(4,4)$, which is also disjoint with the preceding ones, and the answer to part (2) is $3 \times \frac{1}{16} = \frac{3}{16}$.

In part (3), the only favorable outcome is the pair $(4,4)$, thus the answer is $p((4,4)) = 1/16$. The events in parts (1) and (3) are disjoint and their union is the event in part (2), that is why the answer in part (2) is the sum of those in parts (1) and (3). □

Problem 1.6.3. This weekend Kathy either goes to the movies, with the probability of this event 0.7, or to the restaurant with the probability 0.5.
(1) Given this information, is it possible to conclude that these two events are mutually exclusive?
(2) What is the smallest and the largest possible probability that this weekend Kathy will have both these pleasures?

(3) How can we change the problem to be able to determine precisely the probability that this weekend Kathy gets at least one of these pleasures? Both these pleasures?

Problem 1.6.4. Two dice are rolled simultaneously. What is the probability to get at least one number greater than 4?

Consider again Problem 1.4.18, where we found the number of n-arrangements with repetition from the set $A = \{0, 1\}$, containing an even number of 0s, but now we state the question in probabilistic terms.

Problem 1.6.5. Given all 2^n n-arrangements with repetition from the set $A = \{0, 1\}$, we choose at random one of them, assuming that every arrangement has equal chances to occur. What is the probability to pick an arrangement containing an even number of 0s?

Solution. The sample space consists of 2^n arrangements. According to the solution of Problem 1.4.18, 2^{n-1} of them (exactly half of the sample space), are favorable outcomes for our problem. Therefore, the probability we sought, is $p = 2^{n-1}/2^n = 1/2$. □

Analyzing the solution, we observe an important feature of all similar problems:

To solve a probabilistic problem with the finite sample space, we have to solve two enumerative combinatorial problems.

Problem 1.6.6. All permutations of the letters of word MISSISSIPPI are written on balls, and one of these balls is chosen at random. What is the probability that we pick up the ball with the word MISSISSIPPI?

Solution. The sample space consists of all permutations with repetition of the letters of word MISSISSIPPI and by Theorem 1.5.1 contains $C(11; 1, 4, 4, 2)$ elements. Among them there is only one favorable outcome, thus, the probability is $1/C(11; 1, 4, 4, 2) = \frac{4!4!2!}{11!} \approx 0.000029$. □

Problem 1.6.7. Among all permutations with repetition of the letters of word DAD, one is chosen at random. What is the probability to find the chosen combination of letters in an English dictionary?

Solution. The sample space consists of $3!/2! = 3$ permutations with repetition, ADD, DAD, DDA, but only the first two strings are meaningful English words, that is, are favorable outcomes in our problem. Therefore, the probability we sought, is $p = 2/3$. □

Any probability distribution on a sample space S puts a number $p(s)$, $0 \leq p(s) \leq 1$, into a correspondence to a point s of the sample space, therefore, this distribution constitutes a function $f : S \to \mathbf{R}$. Since the domain of this function consists of the outcomes of a random experiment, the values of the function are also random. Such functions are called *random variables*. An initial probability distribution is also a random variable. In many problems it may be advantageous to change the values of a given probability distribution, as long as we preserve axioms (PA1)–(PA3).

Definition 1.6.4. Given a random experiment with a sample space S, any real-valued function

$$f : S \to \mathbf{R}$$

with the domain S is called a *random variable* or a *random function* whenever it satisfies the three properties similar to the probabilistic axioms (PA1)–(PA3):
(RA1) $f(s) \geq 0$ for any point $s \in S$.
(RA2) If $E = \{s_1, s_2, \dots, s_k\} \subset S$, then $f(E) = f(s_1) + f(s_2) + \cdots + f(s_k)$.
(RA3) $f(S) = 1$.

In particular, any probability distribution is a random function.

Problem 1.6.8. Consider a sample space $S = \{1, 2, \dots, n\}$, where n is a given natural number. Let f be a linear function, $f(s) = cs, c$ being a real constant. Find the coefficient c so that the function f is a random variable on S.

Solution. We must verify the properties (RA1)–(RA3); (RA1) is clear if $c \geq 0$, (RA2) is a rule of computing $f(E)$ through the values $f(s), \forall s \in S$, and we only have to compute the normalization constant c by making use of (RA3). We have

$$1 = f(S) = f(1) + \cdots + f(n) = c \cdot 1 + c \cdot 2 + \cdots + c \cdot n = c\frac{n(n+1)}{2}$$

by Problem 1.1.6, thus, for f to be a random variable, we must have $c = \frac{2}{n(n+1)}$. □

The equation $p(E_1 \cup E_2) = p(E_1) + p(E_2)$ tells us that two events are mutually exclusive. Another important mutual characteristic of a pair of events E_1, E_2 is their (stochastic) dependence or independence. It turns out that this property is connected with the equation $p(E_1 \cap E_2) = p(E_1) \cdot p(E_2)$, which is not always valid. Intuitively,

> Two events are independent,
> if the occurrence or non-occurrence of either of them (1.6.3)
> does not affect the probability of the occurrence of the other event.

To define the dependence/independence in more precise analytic terms, it is convenient to connect it with another important concept, namely, with the conditional probability of an event. First, we again model this notion by using an example.

Example 1.6.2. Based on many-year statistic, the probability for a freshman to graduate in four years from The Liberal College is 0.85, while for the freshman majoring in sciences this probability is only 0.70. The sample space in this problem consists of all students ever graduated from the college. In the problem we have two probabilities—for all students the probability is 0.85 while for the science majors it is 0.70. The second number is different from the first one, because in computing it we have used some additional information on the students' majors, actually we reduced the sample space by removing all non-science majors. Since the second probability was computed under an extra condition, it is called the conditional probability.

To arrive at a definition, we sketch a computation of the conditional probability of an event E, given another event (a condition) C, in terms of favorable outcomes. Computing the probability $p(E)$, we have to take into account all outcomes favorable to E and relate them to the whole sample space S. However, when computing the conditional probability we certainly know that the event C has occurred, thus, now we consider only those favorable outcomes of E, which are favorable to C as well. Moreover, we must relate them not to the entire original sample space S, but only to the subset of outcomes favorable to C, hence, we must reduce the original sample space. If we express all these quantities in terms of the size $|S|$ of the sample space and of the probabilities $p(C)$ and $p(E \cap C)$, we derive formula (1.6.4). It is convenient to reverse this reasoning and use (1.6.4) as a definition of the conditional probability.

Definition 1.6.5. Consider a random experiment with the sample space S, a generic event E, and a specified event (condition) C, such that $p(C) > 0$. The *conditional probability* $p(E|C)$ of an event E given the event C, is defined by

$$p(E|C) = \frac{p(E \cap C)}{p(C)}. \tag{1.6.4}$$

It is often convenient to rewrite this formula as

$$p(E|C)p(C) = p(E \cap C).$$

Problem 1.6.9. What is the probability to get a 3 in one roll of a die given that the outcome is odd?

Solution. Introduce the event $E_3 = \{x = 3\}$ and the condition $C = \{x \text{is odd }\}$; we know that $p(E_3) = 1/6$ and $p(C) = 1/2$. The intersection of these events is $E_3 \cap C = E_3$, thus, $p(E_3 \cap C) = 1/6$. By (1.6.4), the conditional probability is $p(E_3|C) = (1/6)/(3/6) = 1/3$. □

Problem 1.6.10. What is the probability to get a 2 in one roll of a die given that the outcome is odd?

Solution. It is clear without computations that, if the outcome is odd, it cannot be 2, but let us formally compute the result. Let $E_2 = \{x = 2\}$ and $C = \{x \text{ is odd }\}, p(C) = 1/2$. The intersection of the two events is empty, $E_2 \cap C = \{2\} \cap \{1, 3, 5, \} = \emptyset$, thus, $p(E_2 \cap C) = 0$ and the conditional probability is $p(E_2|C) = 0/(1/2) = 0$. □

Now we can define the independence of two events in terms of conditional probability.

Definition 1.6.6. Two events E and C are called (stochastically) *independent* if

$$p(E|C) = p(E), \tag{1.6.5}$$

otherwise the events are called *dependent*.

Comparing (1.6.4) with (1.6.5), we see that two events are independent if

$$p(E \cap C) = p(E)p(C), \tag{1.6.6}$$

thus, equation (1.6.6) formalizes our "intuitive" definition (1.6.3). It is worth noting that the independence is a symmetric property, which is obvious from (1.6.6), but not from (1.6.3).

Problem 1.6.11. A card is drawn at random from a regular deck containing 52 cards. Are the events A—"To pick an Ace" and C—"To pick a club" dependent or independent?

Solution. The deck contains 4 Aces, so that $p(A) = 4/52 = 1/13$. Calculate the conditional probability $p(A|C)$. Obviously, $p(A \cap C) = 1/52$ and $p(C) = 13/52$, therefore $p(A|C) = \frac{p(A\cap C)}{p(C)} = 1/13$. Since $p(A|C) = p(A)$, we conclude that these events are independent. □

Problem 1.6.12. To win the jackpot in the New York Lottery Mega Millions game, one must guess correctly 5 numbers among $1, \dots, 56$ and one more number from $1, \dots, 46$. What is the probability to win the jackpot if you have one ticket?

Solution. Since the order is not important, there are $C(56, 5) = 3\,819\,816$ ways to choose five numbers and $C(46, 1) = 46$ ways to select the Mega Ball number. Since the last choice is independent from the first five numbers (Why?), there are $C(56, 5) \cdot C(46, 1) = 175\,711\,536$ different tickets, and this is the cardinality of the sample space. Therefore, the probability we look for, is $1/175\,711\,536 \approx 5.69 \times 10^{-9}$. □

Thus, if we have enough funds and time to buy 175 711 536 $1-tickets, we definitely get the jackpot. Considering the appropriate taxes, not to mention a slight possibility that someone else has a winning ticket, we can estimate how large the jackpot is to be to pay off such an expense.

If we occasionally buy a lottery ticket, we cannot predict the future–the chances are slim, but who knows... Sometimes people win the jackpot. However, if we play any game of chance systematically, we may want to estimate our chances in the long run. The mathematical instrument for such estimations is called the mathematical expectation or the expected value of a random variable.

Definition 1.6.7. Consider a probability distribution $p(s)$ on the sample space S and a random function $f(s), s \in S$. The *mathematical expectation* or the *expected value* of the random variable f is the sum

$$E(f) = \sum_{s \in S} p(s)f(s). \tag{1.6.7}$$

Problem 1.6.13. Find the expected value of the net gain in the preceding problem if the jackpot was \$10 000 000.

Solution. The sample space has only two points, $s_1 = W$ with the probability $p(s_1) = 1/175\,711\,536$ and $s_2 = L$ with $p(s_2) = 1 - 1/175\,711\,536$. Corresponding gains are $f(s_1) = \$10\,000\,000 - \$1$ and $f(s_2) = -\$1$. Therefore,

$$E(f) = \$9\,999\,999 \cdot 1/175\,711\,536 - \$1 \cdot (1 - 1/175\,711\,536) \approx -\$0.94,$$

and in the long run we should expect to lose about 94 cents from each dollar spent. □

Problem 1.6.14. Find the expected value of the net gain in the preceding problem if the jackpot was \$100 000 000.

Problem 1.6.13 gives an example of a *binomial distribution*, that is, a random experiment with exactly two outcomes, usually called *a success* and *a failure*, whose probabilities do not change in time. We follow the tradition and denote the probability of success by $p = p(\text{success})$ and the probability of failure by $q = p(\text{failure})$, thus, $0 \le p, q \le 1$ and $p + q = 1$. If we repeat a binomial experiment n times, assuming all the outcomes being independent (such series is called Bernoulli's trials), then a typical problem is to compute the probability of getting r successes in these n trials. Since there are $C(n, r)$ ways to select r "successful" trials among n, the probability of getting r successes is, by the product rule,

$$p(r, n) = C(n, r)p^r q^{n-r}. \tag{1.6.8}$$

We used (1.6.8) in the solution of Problem 1.6.13 with $n = 1, r = 1, p = 1/175\,711\,536$ and $q = 1 - 1/175\,711\,536$.

Consider again the definition of the conditional probability (1.6.4). Since the operation of intersection of two sets is commutative, it implies the following property

$$p(E|C)p(C) = p(C|E)p(E). \tag{1.6.9}$$

Thus, we can express the conditional probability of two events through their conditional probability in reversed order, that is, as

$$p(C|E) = \frac{p(E|C)p(C)}{p(E)}$$

if $p(E) > 0$. This property can easily be extended to the case of several conditions.

Theorem 1.6.2. *Let events $C_1, C_2, \dots, C_k$ make a partition of the sample space S, that is, all C_j are non-empty, pairwise disjoint, and exhaust the sample space S. Then the following equation, called Bayes's formula, is valid for any event E with $p(E) > 0$ and any $j, 1 \le j \le k$,*

$$p(C_j|E) = \frac{p(E|C_j)p(C_j)}{\sum_{j=1}^{k} p(E|C_j)p(C_j)}. \tag{1.6.10}$$

Proof. By Problem 1.1.12, we have

$$E = E \cap S = E \cap \left(\cup_{j=1}^{k} C_j\right) = \cup_{j=1}^{k} (E \cap C_j).$$

The intersections $E \cap C_j, 1 \le j \le k$, are also mutually exclusive, thus $p(E) = \sum_{j=1}^{k} p(E \cap C_j)$. Combining the latter with the formula for the conditional probability $p(C_j|E) = p(E \cap C_j)/p(E)$, we deduce (1.6.10). □

Problem 1.6.15. Let us note that all $p(C_j) \ne 0$, since we have assumed $C_j \ne \emptyset$. Is it possible that the denominator in (1.6.10) is zero?

Problem 1.6.16. A die was randomly selected among a set, containing 999 999 regular dice and one die with all faces marked by 1, and was rolled 10 times. What is the probability that the fake die was chosen, given that a 1 was observed in all 10 trials?

Solution. Consider three events,

$$\begin{aligned} C &= \{\text{Observe a 1 in 10 consecutive trials}\}, \\ D &= \{\text{Choose a fair die}\}, \quad \text{thus } p(D) = 1 - 10^{-6}, \\ F &= \{\text{Choose a fake die}\}, \quad \text{thus } p(F) = 10^{-6}. \end{aligned}$$

By Bayes's formula,

$$\begin{aligned} p(F|C) &= \frac{p(C|F)p(F)}{p(C|D)p(D) + p(C|F)p(F)} \\ &= \frac{1 \times 10^{-6}}{(1/6)^{10} \times (1 - 10^{-6}) + 1 \times 10^{-6}} \approx 0.94. \end{aligned}$$

The result is so close to 1 that it may look counterintuitive, and it is useful to compare it with the negligible probability to observe 10 consecutive 1s in rolling a fair die, which is $(1/6)^{10} \approx 1.65 \cdot 10^{-8}$. Compare it also with the result of Exercise 1.6.35. □

Now we consider a classical *birthday problem*.

Problem 1.6.17. What is the probability that among s members of the Combi Club at least two have the same birthday?

Solution. To simplify computations, we consider a non-leap year with 365 days and assume that for any day of the year the probability that someone was born this day, is the same; thus, this probability is 1/365. Moreover, as we always do in problems involving people, we suppose that the birthdays of all the people involved are independent, in particular, any day of a year has equal probability to be someone's birthday.

It is easier in this problem to compute the probability of the complementary event, that is, the probability that no two members have the same birthday. First of all we notice that for any member there are 365 options to fix the birthday, hence the sample space contains 365^s points. To find the number of favorable outcomes, we choose s days from 365 for s birthdays–this can be done in $C(365, s)$ ways. However, when we distribute members' birthdays among these s days, we can permute them in $s!$ ways, generating different favorable outcomes. Hence, there are $s!C(365, s) = P(365, s)$ favorable outcomes, so that the probability of the complementary event is $P(365, s)/365^s$, and the probability that at least two members have the same birthday is $1 - P(365, s)/365^s$. We see that, if $s \geq 366$, then this probability is 1, which is obvious. An easy numerical experiment shows that this probability is increasing and becomes bigger than 1/2 for $s = 23$. □

Remark 1.6.1. It is instructive to rephrase this problem in terms of placing balls in urns; see Exercise 1.5.24.

Exercises 1.6.

Exercise 1.6.1. Describe the sample space if we simultaneously toss two indistinguishable coins, that is, the outcomes (H, T) and (T, H) must be identified. What is the sample space if we roll simultaneously 3 identical dice?

Exercise 1.6.2. Describe the sample space if we simultaneously flip a coin and roll a die.

Exercise 1.6.3. Describe the sample space if we flip a coin six times. What is the probability that at least one head and at least two tails will appear in the six tosses? What is the probability that a streak of at least four consecutive tails will appear in the six tosses?

Exercise 1.6.4. Simultaneously toss a coin and roll a die. What is the probability to get a head and a multiple of 3?

Exercise 1.6.5. Simultaneously roll a die and draw a card from a standard desk of 52 cards. What is the probability to get an even number and a red face card?

Exercise 1.6.6. A lottery ticket contains six boxes–two for letters followed by four for digits. If there is only one winning ticket, what is the probability to win the lottery?

Exercise 1.6.7. A 9-digit natural number is chosen at random. What is the probability that all its digits are different?

Exercise 1.6.8. A woman can give birth to a girl, a boy, two girls, a girl and a boy, two boys, etc. Consider this as a random experiment with outcomes to be the number of children and the gender composition of the children born. Describe the sample space if the order at birth is important, that is, we consider the pairs "boy-girl" and "girl-boy" as different. What is the sample space if the order does not count?

Exercise 1.6.9. Assuming that a new-born baby has equal chances to be a girl or a boy, what are the probabilities to have no boy, one boy, two boys, three boys, ..., n boys in a family with n children? Compare with problems **1.4.9**.

Exercise 1.6.10. A die is rolled once. Find the complementary event to the following combined events.
(1) "To get an odd number AND To get a number less than 3".
(2) "To get an odd number OR To get a number less than 3".

Exercise 1.6.11. Let E be any event, $\overline{E}$ its complement, and $p = p(E)$. What are the events $E \cap \overline{E}$, $E \cup \overline{E}$, $\overline{E \cap \overline{E}}$, $\overline{E \cup \overline{E}}$, and what are their probabilities, in terms of p?

Exercise 1.6.12. 7 people get in an elevator on the first floor of an 11-story building. What is the probability that no two of them get out of the elevator at the same floor?

Exercise 1.6.13. A $5 \times 5 \times 10$ wooden parallelepiped with red sides cut into 250 unit cubes. What is the probability that a randomly chosen unit cube has no red face? One red face? Two red faces? Three red faces? Four or more red faces?

Exercise 1.6.14. A die has 4 blue and 2 red faces. What is the probability that the two red faces have a common edge? What is the probability that the two blue faces have a common edge?

Exercise 1.6.15. Six faces of a regular die are marked by letters *A, B, A, C, U, S*. Find the probability that on six rolls of the die, the letters shown can be rearranged to spell "*ABACUS*".

Exercise 1.6.16. There are d dolphins in the ocean. d_1 of them were caught, marked, and released back. Next time, d_2 dolphins were caught and checked. Assuming independence, compute the probability that m marked species were caught the second time.

Exercise 1.6.17. A gentleman has 10 dress shirts and 10 ties, one matching tie to every shirt. Preparing for a long meeting, he selects at random 2 shirts and 2 ties. What is the probability that he gets exactly one matching pair of a shirt and a tie? At least one matching pair? Two matching pairs? If he selects at random 5 shirts and 5 ties, what is the probability to have exactly 2 matching pairs?

Exercise 1.6.18. What is the probability to get (at least) two consecutive tails if a fair coin is tossed 12 times?

Exercise 1.6.19. What is the probability that in a random permutation of numbers $1, 2, 3, \ldots, 1000$ at least one number occupies its own place (for example, the 5 is the 5th number in the permutation)?

Exercise 1.6.20. Among the whole numbers, some can be written down without the digit 1, like 527, while the others contain a 1, like 21 345. If you randomly pick a whole number from 0 through 999 999 inclusive, what is more probable, to pick a number with or without a 1 in its decimal representation? Does the probability change if we consider numbers from 1 to 999, or from 1 through 999 999 999?

Exercise 1.6.21. The U. S. Senate consists of 100 Senators, two from each state. If 10 senators are chosen at random, what is the probability that this cohort contains a senator from New York State?

Exercise 1.6.22. An urn contains 3 black, 3 white, and 3 yellow balls. n balls are taken at random without replacement. For each $n, n = 1, 2, \ldots, 9$, find the probability that among the n selected balls there are balls of all three colors?

Exercise 1.6.23. An urn contains 8 balls with the letters of the word STALLION. If 4 balls are chosen at random without replacement, what is the probability that either the word TOLL or the word LION can be composed of these balls?

Exercise 1.6.24. Two fair right indistinguishable tetrahedrons, with faces marked 1 through 4, are tossed.
(1) What is the probability that the sum of the numbers on the bottom faces is 7?
(2) What is the probability that the sum of these numbers is greater than or equal to 7?

Exercise 1.6.25. Let m and n be natural numbers. Consider four points $O(0, 0)$, $A(m, 0)$, $B(m, n)$, and $C(0, n)$ in the coordinate plane, and choose a random rectangle R with sides parallel to the coordinate axes and with vertices at points with integer coordinates inside or at the boundary of the rectangle $OABC$. What is the probability that R is a square?

Exercise 1.6.26. Let $S = \{1, 2, \ldots, n\}$, where n can be any natural number, and $f(s) = cs^2$, c is constant. Find c so that f is a random variable on S.

Exercise 1.6.27. Let $S = \{1, 2, \ldots, n\}$, where n can be any natural number, and $f(s) = c/s$, c is constant. Find c so that f is a random variable on S.

Exercise 1.6.28. Let $S = \{1, 2, 3, \ldots\}$, that is, in this problem the sample space is infinite, and $f(s) = c/s^2$, c is constant. Find c so that f is a random variable on S.

Exercise 1.6.29. What is the probability to get at least one 3 in a roll of two dice if the sum is odd?

Exercise 1.6.30. Two hunters simultaneously shoot a wolf. Under the given conditions, the probability to kill the animal for each of them is 1/3. What is the probability for the wolf to survive?

Exercise 1.6.31. Six cards are drawn at random from a regular deck of 52 cards. What is the probability that
(1) The Queen of spades was chosen among these 6 cards?
(2) All 4 suits will appear among these 6 cards?

Exercise 1.6.32. Several cards are drawn at random from a regular deck of 52 cards. We want to guarantee with the probability more than 1/2 that at least 2 cards of the same kind appear among the cards chosen. What is the smallest number of cards that must be drawn for that?

Exercise 1.6.33. If $p(A) = 0.55$, $p(B) = 0.75$, and the events A and B are independent, what are the conditional probabilities $p(A|B)$ and $p(B|A)$?

Exercise 1.6.34. In a certain population, 30 % of men and 35 % of women have a college degree; it is also known that 52 % of the population are women. If a person chosen at random in this population has a college degree, what is the probability that the person is a woman?

Exercise 1.6.35. The Student Government at The Game College sold 250 lottery tickets worth \$1 each. There are one \$100 prize, one \$50, and three \$10 prizes. If a student bought 2 tickets, what is the expected value of her net gain?

Exercise 1.6.36. Among 10 000 coins all but one are fair, and one has tails on both sides. A randomly chosen coin was thrown 12 times.
(1) What is the probability that this coin was false, if a tail was observed in all 12 trials?
(2) What is the probability that this coin was fair, if a tail was observed in all 12 trials?

Exercise 1.6.37. Every juror makes a right decision with the probability p. In a jury of three people two jurors follow their instincts, but the third juror flips a fair coin, and then the verdict follows the majority of jurors. What is the probability that the jury makes the right decision? Does this probability change if the jury consists of four jurors, and only one among them flips a coin?

Exercise 1.6.38. Assuming independence, what is the probability that at least two of the first 43 Presidents of the USA have the same birthday?

Exercise 1.6.39. We roll a fair die 6 times. If a 1 occurs first, or if a 2 occurs second, or if a 3 occurs third, ..., or if a 6 occurs sixth, we get \$1. What is the expected value of our gain?

Remark 1.6.2. This result illustrates an important theorem that the expected value of the finite sum of random variables is equal to the sum of their expected values.

Exercise 1.6.40. The President of the Combi Club introduced the following game. A participant pays \$1 and selects at random an integer number between 1 and 1 000 000 inclusive. If the decimal representation of the number contains a 1, the participant gets \$2, otherwise the participant loses the game. What is the expected value of the game?

Exercise 1.6.41. State the inclusion–exclusion formula (1.1.4) in probabilistic terms of this section.

Exercise 1.6.42. At four class tests, a student scored 76, 81, 89, and 92. At the fifth test, she can equally likely get scores from 75 through 95 inclusive. What is the probability that her average will be 85? At least 85? What is the probability that her average will be 85 if her fifth score is 85?

Exercise 1.6.43. In a class of 20 students, each two have a common grandfather. Prove that among them at least 14 students have the same grandfather.

Exercise 1.6.44. The spring in an old-fashioned watch breaks at a random moment. What is the probability that the hour hand will show time after 2 A. M. but before 4 A. M.?

Exercise 1.6.45. The Combi Club has a round roulette table with a rotating pointer. The table is divided in three sectors, one half-circle marked by the digit 3, and two quarter-circles marked by 1 and by 2, respectively. You pay \$1 to make a single spin and get back the reward in \$ equal to the number in the sector the pointer stops. Using negative numbers for a loss, find a probability distribution for the net gain if the game was played once. What is the average (expected) gain for each play? What would be the fair price of the game? Solve this problem if you pay \$2 for a spin; \$1.50 for a spin. Explore the similar problem if the table is divided in 6 equal sectors marked by the digits $1, 2, 3, 4, 5, 6$.

Exercise 1.6.46. A multiple-choice exam consists of 20 questions, with 4 possible answers for every question. If a student randomly guesses the answer to each question, find the probabilities that she gets *zero*, *one*, *two*, *three*, *four*, ..., 20 correct answers. What is the expected number of questions guessed correct? Solve the same problem if each question has 5 possible answers.

Exercise 1.6.47. Assuming that any day has equal chances to be the birthday of a person chosen at random from a very big population, find the probability that two randomly selected people both have birthdays on Sunday; find the probability that at least one of two randomly selected people will have a birthday on Sunday; find the probability that two randomly selected people have birthdays on the same day of a week.

Exercise 1.6.48. Each member of the Combi Club is required to take a course in Combinatorial Analysis (course C) and in Probability Theory (course P). The Registrar Office reports that 80 % pass course C, 75 % pass course P, and 90 % pass at least one of the courses. Find the probability of passing both courses. What is the probability that a person who passes course C will also pass course P? Are passing course C and passing course P independent events?

2 Basic graph theory

Graphs and, in particular, trees are graphical mathematical models useful in many problems. This chapter is devoted to a brief introduction to graph theory. The first two sections introduce the vocabulary. Graphs are defined in terms of sets and mappings in the spirit of Section 1.1, but we shall soon resort to more intuitive and transparent language of geometric diagrams. In the next three sections we study important special classes of graphs—trees, planar graphs, Eulerian graphs—and use the methods and results of Chapter 1 to solve some problems of graphical enumeration. More such problems are considered in Chapters 4 and 5. Our exposition was strongly influenced by Wilson's beautiful book [58]. For a deeper study of the subject we recommend the course of B. Bollobás [9]; we mostly follow the latter in terminology.

2.1 Vocabulary

In this section we create the basic vocabulary of the graph theory and consider elementary properties of graphs. Recall that 2_2^V denotes the set of all two-element subsets of a set V.

Coffee-time browsing

– http://graphjam.com/ (Drawing graphs)
– www.edu.pe.ca/kish/Grassroots/math/euclid.htm (Euclid's biography)

Definition 2.1.1. A *graph* G is a *triple* $G = (V, E, f)$, where $V \neq \emptyset$ is a non-empty set, whose elements are called *vertices* of the graph G, E is a set, maybe empty, whose elements are called *edges* of G, and the mapping

$$f : E \to 2_2^V \cup V$$

is called the *incidence function* of the graph. We write here $2_2^V \cup V$ instead of $2_2^V \cup 2_1^V$ to simplify notations. If $E = \emptyset$, all vertices are said to be *isolated*; in this case f is the empty mapping. The number of vertices $p = |V|$ is called the *order* of the graph G and the number of edges $q = |E|$ is called the *size* of G.

Thus, for any edge $e \in E$, its image $f(e)$ contains either one vertex v if $f(e) \subset V$, or two vertices v_1, v_2 if $f(e) = \{v_1, v_2\} \subset 2_2^V$; these vertices are called the end vertices of the edge e. If a vertex v is an end vertex of an edge e, the vertex and the edge are called *incident* to one another. If two edges e_k and e_l have the same pair of end vertices, $f(e_k) = f(e_l) = \{v_i, v_j\}$, these edges are called parallel or multiple. If $f(e) \in V$, that is, the edge e has the only end vertex, e is called a loop. Vertices having only one incident edge, which is not a loop, are called pendants or leaves.

https://doi.org/10.1515/9783110751185-002

Example 2.1.1. Consider a graph $G = (V,E,f)$ (Fig. 2.1), where $V = \{v_1,\dots,v_5\}$, $E = \{e_1,\dots,e_6\}$, and

$$f(e_1) = \{v_1,v_2\},\quad f(e_2) = f(e_3) = \{v_1,v_3\},$$
$$f(e_4) = \{v_2,v_4\},\quad f(e_5) = \{v_3,v_4\},\quad f(e_6) = v_4. \tag{2.1.1}$$

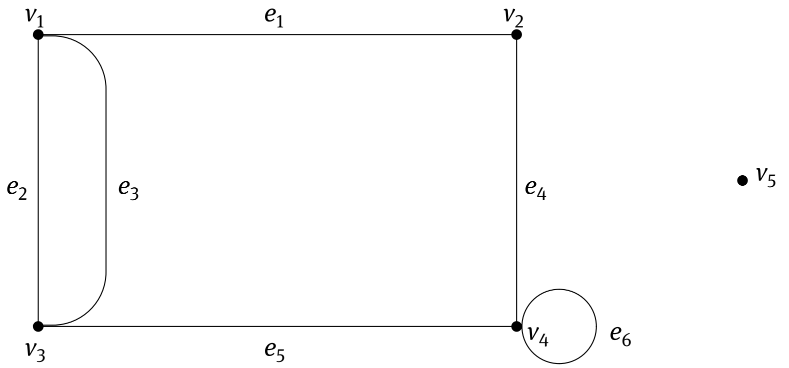


Figure 2.1: Diagram g represents the graph G in Example 2.1.1.

This graph has no pendant vertex, the edges e_2 and e_3 are parallel, the edge e_6 is a loop, the vertex v_5 is isolated.

We often suppress the symbol f and denote graphs by $G = (V,E)$. In this text we mostly consider *undirected* graphs, which means that the two end vertices of an edge are not ordered. Moreover, we deal only with *finite* graphs, always assuming $|V| < \infty$ and $|E| < \infty$.

We consider graphs with parallel (multiple) edges, called sometimes *multigraphs*, and/or with loops, called *pseudographs*. Some authors use the term “graph” only for the so-called *simple graphs*, that is, for graphs without parallel edges and loops. Hereafter, we call a graph G *simple* if G has no parallel edges nor loops. To define simple graphs in set-theory terms, it suffices to consider the set of edges E as a subset of 2^V_2.

It is useful to visualize abstract concepts. All the more this is true in the case of graphs. We call the graphs in the sense of Definition 2.1.1 abstract graphs and represent them by drawing *geometric graphs* or *diagrams*. These are sets of points in the plane (one can even imagine points in R^n) representing the vertices of a graph, some of them connected by smooth arcs, maybe line segments, representing the edges of the graph. For each edge only its end points belong to V, in other words, no interior point of any edge belongs to V. It should be noted that an actual shape of an edge makes no difference in our considerations. For any such a diagram, we can always easily restore the *incidence function* of the graph presented. Such diagrams are called *embeddings* of a graph into the Euclidean plane R^2. For example, Fig. 2.1 depicts diagram g corresponding to graph G defined by (2.1.1).

Definition 2.1.2. A *directed graph (digraph)* $\overline{G}$ is a triple $\overline{G} = (V, E, \bar{f})$, where V and E have the same meaning as above and the incidence function is now the mapping

$$\bar{f} : E \to V \times V.$$

If $\bar{f}(e) = (v_1, v_2)$ for an edge $e \in E$, then v_1 is called the *initial vertex* of e and v_2 the *end vertex* of e.

Let a digraph $\overline{G}$ have vertices v_1 and v_2 and edges e_1 and e_2, such that $\bar{f}(e_1) = (v_1, v_2)$ and $\bar{f}(e_2) = (v_2, v_1)$. Since the ordered pairs $(v_1, v_2) \neq (v_2, v_1)$, this implies that the oriented edges $e_1 \neq e_2$.

Example 2.1.2. Consider a digraph $\overline{G} = (V, E, \bar{f})$ shown in Fig. 2.2, where $V = \{v_1, \dots, v_5\}$, $E = \{e_1, \dots, e_5\}$, and

$$\bar{f}(e_1) = (v_1, v_2), \quad \bar{f}(e_2) = (v_3, v_1), \quad \bar{f}(e_3) = (v_1, v_3),$$
$$\bar{f}(e_4) = (v_2, v_4), \quad \bar{f}(e_5) = (v_3, v_4).$$

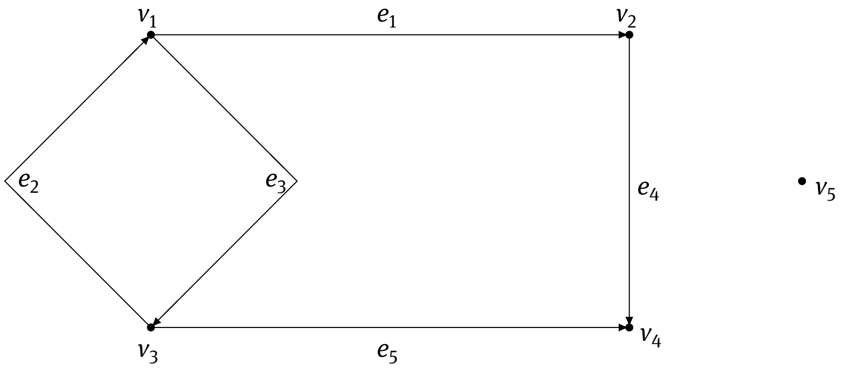


Figure 2.2: Digraph $\overline{G}$.

We slightly abuse the language and do not distinguish an abstract graph and its geometric realization like the graph G in (2.1.1) and the corresponding diagram g in Fig. 2.1. Usually this does not lead to any misunderstanding. If we need to emphasize this distinction, we reserve capital letters (G) for graphs and small ones (g) for the corresponding geometric diagrams. The same graph (incidence relation (2.1.1)) can be drawn in infinitely many other ways; compare, for instance, the diagram g (Fig. 2.1) and the diagram g' (Fig. 2.3).

Definition 2.1.3. A graph $G = (V, E)$ is said to be *regularly embedded* in R^n if its edges have no common points except for the end vertices. A graph is called *planar* if it can be regularly embedded in R^2.

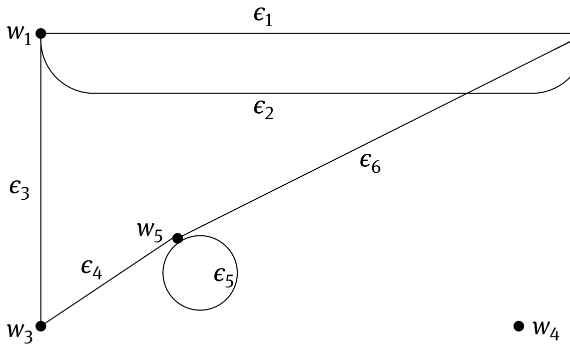


Figure 2.3: Diagram g'.

Example 2.1.3. The diagram (graph) g (Fig. 2.1) is regularly embedded in R^2, but g' (Fig. 2.3) is not, since the edges ϵ_2 and ϵ_6 intersect at a point which is not a vertex of the graph. However, this graph is planar.

Problem 2.1.1. Draw a regular plane embedding of the graph g' (Fig. 2.3), that is, its embedding into $\mathbf{R}^2$.

Lemma 2.1.1. *Every finite (and even countable) graph can be regularly embedded in R^3.*

Proof. Indeed, let a graph G be of order p and size q. Consider a line L in R^3 and a bundle of q different half-planes bounded by L; any half-plane corresponds to one and only one edge (Fig. 2.4). Select p different points on L, one for each vertex. If two vertices v_1 and v_2 of the graph are incident to an edge e_1, we connect the corresponding points in L by a half-circle located in the half-plane corresponding to e_1. If e_2 is a loop at a vertex v_3, we draw a circle in the half-plane corresponding to e_2, such that this circle is tangent to L at v_3. It is obvious that these q circles and half-circles have no points in common but maybe their end vertices. □

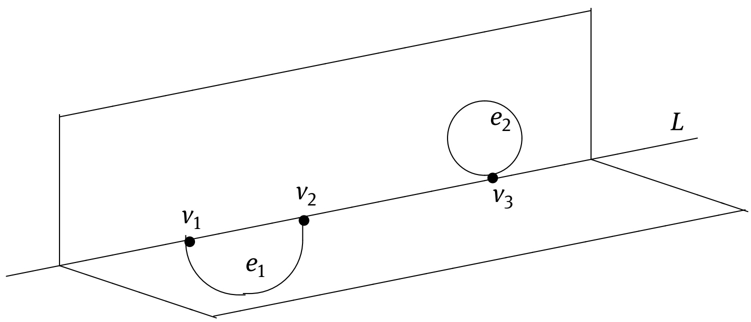


Figure 2.4: Regular embedding of a graph in R^3.

Definition 2.1.4. The *degree*, $\deg(v)$, of a vertex $v \in V$ is the total number of edges incident to this vertex; by definition, every loop must be counted twice.[1] A vertex of an even (odd) degree is for short called an *even (odd)* vertex. If $V = \{v_1, \ldots, v_p\}$, then the sequence $(\deg(v_1), \ldots, \deg(v_p))$ is called the *degree sequence* of a graph; unlike the set $\{\deg(v_1), \ldots, \deg(v_p)\}$, this sequence depends on the numbering of the vertices.

Example 2.1.4. In diagram g (Fig. 2.1), $\deg(v_1) = \deg(v_3) = 3$, $\deg(v_2) = 2$, $\deg(v_4) = 4$ since v_4 is incident to edges e_4, e_5 and to the *loop* e_6, and $\deg(v_5) = 0$ since v_5 is an isolated vertex. The degree sequence is $(3, 2, 3, 4, 0)$.

Definition 2.1.5. If any two vertices of a simple graph are adjacent, the graph is called *complete*. A complete graph of order p is denoted by K_p.

Problem 2.1.2. What is the degree of any vertex in K_p? What is the degree sequence of K_p?

Definition 2.1.6. A simple graph $G = (V, E)$ is called *bipartite* if $V = V_1 \cup V_2$ with $V_1 \cap V_2 = \emptyset$, and each edge connects a vertex in V_1 with a vertex in V_2. A simple bipartite graph $G = (V_1 \cup V_2, E)$ is called complete if each vertex in V_1 is connected with every vertex in V_2. If $|V_1| = m$ and $|V_2| = n$, the complete bipartite graph is denoted by $K_{m,n}$.

Problem 2.1.3. What are the degrees of vertices of $K_{m,n}$?

Lemma 2.1.2. *In any graph of size q,*

$$\sum_{v \in V} \deg(v) = 2q. \tag{2.1.2}$$

Proof. We again do double counting, computing twice the total number of the end vertices. First, we just sum up over all the end vertices and second, we take into consideration that each edge has two ends. This reasoning shows, in particular, why it is often convenient to consider loops as having two ends. □

Lemma 2.1.2 is called the *handshaking lemma*, because if we depict the participants of a party by vertices of a graph such that two vertices of the graph are connected by an edge if and only if the two corresponding people exchanged a handshake, then the lemma just states that the total number of shaken hands is even.

Corollary 2.1.1. *In any graph the number of odd vertices is even.*

Proof. Indeed, if we split the left-hand side of (2.1.2) as

$$\sum_{v \in V} \deg(v) = \sum_{v \in V : \deg(v) \text{ is even}} \deg(v) + \sum_{v \in V : \deg(v) \text{ is odd}} \deg(v),$$

1 However, there are problems where it is more convenient to count a loop just once.

then the first sum on the right is even and the total sum is even, thus the second sum on the right must be even. But this sum contains only odd addends, so there must be an even number of such vertices. □

Definition 2.1.1 of abstract graphs is convenient, for it includes graphs with loops and parallel edges. However, it implies, for example, that graphs $G = (V, E, f)$ and $G_1 = (V_1, E, f_1)$ with $V \neq V_1$ are different even if $|V| = |V_1|$, since the incidence functions have different domains. Moreover, diagrams g and g' (Figs. 2.1 and 2.3) have different appearance, despite the fact that they realize the same incidence relationship among five points. In many problems it is natural to identify such graphs. To this end the following definition is useful.

Definition 2.1.7. Two graphs $G = (V, E, f)$ and $G_1 = (V_1, E_1, f_1)$ are called *isomorphic*, denoted by $G \cong G_1$, if there are two one-to-one correspondences,

$$\varphi : V \to V_1$$

and

$$\psi : E \to E_1$$

compatible with the incidence functions in the sense that

$$f_1(\psi(e)) = \varphi(f(e)), \quad \forall e \in E.$$

Example 2.1.5. Diagrams (graphs) g and g' (Fig. 2.1 and 2.3) are isomorphic to one another; the one-to-one correspondences between them can be established as follows:

$$\varphi(v_1) = w_1, \quad \varphi(v_2) = w_3, \quad \varphi(v_3) = w_2, \quad \varphi(v_4) = w_5, \quad \varphi(v_5) = w_4,$$
$$\psi(e_1) = \epsilon_3, \quad \psi(e_2) = \epsilon_1, \quad \psi(e_3) = \epsilon_2, \quad \psi(e_4) = \epsilon_4, \quad \psi(e_5) = \epsilon_6, \quad \psi(e_6) = \epsilon_5.$$

Definition 2.1.8. A graph of order p is called *labeled* if its vertices are labeled by the first p natural numbers, thus $V = \{1, 2, \ldots, p\}$. Labeled graphs are called *isomorphic* if the bijection φ in Definition 2.1.7 preserves not only the incidence relation but also the labeling, that is $\varphi(i) = i$, $1 \leq i \leq p$.

For example, if we identify $v_i \equiv i$ and $w_i \equiv i$ for $1 \leq i \leq 5$, then it turns out that the diagrams g and g' (Fig. 2.1 and 2.3), which are isomorphic in the sense of Definition 2.1.7, are not isomorphic as the labeled graphs.

We end this section by solving a few problems of enumeration of labeled graphs. Fix a set E and consider all labeled graphs $G = (V, E, f)$, where $V = \{1, 2, \ldots, p\}$, of order p and size $q = |E| \geq 0$. To count all different graphs, we have to enumerate various incidence functions $E \to 2_2^V \cup V$; by Lemma 1.1.2, $|2_2^V \cup V| = \frac{p(p-1)}{2} + p = \frac{p(p+1)}{2}$. Then, by Theorem 1.1.6, there are

$$\left(\frac{p(p+1)}{2}\right)^q \tag{2.1.3}$$

such graphs.

If we want to count the graphs without loops, but with parallel edges, it is enough to consider incidence functions $f: E \to 2_2^V$; in this case there are $(\frac{p(p-1)}{2})^q$ (possibly pairwise isomorphic) graphs. If we want to exclude both loops and parallel edges, we should consider only *injective* mappings $f: E \to 2_2^V$; to ensure their existence, we assume $p(p-1)/2 \geq q$. By Theorem 1.1.7 and Corollary 1.1.1, there are $2^{p(p-1)/2}$ simple graphs of order p with any $q, 0 \leq q \leq p(p-1)/2$, and by Theorem 1.1.10 among them there are

$$\frac{(\frac{p(p-1)}{2})!}{(\frac{p(p-1)}{2} - q)!} \tag{2.1.4}$$

(possibly isomorphic) simple graphs of order p and size q.

However, this number counts separately graphs with *different labeling* of edges. If we want to identify such graphs, that is, to make the edges indistinguishable, we must identify incidence functions corresponding to different ordering of edges. Therefore, instead of injective mappings we have to consider combinations of $p(p-1)/2$ elements taken q at a time; by (1.4.1), there are

$$\frac{(\frac{p(p-1)}{2})!}{q!(\frac{p(p-1)}{2} - q)!} \tag{2.1.5}$$

such graphs.

Consider, for instance, simple graphs of order $p = 3$ and size $q = 2$; if $V = \{v_1, v_2, v_3\}$, $E = \{e_1, e_2\}$, then by (2.1.4) there are $3!/1! = 6$ such abstract graphs $G_i(V, E, f_i), 1 \leq i \leq 6$, whose incidence functions are

$$\begin{aligned}
G_1 &: f_1(e_1) = \{v_1, v_2\}, \quad f_1(e_2) = \{v_2, v_3\},\\
G_2 &: f_2(e_1) = \{v_2, v_3\}, \quad f_2(e_2) = \{v_1, v_2\},\\
G_3 &: f_3(e_1) = \{v_1, v_3\}, \quad f_3(e_2) = \{v_1, v_2\},\\
G_4 &: f_4(e_1) = \{v_1, v_2\}, \quad f_4(e_2) = \{v_1, v_3\},\\
G_5 &: f_5(e_1) = \{v_2, v_3\}, \quad f_5(e_2) = \{v_1, v_3\},\\
G_6 &: f_6(e_1) = \{v_1, v_3\}, \quad f_6(e_2) = \{v_2, v_3\}.
\end{aligned}$$

Now, if we draw corresponding diagrams, omitting edge labels and using the standard vertex labels $v_j = j, j = 1, 2, 3$, we get only three different diagrams (Fig. 2.5) as given by (2.1.5) with $p = 3$ and $q = 2$—cf. also Theorem 2.3.2. These labeled graphs are non-isomorphic since the vertices of degree 2 have different labels. If in any of these

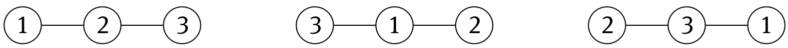


Figure 2.5: Non-isomorphic labeled simple graphs with three vertices and two edges.

graphs we switch the labels of two pendant vertices, we get a labeled graph isomorphic to the initial one.

However, if we erase the vertex labels, thus considering non-labeled graphs, all the three diagrams in Fig. 2.5 become identical, thus, there is only one non-labeled simple graph of order $p = 3$ and size $q = 2$. We solve more problems of graphical enumeration in the sequel sections.

Exercises 2.1.

Exercise 2.1.1. Draw all simple graphs with $1 \leq p \leq 5$ vertices and $0 \leq q \leq p(p-1)/2$ edges.

Exercise 2.1.2. Draw the complete graphs K_1–K_6.

Exercise 2.1.3.
(1) Is it possible to organize a tournament with 40 participating teams, if each team must play precisely three games?
(2) Answer the same question if there are 13 teams and every team must play 5 games.

Exercise 2.1.4. Are there graphs of order $p = 6$ with the following degree sequences? Draw, if there are any, all non-isomorphic diagrams with the given degree sequence.
(1) $(2, 2, 3, 4, 6, 7)$,
(2) $(2, 2, 3, 4, 6, 8)$.

Exercise 2.1.5. Is there a graph of order 6 with each vertex of degree 3?

Exercise 2.1.6. Prove that if $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are two isomorphic graphs, then $|V_1| = |V_2|$, $|E_1| = |E_2|$, and the degrees of the corresponding vertices are equal. Are these necessary conditions also sufficient for two graphs to be isomorphic?

Exercise 2.1.7. What graphs in Fig. 2.6 are pairwise isomorphic?

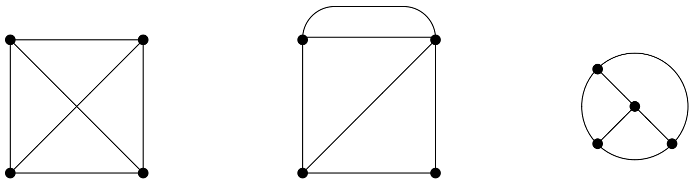


Figure 2.6: Are there isomorphic graphs here?—Cf. Exercises 2.1.6–2.1.7.

Exercise 2.1.8. Draw all non-isomorphic simple graphs of orders 1 through 4.

Exercise 2.1.9. 30 teams compete in a tournament, where each team must play every other team exactly once. Prove that at any time there are two teams that have played the same number of games.

Exercise 2.1.10. How many are there non-isomorphic planar graphs with $2n$ vertices and n edges if all edges are segments of straight lines and do not have common end points?

Exercise 2.1.11.
(1) Every person who now lives on Earth, had pairwise discussions with several other people. Prove that the total number of people having an odd number of conversations, is even.
(2) Prove that any polyhedron has an even number of faces with an odd number of edges.
(3) Prove that any polyhedron has an even number of vertices where an odd number of edges meet.

Exercise 2.1.12. Given a simple graph with v vertices each of degree d, what are restrictions on d and on the parity of the product $d \cdot v$?

Exercise 2.1.13. At how many points do all edges of the complete bipartite graph $K_{m,n}$ intersect, if no three edges intersect at a point?

Exercise 2.1.14. No three edges of the complete graph K_n intersect at a point, except maybe at a vertex of the graph. In how many parts do all the edges split the interior of the graph? Think of cutting a birthday cake.

Exercise 2.1.15. At a meeting of the Combi Club some members are friends and some are not. Assuming that the binary relation of being friends is symmetric (Definition 1.1.11, Part (2)) prove that there are at least two people at the meeting who have the same number of friends in the audience.

Exercise 2.1.16. The inhabitants of planet Triplan exchange handshakes only in triples, that is, by simultaneously connecting three limbs of three inhabitants. State and prove an analogue of the handshaking lemma (Lemma 2.1.2) for the Triplan world.

Exercise 2.1.17. Prove that there are $C(\frac{p(p-1)}{2} + q - 1, q)$ labeled graphs of order p and size q with parallel edges but without loops, and there are $C(\frac{p(p+1)}{2} + q - 1, q)$ labeled graphs of order p and size q with parallel edges and loops.

2.2 Connectivity in graphs

In this section we study graphs as mathematical models for problems concerning with connectivity between different objects.

Definition 2.2.1. Consider a graph $G = (V, E, f)$, a subset $V_1 \subset V$, $V_1 \neq \emptyset$, of its set of vertices, and a subset $E_1 \subset E$ of its set of edges, such that all end vertices of edges in E_1 belong to V_1. The graph $G_1 = (V_1, E_1, f_1)$ is called a *subgraph* of G if its incidence function f_1 is the *restriction* (see Definition 1.1.15) of the incidence function f, $f_1 = f|_{E_1}$.

The definition means that we pick one or more vertices of the given graph and several, maybe no, edges of the given graph, connected with the selected vertices. We consider examples after the next definition.

Definition 2.2.2. Any subgraph G_1 of G with $V_1 = V$ is called a *spanning subgraph* or a *factor* of G.

A spanning subgraph always exists, for instance, the graph itself is its own spanning subgraph, but in general it is not unique. Subgraphs can be thought of as derived from the given graph by removing some edges and vertices; after removal of a vertex we must remove all edges incident to this vertex.

Example 2.2.1. This example again refers to the graph $G = (V, E, f)$ (Example 2.1.1 and Fig. 2.1). Consider the subset

$$V_1 = \{v_1, v_2, v_3, v_4\} \subset V$$

and the set

$$E_1 = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

of all edges of g whose end vertices belong to V_1. Thus, in this example $E_1 = E$, and the graph $G_1 = (V_1, E_1, f_1)$ with f_1 being the restriction of f onto E_1, is a subgraph of G. If we consider the same subset $V_2 = V_1 = \{v_1, v_2, v_3, v_4\} \subset V$ and the empty set of edges $E_2 = \emptyset$, the graph $G_2 = (V_2, E_2, f_2)$ with the "empty" incidence function f_2 is another subgraph of G with the same set of vertices as G_1. If we choose the same subset of vertices $V_3 = V_1$ and the set of edges $E_3 = \{e_1, e_2\}$, we get yet another subgraph of G with the same set of vertices.

If we start with another subset of vertices, say $V_2 = \{v_5\} \subset V$, the corresponding subgraph is $G_2 = (V_2, \emptyset, f_2)$, where $f_2 = f|_\emptyset$ is the empty mapping. The subset $V_3 = \{v_2, v_4\} \subset V$ generates a subgraph $G_3 = (V_3, E_3, f_3)$, where $E_3 = \{e_4, e_6\}$. The graph $G_4 = (V_4, E_4, f_4)$, where $V_4 = V$, $E_4 = \{e_1, e_2, e_4, e_5\}$, and $f_4 = f|_{E_4}$, is a spanning subgraph of G.

Graph theory provides a convenient language for formalizing the concept of *connectivity* of different objects.

Definition 2.2.3. A sequence of intermittent vertices v_{i_l} and edges e_{j_k},

$$(v_{i_0}, e_{j_1}, v_{i_1}, e_{j_2}, v_{i_2}, \ldots, e_{j_k}, v_{i_k}),$$

is called a *walk* of length k connecting its end vertices v_{i_0} and v_{i_k}. A walk is called a *trail* if all its edges are different. If all vertices of a trail, except maybe for its end vertices, are different, the trail is called a *path*. A trail is called a *circuit* if $v_{i_0} = v_{i_k}$. If the end vertices of a path coincide, the path is called a *cycle*. Thus, any loop (v_0, e_1, v_0) is a cycle.

The corresponding objects in a digraph are called a directed walk, directed trail, etc. A directed cycle is called a contour. It is important that all edges in a directed walk must have the same orientation, that is, the end vertex of a preceding edge must be the initial vertex of the sequel edge.

Hence, a trail can contain repeating vertices, that is, have self-crossings, a circuit is a closed trail, and a cycle is a closed path. If $k = 0$, the walk consists of one vertex and has zero length. The edge e_{j_m} is obviously a loop if $v_{i_{m-1}} = v_{i_m}$. If $e_{j_{n-1}} = e_{j_n}$, the edge e_{j_n} is passed twice. A circuit is a closed trail and we do not single out any its vertex as an end vertex. It follows from these definitions that any open path in a simple graph of order n contains at most $n-1$ edges. A path in a graph can be considered as a subgraph of the graph.

Example 2.2.2. Consider a walk $(v_1, e_2, v_3, e_3, v_1, e_1, v_2, e_1, v_1, e_3, v_3)$—see graph $G = (V, E, f)$ in Fig. 2.1. This walk is not a trail since it contains repeating edges, say, e_1. However, the walk $(v_1, e_2, v_3, e_3, v_1, e_1, v_2)$ is a trail but not a path, because it has a repeating vertex v_1. The walk $(v_4, e_5, v_3, e_3, v_1, e_1, v_2, e_4, v_4, e_6, v_4)$ is a circuit, and the path $(v_4, e_5, v_3, e_3, v_1, e_1, v_2, e_4, v_4)$ is a cycle. It is often possible, without any ambiguity, to write down walks with only edges and skip some vertices. For instance, the latter cycle can be represented as (e_5, e_3, e_1, e_4).

The following simple properties of graphs are useful in many instances.

Lemma 2.2.1.
(1) *If two vertices of a graph are connected by a walk W of length n, then they are connected by a path of length at most n.*
(2) *Each circuit contains a cycle.*

Proof. If the walk W is not a path, then while traversing it, we must eventually arrive at some vertex v' the second time, otherwise the graph would be infinite. Thus, W contains a circuit starting and ending at v'. Removing from W all elements of this circuit, except for the vertex v', we get a walk W', which must be shorter than W, otherwise W would be a path. Since W is finite, repeating this process several times we remove from W all repeating edges and vertices and derive the path we look for. The same reasoning yields the second statement of the lemma. □

Consider again the diagram g in Fig. 2.1. If we depart from vertex v_1 and travel through g using its edges as roadways, we can reach any other vertex except for the isolated vertex v_5, and the same is true if we depart from the vertices v_2–v_4. However, from v_5 we can reach only v_5 itself, using zero path. This observation leads to the next definition.

Definition 2.2.4. Given a vertex $v \in V$, denote by $C(v)$ the set of all vertices of a graph $G = (V, E)$ connected with v by a path[2] in G. The subgraph G' of G, generated by $C(v)$, is called a *connected component* of G; that is, a connected component consists of all vertices in $C(v)$ and all edges of G whose end vertices also are in $C(v)$.

A graph consisting of only one connected component, is called a connected graph, thus, every connected component is a connected graph.

An edge e in a graph G is called a cut-edge or a bridge if its removal increases the number of connected components in the graph.

We denote the number of connected components in a graph G by $cc(G)$.

Example 2.2.3. The diagram g (Fig. 2.1) has two connected components, namely, g_1 and $g_2 = (V_2, \emptyset, \emptyset)$, where $V_2 = \{v_5\}$—see Fig. 2.7.

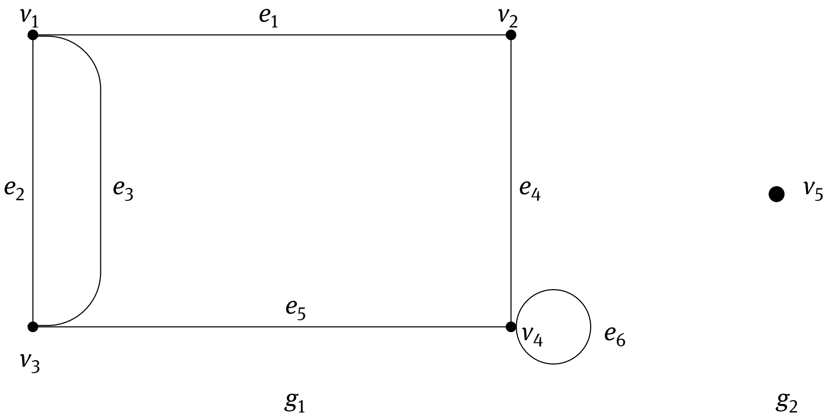


Figure 2.7: Diagrams g_1 and g_2—see graph G_1.

Lemma 2.2.2. *If the degree of every vertex of a finite graph is at least 2, the graph contains a cycle.*

Proof. Without loss of generality, we can consider a simple connected graph. Pick any vertex v and any edge e incident to v. Another end vertex of e also has the degree at least 2, so that it has at least one incident edge distinct from e. Continuing this way, after several steps we must meet some vertex the second time, since the graph is finite. The part of our walk, which starts and ends at this repeating vertex, is the cycle we sought for. □

Problem 2.2.1. Why is the other end vertex of e distinct from v?

2 Or, which is equivalent due to Lemma 2.2.1, by a walk.

Lemma 2.2.3. *Consider a graph $G = (V, E)$ and let $G' = (V, E \setminus \{e\})$ be a subgraph of G derived by removal of an edge $e \in E$. If the edge e is a bridge in G, then $cc(G') = cc(G) + 1$.*

Proof. The removal of a bridge e can affect only the connected components containing e. Thus without loss of generality, we can consider a connected graph $G, cc(G) = 1$, and prove that $cc(G') = 2$. Now, if we assume that G' has at least three connected components, $cc(G') \geq 3$, then there are three vertices, say $v_1, v_2, v_3 \in V$, belonging to these three different connected components. Since G is connected, there are three paths connecting v_1, v_2, and v_3 pairwise. Moreover, since e is a bridge, each of these three paths must contain the edge e. Then it is easy to see for ourselves that we can always rearrange these paths and construct a new path Γ, which connects two certain vertices among the three vertices v_1, v_2, v_3, and does not contain the bridge e, contrary to our assumption on the existence of three connected components in G'. □

Problem 2.2.2. There is a missing step in the end of the latter proof; restore it, namely, give a more precise construction of the path Γ. Hint: consider several possible cases.

Lemma 2.2.4. *For any finite graph $G = (V, E)$ of order p and size q, the inequality $p - cc(G) \leq q$ is valid.*

Proof. We prove the inequality by mathematical induction on the size q of the graph. If $q = 0$, then each vertex is isolated, so that $cc(G) = p$ and the conclusion is clear. Suppose now that the inequality holds for any graph of size less than $q, q > 0$, and prove the statement for graphs of size q.

Consider separately two possible cases: either each edge of G is a bridge or there is an edge $e \in E$, which is not a bridge. In the latter case, we remove e and get a connected graph G' with $q - 1$ edges; by the inductive assumption $p - cc(G') \leq q - 1$ and since $cc(G) = cc(G')$ (e is not a bridge!) we immediately derive that

$$p - cc(G) = p - cc(G') \leq q - 1 < q.$$

In the former case, we remove an arbitrary edge $e' \in E$ and denote the remaining subgraph by G'. By Lemma 2.2.3, $cc(G') = cc(G) + 1$. Now by the inductive assumption $p - cc(G') \leq q - 1$, or $p - cc(G) - 1 \leq q - 1$, whence $p - cc(G) \leq q$. □

Geometrical diagrams visualize graphs; however, on many occasions, for instance, to store graphs in computer memory, it is more convenient to represent them analytically rather than geometrically. First we recall a few standard definitions.

Definition 2.2.5. An $m \times n$ *matrix* $M = (a_{i,j})$, $1 \leq i \leq m$, $1 \leq j \leq n$, is a rectangular array of m rows and n columns, where $a_{i,j}$ stands for the entry at the crossing of the ith row and jth column. If $m = 1$, a matrix consists of one row and can be considered as a row vector (n-tuple) of length n. A matrix with $m = n$ is called a square matrix. A square matrix is called symmetric if $a_{i,j} = a_{j,i}, \forall i, j$.

The matrix $M^* = (b_{i,j})$, $1 \leq i \leq m$, $1 \leq j \leq n$, with $b_{i,j} = a_{j,i}$, is called the transpose of the matrix M.

Definition 2.2.6. The *dot product* of two n-vectors $\mathbf{v}_1 = (a_{1,1}, a_{1,2}, \dots, a_{1,n})$ and $\mathbf{v}_2 = (a_{2,1}, a_{2,2}, \dots, a_{2,n})$ is the sum of pairwise products of their corresponding components,

$$\mathbf{v}_1 \bullet \mathbf{v}_2 = a_{1,1}a_{2,1} + a_{1,2}a_{2,2} + \cdots + a_{1,n}a_{2,n}.$$

Definition 2.2.7. Consider a $p \times q$ *matrix* $A = (a_{i,j})$ and a $q \times r$ matrix $B = (b_{i,j})$. Their *product* is a $p \times r$ matrix $C = A \times B = (c_{i,j})$, where each entry is the dot product of a q-vector $\mathbf{a}_i = (a_{i,1}, a_{i,2}, \dots, a_{i,q})$ by a q-vector $\mathbf{b}_j = (b_{1,j}, b_{2,j}, \dots, b_{q,j})$, that is, $c_{i,j} = \mathbf{a}_i \bullet \mathbf{b}_j = \sum_{k=1}^{q} a_{i,k}b_{k,j}$.

Problem 2.2.3. Let

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(1) Find the transposes A^*, B^*, C^*.
(2) Calculate the products $A \times B$, $B \times A$, and $A \times C$. Does the matrix product $C \times A$ exist? Is the matrix multiplication commutative?
(3) Find the transposes $(A \times B)^*$ and $(B \times A)^*$. Is there any relationship between A^*, B^*, and $(A \times B)^*$ or $(B \times A)^*$? Prove this relationship.

Definition 2.2.8. Let $G = (V, E)$ be a graph of order p with the set of vertices $V = \{v_1, \dots, v_p\}$. The *adjacency matrix* of G, corresponding to the given numbering of the vertices, is a $p \times p$ square matrix

$$A(G) = (a_{i,j}), \quad 1 \le i, j \le p,$$

where $a_{i,j}$ is the number of edges connecting the vertices v_i and v_j. Thus, $a_{i,i}$ is the number of loops at the vertex v_i. This matrix is symmetric as long as we consider only undirected graphs.

If we compute the column sum $\rho(v_j) = \sum_{i=1}^{p} a_{i,j}$ or the row sum $\rho(v_j) = \sum_{i=1}^{p} a_{j,i}$, $1 \le j \le p$, in the adjacency matrix, which is the same due to the symmetry of $A(G)$, then we find the number of edges incident to v_j, but unlike the degree of vertex v_j, each loop in this sum is counted only once. Thus,

$$\deg(v_j) = \left(\sum_{i=1, i \ne j}^{p} a_{i,j} \right) + 2a_{j,j} = \left(\sum_{i=1}^{p} a_{i,j} \right) + a_{j,j}.$$

Example 2.2.4. The adjacency matrix of the graph G (Fig. 2.1) is

$$A(G) = \begin{pmatrix} 0 & 1 & 2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix $A(G)$ has a 2×2 block structure,

$$A(G) = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

where $A_1 = A(g_1)$ is the 4×4 adjacency matrix of the graph g_1 and $A_2 = A(g_2)$ is the 1×1 adjacency matrix of the graph $g_2 = (\{v_5\}, \emptyset, \emptyset)$ (Fig. 2.7); the latter graph consists of one isolated vertex v_5, whence A_2 is the 1×1 matrix $A_2 = (0)$. All other elements of $A(G)$ are zeros situated in two rectangular blocks. The two graphs g_1 and g_2 are the two connected components of the graph G. Any adjacency matrix has such a block structure.

Problem 2.2.4. Prove that if a graph has k connected components, then its adjacency matrix is a $k \times k$ block matrix, these k blocks are the adjacency matrices of the connected components of the graph and are situated along the main diagonal; all the other elements of the adjacency matrix being zeros. The size of each block is equal to the order of the corresponding connected component of the graph.

Let us take another look at the graph G (Fig. 2.1)—the vertices v_1 and v_3 have two incident edges, e_2 and e_3. Hence there are two walks from v_1 to v_3, and we can read this off from the adjacency matrix $A(G)$, since $a_{1,3} = a_{3,1} = 2$. All other entries of $A(G)$ can be interpreted the same way. For instance, we see that v_4 is connected by walks of length 1 with v_2, v_3, and with itself, thus there is a loop at v_4. In many instances it is necessary to count walks of lengths bigger than 1. To approach this problem, we compute the square $A^2(G) = A(G) \cdot A(G)$ of the adjacency matrix $A(G)$,

$$A^2(G) = \begin{pmatrix} 5 & 0 & 0 & 3 & 0 \\ 0 & 2 & 3 & 1 & 0 \\ 0 & 3 & 5 & 1 & 0 \\ 3 & 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Consider here the very first entry, $a_{1,1} = 5$. By Definition 2.2.6,

$$a_{1,1} = (0, 1, 2, 0, 0) \bullet (0, 1, 2, 0, 0) = 1 \cdot 1 + 2 \cdot 2 = 5. \tag{2.2.1}$$

This shows that there are five walks of length 2 starting and ending at v_1. Indeed, we can go from v_1 to v_2 and come back using the edge e_1, which gives $1 \cdot 1 = 1$ walk; or go from v_1 to v_3 using either e_2 or e_3 and come back using again either e_2 or e_3, which gives $2 \cdot 2 = 4$ more walks. We see for ourselves that each term in (2.2.1) represents a walk of length 2 from v_1 to itself. The five closed walks of length 2 starting at v_1 and returning back to v_1, are

$$(v_1, e_1, v_2, e_2, v_1), \quad (v_1, e_2, v_3, e_2, v_1), \quad (v_1, e_2, v_3, e_3, v_1), \quad (v_1, e_3, v_3, e_2, v_1),$$

and $(v_1, e_3, v_3, e_3, v_1)$. The other entries of $A^2(G)$ have the same meaning.

Obviously, this reasoning holds true in general and leads to the following statement.

Theorem 2.2.1. *The number of walks of length r from v_i to v_j in a graph $G = (V, E)$, $V = \{v_1, \dots, v_p\}$, is equal to the element $a_{i,j}^{[r]}$ of the matrix $(A(G))^r$, where $A(G)$ is the adjacency matrix of G. The matrices $(A(G))^r, r = 2, 3, \dots$, here exist since $A(G)$ and all its powers are square matrices.*

Problem 2.2.5. Prove this theorem by mathematical induction.

The graph-theory language is often useful in solving various problems that seem unrelated to graphs. For instance, the information in the following problem can be conveniently represented by a graph.

Problem 2.2.6. The organizing committee of The Combi Club Annual Meeting observed that among each four participants there is a person who knows three other people in this quartet. Show that every quartet of the participants contains someone who is familiar with all the other participants of the meeting.

Solution. Consider a graph with vertices corresponding to the participants, where two vertices are adjacent if and only if the corresponding participants *do not know* one another. We can clearly assume that the graph has at least 4 vertices. This graph has no parallel edges nor loops. We have to prove that if the graph contains no quadruple of vertices, such that each of these four vertices is adjacent at least with one of the three other vertices in the quadruple, then each quadruple of vertices contains an isolated vertex.

If the graph contains the only edge, the conclusion is obvious. So that we suppose that there are at least two edges, say a and b. If they are not adjacent, the quadruple of their vertices contradicts our assumptions, thus, a and b have a common vertex v_0. Similar reasoning shows that, except for a and b, the graph can have at most one edge c, and this edge must have a common vertex with the edges a and b. Now, c cannot be incident to v_0, since that would imply $\deg(v_0) = 3$. Therefore, the edges a, b, and c make a triangle, and all the other vertices are isolated. □

Exercises 2.2.

Exercise 2.2.1. Find all subgraphs of the graph G, given by incidence function (2.1.1), or which is the same, by the diagram g in Fig. 2.1. Are there isomorphic graphs among them?

Exercise 2.2.2. Are any two of the three diagrams in Fig. 2.8 isomorphic?

Exercise 2.2.3. Find all connected components of the graphs in Fig. 2.3, 2.7 and Example 2.2.3.

Exercise 2.2.4. Does there exist a connected graph with 5 vertices, such that each of its edges is a bridge, and moreover, removal of every edge generates exactly two connected components?

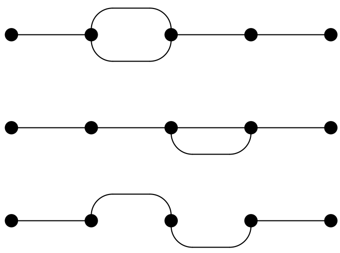


Figure 2.8: Diagrams to Exercise 2.2.2.

Exercise 2.2.5. List all walks of length 3 or less in the graphs in Fig. 2.7 and Example 2.2.3. Classify them as trails, paths, circuits, or cycles.

Exercise 2.2.6. Is it true or false that a graph of size at least 1 has walks of any finite length?

Exercise 2.2.7. Prove that the connectivity is an equivalence relation on the set V of the vertices of a graph. What are its equivalence classes?

Exercise 2.2.8.
(1) How many cycles are in the complete graphs K_4, K_5, K_n?
(2) How many paths of length 3 are in K_5?

Exercise 2.2.9. Prove that a graph $G = (V, E)$ of order p and size q satisfying the inequality $q + 2 \leq p$ cannot be connected, that is, $cc(G) \geq 2$.

Exercise 2.2.10. Prove that the size of a connected graph of order p is at least $p - 1$.

Exercise 2.2.11. Find the adjacency matrices of the graphs in Figs. 2.1, 2.3, and 2.7.

Exercise 2.2.12.
(1) How many walks of length 2 connect any two vertices in the complete graph K_n?
(2) In graph $K_{3,3}$?

Exercise 2.2.13. For the incidence matrix of the graph G (Fig. 2.1) compute $\sum_{i=1}^{5}(\sum_{j=1}^{5} a_{i,j})$ and $\sum_{j=1}^{5}(\sum_{i=1}^{5} a_{i,j})$.

Exercise 2.2.14. Prove that the vertices v_i and v_j in a graph G of order p are connected if and only if the (i, j)th-element of the matrix

$$A(G) + A^2(G) + \cdots + A^{p-1}(G)$$

is not zero.

Exercise 2.2.15. Prove that the vertices of a bipartite graph can be renumbered so that its adjacency matrix is a 2×2 block matrix $\begin{pmatrix} 0 & A_1 \\ A_2 & 0 \end{pmatrix}$, where A_1, A_2 are square matrices and both zero blocks represent rectangular matrices with all zero entries. What are the sizes of the blocks A_1 and A_2?

Exercise 2.2.16. Prove that if two graphs have the same adjacency matrix, they are isomorphic. Is the converse statement true?

Exercise 2.2.17. Prove that two graphs $G = (V, E)$ and $G' = (V', E')$ with the adjacency matrices $(a_{i,j})$ and $(a'_{i,j})$ are isomorphic if and only if they have the same order p and there exists a p-permutation σ such that $a_{i,j} = a'_{\sigma(i),\sigma(j)}$.

Exercise 2.2.18. Each edge of the complete graph K_n is colored in one of $n - 1$ colors, such that for every vertex all its incident edges have different colors. For what n is it possible?

Exercise 2.2.19. Each student club on a campus has a two-color flag. For these flags, the college bought fabrics of 8 different colors. It is known that every color meets on the flags with at least four other colors. Prove that the college can select no more than four clubs whose flags represent all the 8 colors.

Exercise 2.2.20. To connect every pair among 25 cities by a direct flight, it is necessary to have $C(25, 2) = 300$ flights. Suppose now that 25 cities are connected by only 277 direct flights with at most one direct flight between any two cities. Prove that if transfers are allowed, then any city can be reached by air from any other city with at most one transfer. Prove that if a graph with v vertices has at least $C(v - 1, 2) + 1$ edges, then the graph is connected. This bound is very simple and crude—how far can you improve this estimate?

Exercise 2.2.21. Prove that if a graph of order p has a directed path of length at least p, then this path must pass through some vertex at least twice, therefore, the graph contains a circuit.

2.3 Trees

In this section we study properties of a special kind of graphs important in many applications.

Coffee-time browsing

- http://www.neatorama.com/2007/03/21/10-most-magnificent-trees-in-the-world/ (That's the Tree!)
- http://en.wikipedia.org/wiki/Joseph_Kruskal (Joseph Kruskal's biography)
- http://www.princeton.edu/main/news/archive/S16/79/27A47/ (Another Kruskal)
- http://en.wikipedia.org/wiki/Kruskal's_algorithm (Kruskal's algorithm)
- http://dictionary.reference.com/browse/algorithm (Algorithms)
- everything2.com/title/Arthur+Cayley (Cayley's biography)
- www.gap-system.org/~history/Biographies/Prufer.html (Prufer's biography)

Definition 2.3.1. A graph without cycles, or *acyclic graph*, is called a *forest*.
A connected forest is called a *tree* (Fig. 2.9).

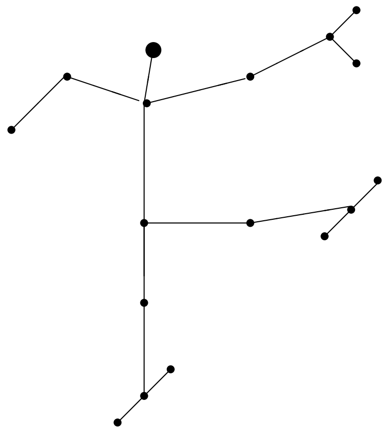


Figure 2.9: An example of a tree.

A *rooted* tree is a tree, which has a singled out vertex, called the root.

Thus, a forest is a family of trees and a tree is a connected graph without cycles.
The following theorem lists several important equivalent properties of the trees.

Theorem 2.3.1. *Let $G = (V, E)$ be a finite graph of order $|V| = p$. Then the following statements are equivalent.*

(1) *G is a tree.*
(2) *G is a connected graph and each of its edges is a bridge.*
(3) *G is an acyclic graph and its size is $|E| = p - 1$.*
(4) *G is a connected graph and $|E| = p - 1$.*
(5) *For any pair of vertices of G there is the unique path connecting them.*
(6) *G is acyclic but any new edge added to G generates precisely one cycle.*

Proof. We establish the following chain of implications,

$$(1) \Rightarrow (2); \quad (1) \,\&\, (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (6) \Rightarrow (1).$$

First we prove that (1) implies (2). Since a tree is connected by definition, it suffices to prove that every edge is a bridge. On the contrary, if we assume that some edge $e \in E$ is not a bridge, then we can remove it and still get a connected graph G'. Hence the end vertices of e are connected in G' by a walk, which clearly does not contain e. Now, the addition of e to the latter walk would generate a cycle in G, which is impossible since G is a tree and cannot contain cycles.

Next we prove that (1) and (2) together imply (3). Indeed, if we remove any edge from a cycle, the remaining graph is still connected, implying that G must be acyclic. To prove that G has $p-1$ edges, we apply mathematical induction on $p = |V|$. If $p = 1$, this single vertex must be isolated and the assertion is trivial. Suppose the assertion holds true for all trees of order less than some $p > 1$ and consider a graph $G = (V, E)$ of order p. Let $e \in E$, then e is a bridge by the assumption. Thus, if we remove e and denote the remaining graph by G', then, by Lemma 2.2.3, $cc(G') = cc(G) + 1 = 2$. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be two connected components of G'. They cannot be empty, they cannot have cycles, and by the inductive assumption, $|E_1| = |V_1| - 1$, $|E_2| = |V_2| - 1$. Adding up these equations and considering that $E = E_1 \cup E_2 \cup \{e'\}$, $V = V_1 \cup V_2$, where both unions are disjoint, we arrive at the conclusion.

To prove that (3) implies (4), we again assume that, on the contrary, G is not connected, that is, it consists of $k \geq 2$ connected components. Each of these components is a tree and by the assumption for each of them $|E'| = |V'| - 1$. Adding up $k \geq 2$ such equations leads to $|E| = |V| - k = p - k < p - 1$, which contradicts the premise.

The implication (4) $\Rightarrow$ (5) follows readily if we notice that if there are two vertices connected by two different paths, then these paths together make up a cycle. Removing any edge e off this cycle, we get a *connected* graph $G'(V, E')$, where $E' = E \setminus \{e\}$, such that $|V| = p$ and $|E'| = p - 2$ in contradiction with Lemma 2.2.4.

Next we prove the implication (5) $\Rightarrow$ (6). If we can add an edge and generate two cycles, this would mean that the end vertices of the new edge were connected by two paths in the original graph, which is impossible.

Finally, to prove that (6) $\Rightarrow$ (1), we have to prove that the graph is connected, which is obvious; indeed, a new cycle connects any two of its vertices twice, thus, one connection must have existed before we added the edge. □

We study some other properties of trees. Compare the graphs G_1 (Fig. 2.7) and $T = (\{v_1, v_2, v_3, v_4\}, \{e_1, e_2, e_5\}, f_T)$ (Fig. 2.10).

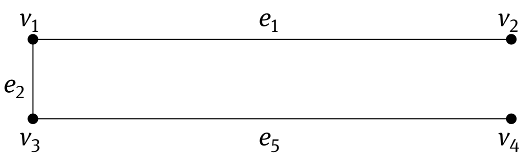


Figure 2.10: Graph T.

Problem 2.3.1. Write down explicitly the incidence function f_T of the graph T in Fig. 2.10.

Solution. From Fig. 2.10 we observe immediately that $f_T(e_1) = \{v_1, v_2\}$, $f_T(e_2) = \{v_1, v_3\}$, and $f_T(e_5) = \{v_3, v_4\}$. □

The graph T is a connected spanning subgraph of the graph g_1 in Fig. 2.7—it contains all the vertices of g_1 and some of its edges.

Definition 2.3.2. If a spanning graph of a graph G is a tree, this tree is called a *spanning tree* of G.

Thus, the tree T (Fig. 2.10) is a spanning tree of the graph g_1. A graph may have several spanning trees. The next statement follows immediately from Theorem 2.3.1.

Corollary 2.3.1. *Every graph has a spanning forest. Every connected graph has a spanning tree.* □

In many applications it is useful to supply edges of a graph with an additional piece of information, usually called the *weight* of this edge. The weight can be a number like the length of an edge, or a symbol like a traffic sign indicating whether this is a one-way or two-way street. If every edge of a graph carries a weight, the graph is called *weighted*. Weighted graphs have *weighted spanning trees*.

Example 2.3.1. Consider a connected weighted graph G_5 (Fig. 2.11), where the weights are $w_1 = 2$, $w_2 = 5$, $w_3 = 1$, $w_4 = 3$. This graph has three different spanning trees shown in Fig. 2.12. These trees have different weights, namely $W(T_1) = w_2 + w_3 + w_4 = 9$, $W(T_2) = w_1 + w_3 + w_4 = 6$, and $W(T_3) = w_1 + w_2 + w_4 = 10$; among them the tree T_2 has the smallest weight—it is the *minimum* spanning tree of the graph G_7.

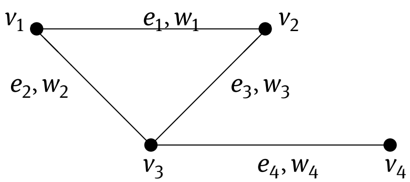


Figure 2.11: Graph G_7.

There are several algorithms for finding a minimum spanning tree in a graph. We present the well-known algorithm of Kruskal. Connectedness of a graph is, obviously, a necessary condition for the existence of a spanning tree.

Kruskal's algorithm for finding a minimum spanning tree

Given a connected weighted graph $G = G(V, E)$ with n vertices, find a minimum spanning tree in G. We assume that all weights are nonnegative numbers.

1. Select an edge e with the smallest weight. If the graph has several edges with the same minimum weight, we can choose any of them. The edge e and its end vertices form the initial subgraph (subtree) T_1 of G.

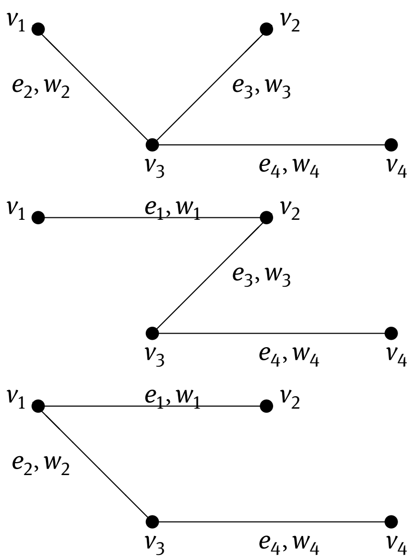


Figure 2.12: Spanning trees T_1, T_2, T_3.

2. For $m = 1, 2, \ldots, n-2$, select an unused edge with the smallest weight, such that this edge does not make a cycle with the edges selected earlier. In particular, we can use an edge with the same weight as the one in the previous step. Append the edge chosen and, if necessary, its end vertices to the subgraph T_m generated at the previous step, to built the next subgraph T_{m+1}.
3. Repeat Step 2 $n-2$ times. The subtree T_{n-1}, where n is the order of the given graph G, is a minimum spanning tree of G. □

Remark 2.3.1. Not every graph among $T_2, \ldots, T_{n-2}$ is to be a tree, some of them can be forests, but T_{n-1} is a tree.

Problem 2.3.2. Prove that Kruskal's algorithm generates a minimum spanning tree in any connected graph.

Problem 2.3.3. Find a minimum spanning tree for the complete graph K_8, where the weights of the edges are given in the following symmetric Table 2.1.

Solution. In this problem, the vertices are denoted by c_i, c_j, etc.; the (i,j)-entry of the table is the weight, $w_{i,j} = w_{j,i}$, of the edge incident to the vertices c_i and c_j. The reader may notice that the weights are all the integer numbers from 1 through 28 inclusive.

The following figures exhibit all the consecutive steps of Kruskal's algorithm being applied to Problem 2.3.3. The smallest weight is $w_{3,4} = 1$, thus we start with the graph with 8 isolated vertices $c_1, \ldots, c_8$ (Fig. 2.13) and first connect the vertices c_3 and c_4 by

Table 2.1: The weights of the complete graph K_8 in Problem 2.3.3.

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8
c_1	0	5	10	7	22	27	25	13
c_2	5	0	8	12	28	23	17	6
c_3	10	8	0	1	9	19	3	26
c_4	7	12	1	0	4	14	2	21
c_5	22	28	9	4	0	11	16	18
c_6	27	23	19	14	11	0	15	20
c_7	25	17	3	2	16	15	0	24
c_8	13	6	26	21	18	20	24	0

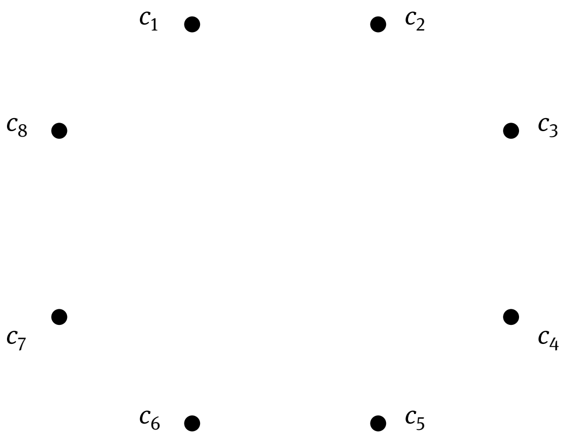


Figure 2.13: The initial graph without edges. All vertices are isolated.

an edge of weight 1 (Fig. 2.14). The second smallest weight is $d_{4,7} = 2$. Adding an edge of weight 2 connecting the vertices c_4 and c_7, we get the graph shown in Fig. 2.15.

The next smallest weight is $d_{3,7} = 3$. However, we cannot connect the vertices c_3 and c_7, because such an edge would form a cycle with the two previously included edges (Fig. 2.16), which is forbidden by Part 2 of the algorithm.

Thus, we look for the next smallest weight, $w_{4,5} = 4$, and at the next step we connect the vertices c_4 and c_5 by the edge of weight 4 (Fig. 2.17). Figures 2.18–2.21 show the sequel subgraphs leading to a minimum spanning tree of weight 20 (Fig. 2.21). Keep in mind, that not all of these graphs are trees, for example, Fig. 2.19–2.20. □

Problem 2.3.4. Find a minimum spanning tree in graph G_8 (Fig. 2.22).

In the end of this section we again consider a problem of graph enumeration and prove Cayley's formula on the enumeration of labeled trees.

Theorem 2.3.2. *There are p^{p-2} non-isomorphic labeled trees with $p \geq 1$ vertices.*

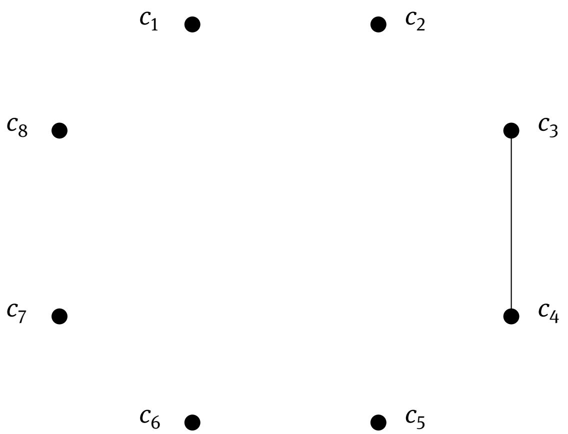


Figure 2.14: First step of Kruskal's algorithm. The first (non-spanning) subgraph with only one edge is formed.

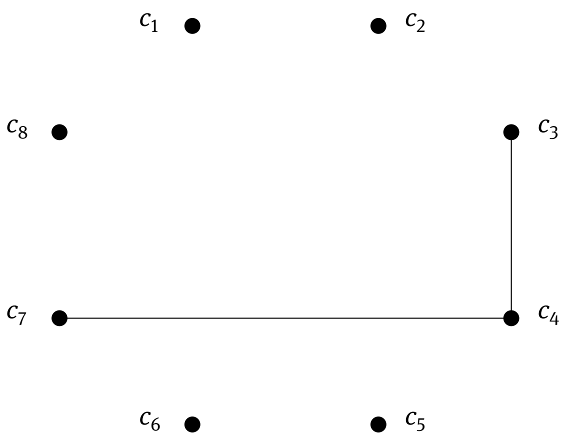


Figure 2.15: A subtree with two edges.

Proof. If $p = 1$, then $p^{p-2} = 1$, and the statement is obvious, since the unique labeled tree with one vertex is this isolated vertex labeled by 1. Now, let $p \geq 2$ and T be a labeled tree of order p. Delete the end vertex with the smallest label, record the label of the adjacent vertex, and repeat this step until only two vertices remain. This procedure generates a sequence of $p - 2$ natural numbers ranging from 1 through p with possible repetitions. The sequence is called the *Prüfer code* of the tree T. By Theorem 1.1.6, there are p^{p-2} such sequences, and, since there is an obvious one-to-one correspondence between such codes and the labeled trees of order p, the proof is complete. □

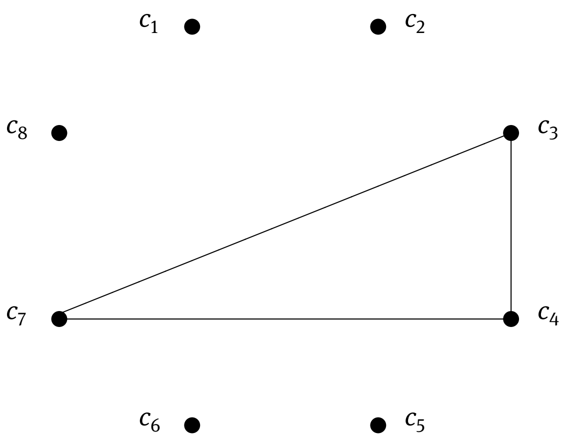


Figure 2.16: This subgraph with three edges is not a tree.

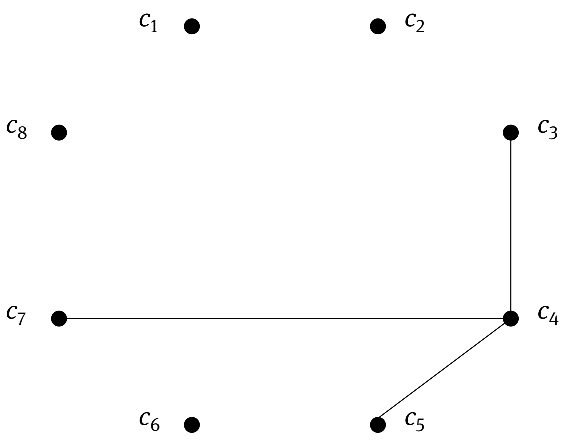


Figure 2.17: The subgraph with three edges.

Example 2.3.2. For instance, if $p = 3$, then $p^{p-2} = 3$. All non-isomorphic labeled trees with 3 vertices are shown in Fig. 2.5.

Corollary 2.3.2. *Let* $1 \le d_1 \le d_2 \le \cdots \le d_p$ *be the degree sequence of a tree of order p. The number of labeled trees of order p with this degree sequence is given by the multinomial coefficient* (1.5.1),

$$C(p-2; d_1-1, d_2-1, \ldots, d_p-1) = \frac{(p-2)!}{(d_1-1)!\cdots(d_p-1)!}.$$

Proof. Indeed, $\sum_{i=1}^{p} d_i = 2p - 2$ by Lemma 2.1.2. If v_i is a pendant vertex then $d_i = 1$, and it is clear from the proof that this label does not appear in the Prüfer code at all. If $d_i \ge 2$, then together with the removal of this vertex we must remove $d_i - 1$ adjacent

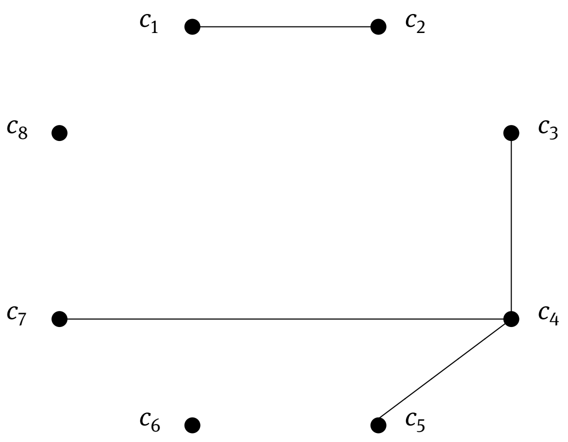


Figure 2.18: This subgraph is not a tree, since it is not connected.

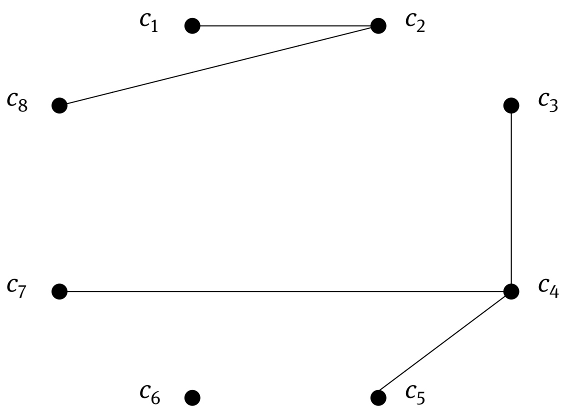


Figure 2.19: This subgraph with 5 edges also is a forest, not a tree.

vertices, hence v_i appears in the Prüfer code $d_i - 1$ times, and the same is true for any other vertex, which proves the corollary. □

Problem 2.3.5. Label the tree in Fig. 2.9 and compute its Prüfer code. Repeat this with another labeling of the same tree and compare their Prüfer codes.

Problem 2.3.6. Restore a labeled tree if its Prüfer code is 133132.

Solution. We do not distinguish vertices and their labels. From the proof of Theorem 2.3.2, we see that, since the length of the code is 6, the tree must have $6 + 2 = 8$ vertices. The vertices (labels) 1, 2, and 3 are present in the code, thus, they were not removed at the first deletion step, hence the very first vertex removed was 4, and this vertex was connected to 1. The next smallest vertex was 5 and it was connected to 3. The next vertex, which is 6, was connected to 3 again. The vertex removed after that

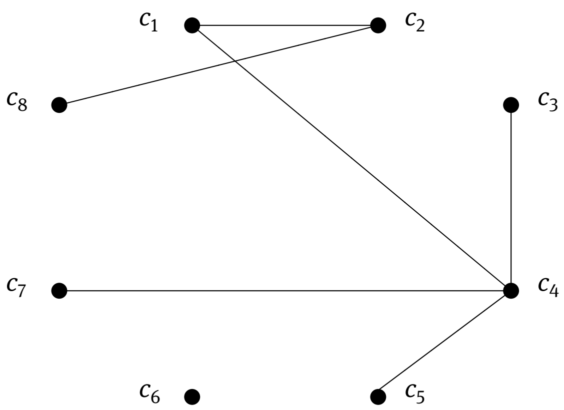


Figure 2.20: Second to last step of the algorithm. Two subtrees merge into a tree with 6 edges. This subtree is not a spanning tree yet, since the vertex c_6 is still isolated.

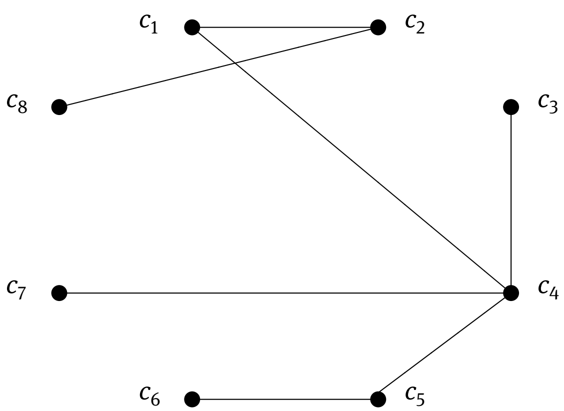


Figure 2.21: The minimum spanning tree of the initial graph in Problem 2.3.3, its weight is $w(T) = 36$.

was 7, and it must have been connected to 1. The vertex 1 does not appear in the code after that; thus now it is the smallest and we remove it, keeping in mind that it is adjacent to 3. At this stage only three vertices, 2, 3, and 8 remain, but we cannot remove 2 now, hence, we have to remove 3, which is, apparently, adjacent with 2. Thus, vertices 8 and 2 are connected—see Fig. 2.23. □

Exercises 2.3.

Exercise 2.3.1. Draw all non-isomorphic trees with 5 vertices and those with 6 vertices.

Exercise 2.3.2. What graphs coincide with their spanning trees?

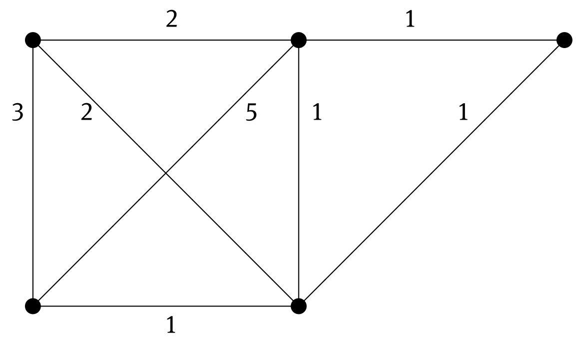


Figure 2.22: Graph G_8.

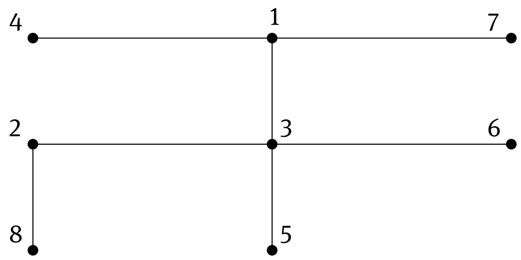


Figure 2.23: This tree has the Prüfer code 133132—Problem 2.3.6.

Exercise 2.3.3. Draw a diagram having only one spanning tree.

Exercise 2.3.4. Forty-one points in the plane are connected by straight segments, such that any two points are connected by either a segment or a broken line, and for any two points this broken line is unique. Prove that there are precisely 40 segments connecting the points.

Exercise 2.3.5. Find all spanning trees of the graph G_1 (Fig. 2.7) and those of the tree T (Fig. 2.10).

Exercise 2.3.6. How many non-isomorphic spanning trees does the bipartite graph $K_{3,3}$ have?

Exercise 2.3.7. There are n towns connected by highways without intersections, such that from each town a driver can reach every other town, and there is the only route between any two towns. Prove that the number of highways is $n - 1$.

Exercise 2.3.8. How many are there non-isomorphic trees with n vertices if the degree of any vertex is no more than 2?

Exercise 2.3.9. Prove that in any simple finite graph $G = (V, E, f)$

$$2q \leq (p - cc(G))(p - cc(G) + 1).$$

Exercise 2.3.10. Draw the labeled trees with the Prüfer codes 234, 3123, 4444, 7485553. Is there a labeled tree with the Prüfer code 126?

Exercise 2.3.11. Prove that there is a one-to-one correspondence between the non-isomorphic labeled trees and Prüfer codes—see Theorem 2.3.2.

Exercise 2.3.12. A tree has p vertices. What is the largest possible number of its pendant vertices?

Exercise 2.3.13. Prove that in any tree of order $p \geq 2$ there are at least two pendant vertices. Moreover, a stronger statement holds true—any acyclic graph of order $p \geq 2$ has at least two pendant vertices.

Exercise 2.3.14. Prove that a graph is a forest if and only if for any two distinct vertices there is at most one path connecting them.

Exercise 2.3.15. Generalize Theorem 2.3.1 to forests: If a forest of t trees has v vertices and d edges, then $v = d - t$.

Exercise 2.3.16. A forest has 67 vertices and 35 edges. How many connected components does it have?

Exercise 2.3.17 (Cayley's second formula). For $1 \leq k \leq n$, prove that there are $k(n+k-1)^{n-2}$ labeled forests with $n + k - 1$ vertices and k connected components, such that k distinguished vertices belong to different connected components.

Exercise 2.3.18. Prove that $1 \leq d_1 \leq d_2 \leq \cdots \leq d_p$ is the degree sequence of a tree of order p if and only if $\sum_{i=1}^{p} d_i = 2p - 2$.

Exercise 2.3.19. Prove that every sequence of integer numbers

$$1 \leq d_1 \leq d_2 \leq \cdots \leq d_p,$$

such that $\sum_{i=1}^{p} d_i = 2p - 2k$, $k \geq 1$, is the degree sequence of a forest with k connected components.

Exercise 2.3.20. Prove that $F(p)$, the number of forests of order p, satisfies the recurrence relation

$$F(n) = \sum_{k=1}^{n} C(n-1, k-1) k^{k-2} F(n-k).$$

Exercise 2.3.21. Theorem 2.3.1 claims that a tree of order p has $p - 1$ edges. For non-acyclic graphs this conclusion clearly fails. Nevertheless, prove that a connected graph of order p must have at least $p - 1$ edges.

Exercise 2.3.22. How many edges are to be removed from a connected graph with 12 vertices and 15 edges to generate a spanning tree of the graph? Does this number depend upon the order in which the edges are being removed?

Exercise 2.3.23. There are 300 cities in a state and 3 000 highways connecting them, such that each city is connected with at least one other city. How many of the highways can simultaneously be closed for repair if no city should be completely isolated from the others?

Exercise 2.3.24. Show that a graph is connected if and only if it has a spanning subtree.

2.4 Eulerian graphs

This section is concerned with edge traversal problems.

Coffee-time browsing

- www-history.mcs.st-and.ac.uk/Biographies/Euler.html (Euler's biography)
- http://en.wikipedia.org/wiki/Seven_Bridges_of_K%C3%B6nigsberg (The Seven Bridges of Königsberg)
- http://mathforum.org/kb/message.jspa?messageID=3648262&tstart=135 (Fleury's algorithm)
- home.att.net/~numericana/answer/graphs.htm (Who is Fleury?)
- http://www.absoluteastronomy.com/topics/Eulerian_path (Eulerian graphs)
- www.gap-system.org/~history/Mathematicians/Hamilton.html (Hamilton's biography)

The next problem should remind the reader of an old puzzle.

Problem 2.4.1. Can you draw either of the two graphs in Fig. 2.24 without traversing an edge twice and without interruption the drawing (that is, your pencil must not leave the paper)?

Definition 2.4.1. A circuit (a trail) in a graph is called *Eulerian* if it contains all edges of the graph. A graph is called Eulerian if it contains an Eulerian circuit. A graph is called *semi-Eulerian* if it contains an Eulerian trail.

The results of this section essentially depend on the *parity* of the vertex degrees of a graph, that is, whether the degree is even or odd. We call a vertex even (odd) if its degree is even (odd).

Problem 2.4.2. Is there a graph with just one odd vertex?

Theorem 2.4.1. *A connected graph is Eulerian if and only if it has only even vertices. A connected graph is semi-Eulerian if and only if it contains exactly two odd vertices.*

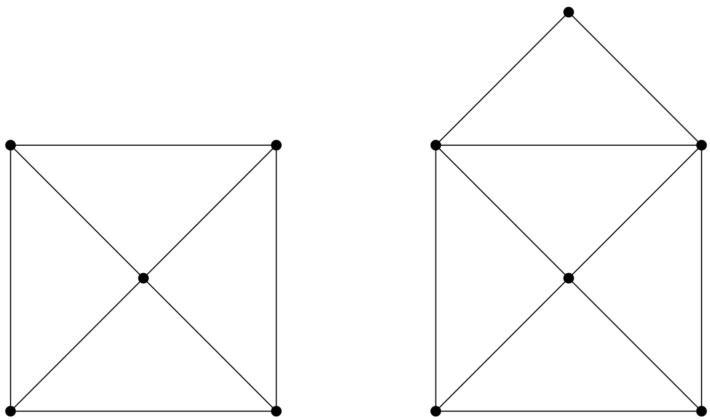


Figure 2.24: Is any of these "envelopes" Eulerian? Semi-Eulerian?

Problem 2.4.3. Is any of graphs in Fig. 2.24 Eulerian? Is any of them semi-Eulerian?

Proof of Theorem 2.4.1. The necessity of these conditions, including the connectedness, is obvious. Indeed, if we begin to traverse an Eulerian circuit and remove every edge traversed, after passing through any vertex its degree decreases by 2, so that the parity of any vertex's degree does not change. After completing the route we must arrive at the initial vertex after traversing and removing behind us every edge. Thus, the degree of each vertex gradually reduces to zero, implying that initially the degree was even. In the case of semi-Eulerian graphs the same argument works if we begin at either one of the two odd vertices; we must finally arrive at the other odd vertex.

We prove the sufficiency by induction on the size q of the graph. Begin again with the Eulerian case. If $q = 1$, the graph consists of one vertex with an attached loop, so the statement is obvious. Suppose now that, for all connected graphs of the size $|E| < q$, $q \geq 2$, with all even vertices the statement is valid and consider a connected graph $G = (V, E)$, $|E| = q$. By Lemma 2.2.2, the graph G contains a cycle C. If this cycle includes all the edges of G, there is nothing more to prove. Otherwise, we remove all edges of the cycle C from G, which can result in decomposing G into several connected components, $G_1, \dots, G_l$, of smaller sizes.

Since G is connected, every component G_i contains some vertex $v_i \in C$, whose degree in G must be at least 4, so that its degree in G_i is at least 2. By the inductive assumption, each of $G_i, 1 \leq i \leq l$, has an Eulerian circuit C_i, and we conclude that $v_i \in C_i$ as well. It is now obvious how to assemble all cycles $C, C_1, \dots, C_l$ in an Eulerian cycle in the graph G.

To consider the case of semi-Eulerian graphs, we connect the two existing odd vertices by an additional edge, thus making the graph Eulerian, and apply the statement we have just proved. □

Exercises 2.4.

Exercise 2.4.1. Examine graphs in Figs. 2.1–2.23—which of them are Eulerian? Semi-Eulerian? Find, if any, semi-Eulerian trails or Eulerian circuits in these graphs.

Exercise 2.4.2. In the proof of Theorem 2.4.1 we claim that each component G_i has a vertex such that its degree in G is at least 4. Prove this claim.

Exercise 2.4.3. Draw the floor plans of the buildings on your campus, where you are (were) taking your classes. Draw a graph representing each room with a vertex, such that two vertices are connected by an edge if the corresponding rooms have a common wall. Are these graphs Eulerian? Semi-Eulerian? Find, if there are any, semi-Eulerian trails or Eulerian cycles in these graphs.

Exercise 2.4.4. Prove that the following procedure, called *Fleury's algorithm*, returns an Eulerian circuit in any Eulerian graph:

Start at any vertex and pass any edge incident to this vertex. Remove the edge passed and go through any other edge incident to the vertex reached, subject to the only restriction: a bridge can be used only if there is no other edge available.

Exercise 2.4.5. Apply Fleury's algorithm to those graphs in Figs. 2.1–2.24, which are Eulerian or semi-Eulerian, and find semi-Eulerian trails or Eulerian circuits in those graphs.

To consider vertex traversal problems, we introduce Hamiltonian graphs.

Definition 2.4.2. A path (circuit) without repeating vertices in a graph G is called *Hamiltonian* if it contains every vertex of G. A graph is called Hamiltonian if it has a Hamiltonian path.

Exercise 2.4.6. Prove the following *necessary* condition of the existence of a Hamiltonian circuit: If a graph G contains a Hamiltonian circuit, then it contains a connected spanning subgraph H, which has the equal order and size, and the degree of every vertex of H is 2.

In the opposite direction, prove the following *sufficient* condition: If in a simple graph $G = (V, E)$ of order $p \geq 3$,

$$\deg(v) \geq p/2, \quad \forall v \in V,$$

then G has a Hamiltonian circuit.

Exercise 2.4.7. Find a Hamiltonian circuit in the complete graph K_5.

Exercise 2.4.8. Is any of the graphs in Fig. 2.1–2.24 Hamiltonian?

Exercise 2.4.9. Extend Definition 2.4.1 and Theorem 2.4.1 to digraphs. Is the digraph $\overline{G}$ (Fig. 2.2) Eulerian or semi-Eulerian? Find, if there are any, the corresponding directed paths.

2.5 Planarity

We prove in this section a remarkable property of planar graphs, also related to the name of Euler, called *Euler's formula* or *Euler's polyhedron theorem*.

Coffee-time browsing
- http://thales.cica.es/rd/Recursos/rd99/ed99-0289-02/biografias/cjordan.html (Jordan's biography)
- www.gap-system.org/~history/.../Kuratowski.html (Kuratowski's biography)

Definition 2.5.1. A regular embedding of a planar graph in $\mathbf{R}^2$ is called a *plane graph*. Therefore, the edges of a plane diagram cannot have common points, except maybe for the vertices of the graph.

For example, the diagram g' (Fig. 2.3) is not plane, since the edges ϵ_2 and ϵ_6 intersect at a point, which is not a vertex of the graph, but g' is planar because it can easily be redrawn without this intersection.

Consider a connected plane graph consisting of two vertices connected by an edge. If we choose two arbitrary points in the plane outside of this graph, it is possible to connect them with a smooth curve having no common point with the graph. Consider now a connected graph with cycles, say g_1 (Fig. 2.7). The edges of g_1 split the entire plane into four separate regions, among them there is one unbounded external domain, and three bounded—one interior to the loop e_6, one between the parallel edges e_2 and e_3, and the last one bounded by the edges e_1, e_3, e_4, e_5. These regions are called *faces* of the graph g_1. Any two points inside a face can be connected by a smooth curve lying completely inside the face, and it is (almost) obvious[3] that if two points lie in different faces, they cannot be connected by a smooth plane curve unless this curve intersects at least one edge of the graph.

Theorem 2.5.1 (Euler). *Let G be a connected plane graph of order p and size q. If G has f faces, including an unbounded face, then*

$$p - q + f = 2. \tag{2.5.1}$$

The expression $p - q + f$ is called the Eulerian characteristic of a graph.

Proof. We carry out mathematical induction on the number of faces f. If the graph G has at least one bounded face, then the boundary of the latter consists of edges of G, and these edges form a cycle—cf. Figure 2.7. Thus, if $f = 1$, then this face must be unbounded, therefore G is a tree, hence $p = q + 1$ by Theorem 2.3.1 and the conclusion follows straightforwardly as $p - q + f = q + 1 - q + 1 = 2$.

3 This statement is a deep topological theorem by C. Jordan.

Now fix $f \geq 2$, assume that (2.5.1) is valid for all plane connected graphs with less than f faces, and consider a graph G with f faces. At least one of these faces is bounded, thus the boundary of this face is a cycle, so that the edges making up this cycle, are not bridges. Moreover, by virtue of planarity such an edge must belong to the boundaries of two faces. The deletion of any such edge generates a plane connected graph G' of the same order p, of size $q-1$, and with $f-1$ faces. Applying the inductive assumption to the graph G', we get the equation $p-(q-1)+(f-1)=2$, which immediately reduces to (2.5.1). □

For example, if G is a tree with p vertices, then $q = p-1$, $f = 1$, thus $p-(p-1)+1 = 2$, in agreement with (2.5.1).

Corollary 2.5.1. *In a connected simple plane graph G of order $p \geq 3$ and size q,*

$$q \leq 3p - 6. \tag{2.5.2}$$

Proof. Since G is simple, it cannot have parallel edges, thus every face is bounded by at least three edges, and every edge adjoins two faces, hence $3f \leq 2q$. Combining this inequality with (2.5.1) we deduce (2.5.2). □

Problem 2.5.1. Prove that any cycle in a bipartite graph has an even length, that is, it consists of an even number of edges.

Corollary 2.5.2. *The complete graph K_5 and the complete bipartite graph (the Thomsen graph) $K_{3,3}$ shown in Fig. 2.25, are not planar.*

Figure 2.25: Graphs K_5 and $K_{3,3}$.

Proof. If the graph K_5 were planar, (2.5.2) would immediately lead to contradiction. As for $K_{3,3}$, every cycle in it has an even length by Problem 2.5.1, thus its length is at least 4, therefore the inequality $3f \leq 2q$ of the preceding corollary can be strengthened to $4f \leq 2q$, which together with (2.5.1) and (2.5.2) gives $q \leq 2p - 4$ for any bipartite graph. In the case of $K_{3,3}$, $p = 6$, $q = 9$, and the latter becomes $9 \leq 8$. This contradiction shows that $K_{3,3}$ cannot be planar. □

It turns out that all the non-planarity in our world is due to these two simple non-planar graphs, as follows from the following result. We state it without proof, which can be found, for example, in [9, p. 24].

Theorem 2.5.2. (Pontryagin–Kuratowski[4]) *A graph G is planar if and only if it does not contain a subgraph that can be derived from $K_{3,3}$ or K_5 by subdividing some of their edges by inserting additional vertices.* □

Exercises 2.5.

Exercise 2.5.1. Is there a planar graph with 6 vertices, each of them of degree 3? Of degree 4? Of degree 5?

Exercise 2.5.2. A diagram with 9 vertices and 9 edges is embedded into a plane. Each vertex has the degree of 3 and every edge is incident to 3 vertices. Draw the diagram.

Exercise 2.5.3. A connected planar graph of order 6 has 5 vertices of degree 3 and a vertex of degree 1. In how many regions does the graph divide the plane? Draw the corresponding diagram.

Exercise 2.5.4. Five neighboring cities decided to build highways connecting each of them with every other city. Is it possible to build this road network without intersections or over- and underpasses?

Exercise 2.5.5. Give an example of a connected but not complete graph.

Exercise 2.5.6. Draw complete bipartite graphs $K_{1,1} - K_{3,4}$.

Exercise 2.5.7. Prove that any tree of order $p \geq 2$ is bipartite.

Exercise 2.5.8. The converse of the preceding problem is false—not every bipartite graph is a tree. Find an example of such a bipartite graph. For which m and n the complete bipartite graphs $K_{m,n}$ are trees?

Exercise 2.5.9. Prove that a simple graph G is bipartite if and only if it does not contain an odd circuit. Moreover, G is bipartite if and only if it does not contain an odd cycle.

Exercise 2.5.10. Let δ be the least vertex degree in a simple graph of order p. Prove that if $\delta \geq (p-1)/2$, then the graph is connected. Compare this conclusion with Corollary 2.5.1.

Exercise 2.5.11. Is any of graphs in Fig. 2.26 planar?

---

4 According to some sources, this theorem was proved by Pontryagin a few years before Kuratowski, but that proof was not published.

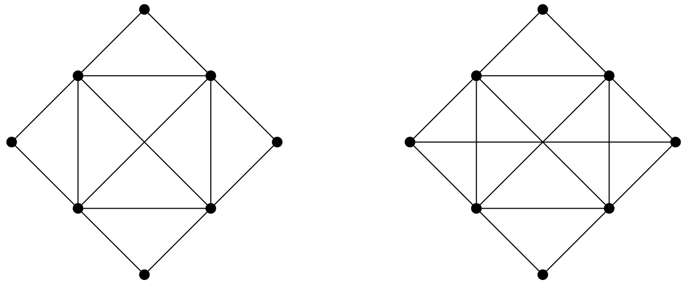


Figure 2.26: Is either of these graphs planar?

Exercise 2.5.12. Prove that if several lines in a plane divide the plane in p parts, then $t - l + p = 1$, where t is the number of intersection points of these lines, and l is the number of pieces the lines are split by the points of intersection.

Exercise 2.5.13. Give examples of graphs whose Eulerian characteristic is not 2.

2.6 Graph coloring

This section is concerned with vertex and edge coloring of simple graphs, and some of its applications.

Coffee-time browsing

– http://www-history.mcs.st-and.ac.uk/Biographies/Euler.html (Euler's biography)

A graph is a pair of sets, the edges E and the vertices $V \neq \emptyset$, and the incidence function. Nevertheless, there are many problems, where we want to add some more information to edges and/or vertices. For instance, in the next chapter about the graph clustering we supply the edges with weights, called their dissimilarities. In this chapter an additional piece of information, added to a vertex, is called a *color* of the vertex. It is possible to color the edges as well, but in this brief introduction to graph coloring we do not consider that option.

Definition 2.6.1. Consider a graph $G(V, E, f)$ of order $|V| = k$, a finite set $\mathrm{Col} = \{c_1, \ldots, c_k\}$, whose elements are called colors, or any abstract symbols, and a map $\mathrm{col} : V \to \mathrm{Col}$, called the chromatic function of the colored graph $G = (V, E, f, \mathrm{col})$, satisfying the following condition: If two vertices v_i and v_j are adjacent, i. e., connected with an edge, then $\mathrm{col}(v_i) \neq \mathrm{col}(v_j)$.

For instance, different states or counties on maps are usually painted in different colors to easier distinguish them, and a natural question occurs immediately, whether we have enough colors to paint a map or a globe appropriately? The latter usually means that any two neighboring domains must be painted differently. This question immediately translates into a graph problem. Indeed, given a map, we design the graph, such that the every vertex corresponds to a domain on the map, and two vertices are adjacent iff the corresponding domains have a common piece of a boundary, which does not degenerate to a point.

Thus, can we color all the vertices of a graph, so that any two adjacent vertices have different colors? Surely, it is possible if there are enough paints, and the problem becomes: what is the minimal number of paints to be used?

Definition 2.6.2. The smallest number of the required different colors (recall that any two adjacent vertices are painted differently), is called the chromatic number $\chi(G)$ of the graph G. Thus, for any $k \geq \chi(G)$ the graph is k-colorable, i. e., it can be colored in k colors.

Problem 2.6.1. Find the chromatic numbers of the graphs in Fig. 2.24 and 2.26.

Problem 2.6.2. Compute the chromatic number of a bipartite non-planar graph, $\chi(K_{3,3})$, Fig. 2.29, and of any other bipartite graph.

Problem 2.6.3. Prove that a graph is bipartite iff it is 2-colorable, and iff it contains a cycle of even length, i. e., containing an even number of edges.

Example 2.6.1. Consider, for instance, the graph G_7; see Fig. 2.11. We reproduce it in Fig. 2.27, and let the vertex marks mean different colors. Since the three vertices v_1, v_2, v_3 are pairwise adjacent, they cannot be colored in only two colors, three colors are necessary for a triangle, and any two of these colors, except for v_3, can be used for v_4. Thus, $\chi(G_7) = 3$.

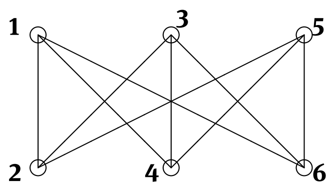


Figure 2.27: The bipartite graph $K_{3,3}$.

Problem 2.6.4. Find the chromatic number of the following spanning subgraph G_{71} of the graph G_7.

Since G_{71} does not contain any triangle, but contains adjacent vertices, we need 2 colors, $\chi(G_{71}) = 2$. □

Problem 2.6.5. Prove that, for any graph $G = (V, E, f)|, \chi(G) \leq |V|$.

Problem 2.6.6. For what graphs G, the chromatic numbers are $\chi(G) = 1$? $\chi(G) = 2$?

Problem 2.6.7. Compute the chromatic numbers $\chi(T)$ for the trees and for the forests.

Problem 2.6.8. Let K_n be the complete graph with n vertices. Prove that $\chi(K_n, \lambda) = \lambda(\lambda - 1)(\lambda - 2)\cdots(\lambda - n + 1)$. We recall that the vertices of a graph are distinguishable, hence when we permute the colors, we get another coloring of a graph.

We have shown above that the chromatic numbers $\chi(G_{71})$ and $\chi(G_7)$ are $\chi(G_{71}) = 2$ and $\chi(G_7) = 3$, respectively. Moreover, we see that if a graph has at least two adjacent vertices, its chromatic number is at least 2. A natural question is for what graphs the chromatic number is 3? Of course, at the very least, for that to be valid, a graph must contain a loop with 3 edges. And indeed, we can simplify the graph G_7 in Fig. 2.30, which has $\chi(G) = 3$, since if any two of its three vertices are painted, the remaining vertex cannot have the same color. We have proved here half of the next problem.

Problem 2.6.9. A graph G has the chromatic number $\chi(G) = 1$ iff the graph has no edge, only any number of isolated vertices. □

The graphs above have the chromatic numbers 1, 2 and 3. Can we continue, i. e., can we design a graph, such that its chromatic number is any positive integer? To begin with, we construct a graph with $\chi(G) = 4$; we will need this example later on.

Problem 2.6.10. Design a graph with $\chi(G) = 4$.

Solution. Analyzing the triangle in Fig. 2.30, we can try a cycle with 4 edges; see Fig. 2.31. However, we immediately see that the chromatic number of a square is 2.

To prevent the "diagonal-opposite" vertices of a square to have the same color, we connect them with a diagonal; see Fig. 2.32. The intersection of the diagonals is not a vertex of the graph.

It is now clear how to proceed. One of the possible solutions is the complete graph K_5, Fig. 2.25. □

Problem 2.6.11. The complete graph K_5 has 5 vertices, 5 edges and $\chi(G) = 5$. Does there exist a graph with the same chromatic number, but with smaller order or size?

Problem 2.6.12. What is the chromatic number of the complete graph K_n, i. e., the graph with n vertices, such that every two of them are connected with an edge?

However, the complete graph K_n with $n \geq 5$ is not planar. To compute the chromatic numbers of arbitrary graphs, the next formula, which immediately follows from the inclusion–exclusion theorem, sometimes can be useful.

Definition 2.6.3. The number of various paintings of a graph G into λ distinct colors is denoted as $\pi(G, \lambda)$. Thus, $\chi(G) = \min\{\pi(G, \lambda) : \lambda \geq 1\}$.

Theorem 2.6.1. *Let G be a graph of order $p = |V|$ and of size $q = |E|$. Fix a certain numbering $(e_1, e_2, \dots, e_q)$ of the edges of G. Then*

$$\pi(G,\lambda) = \lambda^p - \sum_i \mu(e_i) + \sum_{i\neq j} \mu(e_i, e_j) - \sum_{i,j,k} \mu(e_i, e_j, e_k) + \cdots$$

where every sum is taken over all the different pairs, triples, etc., of indices, and $\mu(e_i, e_j, \dots, e_l)$ are the numbers of the colorings of the graph G, such that the end vertices of every of the edges $e_i, e_j, \dots, e_l$ are painted the same color (each for every edge, but may be different for different edges). □

The formula of Theorem 2.6.1 can be transformed if we introduce the quantity $p(s, c)$ to be the number of spanning subgraphs of G of size s, i. e., with s edges and with c connected components.[5]

Theorem 2.6.2.

$$\pi(G,\lambda) = \sum_{s,c} (-1)^s p(s,c)\lambda^c.$$

Proof. Consider k edges $e_{i_1}, \dots, e_{i_k}$ and a spanning subgraph G_1 of G, containing these and only these k edges; let it have c connected components. If we paint any connected component of G_1 in a certain color, then the edges $e_{i_1}, \dots, e_{i_k}$ have the same painted end points, and vice versa, for any painting the ends of the edges $e_{i_1}, \dots, e_{i_k}$ are pairwise equally colored, then every connected component of G is colored in one color. If the number of colorings is c and the color λ is fixed, then there are λ^c paintings, and the generic term of formula in Theorem 2.6.1 is

$$\sum_{i,j,k} \mu(e_i, e_j, e_k) = \sum_c p(k,c)\lambda^c,$$

where the sum is taken over all the non-zero $p(k, c)$, and this formula implies the theorem directly. □

To apply Theorem 2.6.1, let us note that in that theorem the number c is a natural number, thus, the function $\pi(G, \lambda)$ is a polynomial in λ. Moreover, every graph has at least one connected component, thus, $p(s, 0) = 0$, therefore, every term of this polynomial contains a power of λ, so that λ is a divisor of this polynomial. Next, if $s = 0$, then, for any c, $p(0, c) = 1$, thus, this polynomial starts with $\lambda^c_{\max}$, where $c_{\max}$ is the largest possible value of the c.

Problem 2.6.13. Compute polynomials [$\pi(G, \lambda)$ and the chromatic number $\chi(G)$ for the triangle G in Fig. 2.28.

5 We follow here [6, Chap. 6].

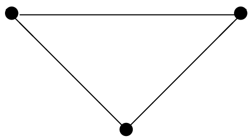


Figure 2.28: This graph has the chromatic number $\chi(G) = 3$.

Solution. For the triangle G, the appropriate subgraphs are either a graph with three isolated vertices without edges, thus, $s = 0$ and $c = 3$, or three subgraphs with one edge and one isolated vertex, thus, $s = 1$ and $c = 2$, or else three subgraphs with two adjacent edges and one isolated vertex, thus, $s = 2$ and $c = 0$. Thus, non-zero coefficients are $p(0,3) = 1$, $p(1,2) = 3$, $p(2,1) = 3$, $p(3,1) = 1$, and finally, $\pi(G,\lambda) = \lambda^3 - 3\lambda^2 + 3\lambda - \lambda$, which can be factorized as $\pi(G,\lambda) = \lambda(\lambda-1)(\lambda-2)$. It follows from the latter equation that the chromatic number of the triangle is $\chi(G) = 3$, since for smaller natural numbers, $\pi(G,1) = \pi(G,2) = 0$. □

Problem 2.6.14. Compute the polynomials $\pi(G,\lambda)$ and the chromatic numbers $\chi(G)$ for the graph in Fig. 2.29, for the tree G_{71} in Fig. 2.30, and for the square, Fig. 2.31.

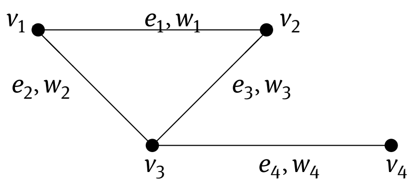


Figure 2.29: Graph G_7.

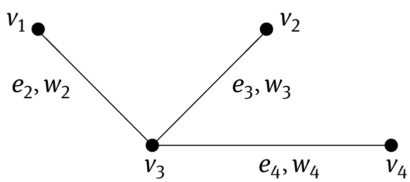


Figure 2.30: Graph G_{71}.

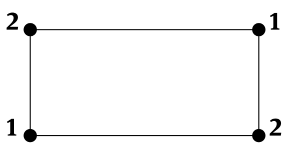


Figure 2.31: This graph has $\chi(G) = 2$. It is 2, and also 3 and 4 colorable.

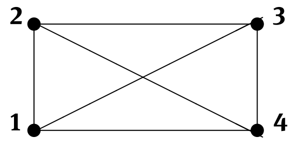


Figure 2.32: This graph has $\chi(G) = 4$.

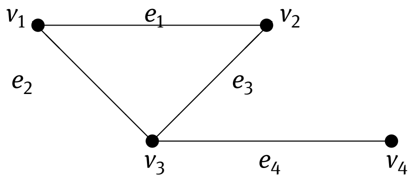


Figure 2.33: Graph G.

Thus, in general the chromatic numbers are unlimited – see, for instance, the examples in Fig. 2.31–2.33, i. e., for any natural number n there exists a graph, namely, the complete graph K_n, such that its chromatic number $\chi(K_n) = n$. However, it may not be the case for certain special classes of graphs. For instance, the famous problem of the graph theory, the problem of four colors states: "The chromatic number of a connected planar graph does not exceed 4." This has been a hypothesis for about a century, until it was proved by Kenneth Appel and Wolfgang Haken in 1976. Their proof started a vivid discussion about the validity of computer-supported proofs, because even after essential further simplification, the proof involves the massive computer item-by-item examination of graphs. Without entering that discussion, we end this chapter by proving less precise result, which can be done without computers. We prove the five-colors claim, but first consider a simpler problem.

Problem 2.6.15. Prove that any planar graph can be colored in at most 6 colors.

These statements deal with planar graphs. Moreover, all the following graphs in this chapter are assumed to be connected and to have no parallel edges nor loops. Indeed, our proofs can be carried through any connected component of a planar graph. Also, if a planar graph has several edges, i. e., edges with the same end points, the presence of the parallel edges or loops is irrelevant for the painting of the end points.

To prove the main result of this section, we need the following lemma.

Lemma 2.6.1. *If G is a planar graph with p vertices, then it has at least one vertex v such that its degree* $\deg(v) \leq 5$.

Proof. Consider any connected component of G, and assume that every vertex of that component has the degree at least 6. Then, by Corollary 2.5.1, $q \leq 3p - 6$, but by the Handshaking Lemma 2.1.2, $2q = 6p$, or $q = 3p$, implying the contradiction $3q \leq 3q - 6$. Hence, $\deg(v) \leq 5$. □

Theorem 2.6.3. *A planar graph is 5-colorable, i. e., its vertices can be painted in at most 5 colors, so that any two adjacent vertices have different colors.*

Proof. We do induction over the vertices over a connected graph (connected component) without parallel edges or loops. The base of induction is clear, since any planar graph with $p \le 5$ is colorable in at most 5 colors. Consider a graph G with $p+1$ vertices and remove a vertex v, such that by the previous Lemma 2.6.1, $\deg(v) \le 5$. Removing the vertex v, we conclude that the graph $G - v$ is 5-colorable. There are two options. If the vertices adjacent to v are colored in at most 4 colors, say, the color c_0 is not used, we use this color c_0 to paint v, and the theorem is proven.

Assume now that the vertex v is adjacent to exactly 5 vertices v_1, v_2, v_3, v_4, v_5, painted in 5 different colors c_1, c_2, c_3, c_4, c_5, and the edges e_1, e_2, e_3, e_4, e_5, connecting the initial vertex v with these 5 vertices, make no triangle. Identifying these two edges (which do not form a triangle) and their second vertices, different from v, we have a new graph of smaller order, and properly painted. Thus, we accomplish a step of induction in this case either. Finally, assume that the vertex v is again adjacent to exactly 5 vertices v_1, v_2, v_3, v_4, v_5, painted in 5 different colors c_1, c_2, c_3, c_4, c_5, however, the vertices v_1, v_2, v_3, v_4, v_5 make a convex pentagon, containing the vertex v inside. Due to planarity, the complete graph K_5 cannot occur here, thus, among the five vertices v_1, v_2, v_3, v_4, v_5, there are two non-adjacent vertices. To finish the proof, we delete the vortex v together with the incident edges, and glue (identify) these two non-adjacent vertices. Let us note that the new graph has less than p edges, hence, at most p vertices. The end of the proof in this case is also clear. □

Exercises 2.6.

Exercise 2.6.1. Prove that $\chi(G) \le |V|$.

Exercise 2.6.2. Prove that $\chi(K_p, \lambda) = \lambda(\lambda-1)(\lambda-2)\cdots(\lambda-p+1)$.

Exercise 2.6.3. Prove that if v and u are non-adjacent vertices of a graph G, then $\chi(G,\lambda) = \chi(G_{y-z},\lambda) + \chi(G_{y=z},\lambda)$, where in $\chi(G_{y-z},\lambda)$ an edge from y to z is added, and in $G_{y=z}$ the vertices y and z are identified.

Exercise 2.6.4. Prove that the chromatic function of a graph of degree p is a polynomial of degree p with integer coefficients.

Exercise 2.6.5. Prove that the complete graph K_n is a subgraph of a graph G, then $|\chi(K_n,\lambda)$ is a divisor of the chromatic polynomial $|\chi(G,\lambda)$.

3 Hierarchical clustering and dendrogram graphs

Given a set of objects, the cluster analysis aims at splitting this set into separate groups according to a certain prescribed measure of the proximity of the given objects. In this chapter we apply the graph theory to develop simple clustering algorithms. These algorithms essentially use the notion of a spanning tree, which was developed in Section 2.3.

3.1 Introduction

In this section we introduce basic terminology of hierarchical cluster analysis.

Coffee-time browsing

– home.dei.polimi.it/matteucc/Clustering/tutorial_html/ (Clustering)

A student wants to put some money in mutual funds. To make the right choice, she considers many different funds analyzing their characteristics such as long-term and short-term performance, the manager's philosophy of investing, administrative costs, and other features. Comparing various funds, she can pick up a few funds that look suitable for her goals. The things under consideration, like mutual funds, are called *objects* or *entities*. Certain properties of objects, like performance or attitude to risk, are called *features*, or *variables*, or *attributes*. However, if every object is characterized by several variables, it is difficult to compare different objects, and we want to have a kind of a "common denominator" to be able to measure the *similarity* of the objects.

We can separate all mutual funds under consideration into several groups containing similar funds. Such a classification is useful in many occasions. For instance, if the investor learns on a new fund within a short time after its inception, it is difficult, without any information, to make a prediction of the fund's future performance based on its own short history. However, if the student can include the fund in a group of similar funds, she can apply the information on the whole group to each element of the group and make more reliable predictions. Furthermore, if we have many similar objects, it is often just impossible to study every one of them separately, but we can study a representative of each group of similar objects and apply the information found to every item.[1]

To perform such analysis, we must first separate the objects into smaller groups, called *clusters* (overlapping groups are sometimes called *clumps*). This process, called *clustering*, is an essential part of the *cluster analysis*. In this chapter we discuss some basic concepts and algorithms of this subject. For more on the cluster analysis the reader can consult, for example, [14, 17, 21, 28, 32, 33, 41].

1 This is similar to partitioning a set into the disjoint equivalence classes and studying the representatives of these classes instead of the entire original set.

https://doi.org/10.1515/9783110751185-003

Obviously, the objects combined in a group should be similar, should have some common features. In the cluster analysis, however, it is often more convenient to measure the dissimilarity rather than the similarity of various objects. The more two objects have in common, the less is their *dissimilarity*. Ultimately, the similarity of identical objects is infinite and their dissimilarity is zero. We do not discuss here how to assign the (dis)similarity values to multivariate objects, because it essentially depends upon particularities of a specific problem. We assume that the dissimilarities are assigned in advance—given a set of objects to be explored, we are provided with a *table* of their dissimilarities, called the *dissimilarity table* or *dissimilarity matrix*.

Definition 3.1.1. A square symmetric matrix (table) with non-negative elements, whose main diagonal contains only zeros, is called a *dissimilarity matrix (table)*.

Example 3.1.1. Table 3.1 contains the average altitudes above the sea level of 15 states in the U. S. If we are interested in the altitudes only, we can consider the absolute values of the differences between the altitudes as a measure of the dissimilarity between two states. Even though this difference is not the real geographical distance, similar quantities, subject to certain conditions, are called in mathematics *distances* or *metrics*. In this sense, the dissimilarity between Alabama and Delaware is $50 - 6 = 44$, the dissimilarity between Florida and Georgia is $|10 - 60| = 50$, and the dissimilarity between Florida and Louisiana is 0—unlike the mathematical metric, the dissimilarity of two different objects can be 0. Table 3.1 can be transformed into the dissimilarity table, Table 3.2, where the main diagonal contains only zeros since each object is absolutely similar to itself. We have completed only the upper triangle, because the table is symmetrical with respect to the main diagonal.

Table 3.1: The rounded average altitudes above the sea level of 15 southern states in the U. S. (in tens of feet).

AL	DE	FL	GA	KY	LA	MD	MO	MS	NC	SC	TN	TX	VA	WV
50	6	10	60	75	10	35	80	30	70	35	90	170	95	150

We can construct different clusterings depending upon the level of dissimilarity we are willing to accept—this level is called a *threshold value* or just a *threshold*. That is, we can form different partitions (clusterings) of the 15 states. For instance, the following is a partition of these states into ten clusters with a threshold value of 5, that is, the maximum distance between any two objects in each of the following clusters does not exceed 5:

$$\{\mathrm{DE},\mathrm{FL},\mathrm{LA}\},\ \{\mathrm{MD},\mathrm{MS},\mathrm{SC}\},\ \{\mathrm{AL}\},\ \{\mathrm{GA}\},$$
$$\{\mathrm{NC}\},\ \{\mathrm{KY},\mathrm{MO}\},\ \{\mathrm{TN}\},\ \{\mathrm{VA}\},\ \{\mathrm{TX}\},\ \{\mathrm{WV}\}.$$

Table 3.2: Dissimilarity table for the average altitudes.

	AL	DE	FL	GA	KY	LA	MD	MO	MS	NC	SC	TN	TX	VA	WV
AL	0	44	40	10	25	40	15	30	20	20	15	40	120	45	100
DE		0	4	54	69	4	29	74	24	64	29	84	164	89	144
FL			0	50	65	0	25	70	20	60	25	80	169	85	140
GA				0	15	50	25	20	30	10	25	30	110	35	90
KY					0	65	40	5	45	5	40	15	95	20	75
LA						0	25	70	20	60	25	80	160	85	140
MD							0	45	5	35	0	55	135	60	115
MI								0	50	40	5	60	140	65	120
MO									0	10	45	10	90	15	70
NC										0	35	20	100	25	80
SC											0	55	135	60	115
TN												0	80	5	60
TX													0	75	20
VA														0	55
WV															0

We can also set up another clustering with the same dissimilarity level of 5 but now with nine clusters,

$$\{\text{DE}, \text{FL}, \text{LA}\},\ \{\text{MD}, \text{MS}, \text{SC}\},\ \{\text{AL}\},\ \{\text{GA}\},$$
$$\{\text{NC}, \text{KY}\},\ \{\text{MO}\},\ \{\text{TN}, \text{VA}\},\ \{\text{TX}\},\ \{\text{WV}\}.$$

We see that in general this procedure is not unique. If we select bigger threshold level of 10, then the corresponding clustering may be the following one, containing just eight clusters,

$$\{\text{DE}, \text{FL}, \text{LA}\},\ \{\text{MD}, \text{MS}, \text{SC}\},\ \{\text{AL}, \text{GA}\},$$
$$\{\text{NC}, \text{KY}, \text{MO}\},\ \{\text{TN}\},\ \{\text{VA}\},\ \{\text{TX}\},\ \{\text{WV}\}.$$

It is clear also and we see that in the example above that, if we increase the threshold value, some clusters may merge (amalgamate) into bigger ones.

Compare these clusterings. While deriving the second clustering, we had to relocate some objects, and the group {TN, VA} of the second clustering does not belong completely to any cluster in the third clustering. On the other hand, every cluster of the first clustering is contained completely in some cluster of the third one. A process that makes a series of consecutive clusters such that every cluster of the preceding level is a subset of a cluster on the next level, is called the *hierarchical clustering*. We begin with the *completely disjoint* clustering, where every object forms its own single-element cluster. Then step by step, we merge (*amalgamate*) two or more clusters with the smallest dissimilarity into larger ones, until we reach a threshold value. Such algorithms are called *agglomerative*.

We can also proceed in the opposite direction. Namely, we can depart from a *conjoint clustering*, when the initial cluster contains all the objects under consideration, and split it repeatedly into smaller groups, until we reach either the threshold value or the completely disjoint clustering. Such algorithms are called *divisive*.

Problems, involving classification of real data, cannot be reduced only to applying a mathematical clustering algorithm. Before that, the data must be collected and consistently presented, and the dissimilarity values must be assigned. After building the clusters, they must be *validated* and *assessed*. The results have to be properly *interpreted*. All these are crucial issues, because any algorithm generates some clustering, but without further considerations we cannot conclude whether the clusters derived reflect the real structure of data or this is just an artifact of the algorithm. We leave out all these issues along with the problem of computer implementation.

In this chapter we consider only agglomerative hierarchical algorithms for clustering discrete sets of data. These algorithms are based on the properties of the graph representing the initial collection of objects. In Section 3.2, using a small model example, we develop a simple single-link hierarchical clustering algorithm. It is called *single-link*, because at every step we connect two existing clusters by a single edge (link) of the underlying graph. Section 3.3 is devoted to Hubert's single-link algorithm. We discuss relations of the single-link hierarchical clustering algorithm with minimum spanning trees. Section 3.4 is devoted to another hierarchical clustering algorithm—Hubert's complete-link algorithm. In Section 3.5 we apply the single-link algorithm to a more realistic problem. We also validate the clustering derived in this example, by making use of Pearson's coefficient of correlation, thus demonstrating the quality of the clustering developed.

Exercises 3.1.

Exercise 3.1.1. Construct other clusterings of these 15 states (Table 3.1) with the same threshold levels of 10 or 5.

Exercise 3.1.2. Find a clustering of these 15 states corresponding to the threshold value 1; corresponding to the threshold value of 164.

Exercise 3.1.3. Construct clusterings of these 15 states consisting of four or five clusters. Find corresponding threshold values.

Exercise 3.1.4. Find a dissimilarity level that generates the unique clustering in this problem.

3.2 Model example

Here we consider a simple but not simplistic model example to introduce a classical agglomerative algorithm of hierarchical clustering.

Coffee-time browsing

– http://www.plantbio.ohiou.edu/instruct/multivariate/Week7Lectures.PDF (Cluster Analysis)

There are eight cities, $c_1, c_2, \ldots, c_8$, in a region. It is necessary to connect all of them by highways. It is possible to build a highway connecting every pair of cities. In the graph theory terms, such a road network can be described as the complete graph K_8. This network contains $C(8,2) = 28$ highways and is rather expensive. On the other hand, it is possible to link every city with only one of the other cities, thus having a minimal number of roads built. Even though it is less expensive to construct, this network, which can be modeled by a spanning tree with 8 vertices and, therefore, with 7 edges, may be inconvenient for the commuters, who will have to waste their time and fuel, because many pairs of the cities do not have a direct connection.

Thus, a local mathematician has offered an intermediate approach, namely, to split all the cities into several groups-clusters. The cities within each cluster are to be connected completely, but any two different clusters should be connected by only one road. A cluster should, obviously, include the cities that are close to each other. *However, the closeness can be measured in many various ways*. A reasonable way to measure the closeness is to use the number of commuters between the cities.

The information about the average number of commuters in both directions between the cities, in thousands of people per day, is contained in Table 3.3. For instance, the amounts of commuters are 24 between the cities c_1 and c_2, 2 between c_1 and c_6 and 6 between c_2 and c_6. Thus, there is a large flow of commuters between c_1 and c_2—in this sense these two cities are close, even though geographically they may be located far away from one another. Thus, they are to be considered similar and should be placed in one cluster. Yet, c_6 is distant from them. However, if we use these quantities, 24, 2, 6, etc., as a measure of closeness (a generalized distance), then the similarity between nearby cities may be greater than the similarity between the distant ones. In this problem and, as we have already mentioned, in clustering analysis generally, it is often more suitable to use the dissimilarities of objects rather than their similarities. We can always convert the commuter data into dissimilarity values, for example, by taking inverses or subtracting from some maximum value.

We consider the total amount of commuters in both directions, so Table 3.3 is symmetrical with respect to the main diagonal and we filled in only the upper triangle of the table. Moreover, since we want to start with a simple example, all the entries are different (the table contains no *ties*) and they are all natural numbers from 1 to $C(8,2) = 28$. The mathematician must now solve the problem of combining the cities into clusters according to this dissimilarity matrix.

We first develop a simple intuitive algorithm, which starts with 1-element clusters and step-by-step combines them until we reach some goal, which should be set in advance. At the initial step, the algorithm treats each given object as a single-element

Table 3.3: The dissimilarity table for the model example.

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8
c_1	0	5	10	7	22	27	25	13
c_2		0	8	12	28	23	17	6
c_3			0	1	9	19	3	26
c_4				0	4	14	2	21
c_5					0	11	16	18
c_6						0	15	20
c_7							0	24
c_8								0

cluster. The set of these clusters is called the *clustering of level zero*. If we use the graph-theory language, we can depict this clustering as a graph having only isolated vertices, with no edge. Then, at every step, the algorithm uses only one edge to merge, *agglomerate* two closest (that is, with the smallest dissimilarity) clusters into a new cluster of the next level. Such an edge connecting two clusters of the same level into a cluster of the next level is called a *link*. That is why this and similar procedures are called *single-link algorithms* or *single linkage*.

We begin with a descriptive version of an agglomerative single-link clustering algorithm, then apply it to the dissimilarity table above and give a more formal treatment of the algorithm. In Section 3.3 we present a version of the algorithm known as Hubert's single-link algorithm [31]. This algorithm is based on the notion of spanning trees.

Denote the consecutive clusterings by boldface capital letters with one index, $\mathbf{C}_0$, $\mathbf{C}_1$, $\mathbf{C}_2$, and so forth. The italic capital letters with double indices, $\mathcal{C}_{k,l}$, denote clusters—the first index, k, indicates the level of clustering and the second index, l, stands for the number of this particular cluster in the clustering of the k^{th} level. Thus, $\mathcal{C}_{3,4}$ denotes the fourth cluster in the third-level clustering $\mathbf{C}_3$. Now we build clusterings for the model example. In our notations, $\{c_i, c_j\}$ is a pair (2-element set) comprising the ith city c_i and jth city c_j, and a number $d(c_i, c_j)$, or $d(i, j)$ for short, at the crossing of the ith row and jth column of the dissimilarity matrix, stands for the dissimilarity of these two cities; due to the symmetry, $d(i, j) = d(j, i)$. First, we rearrange all pairs of the cities in the ascending order of their dissimilarities—see Table 3.4.

At the initial step, we derive a *disjoint clustering*, such that each city forms its own cluster containing only one element. Thus, the clustering of level zero is

$$\mathbf{C}_0 = \{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}, \mathcal{C}_{0,3}, \mathcal{C}_{0,4}, \mathcal{C}_{0,5}, \mathcal{C}_{0,6}, \mathcal{C}_{0,7}, \mathcal{C}_{0,8}\}$$

where $\mathcal{C}_{0,i} = \{c_i\}, i = 1, 2, \ldots, 8$. Then at every step we must determine the dissimilarities between all the existing clusters, both new and old.

Table 3.4: Dissimilarity table, Table 3.3, for the model example rearranged in the ascending order of the dissimilarities.

Pair $\{c_i, c_j\}$	Dissimilarity $d(i,j) = S_0(i,j)$
$\{c_3, c_4\}$	$d(3,4) = 1$
$\{c_4, c_7\}$	$d(4,7) = 2$
$\{c_3, c_7\}$	$d(3,7) = 3$
$\{c_4, c_5\}$	$d(4,5) = 4$
$\{c_1, c_2\}$	$d(1,2) = 5$
$\{c_2, c_8\}$	$d(2,8) = 6$
$\{c_1, c_4\}$	$d(1,4) = 7$
$\{c_2, c_3\}$	$d(2,3) = 8$
$\{c_3, c_5\}$	$d(3,5) = 9$
$\{c_1, c_3\}$	$d(1,3) = 10$
$\{c_5, c_6\}$	$d(5,6) = 11$
$\{c_2, c_4\}$	$d(2,4) = 12$
$\{c_1, c_8\}$	$d(1,8) = 13$
$\{c_4, c_6\}$	$d(4,6) = 14$
$\{c_6, c_7\}$	$d(6,7) = 15$
$\{c_5, c_7\}$	$d(5,7) = 16$
$\{c_2, c_7\}$	$d(2,7) = 17$
$\{c_5, c_8\}$	$d(5,8) = 18$
$\{c_3, c_6\}$	$d(3,6) = 19$
$\{c_6, c_8\}$	$d(6,8) = 20$
$\{c_4, c_8\}$	$d(4,8) = 21$
$\{c_1, c_5\}$	$d(1,5) = 22$
$\{c_2, c_6\}$	$d(2,6) = 23$
$\{c_7, c_8\}$	$d(7,8) = 24$
$\{c_1, c_7\}$	$d(1,7) = 25$
$\{c_3, c_8\}$	$d(3,8) = 26$
$\{c_1, c_6\}$	$d(1,6) = 27$
$\{c_2, c_5\}$	$d(2,5) = 28$

Definition 3.2.1. The *dissimilarity* $\mathrm{diss}(\mathcal{C}_{0,i}, \mathcal{C}_{0,j})$ between two clusters of level zero is defined as the dissimilarity between the corresponding objects, that is,

$$\mathrm{diss}(\mathcal{C}_{0,i}, \mathcal{C}_{0,j}) = d(i,j).$$

It is helpful to visualize the process of clustering by drawing graphs of special kind, called threshold graphs.

Definition 3.2.2. Given a dissimilarity $n \times n$ matrix and a threshold value λ, the *threshold graph* $G(\lambda)$ is a simple weighted graph with n vertices corresponding to n entities under consideration, such that two vertices v_i and v_j are adjacent if and only if $d(i,j) \leq \lambda$. The weight of the edge $e_{i,j}$ connecting two vertices v_i and v_j is the given dissimilarity $d(i,j)$.

Thus if $\lambda = 0$, two vertices are adjacent in $G(0)$ if and only if their dissimilarity is zero; if there is no such a pair of vertices, $G(0)$ contains only n isolated vertices and no edge. On the other hand, if the threshold value λ is greater than or equal to the largest entry of the dissimilarity matrix, then the threshold graph is (isomorphic to) the complete graph K_n, and we denote it by $G(\infty)$.

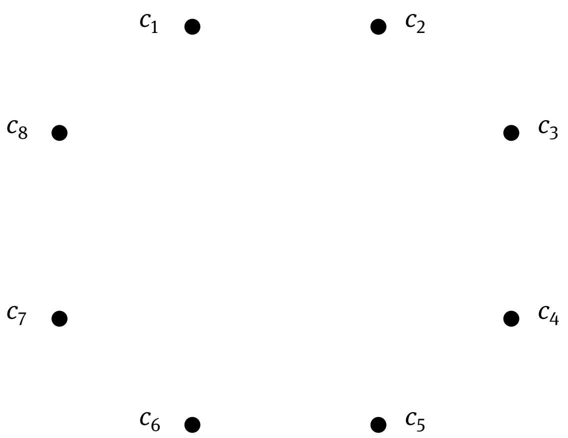


Figure 3.1: The threshold graph $G(0)$ for the model example. It corresponds to the $\mathbf{C}_0$-clustering.

The smallest dissimilarity in the problem is $d(3,4) = 1$. Thus, if the threshold value (an acceptable level of dissimilarity) is less than 1, we cannot combine any two cities in one cluster and have to stop here. In terms of our model example, this means that no cluster has an infrastructure, and we have to build a road between all the pairs of cities (Figure 3.2).

Suppose next that the threshold is at least 1, $\lambda \geq 1$. Then we have to consider all 28 pairwise unions

$$\begin{aligned}
&\mathcal{C}_{0,1} \cup \mathcal{C}_{0,2}, \mathcal{C}_{0,1} \cup \mathcal{C}_{0,3}, \ldots, \mathcal{C}_{0,1} \cup \mathcal{C}_{0,8}, \\
&\mathcal{C}_{0,2} \cup \mathcal{C}_{0,3}, \ldots, \mathcal{C}_{0,2} \cup \mathcal{C}_{0,8}, \ldots, \\
&\vdots \\
&\mathcal{C}_{0,7} \cup \mathcal{C}_{0,8}.
\end{aligned}$$

In the corresponding complete graph (the same Figure 3.2), its 28 edges, having the weights $d(1,2), \ldots, d(7,8)$ and connecting the pairs of vertices, respectively,

$$\{c_1, c_2\}, \{c_1, c_3\}, \ldots, \{c_1, c_8\}, \{c_2, c_3\}, \ldots, \{c_7, c_8\},$$

correspond to these 28 pairwise unions.

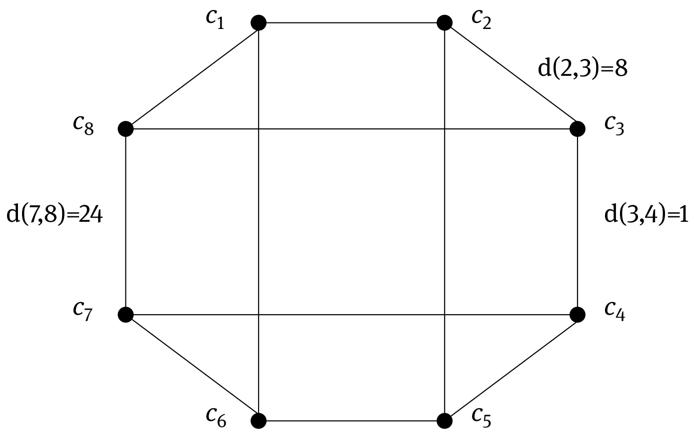


Figure 3.2: The complete threshold graph $G(\infty)$ for the model example; only a few edges and dissimilarities are shown.

Since the lowest weight is $d(c_3, c_4) = 1$, the clusters $\mathcal{C}_{0,3}$ and $\mathcal{C}_{0,4}$ have the smallest dissimilarity $\operatorname{diss}(\mathcal{C}_{0,3}, \mathcal{C}_{0,4}) = d(c_3, c_4) = 1$. In terms of our problem they have the largest flow of commuters between them. Thus, we have to connect them first, and we amalgamate these two clusters of level zero in the cluster $\mathcal{C}_{1,1}$ of the first level. All the other clusters of level zero automatically become clusters of the first level. This way, we get the first-level clustering

$$\mathbf{C}_1 = \{\mathcal{C}_{1,1}, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{1,4}, \mathcal{C}_{1,5}, \mathcal{C}_{1,6}, \mathcal{C}_{1,7}\}$$

consisting of one 2-element cluster $\mathcal{C}_{1,1} = \mathcal{C}_{0,3} \cup \mathcal{C}_{0,4} = \{c_3, c_4\}$ and six 1-element clusters $\mathcal{C}_{1,i} = \mathcal{C}_{0,i-1} = \{c_{i-1}\}$ for $i = 2, 3$ and $\mathcal{C}_{1,i} = \mathcal{C}_{0,i+1} = \{c_{i+1}\}$ for $i = 4, 5, 6, 7$. This clustering is shown in Fig. 3.3, where the connected component with the vertices $\{c_3, c_4\}$ corresponds to the cluster $\mathcal{C}_{1,1}$.

To complete this step of the algorithm, we must define the dissimilarities between new clusters. The dissimilarities between the clusters $\mathcal{C}_{1,2}, \dots, \mathcal{C}_{1,7}$ are the same as those between the corresponding “old” clusters of level zero. The dissimilarity between $\mathcal{C}_{1,1}$ and any cluster $\{c_i\}, i = 1, 2, 5, 6, 7, 8$, is, by Definition 3.2.1, the smaller of $d(c_3, c_i)$ and $d(c_4, c_i)$. For instance,

$$\operatorname{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,4}) = \min\{d(c_3, c_5); d(c_4, c_5)\} = \min\{9; 4\} = 4.$$

It should be repeated that while deriving $\mathbf{C}_1$ from $\mathbf{C}_0$, we have used only one link—the threshold graph corresponding to $\mathbf{C}_1$ contains just one more edge than the graph corresponding to $\mathbf{C}_0$. All the 28 pairs of vertices

$$\{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}\}, \dots, \{\mathcal{C}_{0,7}, \mathcal{C}_{0,8}\},$$

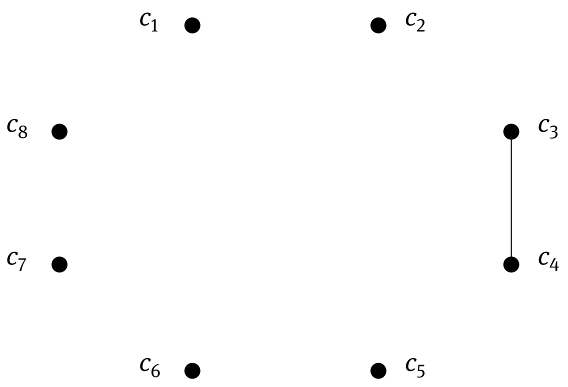


Figure 3.3: The threshold graph $G(1)$ corresponds to the threshold level $\lambda = 1$ and the clustering $\mathbf{C}_1$—only two vertices are connected.

each pair taken together with the incident edge, represent connected two-vertex subgraphs of the graph in Fig. 3.2—we have selected among them a subgraph with the minimal weight and linked two vertices of this subgraph into a cluster. Again in terms of our model example, this means that we have to build a road between c_3 and c_4.

Then we have to connect the other cities with either c_3 or c_4, but not with both, using roads between clusters. Given two clusters, $\mathcal{C}_{1,1} = \{c_3, c_4\}$ and $\mathcal{C}_{1,i} = \{c_{i-1}\}$, $i = 2, 3$, and $\mathcal{C}_{1,i} = \{c_{i+1}\}$, $i = 4, 5, 6, 7$, the decision regarding which city, c_3 or c_4, should be connected with c_i, is based on the dissimilarity between the clusters $\mathcal{C}_{1,1}$ and $\mathcal{C}_{1,i}$. For example, since

$$\operatorname{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,4}) = \min\{d(c_3, c_5); d(c_4, c_5)\} = d(c_4, c_5) = 4,$$

the cluster $\{c_5\}$ must be connected with the vertex c_4. The corresponding road map may look like the one in Fig. 3.4.

If the threshold $\lambda = 1$, we should stop here and the road map is given by the spanning tree in Fig. 3.4. However, if we can accept a larger threshold, we are to continue. To build a second-level clustering, we proceed in the same way. Namely, we consider all pairs of the first-level clusters and look for a connecting link with the smallest dissimilarity.

The edge $\{c_3, c_4\}$, which had been already utilized, cannot be used again. Among the unused edges, the smallest dissimilarity is $d(c_4, c_7) = 2$, and we form the second-level cluster $\mathcal{C}_{2,1}$ as the union of the two first-level clusters containing the cities c_4 and c_7. To make this cluster, we have again used a single link—the edge $\{c_4, c_7\}$. All the other first-level clusters move into the second-level clustering $\mathbf{C}_2$ unchanged, just being renumbered (Fig. 3.5):

$$\mathbf{C}_2 = \{\mathcal{C}_{2,1}, \mathcal{C}_{2,2}, \mathcal{C}_{2,3}, \mathcal{C}_{2,4}, \mathcal{C}_{2,5}, \mathcal{C}_{2,6}\}$$

where $\mathcal{C}_{2,1} = \mathcal{C}_{1,1} \cup \mathcal{C}_{1,6} = \{c_3, c_4, c_7\}$, $\mathcal{C}_{2,i} = \mathcal{C}_{1,i}$, $i = 2, 3, 4, 5$, and $\mathcal{C}_{2,6} = \mathcal{C}_{1,7}$.

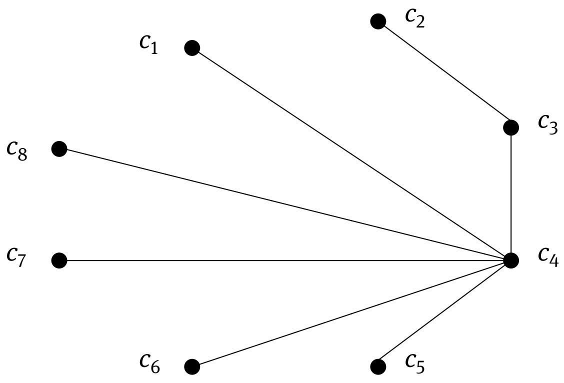


Figure 3.4: A road map corresponding to the first-level clustering $\mathbf{C}_1 = \{\mathcal{C}_{1,1}, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{1,4}, \mathcal{C}_{1,5}, \mathcal{C}_{1,6}, \mathcal{C}_{1,7}\}$.

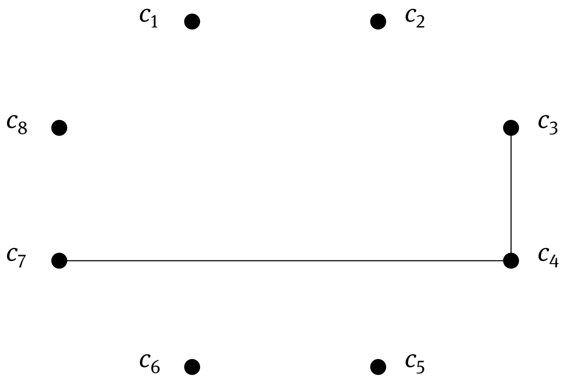


Figure 3.5: The threshold graph $G(2)$ corresponds to the second-level clustering $\mathbf{C}_2$—one more edge is added to $G(1)$; $\mathbf{C}_2$ consists of one 3-element cluster $\{c_3, c_4, c_7\}$ and five 1-element clusters $\{c_1\}$, $\{c_2\}$, $\{c_5\}$, $\{c_6\}$, and $\{c_8\}$.

We can express this in terms of connected subgraphs. In addition to the same two-vertex subgraphs with vertices other than c_3 and c_4, which were considered before, we have to look for connected subgraphs with three vertices. Namely, we consider the subgraphs, which contain the two vertices c_3 and c_4, the incident edge of these two vertices, another vertex, and an edge connecting the latter with either c_3 or c_4. The minimal dissimilarity is now $S_1(4,7) = 2$ and we have to connect clusters $\{c_7\}$ and $\{c_3, c_4\}$ in a cluster of the second level.

This clustering corresponds to the threshold value $\lambda = 2$ and is shown in Fig. 3.5. It is worth noting that the dissimilarity $d(3,7)$ between the objects c_3 and c_7 in the cluster $\mathcal{C}_{2,1}$ is greater than 2, but these vertices can be connected within the cluster by the edges $\{c_3, c_4\}$ and $\{c_4, c_7\}$, such that their weights do not exceed the threshold value. This is an important feature of the single-link algorithms—for any two objects x

and y in a cluster there always exists a sequence of objects in this cluster connecting x and y, such that the dissimilarity of any two neighbors in this sequence does not exceed the threshold value, even though the dissimilarity of x and y may be greater than the threshold.

We continue the construction of the hierarchical clustering for the model example. Suppose that we can accept a value of the threshold greater than 2. The next unused dissimilarity $d(c_3, c_7)$ gives nothing new, because the cities c_3 and c_7 have already been linked in a cluster. Therefore, $d(c_3, c_7)$ does not generate the next clustering (Fig. 3.6).

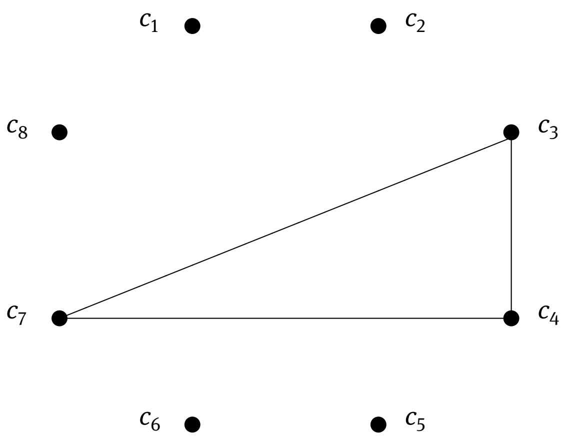


Figure 3.6: The threshold graph $G(3)$ does not generate a new clustering.

Thus, we skip $d(3, 7)$ and use the next bigger dissimilarity $d(4, 5) = 4$, generating the next clustering

$$\mathbf{C}_3 = \{\{c_3, c_4, c_5, c_7\}, \{c_1\}, \{c_2\}, \{c_6\}, \{c_8\}\}$$

which corresponds to the threshold value $\lambda = 4$. Five sets in $\mathbf{C}_3$ represent all five clusters of the third level (Fig. 3.7). Again, the dissimilarity between some vertices in the first cluster $C_{3,1}$ is greater than 4, but for any two vertices there exists a connecting path such that every edge in the path has a weight (dissimilarity) of 4 or less. In formal terms, we consider all the unions $C_{2,a} \cup C_{2,b}$ formed by a single edge and look for the link with the smallest weight, which generates a new cluster.

The next smallest weight to use is $d(1, 2) = 5$, and if we are willing to continue and use the value of the threshold $\lambda = 5$, we have to merge the cities $\{c_1\}$ and $\{c_2\}$ in a cluster. Thus, we derive the next clustering (Fig. 3.8):

$$\mathbf{C}_4 = \{C_{4,1}, C_{4,2}, C_{4,3}, C_{4,4}\} = \{\{c_1, c_2\}, \{c_3, c_4, c_5, c_7\}, \{c_6\}, \{c_8\}\}.$$

A road map corresponding to the clustering $\mathbf{C}_4$, is shown in Fig. 3.9.

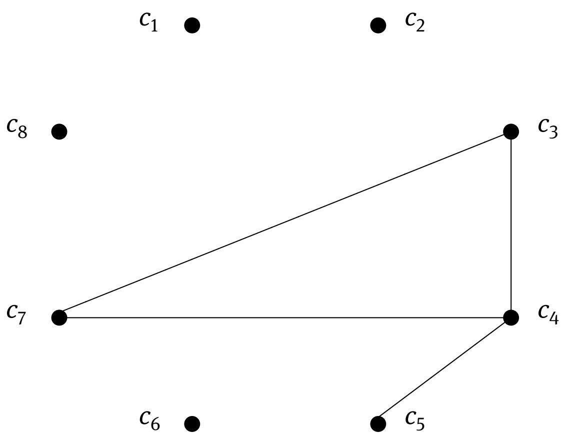


Figure 3.7: The threshold graph $G(4)$ contains one new edge $\{c_4, c_5\}$. It corresponds to $\mathbf{C}_3$-clustering containing one 4-element cluster $\{c_3, c_4, c_5, c_7\}$ and four 1-element ones $\{c_1\}, \{c_2\}, \{c_6\}, \{c_8\}$.

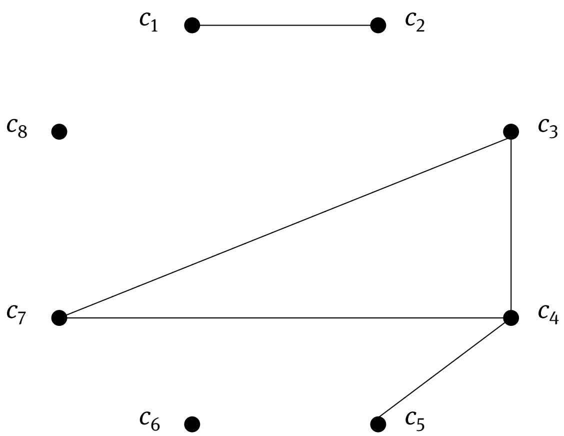


Figure 3.8: The threshold graph $G(5)$ generates the fourth-level clustering $\mathbf{C}_4$ consisting of one 4-element cluster $\{c_3, c_4, c_5, c_7\}$, one 2-element cluster $\{c_1, c_2\}$, and two 1-element clusters $\{c_6\}$ and $\{c_8\}$.

In this way we construct the hierarchy of consecutive clusterings, corresponding to increasing values of the threshold. Now it is the turn of $d(2,8) = 6$, and the fifth-level clustering is (Fig. 3.10)

$$\mathbf{C}_5 = \{\{c_1, c_2, c_8\}, \{c_3, c_4, c_5, c_7\}, \{c_6\}\}.$$

The next unused edge with the lowest weight is $\{c_1, c_4\}$ with $d(1,4) = 7$, and we come up with the clustering (Fig. 3.11)

$$\mathbf{C}_6 = \{\{c_1, c_2, c_3, c_4, c_5, c_7, c_8\}, \{c_6\}\}.$$

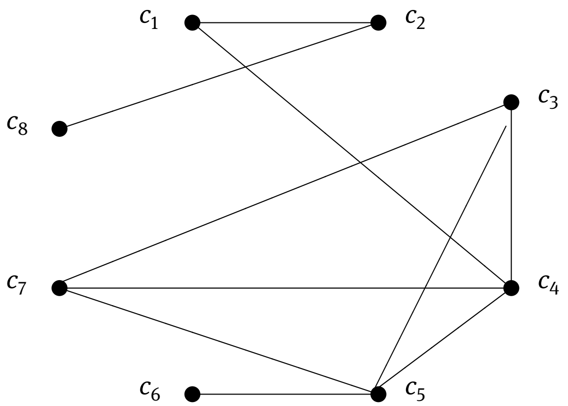


Figure 3.9: The road map corresponding to the clustering $\mathbf{C}_4$. The clusters $\mathcal{C}_{4,1}$ and $\mathcal{C}_{4,2}$ are connected by the edge $\{c_1, c_4\}$, for this edge has the smallest dissimilarity among all the edges connecting the two clusters.

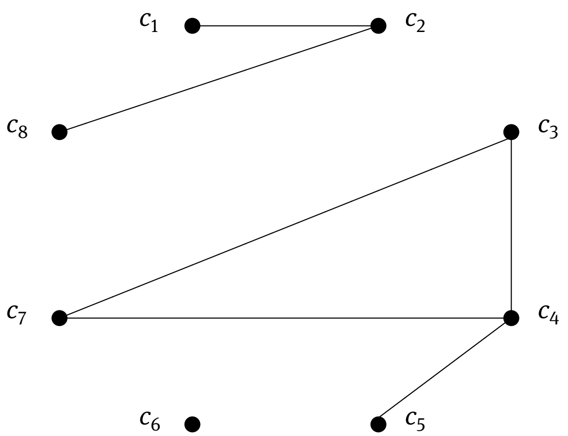


Figure 3.10: The threshold graph $G(6)$ generates the fifth-level clustering $\mathbf{C}_5$, which contains one 4-element cluster $\{c_3, c_4, c_5, c_7\}$, one 3-element cluster $\{c_1, c_2, c_8\}$, and a 1-element cluster $\{c_6\}$.

The edges with weights 8,9, and 10 do not generate new clusters. Finally, by making use of the edge $\{c_5, c_6\}$ with the weight $d(5, 6) = 11$, we get the one-cluster clustering $\mathbf{C}_7 = \{\mathcal{C}_{7,1}\}$, where $\mathcal{C}_{7,1} = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8\}$—see Fig. 3.12; if all the objects are merged in one cluster, the clustering is called *conjoint*.

It is worth noting that, in terms of our model, both $\mathbf{C}_0$ and $\mathbf{C}_7$ result in the road network shown in Fig. 3.2.

Problem 3.2.1. Draw road maps corresponding to all other levels of clustering, $\mathbf{C}_2$, $\mathbf{C}_3$, $\mathbf{C}_5$, $\mathbf{C}_6$, in the model example.

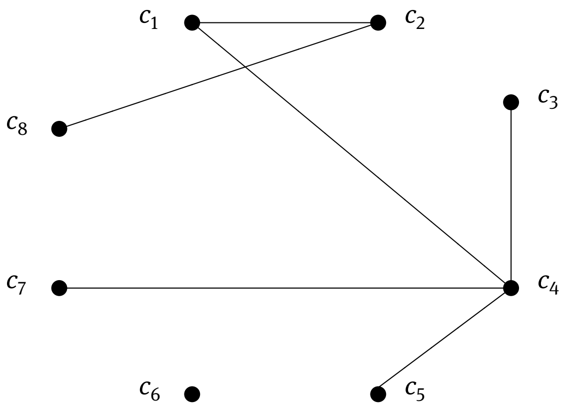


Figure 3.11: The threshold graph $G(7)$ generates the sixth-level clustering $\mathbf{C}_6$, which contains a 7-element cluster $\{c_1, c_2, c_3, c_4, c_5, c_7, c_8\}$ and a 1-element cluster $\{c_6\}$.

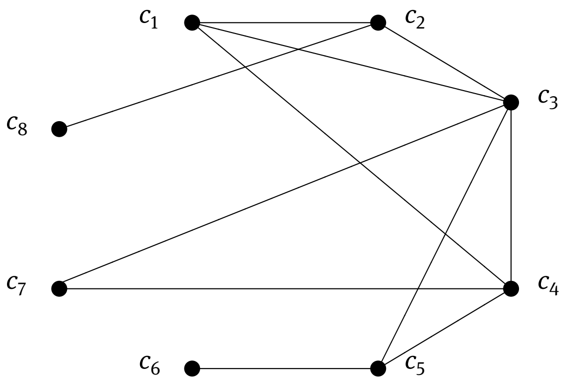


Figure 3.12: The threshold graph $G(11)$; it generates the conjoint clustering $\mathbf{C}_7 = \{\mathcal{C}_{7,1}\}$.

Analyzing our discussion of the model example, we see that the algorithm above can be stated in the following more formal way suitable for computer realization; we have presented it in the *pseudocode* form. In what follows we use the notation

$$a := b,$$

which means that the value b must be assigned to the variable a, or to put it another way, the current value of the variable a must be replaced by the value of b. For example, if $a = 5$ and $b = 0$, then after the command $a := b$ is executed, the value $a = 0$ while the value b remains unchanged, $b = 0$; the initial value of a, namely, $a = 5$ is deleted.

Problem 3.2.2. Starting with the initial value $m = 2$, find the value of the variable m after repeating twice the command $m := m - 1$.

Agglomerative single-link algorithm

Given a set of n objects $X = \{x_1, x_2, \dots, x_n\}$, their dissimilarity table, and a threshold value $\lambda \geq 0$.

1. Rearrange the dissimilarity table in ascending order.
2. Set $m = 0$ and make a completely disjoint clustering of zero level $\mathbf{C}_0 = \{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}, \dots, \mathcal{C}_{0,n}\}$, with 1-element clusters $\mathcal{C}_{0,i} = \{x_i\}, i = 1, \dots, n$.
3. Set $m := m + 1$ and consider the first unused entry, say $d(x_k, x_l)$ in the dissimilarity table. If $d(x_k, x_l) > \lambda$, then stop and return the last derived clustering. Otherwise, there are two possibilities.
   - 3-A The 2-element set $\{x_k, x_l\}$ is a subset of an existing cluster. Then skip $d(x_k, x_l)$ and return to step 3, that is, increase m by 1.
   - 3-B The objects x_k and x_l belong to different existing clusters, say $x_k \in \mathcal{C}_{m-1,a}$ and $x_l \in \mathcal{C}_{m-1,b}, a \neq b$. Form a cluster of the mth level as the union $\mathcal{C}_{m,1} = \mathcal{C}_{m-1,a} \cup \mathcal{C}_{m-1,b}$, renumber all the other clusters of the $(m-1)$th level to the mth level, without changing their elements, and return to step 3. □

Remark 3.2.1. The conjoint clustering can occur before we achieve the threshold level and as we have seen in the example, not every threshold graph generates a new clustering.

Remark 3.2.2. Since we look only for disjoint clusters, a clustering of any level is just a partition of the initial set of objects. Therefore, our algorithm generates a family of nested partitions of the given set. Moreover, we know (see Problems 1.1.18–1.1.19) that every partition of a set generates an equivalence relation on this set and vice versa. This relationship is dealt with in Exercise 3.2.3.

Remark 3.2.3. Part 3-A of this algorithm is quite analogues to the condition of not forming cycles in Part 2 of Kruskal's algorithm (Section 2.3) of constructing the minimum spanning trees.

Exercises 3.2.

Exercise 3.2.1. Given the initial value of the variable $k = 1$, what is the value of k after the command $k := (-1)^k$ is executed 3 times? Four times?

Exercise 3.2.2. Prove that, given n objects, n levels of clustering $\mathbf{C}_0, \dots, \mathbf{C}_{n-1}$ exist, where the last one is the conjoint clustering. Moreover, as far as the dissimilarity table contains $n(n-1)/2$ entries, there are no more than $1 + n(n-1)/2$ threshold graphs (exactly $1 + n(n-1)/2$ if there are no ties).

Exercise 3.2.3. Describe explicitly the equivalence relations corresponding to the partitions of the set $C = \{c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8\}$ generated by the clusterings $\mathbf{C}_0, \ldots, \mathbf{C}_7$ in the model example.

Exercise 3.2.4. Construct dissimilarity tables for a set with n elements such that there are exactly 2, or exactly 3, or exactly $1 + n(n-1)/2$ threshold graphs.

Exercise 3.2.5. Change the $\{c_3, c_7\}$ entry in Table 3.3 to 2 and the $\{c_2, c_8\}$ entry to 7, respectively, so that a new table contains ties. Apply the algorithm of this section to this new dissimilarity table and compare the resulting clusterings with the ones derived above.

Exercise 3.2.6. Using the algorithm of this section, construct all consecutive clusterings of the set $X = \{x_1, x_2, \ldots x_6\}$ with the dissimilarity table, Table 3.5. Which level of clustering corresponds to the threshold level of 3? Of 2?

Table 3.5: The dissimilarity table for Exercise 3.2.6.

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0	6	8	3	4	8
x_2		0	2	4	1	5
x_3			0	6	2	3
x_4				0	9	2
x_5					0	4
x_6						0

3.3 Hubert's single-link algorithm

In this and the following sections we consider two well-known algorithms by Hubert—the single-link and complete-link agglomerative clustering algorithms.

Coffee-time browsing

– www.psych.uiuc.edu/people/faculty/hubert.html (L. Hubert)

In the preceding section we represented the objects and their dissimilarities by weighted graphs. Since the dissimilarity was defined for each pair of objects, these graphs are *complete*. Therefore each cluster, as a set of vertices, is represented by a subgraph of the complete graph corresponding to the initial set of objects. Vice versa, each connected subgraph of this complete graph can be viewed as a cluster consisting of the vertices of this subgraph. Therefore, we will freely interchange the language of objects and their collections (clusters) on one hand, and the language of vertices, graphs, and subgraphs, on the other hand.

Hubert [31] gave versions of a single-link algorithm and a complete-link algorithm based on the concept of a threshold graph. Hubert's single-link algorithm leads to the same clustering as the agglomerative single-link algorithm of the preceding section. In this section we present Hubert's single-link algorithm in more formal *pseudocode form*. First of all, more notation is in order.

As before, we denote clustering of the mth level by

$$\mathbf{C}_m = \{\mathcal{C}_{m,1}, \mathcal{C}_{m,2}, \ldots, \mathcal{C}_{m,n_m}\}, \quad m = 0, 1, 2, \ldots,$$

where n_m stands for the number of clusters contained in the clustering $\mathcal{C}_m$ of mth level. In particular, $n_0 = n$. After $\mathbf{C}_m$ has been derived, we consider all $\frac{1}{2}n_m(n_m - 1)$ pairwise unions of these clusters

$$\mathcal{C}_{m,a} \cup \mathcal{C}_{m,b}$$

where $1 \leq a, b \leq n(m), a \neq b$. The union $\mathcal{C}_{m,a} \cup \mathcal{C}_{m,b}$ contains certain objects, say, the elements $x_i, \ldots, x_j$. Given the union $\mathcal{C}_{m,a} \cup \mathcal{C}_{m,b}$ for fixed a and $b, a \neq b$, we can form several connected subgraphs of the threshold graph $G(\lambda)$ spanned by these vertices $x_i, \ldots, x_j$. Namely, to derive such a subgraph from two clusters $\mathcal{C}_{m,a}$ and $\mathcal{C}_{m,b}$, we consider all possible connections of a vertex from $\mathcal{C}_{m,a}$ with a vertex from $\mathcal{C}_{m,b}$ using only one edge, called *single link*.

For the clusters $\mathcal{C}_{m,a}$ and $\mathcal{C}_{m,b}$ of mth level, denote the smallest dissimilarity between a vertex in $\mathcal{C}_{m,a}$ and a vertex in $\mathcal{C}_{m,b}$ by

$$S_m(a, b) = \min\{d(x_i, x_j) \mid x_i \in \mathcal{C}_{m,a}, x_j \in \mathcal{C}_{m,b}\};$$

clearly, this function is symmetric, that is, $S_m(a, b) = S_m(b, a)$. If the initial dissimilarity matrix contains ties, there may be several edges with the minimal weight, any one of those can be selected. At every step we merge two existing clusters, thus decreasing the number of clusters by 1.

The function $S_m = S_m(a, b)$ is defined on all pairs of clusters $\{\mathcal{C}_{m,a}, \mathcal{C}_{m,b}\}$ of the mth level. Since we only consider finite sets, this function attains its minimum value on a certain pair of clusters, say, $\mathcal{C}_{m,p}$ and $\mathcal{C}_{m,q}$. Let us denote this minimum value of the function $S_m(a, b)$ over all pairs of indices $\{a, b\}$ by $\min_{a,b}\{S_m(a, b)\} = S_m(p, q)$. The subscripts $p = p(m)$ and $q = q(m)$ depend on m, but suppressing this dependence in the notation does not lead to any ambiguity. We use the function $S_m(a, b)$ to present Hubert's single-link algorithm.

Hubert's single-link algorithm

Given a set of n objects $X = \{x_1, x_2, \ldots, x_n\}$, the dissimilarity table, and a threshold value λ.

1. Set $m = 0$ and form the disjoint clustering of zero level,

$$\mathbf{C}_0 = \{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}, \dots, \mathcal{C}_{0,n}\}$$

consisting of n 1-element clusters $\mathcal{C}_{0,k} = \{x_k\}, k = 1, \dots, n$. Define the function S_0 and the dissimilarities between the clusters of level zero by

$$S_0(a, b) = \operatorname{diss}(\mathcal{C}_{0,a}, \mathcal{C}_{0,b}) = d(x_a, x_b).$$

Find the minimum value $S_0^{\min}$ of the function $S_0(a, b)$ over all the pairs (a, b)

$$S_0^{\min} = \min_{a,b} S_0(a, b) = S_0(p, q),$$

attained at the pair (p, q). This pair indicates the clusters of zeroth level, $\mathcal{C}_{0,p}$ and $\mathcal{C}_{0,q}$, to be merged in a cluster of the first level,

$$\mathcal{C}_{1,1} = \mathcal{C}_{0,p} \cup \mathcal{C}_{0,p}.$$

All the other zeroth-level clusters remain the same, we only have to renumber them,

$$\mathcal{C}_{1,r} = \mathcal{C}_{0,s}, \quad r \geq 2,\ s \neq p,\ s \neq q.$$

2. Set $m := m + 1$, calculate the values

$$S_m(a, b) = \min\{d(x_i, x_j) \mid x_i \in \mathcal{C}_{m-1,a},\ x_j \in \mathcal{C}_{m-1,b}\}$$

for each pair of indices $\{a, b\}, a \neq b$, and find the minimum value

$$S_m^{\min} = \min_{a,b}\{S_m(a, b)\} = S_m(p, q)$$

where (p, q) is a pair (a, b) at which the minimum is attained. To build the next clustering $\mathbf{C}_m$, we merge those two clusters $\mathcal{C}_{m-1,p}$ and $\mathcal{C}_{m-1,q}$, whose second indices are p and q from above, into the cluster

$$\mathcal{C}_{m,1} = \mathcal{C}_{m-1,p} \cup \mathcal{C}_{m-1,p}$$

by making use of an edge with the weight $S_m^{\min} = S_m(p, q)$. If there are ties, that is, there exist several edges with the same weight $d_{\min}(p, q)$, we can use either of them. All the other clusters of the level $m-1$ become the clusters of level m without any change, after just renumbering.
3. Update the dissimilarity table as follows. The dissimilarity between every two "old" clusters (promoted from the preceding level) remains the same. The dissimilarity between $\mathcal{C}_{m,1}$ and any cluster $\mathcal{C}_{m,r} = \mathcal{C}_{m-1,s}, s \neq p, s \neq q$, is the smaller of

$$S_{m-1}(p, r) = \operatorname{diss}(\mathcal{C}_{m-1,p}, \mathcal{C}_{m-1,r})$$

and

$$S_{m-1}(q,r) = \mathrm{diss}(\mathcal{C}_{m-1,q}, \mathcal{C}_{m-1,r}),$$

thus, for any $r > 1$,

$$\mathrm{diss}(\mathcal{C}_{m,1}, \mathcal{C}_{m,r}) = \min\{S_{m-1}(p,r); S_{m-1}(q,r)\}.$$

4. Return to step 2 and continue until we reach the threshold value λ or all the objects are merged into one conjoint cluster, whichever occurs first. □

Remark 3.3.1. Thus to find the next clustering, it is necessary to calculate the double minimum

$$S_m^{\min}(p,q) = \min_{a,b}\{S_m(a,b)\} = \min_{a,b}\{\min\{d(x_i,x_j) \mid x_i \in \mathcal{C}_{m,a}; x_j \in \mathcal{C}_{m,b}\}\}.$$

We illustrate this algorithm using the model example from the preceding section, thus in the rest of this section we denote the objects by c_i. The algorithm starts with single-element clusters corresponding to each city $c_1, \dots, c_8$. That is, we set $m = 0$ and form the disjoint clustering

$$\mathbf{C}_0 = \{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}, \dots, \mathcal{C}_{0,n}\}$$

where $\mathcal{C}_{0,k} = \{c_k\}, k = 1, \dots, 8$. This clustering corresponds to the subgraph of the graph $G(\infty)$ with no edge—every vertex is an isolated one. The function $S_0(a,b)$ is shown in the right column of Table 3.4, its minimum value is $S_0^{\min} = S_0(3,4) = 1$.

Now, set $m = 0 + 1 = 1$. To every union $\mathcal{C}_{0,a} \cup \mathcal{C}_{0,b}$ there corresponds the unique connected subgraph of $G(\infty)$, this subgraph contains two vertices and their incident edge. Therefore, at this step $d_{\min} = d(c_3, c_4), p = 3, q = 4, S_0^{\min} = S_0(p,q) = 1$, and we have to join the clusters $\mathcal{C}_{0,3}$ and $\mathcal{C}_{0,4}$ in a cluster $\mathcal{C}_{1,1}$ of the first level. Then we upgrade all other zeroth-level clusters to the first level, and update the dissimilarity table. For example, since $\mathcal{C}_{1,2} = \{c_1\}$, we get

$$\begin{aligned}\mathrm{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,2}) &= \min\{S_0(1,3); S_0(1,4)\}\\ &= \min\{\mathrm{diss}(\mathcal{C}_{0,1}, \mathcal{C}_{0,3}); \mathrm{diss}(\mathcal{C}_{0,1}, \mathcal{C}_{0,4})\}\\ &= \min\{d(c_1,c_3); d(c_1,c_4)\} = \min\{10;7\} = 7.\end{aligned}$$

These computations lead to the updated dissimilarity table of the first level (Table 3.6) and to the *same* first level clustering as in Section 3.2, see Fig. 3.3,

$$\mathbf{C}_1 = \{\mathcal{C}_{1,1}, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{1,4}, \mathcal{C}_{1,5}, \mathcal{C}_{1,6}, \mathcal{C}_{1,7}\}$$

where $\mathcal{C}_{1,1} = \mathcal{C}_{0,3} \cup \mathcal{C}_{0,4} = \{c_3, c_4\}$, $\mathcal{C}_{1,i} = \mathcal{C}_{0,i-1} = \{c_{i-1}\}$ for $i = 2,3$ and $\mathcal{C}_{1,i} = \mathcal{C}_{0,i+1} = \{c_{i+1}\}$ for $i = 4,5,6,7$.

Table 3.6: The updated dissimilarity table of the first level.

diss($\mathcal{C}_{1,a}, \mathcal{C}_{1,b}$)
diss($\mathcal{C}_{1,1}, \mathcal{C}_{1,6}$) = 2
diss($\mathcal{C}_{1,1}, \mathcal{C}_{1,4}$) = 4
diss($\mathcal{C}_{1,2}, \mathcal{C}_{1,3}$) = 5
diss($\mathcal{C}_{1,3}, \mathcal{C}_{1,7}$) = 6
diss($\mathcal{C}_{1,1}, \mathcal{C}_{1,2}$) = 7
diss($\mathcal{C}_{1,1}, \mathcal{C}_{1,3}$) = 8
diss($\mathcal{C}_{1,4}, \mathcal{C}_{1,5}$) = 11
diss($\mathcal{C}_{1,2}, \mathcal{C}_{1,7}$) = 13
diss($\mathcal{C}_{1,1}, \mathcal{C}_{1,5}$) = 14
diss($\mathcal{C}_{1,5}, \mathcal{C}_{1,6}$) = 15
diss($\mathcal{C}_{1,4}, \mathcal{C}_{1,6}$) = 16
diss($\mathcal{C}_{1,3}, \mathcal{C}_{1,6}$) = 17
diss($\mathcal{C}_{1,4}, \mathcal{C}_{1,7}$) = 18
diss($\mathcal{C}_{1,5}, \mathcal{C}_{1,7}$) = 20
diss($\mathcal{C}_{1,1}, \mathcal{C}_{1,7}$) = 21
diss($\mathcal{C}_{1,2}, \mathcal{C}_{1,4}$) = 22
diss($\mathcal{C}_{1,3}, \mathcal{C}_{1,5}$) = 23
diss($\mathcal{C}_{1,6}, \mathcal{C}_{1,7}$) = 24
diss($\mathcal{C}_{1,2}, \mathcal{C}_{1,6}$) = 25
diss($\mathcal{C}_{1,2}, \mathcal{C}_{1,5}$) = 27
diss($\mathcal{C}_{1,3}, \mathcal{C}_{1,4}$) = 28

It is worth noting that after all these discussions, we certainly have a clear geometrical picture of this procedure, but we do not need it for the computations; Hubert's algorithm works analytically, without any appeal to graphs.

Now, set $m = 1 + 1 = 2$. From Table 3.6, $S_1^{\min} = S_1(1,6) = 2$, therefore, at this level $p = 1, q = 6$, and we have to merge the clusters $\mathcal{C}_{1,1} = \{c_3, c_4\}$ and $\mathcal{C}_{1,6} = \{c_7\}$ in the first cluster of the second level,

$$\mathcal{C}_{2,1} = \mathcal{C}_{1,1} \cup \mathcal{C}_{1,6} = \{c_3, c_4, c_7\}.$$

We renumber all the other first level clusters as clusters of the second level and use the same algorithm to calculate the dissimilarities between new clusters; see Table 3.7. We reiterate that again all considerations based on the graph theory, in particular, on the spanning trees, were left behind the scene—see Section 3.2. The whole procedure is based completely on the dissimilarity tables and is convenient for a computer realization.

At the next step we set $m = 3$. In the graph theory terms of the previous section, we looked at the threshold graph $G(2)$, which contained one 3-element and five 1-element clusters. The smallest unused dissimilarity was $d(3,7) = 3$, but adding the corresponding edge to $G(2)$ did not create a new cluster. Therefore, we had to leave out $d(3,7)$ and proceed on to $d(4,5) = 4$. However, now we are using the purely analytical Hubert's

Table 3.7: The updated dissimilarity table of the second level.

diss($\mathcal{C}_{2,a}, \mathcal{C}_{2,b}$)
diss($\mathcal{C}_{2,1}, \mathcal{C}_{2,4}$) = 4
diss($\mathcal{C}_{2,2}, \mathcal{C}_{2,3}$) = 5
diss($\mathcal{C}_{2,3}, \mathcal{C}_{2,6}$) = 6
diss($\mathcal{C}_{2,1}, \mathcal{C}_{2,2}$) = 7
diss($\mathcal{C}_{2,1}, \mathcal{C}_{2,3}$) = 8
diss($\mathcal{C}_{2,4}, \mathcal{C}_{2,5}$) = 11
diss($\mathcal{C}_{2,2}, \mathcal{C}_{2,6}$) = 13
diss($\mathcal{C}_{2,1}, \mathcal{C}_{2,5}$) = 14
diss($\mathcal{C}_{2,4}, \mathcal{C}_{2,6}$) = 18
diss($\mathcal{C}_{2,5}, \mathcal{C}_{2,6}$) = 20
diss($\mathcal{C}_{2,1}, \mathcal{C}_{2,6}$) = 21
diss($\mathcal{C}_{2,2}, \mathcal{C}_{2,4}$) = 22
diss($\mathcal{C}_{2,3}, \mathcal{C}_{2,5}$) = 23
diss($\mathcal{C}_{2,2}, \mathcal{C}_{2,5}$) = 27
diss($\mathcal{C}_{2,3}, \mathcal{C}_{2,4}$) = 28

algorithm and are to browse Table 3.7. From that table, $S_2^{\min} = S_2(1,4) = 4$, therefore, at this level $p = 1, q = 4$, and we have to merge the clusters $\mathcal{C}_{2,1} = \{c_3, c_4, c_7\}$ and $\mathcal{C}_{2,4} = \{c_5\}$ in the first cluster of the third level,

$$\mathcal{C}_{3,1} = \mathcal{C}_{2,1} \cup \mathcal{C}_{2,5} = \{c_3, c_4, c_5, c_7\}.$$

We renumber all the other second-level clusters to the third level and use the same algorithm to calculate the dissimilarities between new clusters; see Table 3.8.

Table 3.8: The updated dissimilarity table of the third level.

diss($\mathcal{C}_{3,a}, \mathcal{C}_{3,b}$)
diss($\mathcal{C}_{3,2}, \mathcal{C}_{3,3}$) = 5
diss($\mathcal{C}_{3,3}, \mathcal{C}_{3,5}$) = 6
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,2}$) = 7
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,3}$) = 8
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,4}$) = 11
diss($\mathcal{C}_{3,2}, \mathcal{C}_{3,5}$) = 13
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,5}$) = 18
diss($\mathcal{C}_{3,4}, \mathcal{C}_{3,5}$) = 20
diss($\mathcal{C}_{3,3}, \mathcal{C}_{3,4}$) = 23
diss($\mathcal{C}_{3,2}, \mathcal{C}_{3,4}$) = 27

Comparing with the algorithm of Section 3.2, we see that Hubert's algorithm at every step leads directly to the next-level clustering without pausing at intermediate threshold graphs, which do not generate the next level of clustering. This way, we build up the single-link clusterings of all higher levels, which are, of course, the same as in Section 3.2, up to the conjoint clustering $\mathbf{C}_7$. We show here only the updated dissimilarity Tables 3.9, 3.10, 3.11 of the sequel fourth, fifth, and sixth levels.

Table 3.9: The updated dissimilarity table of the fourth level.

diss$(\mathcal{C}_{4,a}, \mathcal{C}_{4,b})$
diss$(\mathcal{C}_{4,1}, \mathcal{C}_{4,4}) = 6$
diss$(\mathcal{C}_{4,1}, \mathcal{C}_{4,2}) = 7$
diss$(\mathcal{C}_{4,2}, \mathcal{C}_{4,3}) = 11$
diss$(\mathcal{C}_{4,2}, \mathcal{C}_{4,4}) = 18$
diss$(\mathcal{C}_{4,3}, \mathcal{C}_{4,4}) = 20$
diss$(\mathcal{C}_{4,1}, \mathcal{C}_{4,3}) = 23$

Table 3.10: The updated dissimilarity table of the fifth level.

diss$(\mathcal{C}_{5,a}, \mathcal{C}_{5,b})$
diss$(\mathcal{C}_{5,1}, \mathcal{C}_{5,2}) = 7$
diss$(\mathcal{C}_{5,2}, \mathcal{C}_{5,3}) = 11$
diss$(\mathcal{C}_{5,1}, \mathcal{C}_{5,3}) = 20$

Table 3.11: The updated dissimilarity table of the sixth level.

diss$(\mathcal{C}_{6,a}, \mathcal{C}_{6,b})$
diss$(\mathcal{C}_{6,1}, \mathcal{C}_{6,2}) = 11$

When amalgamating, step by step, the clusters, we are increasing the threshold value and, respectively, generating the threshold graphs. They are the same as before and are shown in Figs. 3.1–3.3, 3.5–3.8, 3.10–3.12.

If we compare these figures with Figs. 2.13–2.21 in Section 2.3, we recognize similar graphs and easily convince ourselves that the steps of the agglomerative clustering algorithms of this and the previous sections correspond to the steps of Kruskal's algorithm of constructing a minimum spanning tree.

To visualize the process of clustering, a special kind of tree-like graphs is useful. These graphs are called *dendrograms*. Below we build the single-link dendrogram corresponding to our model problem; see Fig. 3.13. It is clear from this example how to build a dendrogram for any problem. Different horizontal levels of the dendrogram, top to down, correspond to consecutive clusterings in the problem. Thus, the level

A ↔ **A** gives the clustering

$$\mathbf{C}_3 = \{\{c_1\}, \{c_2\}, \{c_8\}, \{c_3, c_4, c_5, c_7\}, \{c_6\}\},$$

the level **B** ↔ **B** at Fig. 3.13 generates the clustering

$$\mathbf{C}_5 = \{\{c_1, c_2, c_8\}, \{c_3, c_4, c_5, c_7\}, \{c_6\}\}.$$

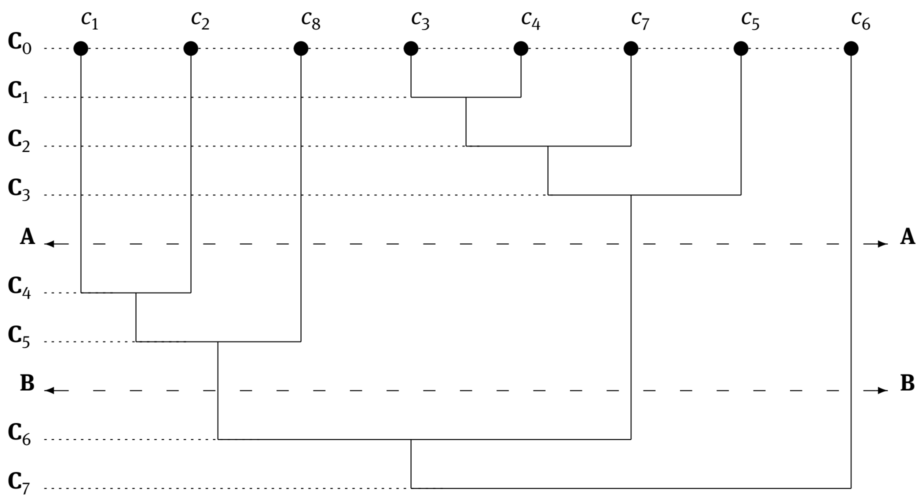


Figure 3.13: The dendrogram for the model example.

Exercises 3.3.

Exercise 3.3.1. Using Hubert's single-link algorithm, build all threshold graphs and clusterings of the set $X = \{x_1, x_2, \dots, x_6\}$, the given dissimilarity table, Table 3.5. Which level of clustering corresponds to the threshold levels of 2? Of 3? Of 4?

Exercise 3.3.2. Draw the dendrogram for the dissimilarity table, Table 3.5.

Exercise 3.3.3. Using Hubert's single-link algorithm, derive a conjoint clustering of the set X with the dissimilarities given in Table 3.12. Draw the corresponding dendrogram.

3.4 Hubert's complete-link algorithm

In this section we consider a different approach to amalgamated clustering, called complete-link clustering. An essential distinction between the single-link and complete-link algorithms is the rule of merging two existing clusters into one of a higher

Table 3.12: The dissimilarity table for Exercise 3.3.3 and Exercise 3.4.2.

X	x_1	x_2	x_3	x_4	x_5
x_1	0	4	1	3	8
x_2		0	2	5	10
x_3			0	6	7
x_4				0	9
x_5					0

level. Instead of *connected subgraphs* of the threshold graph $G(\infty)$ used in the single linkage, now we consider the *maximum complete subgraphs* of $G(\infty)$. Examples show that the single linkage and the complete linkage may result in different clusterings.

Coffee-time browsing

- www.sigkdd.org/explorations/issue4-1/estivill.pdf

We are concerned with another Hubert's clustering algorithm called *complete-link clustering* [31]. We use the same notations as in the previous sections, but consider only dissimilarity matrices without ties.[2] We again start with an informal description of the algorithm and then write down its pseudo-code.

Like the single linkage, the complete linkage uses the same sequence of the threshold graphs. To avoid any ambiguity, we denote complete-link clusterings by $\mathbf{C}_m^{\text{comp}}$. Given a clustering

$$\mathbf{C}_m^{\text{comp}} = \{\mathcal{C}_{m,1}, \mathcal{C}_{m,2}, \dots, \mathcal{C}_{m,n_m}\}$$

of the mth level, $m = 0, 1, 2, \dots$, we consider all pairwise unions

$$\mathcal{C}_{m,a} \cup \mathcal{C}_{m,b}, \quad a, b = 1, 2, \dots, n_m, \ a \neq b.$$

Let the union $\mathcal{C}_{m,a} \cup \mathcal{C}_{m,b}$ contain objects $x_i, \dots, x_j$. While building the single-linkage, we looked for an edge (a single link) with the smallest dissimilarity. Now we are adding edges connecting a vertex in $C_{m,a}$ with a vertex in $C_{m,b}$, in increasing order of their dissimilarities, until we reach a *complete* subgraph of the threshold graph $G(\infty)$ spanned by all vertices $x_i, \dots, x_j$. Only at that point, the union $\mathcal{C}_{m,a} \cup \mathcal{C}_{m,b}$ becomes a cluster of the next, $(m + 1)$st level.

To formalize this procedure, let $T_m(a, b)$ stand for the *maximal* dissimilarity over all edges used in this construction, that is,

$$T_m(a, b) = \max\{d(x_i, x_j) \mid x_i \in \mathcal{C}_{m,a}, x_j \in \mathcal{C}_{m,b}\}.$$

2 Clustering in the presence of ties is discussed, for example, in [32, p. 76].

Similarly to S_m, T_m is a symmetrical function on pairs of clusters of the mth level, but unlike S_m, T_m is the maximal, not minimal dissimilarity. Let $\mathcal{C}_{m,p}, \mathcal{C}_{m,q}$ be a pair of clusters where this function attains its minimum value over all the pairs of clusters of the mth level. Denote this minimum value by

$$T_m^{\min} = \min_{a,b}\{T_m(a,b)\} = T_m(p,q).$$

To build the next clustering $\mathbf{C}_{m+1}$, we merge these two clusters $\mathcal{C}_{m,p}$ and $\mathcal{C}_{m,q}$ and update the dissimilarity table. We give a pseudocode of this algorithm.

Hubert's complete-link algorithm

Given a set $X = \{x_1, x_2, \ldots, x_n\}$, the dissimilarity table, and the threshold value λ.

1. Set $m = 0$ and form the disjoint clustering of level zero,

$$\mathbf{C}_0^{\mathrm{comp}} = \{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}, \ldots, \mathcal{C}_{0,n}\}$$

consisting of n 1-element clusters $\mathcal{C}_{0,k} = \{x_k\}, k = 1, \ldots, n$. Define the function T_0 and the dissimilarities between the clusters of level zero by

$$T_0(a,b) = \mathrm{diss}(\mathcal{C}_{0,a}, \mathcal{C}_{0,b}) = d(x_a, x_b).$$

Find the minimum value $T_0^{\min}$ of the function $T_0(a,b)$ over all the pairs (a,b)

$$T_0^{\min} = \min_{a,b} T_0(a,b) = T_0(p,q)$$

attained at the pair (p,q). This pair determines the clusters of zero level, $\mathcal{C}_{0,p}$ and $\mathcal{C}_{0,q}$, to be merged in a cluster of the first level

$$\mathcal{C}_{1,1} = \mathcal{C}_{0,p} \cup \mathcal{C}_{0,p}.$$

The other zeroth-level clusters remain the same, we only have to renumber them,

$$\mathcal{C}_{1,r} = \mathcal{C}_{0,s}, \quad r \geq 2,\ s \neq p,\ s \neq q.$$

2. Set $m := m + 1$, calculate the values

$$T_m(a,b) = \max\{d(x_i, x_j) \mid x_i \in \mathcal{C}_{m,a},\ x_j \in \mathcal{C}_{m,b}\}$$

for all pairs of clusters of the mth level, and find their minimum value

$$T_m^{\min} = T_m(p,q) = \min_{a,b} T_m(a,b).$$

To form the next clustering $\mathbf{C}_{m+1}^{\mathrm{comp}}$, we define

$$\mathcal{C}_{m+1,1} = \mathcal{C}_{m,p} \cup \mathcal{C}_{m,q}.$$

All the other clusters of the mth level become, after renumbering, the clusters of level $m + 1$ without changes.

3. Update the dissimilarity table as follows. The dissimilarity between every two "old" clusters (promoted from the preceding level) remains the same. The dissimilarity between the "new" cluster $\mathcal{C}_{m,1}$ and any "old" cluster $\mathcal{C}_{m,r}$ with $r \neq p$ and $r \neq q$ is the larger of the two dissimilarities $\text{diss}(\mathcal{C}_{m,p}, \mathcal{C}_{m,r})$ and $\text{diss}(\mathcal{C}_{m,q}, \mathcal{C}_{m,r})$.
4. Continue until we reach the threshold value or all the objects are merged into one conjoint cluster, whichever occurs first. □

Remark 3.4.1. In Step 2 we combine two clusters into a new one only when we reach an edge with the maximal dissimilarity between the entities in the two clusters; so to say, we link them completely. In terms of graphs, we merge two complete subgraphs G_p and G_q by using all edges with one end in G_p and another end in G_q.

Remark 3.4.2. Using the single linkage, we calculate a double minimum of the dissimilarities, first over a fixed pair of clusters and then over all pairs of clusters—see Remark 3.3.1. Unlike that, in the complete linkage we calculate the minimum of maximal values—first we calculate the maximal dissimilarity of the objects over a fixed pair of clusters and then the minimum of these maximums over all pairs of clusters.

We apply this algorithm to our model example with Dissimilarity Table 3.4. The threshold graphs do not depend on the method used, whether it is the single- or complete-linkage method. If some edges in the sequel figures are dashed, this means that this new edge does not generate a new cluster. The subgraph of the threshold graph $G(1)$, generated by two vertices c_3 and c_4 is a complete graph isomorphic to the complete subgraph K_2. This is the same threshold graph $G(1)$ as in the single linkage—see Fig. 3.3. Therefore, first two clusterings are the same as in the single-linkage case,

$$\mathbf{C}_0^{\text{comp}} = \{\mathcal{C}_{0,1}, \mathcal{C}_{0,2}, \mathcal{C}_{0,3}, \mathcal{C}_{0,4}, \mathcal{C}_{0,5}, \mathcal{C}_{0,6}, \mathcal{C}_{0,7}, \mathcal{C}_{0,8}\}$$

where $\mathcal{C}_{0,i} = \{c_i\}, i = 1, \dots, 8$, and

$$\mathbf{C}_1^{\text{comp}} = \{\mathcal{C}_{1,1}, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{1,4}, \mathcal{C}_{1,5}, \mathcal{C}_{1,6}, \mathcal{C}_{1,7}\}$$

where $\mathcal{C}_{1,1} = \mathcal{C}_{0,3} \cup \mathcal{C}_{0,4} = \{c_3, c_4\}$, $\mathcal{C}_{1,i} = \mathcal{C}_{0,i-1} = \{c_{i-1}\}$ for $i = 2, 3$, and $\mathcal{C}_{1,i} = \mathcal{C}_{0,i+1} = \{c_{i+1}\}$ for $i = 4, \dots, 7$.

However, the threshold graph $G(2)$ (Fig. 3.14) does not contain a complete subgraph—its subgraph, spanned by the vertices c_3, c_4, and c_7, is not a complete graph, since vertices c_3 and c_7 are not adjacent. Thus, even though $G(2)$ generates a single-link clustering (cf. Sections 3.2–3.3), it does not generate a complete-link clustering.

Now, the threshold graph $G(3)$ (Fig. 3.15) contains a complete subgraph, isomorphic to K_3, spanned by the vertices c_3, c_4, c_7. Therefore, we merge these three vertices

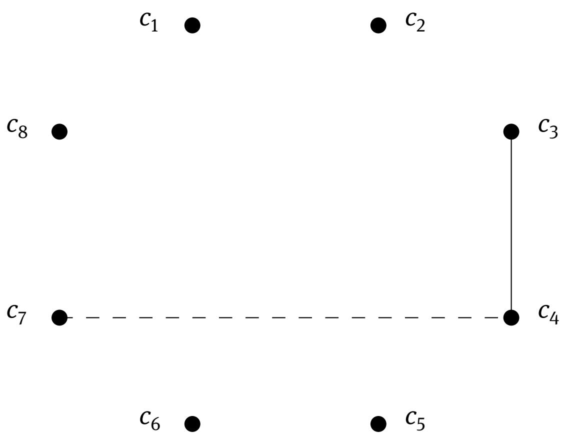


Figure 3.14: The threshold graph $G(2)$. The edge $\{c_4, c_7\}$ is dashed (cf. Fig. 3.5) for it does not generate the next level complete-link clustering.

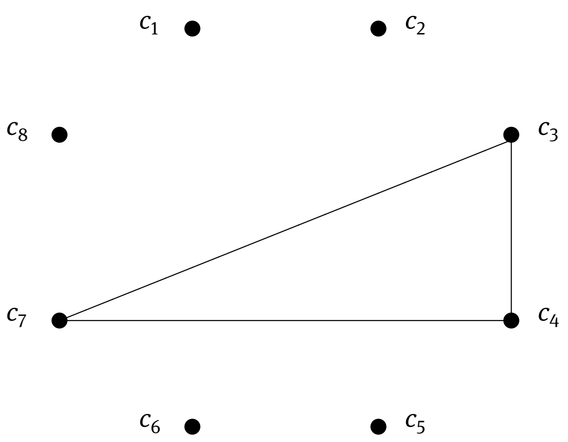


Figure 3.15: The threshold graph $G(3)$ is the same as in Fig. 3.6. The subgraph spanned by the vertices $\{c_3, c_4, c_7\}$ is complete.

into a cluster $\mathcal{C}_{2,1}$, and the next complete-link clustering is

$$\mathbf{C}_2^{\text{comp}} = \{\mathcal{C}_{2,1}, \mathcal{C}_{2,2}, \mathcal{C}_{2,3}, \mathcal{C}_{2,4}, \mathcal{C}_{2,5}, \mathcal{C}_{2,6}\}$$

where $\mathcal{C}_{2,1} = \mathcal{C}_{1,1} \cup \mathcal{C}_{1,6} = \{c_3, c_4, c_7\}$. Five other clusters contain only one vertex each. We notice that only three edges (three links) have been used here. At this step the single-link and the complete-link clusterings still coincide.

The next threshold graph $G(4)$ (Fig. 3.16) is generated by the edge $\{c_4, c_5\}$. However, this edge does not generate a new complete subgraph, and the threshold graph $G(4)$ does not generate the next level of complete clustering.

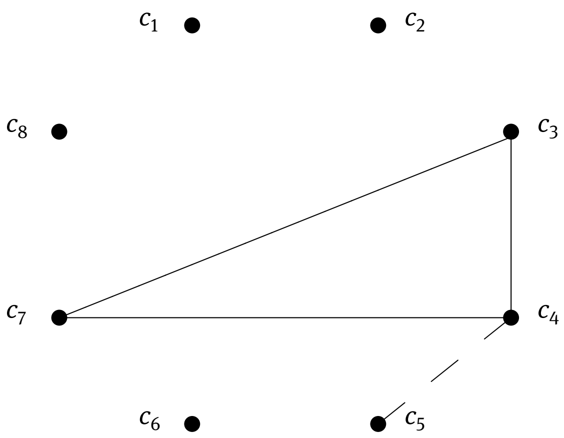


Figure 3.16: Complete linkage: threshold graph $G(4)$, cf. Fig. 3.7. The subgraph spanned by the vertices $\{c_3, c_4, c_5, c_7\}$ is not complete.

The threshold graph $G(5)$ (Fig. 3.17) contains a K_2-isomorphic subgraph with the vertices c_1 and c_2. Hence, it generates a new complete-link clustering

$$\mathbf{C}_3^{\text{comp}} = \{\mathcal{C}_{3,1}, \mathcal{C}_{3,2}, \mathcal{C}_{3,3}, \mathcal{C}_{3,4}, \mathcal{C}_{3,5}\},$$

where $\mathcal{C}_{3,1} = \{c_1, c_2\}$, $\mathcal{C}_{3,2} = \{c_3, c_4, c_7\}$, $\mathcal{C}_{3,3} = \{c_5\}$, $\mathcal{C}_{3,4} = \{c_6\}$, and $\mathcal{C}_{3,5} = \{c_8\}$. Starting at this step, Hubert's complete-link algorithm generates clusterings distinct from the single linkage. The threshold graph $G(6)$ (Fig. 3.18), generated by the edge $\{c_2, c_8\}$ with the dissimilarity $d(2, 8) = 6$, also does not contain a new complete subgraph. The sequel four threshold graphs, $G(7)$–$G(10)$ (Figs. 3.19–3.20) also do not contain new complete subgraphs and generate no new clustering.

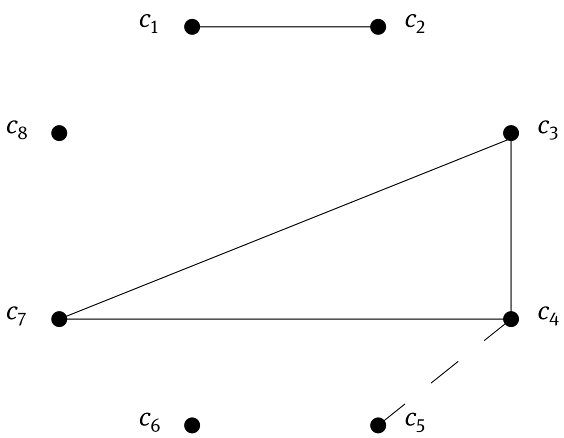


Figure 3.17: Complete linkage: threshold graph $G(5)$, cf. Fig. 3.8.

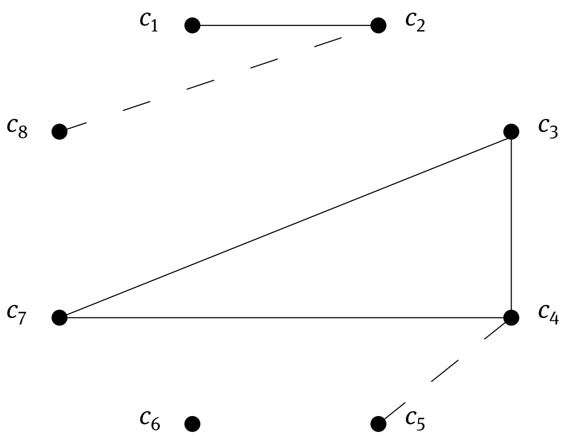


Figure 3.18: Complete linkage: threshold graph $G(6)$, cf. Fig. 3.10.

However, in the threshold graph $G(11)$ (Fig. 3.21) the vertices c_5 and c_6 become connected, and since they belong to no existing cluster, we have to merge them in a cluster of the next level. Thus, we derive the fourth clustering,

$$\mathbf{C}_4^{\text{comp}} = \{\mathcal{C}_{4,1}, \mathcal{C}_{4,2}, \mathcal{C}_{4,3}, \mathcal{C}_{4,4}\}$$

where

$$\mathcal{C}_{4,1} = \{c_5, c_6\}, \quad \mathcal{C}_{4,2} = \{c_1, c_2\}, \quad \mathcal{C}_{4,3} = \{c_3, c_4, c_7\}, \quad \mathcal{C}_{4,4} = \{c_8\}.$$

The next complete-link clustering is generated by the edge $\{c_1, c_8\}$ with the dissimilarity $d(1, 8) = 13$ (Fig. 3.21),

$$\mathbf{C}_5^{\text{comp}} = \{\mathcal{C}_{5,1}, \mathcal{C}_{5,2}, \mathcal{C}_{5,3}\}$$

where

$$\mathcal{C}_{5,1} = \{c_1, c_2, c_8\}, \quad \mathcal{C}_{5,2} = \{c_3, c_4, c_7\}, \quad \mathcal{C}_{4,3} = \{c_5, c_6\}.$$

The threshold graphs $G(14)$–$G(18)$ (Figs. 3.22–3.24) also do not generate a new clustering.

Only the threshold graph $G(19)$ (Fig. 3.24) generates the second to the last complete-link clustering

$$\mathbf{C}_6^{\text{comp}} = \{\mathcal{C}_{6,1}, \mathcal{C}_{6,2}\}$$

with two clusters $\mathcal{C}_{6,1} = \{c_3, c_4, c_5, c_6, c_7\}$ and $\mathcal{C}_{6,2} = \{c_1, c_2, c_8\}$. The final conjoint complete-link clustering $\mathbf{C}_7^{\text{comp}}$ is generated by the threshold graph $G(28)$. We remark

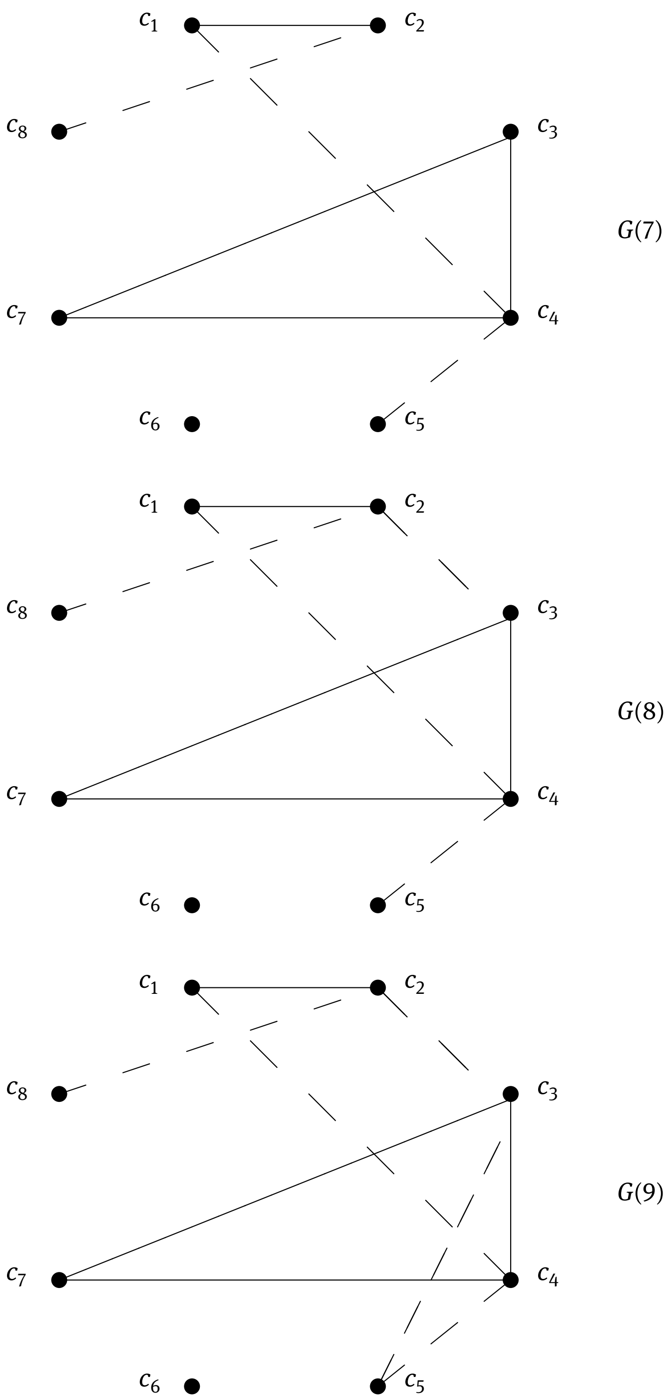


Figure 3.19: Complete linkage: threshold graphs $G(7)$–$G(9)$, cf. Fig. 3.11.

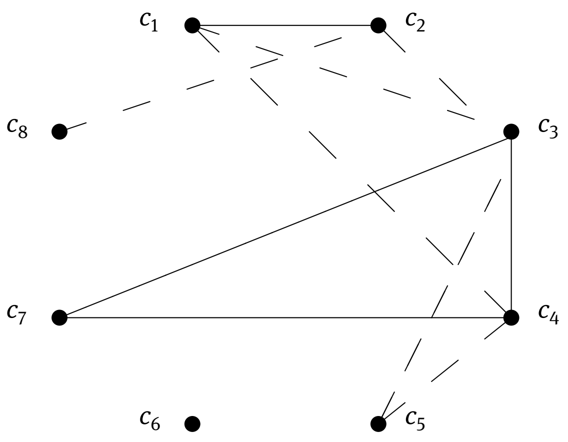


Figure 3.20: Complete linkage: threshold graph $G(10)$.

that in this example only the first three levels of the single-linkage and complete-linkage cases coincide. From the fourth level on, the clusters are different.

Now we translate this construction into formal analytic language and derive the complete-link clustering by making use of the dissimilarity tables. Start with the same dissimilarity table, Table 3.4 of zeroth level. Using the function T_m instead of S_m, we compute the following tables.

From Table 3.13, we see that $T_1^{\min} = T_1(1,6) = 3$, thus we have the same complete-link clustering of the first level

$$\mathbf{C}_1^{\text{comp}} = \{\mathcal{C}_{1,1}, \mathcal{C}_{1,2}, \mathcal{C}_{1,3}, \mathcal{C}_{1,4}, \mathcal{C}_{1,5}, \mathcal{C}_{1,6}, \mathcal{C}_{1,7}\}$$

where $\mathcal{C}_{1,1} = \mathcal{C}_{0,3} \cup \mathcal{C}_{0,4} = \{c_3, c_4\}$, $\mathcal{C}_{1,i} = \mathcal{C}_{0,i-1} = \{c_{i-1}\}$ for $i = 2,3$, and $\mathcal{C}_{1,i} = \mathcal{C}_{0,i+1} = \{c_{i+1}\}$ for $i = 4,\dots,7$.

The next dissimilarity table is Table 3.14, thus, $T_2^{\min} = T_2(2,3) = 5$, and we derive the same complete-link clustering of the second level

$$\mathbf{C}_2^{\text{comp}} = \{\mathcal{C}_{2,1}, \mathcal{C}_{2,2}, \mathcal{C}_{2,3}, \mathcal{C}_{2,4}, \mathcal{C}_{2,5}, \mathcal{C}_{2,6}\}$$

where $\mathcal{C}_{2,1} = \mathcal{C}_{1,1} \cup \mathcal{C}_{1,6} = \{c_3, c_4, c_7\}$.

From the following dissimilarity tables, Tables 3.15–3.18, we observe the corresponding values of the function T_m, $T_3^{\min} = T_3(3,4) = 11$, $T_4^{\min} = T_4(2,4) = 13$, $T_5^{\min} = T_5(2,3) = 19$, and $T_6^{\min} = T_6(1,3) = 28$.

Finally, we draw the dendrogram (Fig. 3.25) for the complete-link clustering in this example—compare it with the one in Fig. 3.13. Are these dendrograms identical?

Exercises 3.4.

Exercise 3.4.1. Using Hubert's complete-link algorithm, build all consecutive threshold graphs and clusterings of the set $X = \{x_1, x_2, \dots, x_6\}$, given dissimilarity table, Table 3.5. Which level of clustering corresponds to the threshold levels of 2? Of 3? Of 4?

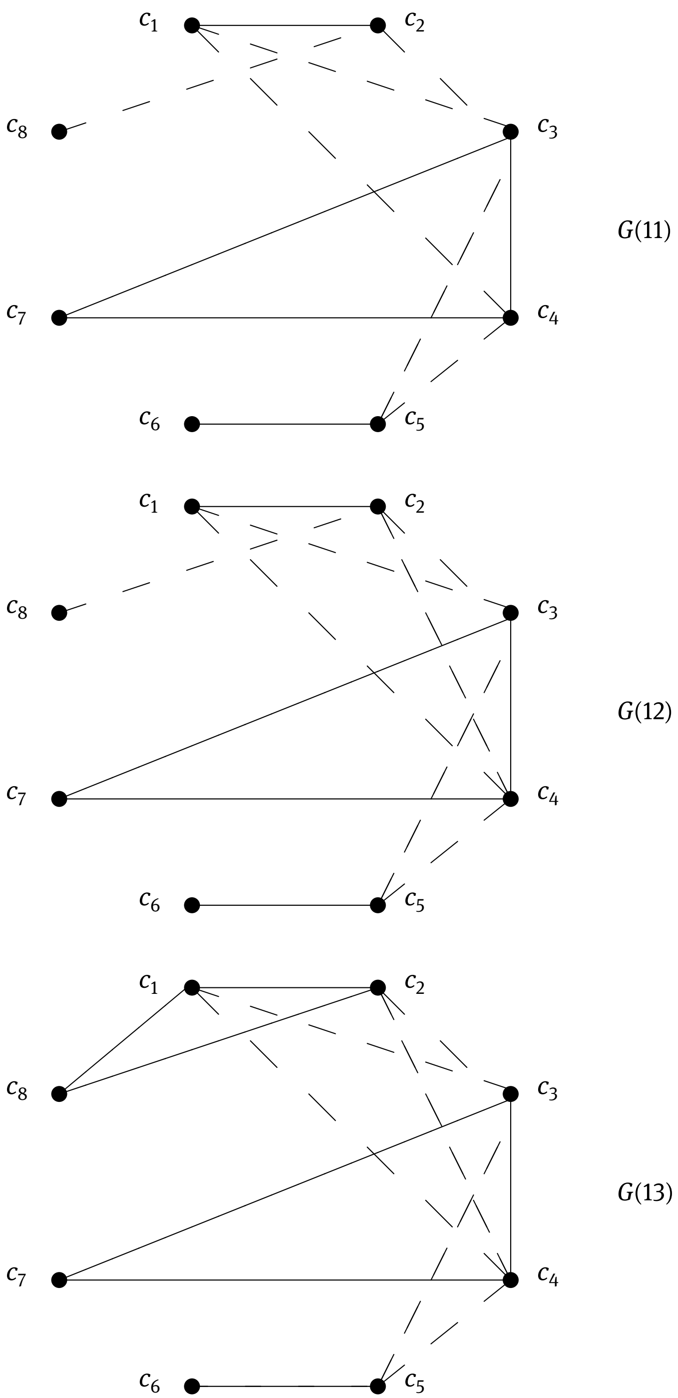


Figure 3.21: Complete linkage: threshold graphs $G(11)$–$G(13)$, cf. Fig. 3.12.

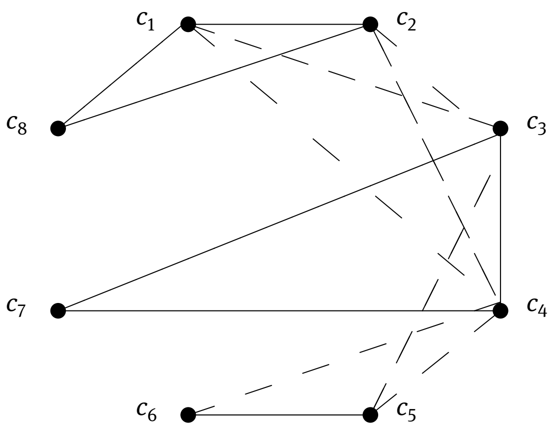


Figure 3.22: Complete linkage: threshold graph G(14).

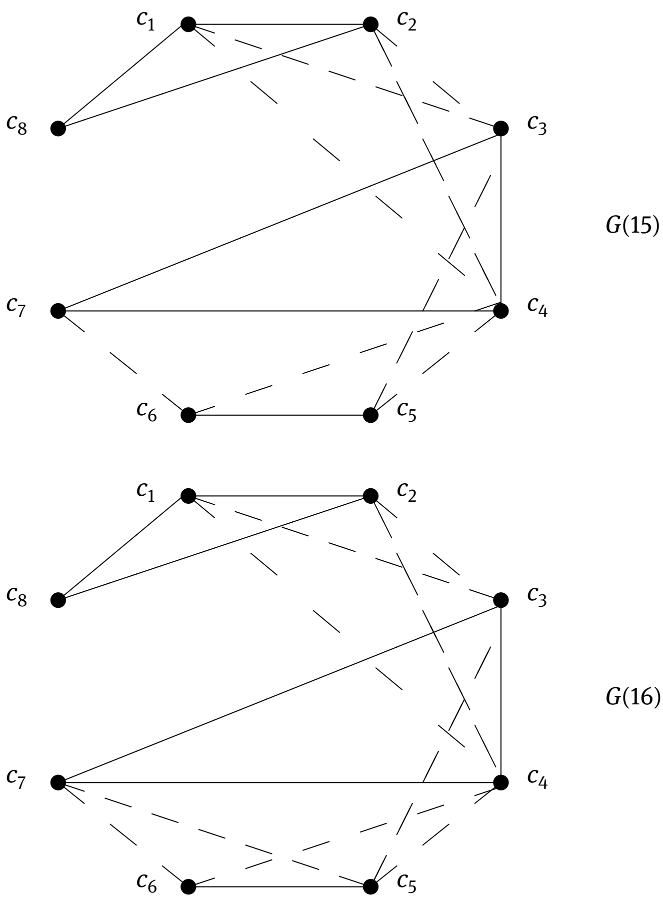


Figure 3.23: Complete linkage: threshold graphs G(15)–G(16).

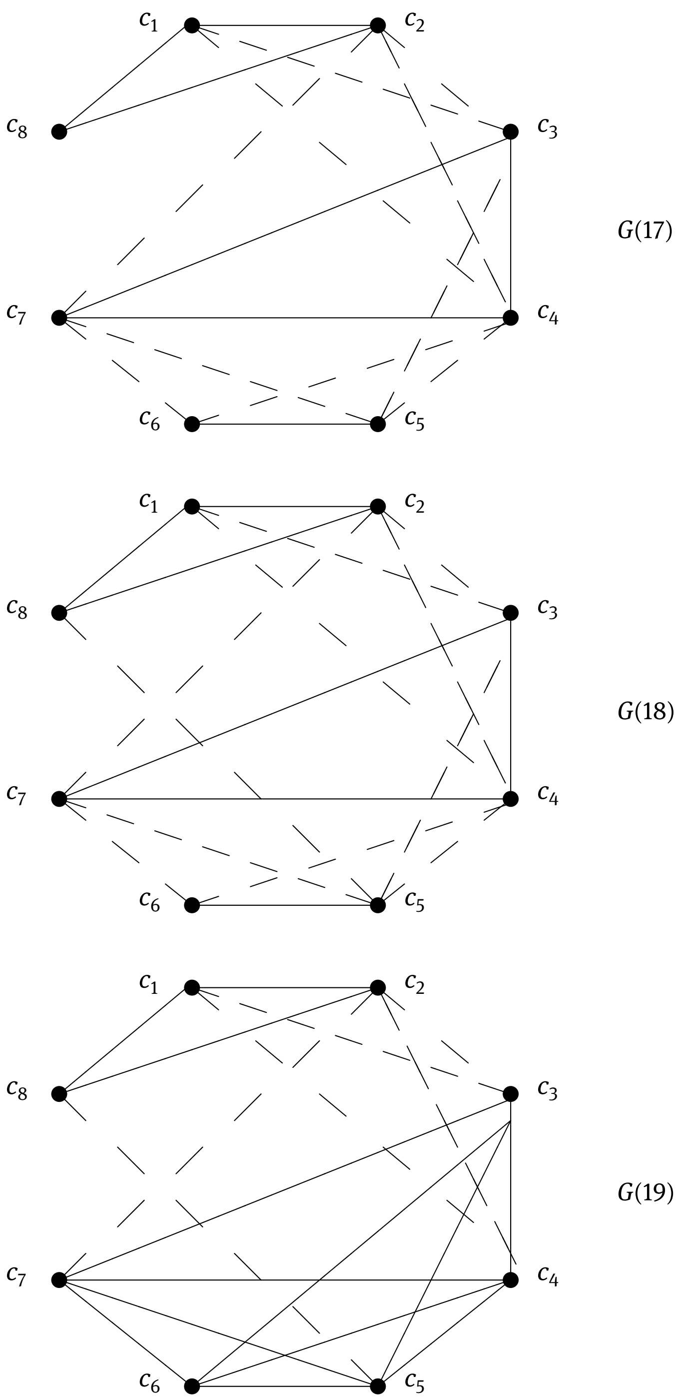


Figure 3.24: Complete linkage: threshold graphs G(17)–G(19).

Table 3.13: Complete linkage: the updated dissimilarity table of the first level.

$\mathbf{diss}(\mathcal{C}_{1,a}, \mathcal{C}_{1,b})$
$\text{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,6}) = 3$
$\text{diss}(\mathcal{C}_{1,2}, \mathcal{C}_{1,3}) = 5$
$\text{diss}(\mathcal{C}_{1,3}, \mathcal{C}_{1,7}) = 6$
$\text{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,4}) = 9$
$\text{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,2}) = 10$
$\text{diss}(\mathcal{C}_{1,4}, \mathcal{C}_{1,5}) = 11$
$\text{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,3}) = 12$
$\text{diss}(\mathcal{C}_{1,2}, \mathcal{C}_{1,7}) = 13$
$\text{diss}(\mathcal{C}_{1,5}, \mathcal{C}_{1,6}) = 15$
$\text{diss}(\mathcal{C}_{1,4}, \mathcal{C}_{1,6}) = 16$
$\text{diss}(\mathcal{C}_{1,3}, \mathcal{C}_{1,6}) = 17$
$\text{diss}(\mathcal{C}_{1,4}, \mathcal{C}_{1,7}) = 18$
$\text{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,5}) = 19$
$\text{diss}(\mathcal{C}_{1,5}, \mathcal{C}_{1,7}) = 20$
$\text{diss}(\mathcal{C}_{1,2}, \mathcal{C}_{1,4}) = 22$
$\text{diss}(\mathcal{C}_{1,3}, \mathcal{C}_{1,5}) = 23$
$\text{diss}(\mathcal{C}_{1,2}, \mathcal{C}_{1,6}) = 25$
$\text{diss}(\mathcal{C}_{1,1}, \mathcal{C}_{1,7}) = 26$
$\text{diss}(\mathcal{C}_{1,2}, \mathcal{C}_{1,5}) = 27$
$\text{diss}(\mathcal{C}_{1,3}, \mathcal{C}_{1,4}) = 28$

Table 3.14: Complete linkage: The updated dissimilarity table of the second level.

$\mathbf{diss}(\mathcal{C}_{2,a}, \mathcal{C}_{2,b})$
$\text{diss}(\mathcal{C}_{2,2}, \mathcal{C}_{2,3}) = 5$
$\text{diss}(\mathcal{C}_{2,3}, \mathcal{C}_{2,6}) = 6$
$\text{diss}(\mathcal{C}_{2,4}, \mathcal{C}_{2,5}) = 11$
$\text{diss}(\mathcal{C}_{2,2}, \mathcal{C}_{2,6}) = 13$
$\text{diss}(\mathcal{C}_{2,1}, \mathcal{C}_{2,4}) = 16$
$\text{diss}(\mathcal{C}_{2,1}, \mathcal{C}_{2,3}) = 17$
$\text{diss}(\mathcal{C}_{2,4}, \mathcal{C}_{2,6}) = 18$
$\text{diss}(\mathcal{C}_{2,1}, \mathcal{C}_{2,5}) = 19$
$\text{diss}(\mathcal{C}_{2,5}, \mathcal{C}_{2,6}) = 20$
$\text{diss}(\mathcal{C}_{2,2}, \mathcal{C}_{2,4}) = 22$
$\text{diss}(\mathcal{C}_{2,3}, \mathcal{C}_{2,5}) = 23$
$\text{diss}(\mathcal{C}_{2,1}, \mathcal{C}_{2,2}) = 25$
$\text{diss}(\mathcal{C}_{2,1}, \mathcal{C}_{2,6}) = 26$
$\text{diss}(\mathcal{C}_{2,2}, \mathcal{C}_{2,5}) = 27$
$\text{diss}(\mathcal{C}_{2,3}, \mathcal{C}_{2,4}) = 28$

Exercise 3.4.2. Using Hubert's complete-link algorithm, build conjoint clustering of the set X with the dissimilarities given in Table 3.12. Draw the corresponding dendrogram.

Table 3.15: Complete linkage: the updated dissimilarity table of the third level.

diss($\mathcal{C}_{3,a}, \mathcal{C}_{3,b}$)
diss($\mathcal{C}_{3,3}, \mathcal{C}_{3,4}$) = 11
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,5}$) = 13
diss($\mathcal{C}_{3,2}, \mathcal{C}_{3,3}$) = 16
diss($\mathcal{C}_{3,3}, \mathcal{C}_{3,5}$) = 18
diss($\mathcal{C}_{3,2}, \mathcal{C}_{3,4}$) = 19
diss($\mathcal{C}_{3,4}, \mathcal{C}_{3,5}$) = 20
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,2}$) = 25
diss($\mathcal{C}_{3,2}, \mathcal{C}_{3,5}$) = 26
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,4}$) = 27
diss($\mathcal{C}_{3,1}, \mathcal{C}_{3,3}$) = 28

Table 3.16: Complete linkage: the updated dissimilarity table of the fourth level.

diss($\mathcal{C}_{4,a}, \mathcal{C}_{4,b}$)
diss($\mathcal{C}_{4,2}, \mathcal{C}_{4,4}$) = 13
diss($\mathcal{C}_{4,1}, \mathcal{C}_{4,3}$) = 19
diss($\mathcal{C}_{4,1}, \mathcal{C}_{4,4}$) = 20
diss($\mathcal{C}_{4,2}, \mathcal{C}_{4,3}$) = 25
diss($\mathcal{C}_{4,3}, \mathcal{C}_{4,4}$) = 26
diss($\mathcal{C}_{4,1}, \mathcal{C}_{4,2}$) = 28

Table 3.17: Complete linkage: the updated dissimilarity table of the fifth level.

diss($\mathcal{C}_{5,a}, \mathcal{C}_{5,b}$)
diss($\mathcal{C}_{5,2}, \mathcal{C}_{5,3}$) = 19
diss($\mathcal{C}_{5,1}, \mathcal{C}_{5,2}$) = 26
diss($\mathcal{C}_{5,1}, \mathcal{C}_{5,3}$) = 28

Table 3.18: Complete linkage: the updated dissimilarity table of the sixth level.

diss($\mathcal{C}_{6,a}, \mathcal{C}_{6,b}$)
diss($\mathcal{C}_{6,1}, \mathcal{C}_{6,2}$) = 28

Exercise 3.4.3. Give an example of a 4-element set with different single-link and complete-link clusterings.

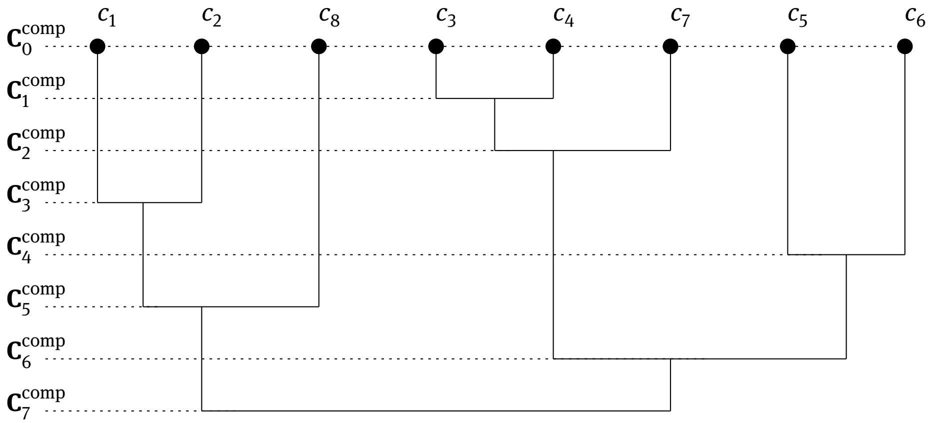


Figure 3.25: Complete linkage: the dendrogram for the model example.

3.5 Case study

In this section we apply the single-link algorithm developed in Sections 3.2–3.3 to a set of real data and use Pearson's correlation coefficient to assess the quality of the derived clustering.

Coffee-time browsing

– www.ucl.ac.uk/stats/department/pearson.html (Carl Pearson's biography and work)
– www.cmh.edu/stats/definitions/correlation.htm (What is correlation?)

Table 3.19 contains the final grades and GPA scores of 15 students in an Introductory Statistics class. The students s_1–s_{15} are listed in alphabetical order. The GPA scores were calculated earlier, so that they do not reflect the grades in this class. Using the final grades, we build single-link clusterings of this 15-element set and compare the results with the students' GPA scores. Our goal in doing this comparison is to assess the validity of the presented clustering algorithm. As a measure of dissimilarity, we have chosen the absolute value of the difference between the final grades, and used this measure to complete the dissimilarity table, Table 3.20.

Table 3.19: The final grades and GPA scores.

Student	s_1	s_2	s_3	s_4	s_5	s_6	s_7	
Final Grade	62	54	71	60	36	81	84	
GPA	1.808	2.369	3.058	2.825	2.460	3.681	3.508	
Student	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}
Final Grade	69	55	70	58	61	60	40	75
GPA	2.793	2.738	3.123	3.100	2.197	2.285	2.113	2.703

Table 3.20: The dissimilarity table for the Statistics Class grades.

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}	s_{13}	s_{14}	s_{15}
s_1	0	8	9	2	26	19	22	7	7	8	4	1	2	22	13
s_2		0	17	6	18	27	30	15	1	16	4	7	6	14	21
s_3			0	11	35	10	13	2	16	1	13	10	11	31	4
s_4				0	24	21	24	9	5	10	2	1	0	20	15
s_5					0	45	48	33	19	34	22	25	24	4	39
s_6						0	3	12	26	11	23	20	21	41	6
s_7							0	15	29	14	26	23	24	44	9
s_8								0	14	1	11	8	9	29	6
s_9									0	15	3	6	5	15	20
s_{10}										0	12	9	10	30	5
s_{11}											0	3	2	18	17
s_{12}												0	1	21	14
s_{13}													0	20	15
s_{14}														0	35

In this problem we have many ties, therefore, some intermediate steps are not unique, but it does not affect our conclusions. The agglomerative single-link algorithm (Section 3.2) gives the following results. There are two elements, whose dissimilarity is zero, thus the first-level clustering $\mathbf{C}_1$ contains one 2-element cluster $\{s_4, s_{13}\}$ and 13 1-element clusters. Next, if the threshold level does not exceed 1, we derive nine clusters of the second level,

$$\mathbf{C}_2 = \{\{s_1, s_4, s_{12}, s_{13}\}, \{s_3, s_8, s_{10}\}, \{s_2, s_9\}, \{s_5\}, \{s_6\}, \{s_7\}, \{s_{11}\}, \{s_{14}\}, \{s_{15}\}\}.$$

There are eight clusters at the next level,

$$\mathbf{C}_3 = \{\{s_1, s_4, s_{11}, s_{12}, s_{13}\}, \{s_3, s_8, s_{10}\}, \{s_2, s_9\}, \{s_5\}, \{s_6\}, \{s_7\}, \{s_{14}\}, \{s_{15}\}\}.$$

Next we have six clusters

$$\mathbf{C}_4 = \{\{s_1, s_2, s_4, s_9, s_{11}, s_{12}, s_{13}\}, \{s_3, s_8, s_{10}\}, \{s_6, s_7\}, \{s_5\}, \{s_{14}\}, \{s_{15}\}\}.$$

At the dissimilarity level of 4, there are only four clusters,

$$\mathbf{C}_5 = \{\{s_1, s_2, s_4, s_9, s_{11}, s_{12}, s_{13}\}, \{s_3, s_8, s_{10}, s_{15}\}, \{s_6, s_7\}, \{s_5, s_{14}\}\}.$$

There is no merger at level 5; however, two of these clusters amalgamate at the dissimilarity level of 6,

$$\mathbf{C}_6 = \{\{s_1, s_2, s_4, s_9, s_{11}, s_{12}, s_{13}\}, \{s_3, s_6, s_7, s_8, s_{10}, s_{15}\}, \{s_5, s_{14}\}\}.$$

At the level of 7 only two clusters remain,

$$\mathbf{C}_7 = \{\{s_1, s_2, s_3, s_4, s_6, s_7, s_8, s_9, s_{10}, s_{11}, s_{12}, s_{13}, s_{15}\}, \{s_5, s_{14}\}\}.$$

Ultimately, these two clusters amalgamate into conjoint clustering at the 14th level.

Now we want to assess the derived clustering. Table 3.21 represents three clusters in $\mathbf{C}_6$. Every chart contains the GPA scores of the students in the corresponding cluster.

Table 3.21: $\mathbf{C}_6$-clustering.

$\mathcal{C}_{6,1}$	**Student**	s_1	s_2	s_4	s_9	s_{11}	s_{12}	s_{13}
	GPA	1.808	2.369	2.825	2.738	3.100	2.197	2.285
$\mathcal{C}_{6,2}$	**Student**	s_3	s_6	s_7	s_8	s_{10}	s_{15}	
	GPA	3.058	3.681	3.508	2.793	3.123	2.703	
$\mathcal{C}_{6,3}$	**Student**	s_5	s_{14}					
	GPA	2.460	2.113					

The real data always have significant variability, thus there is no perfect match. However, we see that at this threshold level the clusters $\mathcal{C}_{6,2}$ and $\mathcal{C}_{6,3}$ demonstrate good uniformity of the GPA scores contained, while $\mathcal{C}_{6,1}$ shows larger variety of scores.

Next, consider the clustering $\mathbf{C}_5$, shown in Table 3.22. Again, we see that there is a noticeable closeness of the GPA scores within the clusters $\mathcal{C}_{5,2}, \mathcal{C}_{5,3}$, and $\mathcal{C}_{5,4}$. In particular, cluster $\mathcal{C}_{5,3}$ contains two highest GPA scores.

Table 3.22: $\mathbf{C}_5$-clustering.

$\mathcal{C}_{5,1}$	**Student**	s_1	s_2	s_4	s_9	s_{11}	s_{12}	s_{13}
	GPA	1.808	2.369	2.825	2.738	3.100	2.197	2.285
$\mathcal{C}_{5,2}$	**Student**	s_3	s_8	s_{10}	s_{15}			
	GPA	3.058	2.793	3.123	2.703			
$\mathcal{C}_{5,3}$	**Student**	s_6	s_7					
	GPA	3.681	3.508					
$\mathcal{C}_{5,4}$	**Student**	s_5	s_{14}					
	GPA	2.460	2.113					

To give a quantifiable assessment of the clusterings derived, we do some statistics. Tables 3.23–3.29 contain the averaged grades and the averaged GPA scores for each cluster at all levels. Finally, Table 3.30 contains Pearson's correlation coefficients for the GPA scores and averaged grades for every level of clustering. We see that every next level of clustering, except for $\mathbf{C}_4$, increases the correlation of the final grades and the GPA scores. This observation validates the clustering algorithm of Sections 3.2–3.3.

Table 3.23: The grades and GPA scores over the entire class—$\mathbf{C}_0$-clustering.

Clustering C_0	**Average grade**	**Average GPA score**
Cluster $\mathcal{C}_{0,1}$	62.4	2.717

Table 3.24: The grades and GPA scores over the entire class—$\mathbf{C}_1$-clustering.

Clustering C_1	**Average grade**	**Average GPA score**
Cluster $\mathcal{C}_{1,1}$	60	2.555
Cluster $\mathcal{C}_{1,2}$	62	1.808
Cluster $\mathcal{C}_{1,3}$	54	2.369
Cluster $\mathcal{C}_{1,4}$	71	3.058
Cluster $\mathcal{C}_{1,5}$	36	2.460
Cluster $\mathcal{C}_{1,6}$	81	3.681
Cluster $\mathcal{C}_{1,7}$	84	3.508
Cluster $\mathcal{C}_{1,8}$	69	2.793
Cluster $\mathcal{C}_{1,9}$	55	2.738
Cluster $\mathcal{C}_{1,10}$	70	3.123
Cluster $\mathcal{C}_{1,11}$	58	3.100
Cluster $\mathcal{C}_{1,12}$	61	2.197
Cluster $\mathcal{C}_{1,13}$	40	2.113
Cluster $\mathcal{C}_{1,14}$	75	2.703

Table 3.25: The grades and GPA scores over the entire class—$\mathbf{C}_2$-clustering.

Clustering C_2	**Average grade**	**Average GPA score**
Cluster $\mathcal{C}_{2,1}$	60.75	2.279
Cluster $\mathcal{C}_{2,2}$	70	2.991
Cluster $\mathcal{C}_{2,3}$	54.5	2.554
Cluster $\mathcal{C}_{2,4}$	36	2.460
Cluster $\mathcal{C}_{2,5}$	81	3.681
Cluster $\mathcal{C}_{2,6}$	84	3.508
Cluster $\mathcal{C}_{2,7}$	58	3.100
Cluster $\mathcal{C}_{2,8}$	40	2.113
Cluster $\mathcal{C}_{2,9}$	75	2.703

Table 3.26: The grades and GPA scores over the entire class—$\mathbf{C}_3$-clustering.

Clustering C_3	**Average grade**	**Average GPA score**
Cluster $\mathcal{C}_{3,1}$	60.2	2.443
Cluster $\mathcal{C}_{3,2}$	70	2.991
Cluster $\mathcal{C}_{3,3}$	54.5	2.554
Cluster $\mathcal{C}_{3,4}$	36	2.460
Cluster $\mathcal{C}_{3,5}$	81	3.681
Cluster $\mathcal{C}_{3,6}$	84	3.508
Cluster $\mathcal{C}_{3,7}$	40	2.113
Cluster $\mathcal{C}_{3,8}$	75	2.703

Table 3.27: The grades and GPA scores over the entire class—$\mathbf{C}_4$-clustering.

Clustering C_4	Average grade	Average GPA score
Cluster $\mathcal{C}_{4,1}$	58.57	2.475
Cluster $\mathcal{C}_{4,2}$	70	2.991
Cluster $\mathcal{C}_{4,3}$	82.5	3.594
Cluster $\mathcal{C}_{4,4}$	36	2.460
Cluster $\mathcal{C}_{4,5}$	40	2.113
Cluster $\mathcal{C}_{4,6}$	75	2.703

Table 3.28: The grades and GPA scores over the entire class—$\mathbf{C}_5$-clustering.

Clustering C_5	Average grade	Average GPA score
Cluster $\mathcal{C}_{5,1}$	58.57	2.475
Cluster $\mathcal{C}_{5,2}$	71.25	2.919
Cluster $\mathcal{C}_{5,3}$	82.5	3.595
Cluster $\mathcal{C}_{5,4}$	38	2.287

Table 3.29: The grades and GPA scores over the entire class—$\mathbf{C}_6$-clustering.

Clustering C_6	Average grade	Average GPA score
Cluster $\mathcal{C}_{6,1}$	58.57	2.475
Cluster $\mathcal{C}_{6,2}$	75	3.144
Cluster $\mathcal{C}_{6,3}$	38	2.287

Table 3.30: Pearson's coefficient of correlation.

Clustering level	C_0	C_1	C_2	C_3	C_4	C_5	C_6
Correlation Coefficient	0.659	0.671	0.788	0.850	0.832	0.925	0.929

Part II: **Combinatorial analysis**

4 Enumerative combinatorics

The methods developed in this chapter allow us to solve more advanced problems with the same question "How many?". Section 4.1 treats the inclusion–exclusion principle. Inversion formulas, including the Möbius inversion and their applications, are studied in Section 4.2. Generating functions are considered in Sections 4.3–4.4, and Section 4.5 is devoted to the Pólya–Redfield enumeration theory.

4.1 The inclusion–exclusion principle

Coffee-time browsing

- http://mathworld.wolfram.com/Inclusion-ExclusionPrinciple.html (Inclusion-Exclusion Principle)
- http://www-history.mcs.st-and.ac.uk/Mathematicians/Eratosthenes.html (Eratosthenes' biography)
- www.math.utah.edu/~pa/Eratosthenes.html (Sieve of Eratosthenes)
- http://www.1911encyclopedia.org/James_Stirling (Stirling's biography)
- mathworld.wolfram.com/Derangement.html (Derangements)
- http://en.wikipedia.org/wiki/Eric_Temple_Bell (Bell's biography)
- planetmath.org/encyclopedia/BellNumber.html (Bell numbers)
- http://en.wikipedia.org/wiki/Carlo_Emilio_Bonferroni (Bonferroni)
- www.answers.com/topic/bonferroni-inequality (Bonferroni inequalities)
- http://en.wikipedia.org/wiki/John_Napier (Napier (Neper) biography)
- http://www.absoluteastronomy.com/topics/Neper (What is neper?)
- http://www-history.mcs.st-and.ac.uk/Mathematicians/Maclaurin.html (Maclaurin's biography)
- http://en.wikipedia.org/wiki/Perfect_totient_number (Totient numbers)

Problem 4.1.1. Each member of the Combi Club plays at least one game, 5 students go to football, 12 to basketball, and 8 to volleyball. How many members are there in the club?

Solution. Denote by S_f, S_b, and S_v the sets of students who play, respectively, football, basketball, and volleyball, and by S the entire membership of the club. Obviously, $|S_f| = 5, |S_b| = 12, |S_v| = 8$ and $S = S_f \cup S_b \cup S_v$. But we cannot apply the sum rule, since a member can play two or three games and the subsets S_f, S_b, S_v do not have to be disjoint.

Unless we have some additional information, this problem has several solutions. For instance, if each student participates in one and only one sport, then the three subsets are mutually disjoint and by the sum rule we have $|S| = 5+12+8 = 25$. However, if all five football players and all eight volleyball players also play basketball, then the

https://doi.org/10.1515/9783110751185-004

entire membership consists of only 12 students. Since $5 + 8 > 12$, the latter option would necessarily imply that at least one club member plays all three games. Thus, if we have no more information, we must conclude that the quantity of club members satisfy the bilateral inequality $12 \le |S| \le 25$ and we cannot say anything more. □

Hence, if we want to make a more specific conclusion, we need certain additional information about intersections of the sets given. The following assertion, which involves these quantities, is equation (1.1.4), which was proved in Section 1.1. We state it here again. Hereafter it is referred to as the *Inclusion–Exclusion Principle*. It is also called the *(Eratosthenes) Sieve Formula*.

Theorem 4.1.1. *If $X_i, 1 \le i \le k$, are finite sets, then*

$$
\begin{aligned}
&|X_1 \cup X_2 \cup \cdots \cup X_k| \\
&\quad = |X_1| + |X_2| + \cdots + |X_k| - |X_1 \cap X_2| - \cdots - |X_{k-1} \cap X_k| \\
&\qquad + |X_1 \cap X_2 \cap X_3| + \cdots + (-1)^{k-1}|X_1 \cap X_2 \cap \cdots \cap X_k|.
\end{aligned} \tag{4.1.1}
$$

□

The right-hand side of this equation consists of k groups of terms. The first group contains the cardinalities of the k given sets. The second group contains $C(k, 2)$ cardinalities of their pair-wise intersections, and all the terms in this group have negative signs. The third group contains $C(k, 3)$ cardinalities of the triple intersections of the sets X_i with the plus sign, and so forth. Since $C(k, k) = 1$, the last group contains one term $(-1)^{k-1}|X_1 \cap X_2 \cap \cdots \cap X_k|$. If all sets X_i are pair-wise disjoint, then the cardinal numbers of all intersections are zero and (4.1.1) reduces to the sum rule (1.2.1). Now let us modify Problem 4.1.1.

Problem 4.1.2. Each member of the Combi Club plays at least one game, 5 people play football, 12 basketball, and 8 play volleyball. In addition, this time we know that two of them are devoted to both football and volleyball, three members go to football and basketball, and four play basketball and volleyball. The best mathematician in the club plays all three sports. How many members are there in the club? How many among them play only volleyball?

Solution. As before, denote by S_f, S_b and S_v the sets of students who play, respectfully, football, basketball, and volleyball, $|S_f| = 5, |S_b| = 12, |S_v| = 8$, and again $S = S_f \cup S_b \cup S_v$. However, now we are given the cardinalities of all the terms in formula (4.1.1) with $k = 3$, and we can straightforwardly apply the inclusion–exclusion principle (4.1.1) with $k = 3$:

$$
\begin{aligned}
|S| &= |S_f| + |S_b| + |S_v| - |S_b \cap S_f| - |S_b \cap S_v| - |S_f \cap S_v| + |S_b \cap S_f \cap S_v| \\
&= 12 + 5 + 8 - 3 - 4 - 2 + 1 = 17.
\end{aligned}
$$

Actually Theorem 4.1.1 contains more information. Thus, applying the same formula (4.1.1) with $k = 2$, we have

$$|(S_b \cap S_v) \cup (S_f \cap S_v)| = |S_b \cap S_v| + |S_f \cap S_v| - |S_b \cap S_f \cap S_v|$$
$$= 4 + 2 - 1 = 5.$$

Therefore, among eight volleyball players five people also play either basketball or football, so $8 - 5 = 3$ students play only volleyball and no other game. □

We leave it to the reader to solve the two following problems.

Problem 4.1.3. In the same club, how many members play only football? Only basketball?

Problem 4.1.4. Each member of the Combi Club participates in at least one sport, 5 students go to football, 12 to basketball and 8 to volleyball. Which additional information should we have to ensure that the club has precisely 13 members? Or exactly 14, or 15, ..., or 24 members?

Problem 4.1.5. How many n-arrangements with repetition from the elements of the set $A = \{0, 1, 2, 3\}$ contain at least one digit 1, at least one digit 2, and at least one digit 3?

Solution. With no restriction, there are 4^n n-arrangements with repetition. Among them there are 3^n arrangements from the set $A_0 = \{1, 2, 3\}$, that is, the arrangements which certainly do not contain 0 and maybe do not contain some other digits either. Similarly, there are 2^n arrangements not containing two digits and there is $1^n = 1$ arrangement consisting only of 0s. Now, by (4.1.1) there are $4^n - 3 \cdot 3^n + 3 \cdot 2^n - 1$ n-arrangements satisfying the problem. For example, if $n = 3$ then there are $4^3 - 3 \cdot 3^3 + 3 \cdot 2^3 - 1 = 6 = 3!$ arrangements, which is clear since all eligible 3-arrangements are precisely the permutations of the three-element set $A_0 = \{1, 2, 3\}$. If $n = 2$ or $n = 1$, $4^n - 3 \cdot 3^n + 3 \cdot 2^n - 1 = 0$, which is also obvious since any arrangement in question must contain at least three numbers 1, 2, 3. □

Theorem 4.1.1 can be stated in other terms; see, for instance, [26, p. 18]. First we introduce some notation. Let q properties P_i, $1 \le i \le q$, be defined for the elements of a finite set X, $|X| = n < \infty$, that is, each element $x \in X$ either possesses or does not possess the property P_i for every i, $1 \le i \le q$. If an element x possesses the property P_i, we denote this by $P_i(x) = 1$; otherwise, if x does not possess this property, we write $P_i(x) = 0$. Therefore, these properties are mappings $P_i : X \to \{0, 1\}$. Let X_i be the subset of X, whose elements have the property P_i, thus these sets are the total preimages of 1, $X_i = P_i^{-1}(\{1\})$, and $n_i = |X_i|$ be its cardinal number. Let $X_{i,j}$ be the subset of X, whose elements possess both properties P_i and P_j, and $n_{i,j} = |X_{i,j}|$ be its cardinal number, etc. Let $X_0 = X \setminus \{X_1 \cup \cdots \cup X_n\}$ be the subset of all elements of X possessing no property P_i, $1 \le i \le q$, and $n_0 = |X_0|$.

The following equation is also called the *Sieve Formula*.

Theorem 4.1.2.

$$n_0 = n - \sum_{i=1}^{q} n_i + \sum_{i_1<i_2} n_{i_1,i_2} - \cdots + (-1)^s \sum_{i_1<i_2<\cdots<i_s} n_{i_1,i_2,\ldots,i_s} + \cdots + (-1)^q n_{1,2,\ldots,q}. \quad (4.1.2)$$

Proof. The conclusion follows immediately from Theorem 4.1.1 being applied to the set $X = X_0 \cup X_1 \cup \cdots \cup X_n$, if we notice that X_0 is disjoint with every set $X_i, 1 \le i \le q$. □

If all sets X_i have equal cardinalities, all sets $X_{i,j}$ also have equal cardinalities, all sets $X_{i,j,k}$ have equal cardinalities, and so forth, then (4.1.2) can be simplified.

Corollary 4.1.1. *If* $n_i = n_1^*, \forall i$, $n_{i,j} = n_2^*, \forall i,j, \ldots, n_{i_1,i_2,\ldots,i_s} = n_s^*$ *for all s-tuples of subscripts, and so forth for* $1 \le s \le q$, *then*

$$n_0 = n - qn_1^* + C(q,2)n_2^* - \cdots + (-1)^s C(q,s)n_s^* + \cdots + (-1)^q n_q^*.$$

This corollary immediately implies the next one.

Corollary 4.1.2. *If* $|X| = n \ge |Y| = m$, *then there are*

$$|\operatorname{Sur}(Y^X)| = m^n - C(m,1)(m-1)^n + \cdots + (-1)^{m-1} C(m,m-1)$$

surjective mappings from X to Y. In particular, if m = n, then a surjective mapping is simultaneously injective and therefore bijective, thus, $|\operatorname{Sur}(Y^X)| = n!$ *and the latter equation becomes*

$$n! = \sum_{k=0}^{n} (-1)^k C(n,k)(n-k)^n.$$

Now we can easily solve the following important problem.

Problem 4.1.6. In how many ways is it possible to place n different balls into m different urns with no urn left empty?

Solution. Considering Definition 1.1.4 of the preimage of an element, the answer is given by the number $|\operatorname{Sur}(Y^X)|$ with $|X| = n \ge |Y| = m$. On the other hand, if all urns are indistinguishable, then any permutations of the urns without changes in their enclosures lead to the same placement of different balls, hence there are

$$S_2(n,m) = \frac{1}{m!}|\operatorname{Sur}(Y^X)| \quad (4.1.3)$$

ways to put n different balls into m identical urns with no empty urn. □

Definition 4.1.1. The numbers $S_2(n,m)$ are called *Stirling numbers of the second kind*. These numbers also count partitions of sets, namely, the number $S_2(n,m)$ is equal to the number of partitions of an n-element set with m nonempty parts, if the order of parts is immaterial.

Properties of the Stirling numbers of the second kind are discussed in Exercises 4.1.22, 4.1.23, 4.2.3, more properties of these numbers can be found in [35]. *The Stirling numbers of the first kind*, $S_1(n, m)$, are defined in the end of this section.

Problem 4.1.7. In how many ways is it possible to paint four walls of a room in three colors, so that any two adjacent walls have different colors?

Solution. There are $A_{\text{rep}}(3, 4) = 3^4$ ways to paint the walls without any restriction. Let us enumerate the corners of the room by digits from 1 through 4 consecutively, starting at any fixed corner in any direction, and consider the following properties $P_i, i = 1, 2, 3, 4$, on the set of all possible colorings:

A coloring has a property $P_i, 1 \le i \le 4$, if two walls adjacent at the ith corner are of the same color.

Then $n_1 = n_2 = n_3 = n_4 = 3 \cdot 3^2$, $n_{1,3} = n_{2,4} = 3 \cdot 3$, $n_{1,2} = n_{2,3} = n_{3,4} = n_{1,4} = 3 \cdot 3$, $n_{1,2,3} = n_{1,2,4} = n_{1,3,4} = n_{2,3,4} = 3$, and $n_{1,2,3,4} = 3$; these parameters were defined before Theorem 4.1.2. Due to (4.1.2) we get

$$3^4 - 4 \cdot 3 \cdot 3^2 + (4 \cdot 3 \cdot 3 + 2 \cdot 3 \cdot 3) - 4 \cdot 3 + 3 = 18$$

different colorings. Similarly, if only two colors are available, there are $16 - 32 + 24 - 8 + 2 = 2$ colorings; though this is clear without calculations. □

Remark 4.1.1. Any coloring in this problem can be represented by a plane graph with four vertices, corresponding to the four walls, such that two vertices are adjacent if and only if they correspond to the neighboring walls. Now Problem 4.1.7 can be stated as a graph coloring problem:

In how many ways is it possible to color the vertices of a cycle of length four in three colors so that any pair of adjacent vertices has different colors?

The inclusion–exclusion principle leads to useful inequalities involving the cardinal numbers of subsets. Consider again equation (4.1.1) with $k = 3$ and $X = X_1 \cup X_2 \cup X_3$. In this case (4.1.1) becomes

$$|X| = |X_1| + |X_2| + |X_3| - |X_1 \cap X_2| - |X_2 \cap X_3| - |X_1 \cap X_3| + |X_1 \cap X_2 \cap X_3|.$$

It is obvious from this that

$$|X| \ge |X_1| + |X_2| + |X_3| - |X_1 \cap X_2| - |X_2 \cap X_3| - |X_1 \cap X_3|.$$

Moreover, since

$$|X_1 \cap X_2| + |X_2 \cap X_3| + |X_1 \cap X_3| \ge |X_1 \cap X_2 \cap X_3|,$$

we clearly have an opposite bound

$$|X| \le |X_1| + |X_2| + |X_3|.$$

Similar inequalities can be derived for any k, therefore, the inclusion–exclusion principle produces a series of alternating upper and lower bounds for $|X|$.

Problem 4.1.8 (Bonferroni's inequalities). Let $X_1, \ldots, X_n$ be nonempty subsets of a finite set X, and T be any subset of $\{1, 2, \ldots, n\}$. Denote

$$\mu = \max\{|T| : T \subset \{1, 2, \ldots, n\} \text{ such that } \cap_{i\in T} X_i \neq \emptyset\}$$

and for $l = 1, 2, \ldots, \mu - 1$,

$$\Delta_l = |\bigcup_{i=1}^{n} X_i| - \sum_{j=1}^{l} (-1)^{j-1} \Big(\sum_{T\subset\{1,\ldots,n\},|T|=j} |\bigcap_{i\in T} X_i| \Big).$$

Then $(-1)^l \Delta_l \geq 0$.

It is instructive to verify these inequalities in some simple case, for example, if $n = 3$, $X = \{1, 2, 3\}$, $X_1 = \{1\}$, $X_2 = \{1, 2\}$, and $X_3 = \{1, 2, 3\}$.

Theorem 4.1.2 can be further generalized if we supply the elements of the set X with *weights*. Consider a mapping $w : X \longrightarrow W$, where a set W can be specified in some convenient way. The image $w(x) \in W$ of an element $x \in X$ is called the *weight* of x. To give an example of weights, let us suppose that X is the inventory of all items in a store. Then the price of any item $x \in X$ can be viewed as the weight of x, and the mapping w assigns to any item x in stock its price $w(x)$. For the time being we do not need any rich algebraic structure on W, it is enough to assume that we can add elements of W, and the operation of addition on W is commutative and associative. We define also the quantities

$$w(P_{i_1}, P_{i_2}, \ldots, P_{i_r}) = \sum_{x\in X_{i_1}\cap X_{i_2}\cap\cdots\cap X_{i_r}} w(x)$$

where the properties P_i and the sets X_i were defined after Problem 4.1.5.

Set also

$$w(r) = \sum w(P_{i_1}, P_{i_2}, \ldots, P_{i_r})$$

where the sum runs over all r-element subsets of the set of properties $\{P_1, \ldots, P_q\}$, that is, $w(r)$ is the sum of the weights of elements possessing *at least* r properties. Moreover, let $w(0) = \sum_{x\in X} w(x)$ and $E(r)$ be the sum of weights of those elements of X that have *exactly* r properties.

Theorem 4.1.3. *In these notations*

$$E(r) = w(r) - C(r+1, r)w(r+1) + \cdots + (-1)^{q-r} C(q, r)w(q)$$

for $0 \leq r \leq q$. *In particular, if* $r = 0$, *then*

$$E(0) = w(0) - w(1) + \cdots + (-1)^q w(q),$$

which implies (4.1.2) *if we choose here all weights* $w(x) \equiv 1$. □

Next we consider a classical application of (4.1.2), called the *derangement problem*. For many other applications of these results see, for instance, [49].

Definition 4.1.2. A permutation $(a_1, \ldots, a_n)$ of the natural segment $\mathbf{N}_n = \{1, 2, \ldots, n\}$ is called the *derangement* if $a_i \neq i$ for all $i = 1, 2, \ldots, n$.

Problem 4.1.9. How many derangements are there among all $n!$ permutations of the natural segment $\mathbf{N}_n = \{1, 2, \ldots, n\}$?

Solution. Denote the number of derangements by D_n and let P_i be the property $a_i = i, 1 \leq i \leq n$, defined on all n-permutations. By Theorem 4.1.2, we get immediately

$$\begin{aligned} D_n &= n! - C(n,1)(n-1)! + C(n,2)(n-2)! - \cdots + (-1)^n C(n,n) \\ &= n!\left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}\right). \end{aligned} \tag{4.1.4}$$

Quite similarly, the number $D_n(r)$ of n-permutations possessing *exactly* r out of n properties $P_1, \ldots, P_n$ is equal to $D_n(r) = C(n,r)D_{n-r}$, thus,

$$D_n(r) = \frac{n!}{r!}\left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^{n-r}}{(n-r)!}\right).$$

It is obvious that $D_n = D_n(0)$. □

Remark 4.1.2. The factorials satisfy recurrence relations $n! = n(n-1)!$ and $n! = (n-1)[(n-1)! + (n-2)!]$. The derangement numbers D_n satisfy similar recurrence relations

$$D_n = nD_{n-1} + (-1)^n = (n-1)[D_{n-1} + D_{n-2}]; \tag{4.1.5}$$

they are called *subfactorials*.

Problem 4.1.10. Prove the recurrence relations (4.1.5).

Remark 4.1.3. Expanding[1] e^{-1} in the Maclaurin series and using the known property of alternating series with monotone decreasing terms [52, p. 607], we get an estimate $|D_n - n!/e| < \frac{1}{n+1}$, that is, for any $n \geq 2$ the derangement D_n can be defined as the nearest integer to $n!/e$, since $\frac{1}{n+1} < \frac{1}{2}$ whenever $n > 2$.

Problem 4.1.11. In how many ways is it possible to place eight rooks on the chessboard, so that none of them could attack another and none of them are on the main white diagonal?

Solution. Formula (4.1.4) immediately gives the answer, $D_8 = 14\,833$; we note that $8!/e \approx 14\,832.9$ □

1 Surely, here $e \approx 2.718281828$ is Napier's number, the base of natural logarithms.

Problem 4.1.12. The Combi Club bought $2n$ tickets and reserved n seats for two ball games. The n tickets to the first game were distributed at random among n students. Then the n tickets to the second game were distributed, also at random, among the same n students. In how many ways is it possible to distribute these $2n$ tickets, so that no student gets the same seat twice?

Solution. Without any restriction, the tickets can be distributed in $(n!)^2$ ways. However, if we want to avoid repetitions of same-seat tickets, the second distribution can be done in D_n ways. Formula (4.1.4) and the product rule result in

$$n! \cdot D_n = n! \cdot \left\{ n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} \right) \right\} \sim \frac{1}{e}(n!)^2$$

different ways to distribute the tickets. □

As another application of Theorem 4.1.2, we again derive two formulas that were proven in Problems 1.4.18–1.4.19.

Problem 4.1.13. Demonstrate formulas

$$2^n = C(n,0) + C(n,1) + \cdots + C(n,n)$$

and

$$1 = C(n,0)2^n - C(n,1)2^{n-1} + \cdots + (-1)^n C(n,n)2^0.$$

Solution. Let us paint a ball using n different colors, not all of which have to be used. Let a property P_i mean that the ith color is applied at the ball. Then $n_0 = 1$, since there is only one way not color the ball at all. Moreover, $n_{i_1,i_2,\ldots,i_k} = 2^{n-k}, 1 \leq k \leq n$, and the total number of colorings is 2^n. Now the second formula follows directly from (4.1.2).

However, if P_i means the property that a coloring contains precisely i colors, then $n_i = C(n,i), n_{i,j} = n_{i,j,k} = \cdots = 0$, and (4.1.2) implies the first formula in the problem. □

Next we apply Theorem 4.1.2 to another graph coloring problem. Given a graph $G = (V,E)$ of order p, we want to paint its p vertices in χ given colors.

Definition 4.1.3. A *coloring* is called *regular* if the end vertices of every edge have different colors. A graph G is called *k-chromatic* if its vertices can be regularly colored in k colors. The smallest such a number is called the *chromatic number* $\chi = \chi(G)$ of the graph G. The number of various regular colorings of a graph G in k colors is denoted by $\pi(G,k)$.

Theorem 1.1.6 immediately implies that for a graph of order p there are k^p, not necessarily regular, k-colorings. To exclude non-regular colorings from this number, we denote by $\mu(e_\alpha, e_\beta, \ldots, e_\delta)$ the number of colorings such that the end vertices of the edge e_α have the same color, the end vertices of e_β also have the same color (maybe different from the color of e_α), etc. Now Theorem 4.1.2 immediately yields the first statement of the following result.

Theorem 4.1.4.

(1) *For any simple graph G of order p*

$$\pi(G,k) = k^p - \sum_{\alpha} \mu(e_\alpha) + \sum_{\alpha\neq\beta} \mu(e_\alpha, e_\beta) - \sum_{\alpha\neq\beta\neq\gamma\neq\alpha} \mu(e_\alpha, e_\beta, e_\gamma) + \cdots . \tag{4.1.6}$$

(2) *If G is a tree, then* $\pi(G,k) = k(k-1)^{p-1}$.

Proof. To prove (2), we notice that the size of the tree G is $p-1$, thus there are $p-1$ ways to select an edge whose end vertices have the same color, and this color can be chosen in k ways. After that we can paint $p-2$ remaining vertices in any of k colors, hence $\sum_\alpha \mu(e_\alpha) = k^{p-2}k(p-1)$. Similarly, $\sum_{\alpha\neq\beta} \mu(e_\alpha, e_\beta) = k^{p-2}(p-1)(p-2)/2$, etc. Substituting these expressions in (4.1.2) and using the binomial expansion (1.4.4) we complete the proof. □

Exercises and Problems 4.1.

Exercise 4.1.1. Consider the following properties on the set of the first 13 whole numbers $S = \{0, 1, 2, \ldots, 12\}$.

P_1: *a number* $x \in S$ *is a multiple of 5 or* $x = 0$

P_2: *a number* $x \in S$ *is a multiple of 11 or* $x = 0$

P_3: *a number* $x \in S$ *is a multiple of 10 or* $x = 0$

P_4: *a number* $x \in S$ *and* $x^2 + x > 5$.

How many elements of the set S satisfy the following properties?

(1) $P_1 \vee P_2 \vee P_3$

(2) $P_1 \wedge P_2 \wedge \overline{P_3}$

(3) $P_1 \wedge P_2 \wedge P_3$

(4) $P_1 \vee (P_2 \wedge P_3)$.

Exercise 4.1.2. Prove the following modifications of the results of this section.

(1) If $X_1, \ldots, X_n$ are subsets of a finite set X and $\overline{Y} = X \setminus Y$ is the complement of Y with respect to X, then

$$\left|\bigcap_{k=1}^{n} \overline{X_k}\right| = |X| + \sum_{\emptyset\neq I\subset\{1,2,\ldots,n\}} (-1)^{|I|} \left|\bigcap_{k\in I} X_k\right|,$$

where the index I runs over all the nonempty subsets of the set $\{1, 2, \ldots, n\}$.

(2) Consider finite sets $X_1, \ldots, X_n$, their union $X = \cup_{k=1}^{n} X_k$, and a function $f : X \to \mathbf{R}$. For any set $Y \subset X$ define $f(Y) = \sum_{x\in Y} f(x)$ and $f(\emptyset) = 0$. Prove the equation

$$f(X) = \sum_{I\neq\emptyset} (-1)^{|I|+1} f(\cap_{k\in I} X_k).$$

Exercise 4.1.3.
(1) If the largest among 66 consecutive odd integers is 213, what is the smallest?
(2) If the largest among several consecutive positive odd integers is 213, what is the largest possible length (that is, the number of elements) of this sequence? Answer the same question if the sequence can contain negative numbers.

Exercise 4.1.4. What is the cardinality of the union of five sets if their cardinalities are, respectively, 17, 23, 41, 45, and 56, each pair of the sets contains six elements, every triple of the sets contains four elements, and any four sets are mutually disjoint?

Exercise 4.1.5. At The Top-Rate College, a_1 students received at least one F grade during a semester, a_2 students received at least two F grades, ..., a_l students received at least l F grades during this semester, while no student had more than l F grades. How many F grades have all the students received during this semester?

Exercise 4.1.6. How many prime numbers do not exceed 300?

Exercise 4.1.7. Prove that there are $[a/n]$ natural numbers not exceeding a and divisible by n, where $[x]$ is the integer part of the number x.

Exercise 4.1.8. How many natural numbers less than 777 are not divisible by 3, by 7, and by 11? How many are not divisible by 4 and by 6?

Exercise 4.1.9. How many seven-digit telephone numbers contain each of the digits 1 and 9 at least once?

Exercise 4.1.10. A paper reports that among 1000 people surveyed, 800 have driver licenses, 750 are from 20 through 30 years old, and 500 have never had a ticket for speeding, while 450 are from 20 through 30 years old and have never had a ticket for speeding. Are these data consistent?

Exercise 4.1.11. Consider three n-families of sets

$$\{A_1, \dots, A_n\},\ \{B_1, \dots, B_n\},\ \{C_1, \dots, C_n\}$$

such that

$$A_1 \cup \dots \cup A_n = B_1 \cup \dots \cup B_n = C_1 \cup \dots \cup C_n \overset{\text{def}}{=} M$$

and

$$|A_i \cap B_j| + |A_i \cap C_k| + |B_j \cap C_k| \geq n, \quad \forall i, j, k.$$

Prove that $|M| \geq \frac{n^3}{3}$.

Exercise 4.1.12. Let F be a forest of order p and size q. Prove that

$$\pi(F, k) = k^{p-q}(k-1)^q.$$

Exercise 4.1.13. A gentleman had 11 daughters. If any girl got married while at least one of her older sisters remained unwed, then these still non-married but older sisters approached their father crying and complaining so bitterly that he had to double their dowry. In how many ways could these sisters arrange their weddings if the gentleman remarked that when the last his daughter got married, he had to double the dowry 11 times?

Exercise 4.1.14. There are 5 people and $n \geq 5$ different pairs of gloves. In how many ways can each of these people choose a right glove and a left one so that no one gets a complete pair of gloves?

Exercise 4.1.15. Ten couples are dining at a round table. In how many ways can they be seated so that no two males, no two females, and no two spouses are sitting alongside?

Exercise 4.1.16. In how many ways can we roll a fair die 12 times, so that a 1 never appears after another 1?

Exercise 4.1.17. The membership of the Combi Club comprises 30 students of five majors. Together they composed 40 problems for the Math Fair. Any two students of the same major composed equal number of problems, while any two students majoring in different subjects composed different number of problems. How many students composed only one problem?

Exercise 4.1.18. Solve again Problem 4.1.7 if we have to paint not only the walls, but also the floor and the ceiling of the room; consider two different cases, if there are three or only two paints available.

Exercise 4.1.19. Prove that, for any $k = 1, 2, \ldots, 9$, there are

$$k^n - C(k,1)(k-1)^n + C(k,2)(k-2)^n - \cdots + (-1)^{k-1}C(k,k-1)$$

n-digit numbers consisting only of the digits $1, 2, \ldots, k$.

Exercise 4.1.20. The quantity of natural numbers that do not exceed a natural number n and are mutually prime with n, is denoted by $\phi(n)$ and is called the Euler (totient) function; $\phi(1) = 1$ by definition.

(1) Evaluate $\phi(2)$, $\phi(3)$, $\phi(4)$, $\phi(5)$, $\phi(6)$, $\phi(7)$.
(2) Prove that

$$\phi(n) = n \prod_{k=1}^{m} \left(1 - \frac{1}{p_k}\right)$$

where $p_1, \ldots, p_m$ are all prime factors of n. Other properties of the totient function are considered in problem Exercise 4.2.2.

Exercise 4.1.21. Use Corollary 4.1.2 to prove that the Stirling numbers of the second kind $S_2(n,m)$ give the number of ways an n-element set can be partitioned into m subsets.

Exercise 4.1.22.

(1) Prove that the Stirling numbers of the second kind satisfy the recurrence relation

$$S_2(n,m) = S_2(n-1,m-1) + mS_2(n-1,m)$$

assuming $S_2(n-1,n) = 0$ and the initial conditions $S_2(n,0) = 0, \forall n$.

(2) Compute $S_2(n,m)$ for $1 \le m \le n \le 4$.

(3) Prove that

$$S_2(n,m) = \frac{1}{m!}\sum \frac{n!}{k_1!k_2!\cdots k_m!}$$

where the sum runs over all positive integer solutions of the equation $k_1 + k_2 + \cdots + k_m = n$.

(4) Verify that $S_2(n,n-1) = C(n,2)$.

(5) Prove that $k^n = \sum_{m=1}^{n} S_2(n,m)(k)_m$, where

$$(k)_m = k(k-1)(k-2)\cdots(k-m+1) = \frac{k!}{(k-m)!}, \quad k \ge n.$$

The latter equation can be written as

$$k^n = \sum_{m=1}^{n} S_2(n,m)C(k,m)m!. \tag{4.1.7}$$

Formula (4.1.7) represents powers of natural numbers through the binomial coefficients, and the Stirling numbers of the second kind.

Exercise 4.1.23.

(1) Let $|X| = 3$ and $|Y| = 4$. How many functions $f : X \to Y$ are injective but not surjective? Surjective but not injective ? Neither injective nor surjective?

(2) Answer the same question if $|X| = 4$ and $|Y| = 3$.

(3) Answer the same question if $|X| = |Y| = 4$.

Definition 4.1.4. The *Stirling numbers of the first kind* $S_1(n,m)$ are the coefficients in the inversion of formula (4.1.7), which represents the binomial coefficients through the powers:

$$n!C(k,n) = \sum_{m=1}^{n} (-1)^{n-m} S_1(n,m)k^m.$$

Exercise 4.1.24. Prove that the number $B(n,m) = \sum_{k=0}^{m} S_2(n,k)$ is the number of placements of n different balls into m indistinguishable urns with empty urns allowed.

Exercise 4.1.25. Compute $\chi(K_2), \chi(K_3), \chi(K_4), \chi(K_{1,2}), \chi(K_{2,3}), \chi(K_{3,3})$, where $\chi(G)$ is the chromatic number of a graph G—see Definition 4.1.3.

4.2 Inversion formulas

In this section we derive inversion formulas such as the Möbius inversion, and apply these results to enumeration of cyclic sequences and bracelets. Other families of inversion formulas are considered in Theorem 4.3.3 and problems thereafter.

Coffee-time browsing

- www.gap-system.org/~history/Biographies/Mobius.html (Mobius' biography)
- www.cut-the-knot.org/do_you_know/moebius.shtml (Mobius strip)
- math.about.com/library/blfermatbio.htm (Fermat's biography)
- www.gap-system.org/~history/.../Fermat's_last_theorem.html (Fermat Last Theorem)
- www.gap-system.org/~history/Biographies/Gauss.html (Carl Gauss, Prince of Mathematics)

Let us revisit Theorem 4.1.1, denoting, for the sake of brevity,

$$X_{1,2} = X_1 \cap X_2, X_{1,2,3} = X_1 \cap X_2 \cap X_3$$

etc. Let X_i^* denote the subset of elements of X possessing *only* the property $P_i, 1 \le i \le q$, $X_{i,j}^*$ denote the set of elements having *exactly two* properties P_i, P_j, and so on. It is clear that we can represent X_1 as

$$X_1 = X_1^* \cup X_{1,2}^* \cup \cdots \cup X_{1,q}^* \cup X_{1,2,3}^* \cup \cdots \cup X_{1,2,\dots,q}^*$$

where all sets on the right are pair-wise disjoint. Consequently,

$$|X_1| = |X_1^*| + |X_{1,2}^*| + \cdots + |X_{1,q}^*| + |X_{1,2,3}^*| + \cdots + |X_{1,2,\dots,q}^*|. \tag{4.2.1}$$

At the same time we can write

$$X_1 = X_1^* \cup X_{1,2} \cup \cdots \cup X_{1,q} \cup X_{1,2,3} \cup \cdots \cup X_{1,2,\dots,q}.$$

Applying Theorem 4.1.1 to the latter, we have

$$|X_1^*| = |X_1| - |X_{1,2}| - \cdots - |X_{1,q}| + |X_{1,2,3}| + \cdots + (-1)^{q-1}|X_{1,2,\dots,q}|. \tag{4.2.2}$$

Equations (4.2.1)–(4.2.2) are *inverse* to one another. Indeed, we can consider (4.2.1) as the equation for the unknown cardinality $|X_1^*|$. Then formula (4.2.2) solves equation (4.2.1) for $|X_1^*|$, that is, expresses $|X_1^*|$ through $|X_{i,\dots,j}|$. Vice versa, (4.2.1) represents $|X_1|$ in terms of $|X_1^*|$, etc. Such transformations are useful in many problems. We now consider the method, called the *Möbius* *inversion*, of inverting finite sums similar to (4.2.1)–(4.2.2). In what follows, we customarily write $d|n$ if d is a natural *divisor* of an integer n; $\sum_{d|n}$ means that the summation index d runs over all divisors of n.

Definition 4.2.1. The *Möbius function* $\mu : \mathbf{N} \to \{-1, 0, 1\}$ is defined as

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1 \text{ has a factor } p^\alpha \text{ with a prime } p \text{ andan integer } \alpha \geq 2, \\ (-1)^r & \text{if } n > 1 \text{ and has } r \text{ different prime factors.} \end{cases}$$

Example 4.2.1. By the definition, $\mu(1) = 1$. Since 2, 3, and 5 are primes, we have $\mu(2) = \mu(3) = \mu(5) = -1$. Next, $4 = 2^2$, thus $\mu(4) = 0$. From these equations we have

$$\sum_{d|2} \mu(d) = \mu(1) + \mu(2) = 0, \quad \sum_{d|3} \mu(d) = \sum_{d|5} \mu(d) = 0$$

and

$$\sum_{d|4} \mu(d) = \mu(1) + \mu(2) + \mu(4) = 0.$$

In the next lemma we prove that these zeros persist.

Lemma 4.2.1.

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \textit{if } n = 1, \\ 0 & \textit{if } n > 1. \end{cases}$$

Proof. If $n = 1$, the statement is obvious. Otherwise, let, for any $n > 1$,

$$n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_r^{\alpha_r}$$

be its prime factorization. Set $n^* = p_1 \cdot p_2 \cdot \cdots \cdot p_r$, thus, n^* contains all the different prime factors of n, though in n^* every factor appears only once. If d divides n but does not divide n^*, then d contains a factor p^α with a prime p and an integer $\alpha \geq 2$. Therefore, $\mu(d) = 0$ and for such a d, $\sum_{d|n} \mu(d) = \sum_{d|n^*} \mu(d)$.

However, for any $k, 0 \leq k \leq r$, the number n^* has $C(r, k)$ divisors d such that d can be written as the product of k different prime factors; if $k = 0$, we set $d = 1$. Thus by Definition 4.2.1, for these d, $\mu(d) = (-1)^k$ and

$$\sum_{d|n} \mu(d) = \sum_{d|n^*} \mu(d) = C(r, 0) - C(r, 1) + \cdots + (-1)^r C(r, r) = 0;$$

see Problem 1.4.18. □

The Möbius function appears in the following equations called the *Möbius inversion formulas*.

Theorem 4.2.1. *Let two infinite sequences* $\{f(m)\}_{m=1}^\infty$ *and* $\{g(m)\}_{m=1}^\infty$ *satisfy countably many equations*

$$f(m) = \sum_{d|m} g(d), \quad m = 1, 2, \ldots. \tag{4.2.3}$$

The system of equations (4.2.3) *has the solution*

$$g(m) = \sum_{d|m} \mu(d) f\left(\frac{m}{d}\right), \quad m = 1, 2, \dots, \tag{4.2.4}$$

where μ *is the Möbius function. Vice versa, simultaneous equations* (4.2.4)*,* $m = 1, 2, \dots$*, imply* (4.2.3) *for all* $m = 1, 2, \dots$*. In other words, the infinite set of simultaneous equations* (4.2.3) *is equivalent to the infinite set of simultaneous equations* (4.2.4)*.*

Proof. If d divides m, then, by (4.2.3),

$$f\left(\frac{m}{d}\right) = \sum_{\delta|(\frac{m}{d})} g(\delta);$$

hence

$$\sum_{d|m} \mu(d) f\left(\frac{m}{d}\right) = \sum_{d|m} \mu(d) \left(\sum_{\delta|(\frac{m}{d})} g(\delta)\right).$$

If $m = d \cdot \delta \cdot m_1$, then for a fixed δ, d runs over the set of divisors of the integer m/δ. Since all sums are finite, we can change the order of summation and get

$$\sum_{d|m} \mu(d) \left(\sum_{\delta|(\frac{m}{d})} g(\delta)\right) = \sum_{\delta|m} g(\delta) \left(\sum_{d|(\frac{m}{\delta})} \mu(d)\right).$$

If $\delta \neq m$, that is, $\frac{m}{\delta} \neq 1$, then $\sum_{d|(\frac{m}{\delta})} \mu(d) = 0$ by Lemma 4.2.1, so that

$$\sum_{\delta|m} g(\delta) \left(\sum_{d|(\frac{m}{\delta})} \mu(d)\right) = g(m).$$

The second part of the theorem can be proved similarly. □

We apply this theorem to calculate the number of special arrangements called *cyclic sequences*. To define them, we consider a set $A = \{a_1, a_2, \dots, a_n\}$ and all n^m m-arrangements with repetition of its elements. Let

$$\alpha_1 = (a_{i_1}, a_{i_2}, \dots, a_{i_m})$$

be any of them. Together with α_1 we consider its circular shifts, that is, m-arrangements

$$\begin{aligned}
\alpha_2 &= (a_{i_m}, a_{i_1}, a_{i_2}, \dots, a_{i_{m-1}}), \\
\alpha_3 &= (a_{i_{m-1}}, a_{i_m}, a_{i_1}, a_{i_2}, \dots, a_{i_{m-2}}), \\
&\vdots \\
\alpha_m &= (a_{i_2}, a_{i_3}, \dots, a_{i_m}, a_{i_1}).
\end{aligned}$$

Two arrangements $\alpha_i, \alpha_j, i \neq j$, can coincide term-wise, however, we distinguish them since they bear different indices i and j.

Definition 4.2.2. The set $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$ is called a *cyclic sequence* of length m corresponding to the arrangement α_1; of course, it also corresponds to any of the arrangements $\alpha_2, \dots, \alpha_m$.

The problem of enumeration of the cyclic sequences is complicated by their periodicity, since it may happen that $\alpha_1 = \alpha_{d+1}$, $\alpha_2 = \alpha_{d+2}$,..., $\alpha_{d-1} = \alpha_{2d-1}$, $\alpha_d = \alpha_{2d}$, etc.; in this case we say that d is a period of the cyclic sequence α. A cyclic sequence can have several periods. If d is the smallest period of a cyclic sequence $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_m\}$, then among m arrangements $\alpha_1, \alpha_2, \dots, \alpha_m$ there are only d different n-arrangements. Thus, d must divide m, and each α_i, $1 \le i \le m$, consists of $\frac{m}{d}$ different d-arrangements with repetition such that each of them generates a cyclic sequence of length d with the minimal period equal to d.

Definition 4.2.3. Given n elements, the number of cyclic sequences of length d from these elements with the minimal period d is denoted by $\mathrm{cyc}_{\mathrm{per}}(n, d)$. The number of cyclic sequences of length m of any period is denoted by $\mathrm{CYC}(n, m)$.

Theorem 4.2.2.

$$\mathrm{CYC}(n, m) = \sum_{d|m} \frac{1}{d}\Big(\sum_{\delta|d} \mu(\delta) n^{d/\delta}\Big). \tag{4.2.5}$$

Proof. Applying (4.2.4) with $f(m) = n^m$, we have

$$\mathrm{cyc}_{\mathrm{per}}(n, d) = \frac{1}{d} \sum_{\delta|d} \mu(\delta) n^{d/\delta}$$

and (4.2.5) follows. □

Problem 4.2.1. A bracelet consists of four geometrically identical beads of two colors. Two bracelets are considered identical if they can be superposed (made indistinguishable) by rotating them on the wrist without flipping, that is, not taking them off. How many different bracelets are there?

Solution. Formula (4.2.5) with $n = 2$ and $m = 4$ gives $\mathrm{CYC}(2, 4) = 6$. These six different bracelets are shown in Fig. 4.1. □

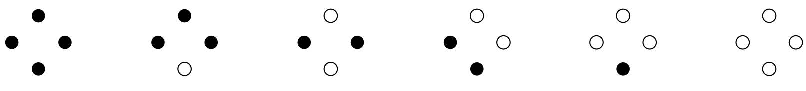


Figure 4.1: Six different bracelets in Problem 4.2.1.

Problem 4.2.2.

(1) How many bracelets are there consisting of six beads ($m = 6$) of three colors ($n = 3$)?

(2) How many geometrically indistinguishable bracelets with $m = 6$ and $n = 3$ do exist if one cannot only rotate but also flip them over?

Problem 4.2.3. (The little Fermat theorem) Show that, for any prime d and natural n,

$$d|(n^d - n).$$

Solution. If d is prime, then in (4.2.5) either $\delta = 1$, leading to $\mu(\delta) = 1$, or else $\delta = d$, resulting in $\mu(\delta) = -1$. Therefore,

$$\sum_{\delta|d} \mu(\delta) n^{d/\delta} = n^d - n.$$

Since a number $\mathrm{cyc}_{\mathrm{per}}(n, d)$ is integer, each addend in (4.2.5) must be integer, thus, d divides the expression in parentheses in (4.2.5), that is, $d| \sum_{\delta|d} \mu(\delta) n^{d/\delta}$, or $d|(n^d - n)$. □

Exercises 4.2.

Exercise 4.2.1. Solve again Problem 4.2.1, assuming that two bracelets are indistinguishable if they can be superposed with one another by rotation or by reflection in the bracelet plane (flipping).

Exercise 4.2.2. Euler's totient function $\phi(n)$ was defined in Exercise 4.1.20.
(1) Prove that $\phi(n)$ is multiplicative, that is $\phi(m \cdot n) = \phi(m) \cdot \phi(n)$ if m and n are mutually prime integers.
(2) Let $d_1, d_2, \ldots, d_k$ be all divisors of n. Prove the *Gauss formula*,

$$\sum_{j=1}^{k} \phi(d_j) = n.$$

(3) Prove that $\phi(n) = \sum_{d|n} \mu(\frac{n}{d}) d$.
(4) Use the latter formula and (4.2.5) to prove the equation

$$\mathrm{CYC}(n, m) = \frac{1}{m} \sum_{d|m} \phi\left(\frac{m}{d}\right) n^d.$$

Exercise 4.2.3. Prove the inversion formulas for the Stirling numbers,

$$\sum_k (-1)^k S_1(n, k) S_2(k, m) = (-1)^n \delta_{mn}$$

and

$$\sum_k (-1)^k S_2(n, k) S_1(k, m) = (-1)^n \delta_{mn}$$

where the Kronecker delta was defined in Exercise 1.4.4(11).

Exercise 4.2.4. Consider two infinite sequences of polynomials $\{P_n(t), n = 0, 1, 2, \ldots\}$ and $\{Q_n(t), n = 0, 1, 2, \ldots\}$ connected by the two sets of equations

$$P_n(t) = \sum_{m=0}^{n} \alpha_{n,m} Q_m(t), \quad n = 0, 1, \ldots,$$

and

$$Q_n(t) = \sum_{k=0}^{n} \beta_{n,k} P_k(t), \quad n = 0, 1, \ldots.$$

For any two sequences $\{u_n\}_{n\geq 0}$ and $\{v_n\}_{n\geq 0}$ of real numbers, prove the inversion formulas

$$u_n = \sum_{m=0}^{n} \alpha_{n,m} v_m \quad (\forall n \geq 0) \Longleftrightarrow v_n = \sum_{k=0}^{n} \beta_{nk} u_k \quad (\forall n \geq 0).$$

Exercise 4.2.5. Deduce from Exercise 4.2.4 the inversion formulas

$$u_n = \sum_{m=0}^{n} C(n, m) v_m \quad (\forall n \geq 0) \Longleftrightarrow v_n = \sum_{k=0}^{n} \delta_{n,k} u_k \quad (\forall n \geq 0).$$

Exercise 4.2.6. Prove the following pairs of inversion formulas.
(1) The equations $a_n = \sum_{k=0}^{n} (-1)^k C(n, k) b_{n-k}, \forall n = 0, 1, 2, \ldots$, are equivalent to the equations $b_n = \sum_{k=0}^{n} C(n, k) a_{n-k}, \forall n = 0, 1, 2, \ldots$.
(2) The equations $a_n = \sum_{k=0}^{n} (-1)^k C(n, k) b_{n-k}, \forall n = 0, 1, 2, \ldots$, are equivalent to the equations $b_n = \sum_{k=0}^{n} (-1)^k C(n, k) a_k, \forall n = 0, 1, 2, \ldots$.
(3) The equations $a_n = \sum_{k=0}^{n} C(n + p, k + p) b_k, \forall n = 0, 1, 2, \ldots$, are equivalent to the equations $b_n = \sum_{k=0}^{n} (-1)^{n-k} C(n + p, k + p) a_k, \forall n = 0, 1, 2, \ldots$.
(4) Apply the inversion formulas above to derive formula (4.1.7).

4.3 Generating functions I. Introduction

In many problems we have to deal with number sequences, for instance, with the combinations $C(m, n)$ or the cyclic sequences $\mathrm{CYC}(m, n)$, whose terms, in turn, depend on one or several integer parameters $m, n, \ldots$. We have to manipulate these sequences, which may result in cumbersome calculations. *The method of generating functions* (GF) is a general way to work out such problems. This method replaces operations on sequences with corresponding operations on certain functions or power series, called the GF of these sequences, which can be simpler and allows us to invoke powerful techniques of algebra and calculus. In this section we develop the method of GF and show on many examples how to derive GF and use them to solve various problems. More applications of the method are considered in the sequel sections.

Coffee-time browsing

- www.gap-system.org/~history/Biographies/Polya.html (Polya's biography)
- www.math.utah.edu/~pa/math/polya.html (How to Solve It?)
- en.wikipedia.org/wiki/John_Howard_Redfield (Redfield's biography)
- http://en.wikipedia.org/wiki/Brook_Taylor (Taylor's biography)
- www.gap-system.org/~history/Biographies/Cauchy.html (Cauchy's biography)
- http://en.wikipedia.org/wiki/Jacques_Hadamard#Biography (Hadamard's biography)
- http://scienceworld.wolfram.com/biography/Abel.html (Abel's biography)
- en.wikipedia.org/wiki/Abel_Prize (Abel Prize)
- http://en.wikipedia.org/wiki/Johann_Heinrich_Lambert (Lambert's biography)
- http://planetmath.org/encyclopedia/LambertWFunction.html (Lambert W function)

In the first example of this section we use the Taylor series of the exponential function. The reader unfamiliar with calculus can interpret the following equation as the statement that the exponential function e^z can be represented for small $|z|$ and for any $n = 1, 2, 3, \dots$ as $e^z \approx 1 + z + z^2/2! + z^3/3! + \cdots + z^n/n! +$ terms that are smaller than $z^n/n!$. For instance, $e^z \approx 1+$ terms which are much smaller than 1; or if we need better accuracy, $e^z \approx 1 + z+$ terms which are much smaller than $|z|$; or $e^z \approx 1 + z + z^2/2!+$ terms which are much smaller than $|z|^2/2$, and so forth. Such understanding is quite adequate for all our purposes. We start with a problem that shows why *the method of generating functions* (GF) is useful.

Example 4.3.1. Consider the obvious equation $e^x \cdot e^x = e^{2x}$ and expand the exponential functions on both sides in the Taylor series[2]

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$$

and

$$e^{2x} = 1 + 2x + \frac{1}{2!}(2x)^2 + \cdots + \frac{1}{n!}(2x)^n + \cdots,$$

deriving the equation

$$\left(1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots\right)\left(1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots\right)$$
$$= 1 + 2x + \frac{2^2}{2!}x^2 + \cdots + \frac{2^n}{n!}x^n + \cdots.$$

Multiplying out term-wise the two series on the left, combining like terms and equating the coefficients of $\frac{x^n}{n!}$ on both sides of the equation, we easily verify the equation (see

2 To justify these manipulations, some elementary calculus is needed.

the solution of Problem 1.4.18)

$$2^n = C(n,0) + C(n,1) + C(n,2) + \cdots + C(n,n).$$

This example demonstrates the essence of the method of GF—direct manipulations with sequences are replaced by transformations of certain functions or corresponding power series. Certainly in Problem 1.4.18 we derived the latter formula in more intuitive way. However, the method of GF gives us a powerful technique for solving various combinatorial, probabilistic, and many other essentially more involved problems, where elementary approaches may not work.

To introduce the method, we have to discuss some preliminaries. For an infinite sequence (a finite sequence can always be augmented by infinitely many zeros on the right)

$$\mathbf{a} = \{a_0, a_1, a_2, \ldots, a_n, \ldots\} = \{a_n\}_{n=0}^{\infty}$$

its GF is a *formal power series*

$$f_{\mathbf{a}}(t) \sim \sum_{n=0}^{\infty} a_n t^n, \tag{4.3.1}$$

where t is an *indeterminate*[3] or a *variable*. The series is called formal, because we do not discuss its convergence at all, the powers $t^n, n = 0, 1, 2, \ldots$, here are just labels, which distinguish different terms of the sequence **a**. When we expand, step-by-step, the sum in (4.3.1) as

$$a_0 + \sum_{n=1}^{\infty} a_n t^n,$$

$$a_0 + a_1 t + \sum_{n=2}^{\infty} a_n t^n,$$

$$a_0 + a_1 t + a_2 t^2 + \sum_{n=3}^{\infty} a_n t^n,$$

etc., (4.3.1) *generates*, one after another, consecutive terms of the sequence **a**. At this point, the noun "function" in the sentence "GF" does not signify a function (mapping) in the sense of Section 1.1. For this reason, we used the tilde sign ~ instead of the equality sign in (4.3.1).

To apply the formal power series, one has to develop some algebraic tools.[4] However, in this book we prefer to avoid the formal algebraic approach and justify the method by making use of convergent power series only. Hereafter we suppose that the series in (4.3.1) has a positive radius of convergence $R_{\mathbf{a}} > 0$, therefore the sum of the

3 Some authors denote the indeterminate by z and call the GF of a sequence $\{a_n\}_{n=0}^{\infty}$ its z-transformation.

4 See, for example, [1, 22, 48].

series $\sum_{n=0}^{\infty} a_n t^n$ exists in the disk $|t| < R_{\mathbf{a}}$ and is a function $f_{\mathbf{a}}(t)$ defined (and holomorphic) in the disk. All GF appearing in this book are represented by convergent power series. Hereafter we write

$$f_{\mathbf{a}}(t) = \sum_{n=0}^{\infty} a_n t^n \tag{4.3.2}$$

where $f_{\mathbf{a}}(t)$ is a function of t in some neighborhood of the point $t = 0$. The actual value of $R_{\mathbf{a}}$ is incidental for our purpose, any positive radius does. The choice of convergent power series narrows down the class of admissible sequences, nevertheless, this class is broad enough for all our applications.

The reader unfamiliar with calculus, can safely skip our discussion of power series and consider the conclusions as operational rules for solving corresponding problems. Moreover, we can arrive at the same results by considering *terminating* series instead of the infinite ones, that is, by making use of polynomials—see Problems 4.3.11 and 4.4.5 below, where we worked out this approach in detail. If we use these *generating polynomials*, we do not have to deal with the convergence issue at all, though computations may be lengthier and more cumbersome, as can be seen in the examples below. That is why hereafter we use the convergent series.

Given a sequence $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$, we need to know whether the corresponding power series (4.3.1) is convergent or divergent. According to the Cauchy–Hadamard criterion [50, p. 195], the series in (4.3.1)–(4.3.2) has a positive radius of convergence[5] $R_{\mathbf{a}} > 0$ if and only if the next quantity is finite,

$$\frac{1}{R_{\mathbf{a}}} = \limsup_{r\to\infty} \sqrt[n]{|a_n|} < \infty \tag{4.3.3}$$

which essentially means that $|a_n|$ has at most exponential growth as $n \to \infty$. Here $|a_n|$ stands for the absolute value (the modulus) of real or complex numbers a_n. For example, the sequences $\{a + bn^k\}_{n=0}^{\infty}$ and $\{a + bk^n\}_{n=0}^{\infty}$ satisfy (4.3.3) for any parameters a, b, k. However, faster growing sequences such as $\{n!\}_{n=0}^{\infty}$, may have $R_{\mathbf{a}} = 0$. A simple sufficient condition for the series (4.3.2) to have a positive radius of convergence is

$$|a_n| \le A_1 + A_2^n \quad \text{for all } n = 0, 1, 2, \ldots \tag{4.3.4}$$

with some positive constants $A_1 > 0$ and $A_2 > 0$.

Recall that the first derivative and indefinite integral of the power functions are given by the formulas $\frac{d}{dt}(t^p) = pt^{p-1}$ and $\int t^p\, dt = \frac{t^{p+1}}{p+1} + \text{const}$; the latter is valid if $p \ne -1$. Inside the disk of convergence, the convergent series can be differentiated and integrated term-by-term, that is if $|t| < R_{\mathbf{a}}$, then

$$\frac{d}{dt} \sum_{n=0}^{\infty} a_n t^n = \sum_{n=0}^{\infty} n a_n t^{n-1}$$

5 The meaning of this was explained at the very beginning of this section.

and

$$\int \left(\sum_{n=0}^{\infty} a_n t^n \right) dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1} t^{n+1} + \text{const}.$$

The term-wise differentiability and integrability of convergent power series are the only properties beyond the precalculus level, we use hereafter in applications of the method of GF.

Throughout we deal mainly with two well-known infinite series. These are the power series of the exponential function

$$e^t = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^n}{n!} + \cdots = \sum_{n=0}^{\infty} \frac{t^n}{n!}, \tag{4.3.5}$$

which is convergent for all (complex) t and the geometric series

$$\frac{1}{1-t} = 1 + t + t^2 + \cdots + \frac{t^n}{n!} + \cdots = \sum_{n=0}^{\infty} t^n, \tag{4.3.6}$$

which is convergent for $|t| < 1$. Differentiating (4.3.6) term-wise $p - 1$ times, we derive the formula

$$\begin{aligned} \frac{1}{(1-t)^p} &= 1 + \frac{p}{1!} t + \frac{p(p+1)}{2!} t^2 + \frac{p(p+1)(p+2)}{3!} t^3 \cdots \\ &+ \frac{p(p+1)\cdots(p+n-1)}{n!} t^n + \cdots = \sum_{n=0}^{\infty} \frac{(p+n-1)!}{n!(p-1)!} t^n; \end{aligned} \tag{4.3.7}$$

the coefficient of t^n in (4.3.7) is

$$C(p+n-1, n) = C_{\text{rep}}(p, n).$$

The same result can be derived without referring to infinite series. Indeed, let us consider a truncated series (4.3.6), that is, a polynomial

$$P_n(t) = 1 + t + t^2 + \cdots + t^n.$$

By the formula for the sum of a finite geometric progression, see Exercise 1.1.8, we derive

$$P_n(t) = \frac{1 - t^{n+1}}{1-t}. \tag{4.3.8}$$

Problem 4.3.1. Prove by mathematical induction that, for $p = 1, 2, \ldots$, the coefficient of $t^k, k \le n$, in the polynomial $(P_n(t))^p$, given by (4.3.8), is $C(p+k-1, k)$.

Solution. Each coefficient of $P_n(t)$ is 1, and also $C(1+k-1,k) = 1$, which establishes the basis of induction. Now suppose that the conclusion is valid for all exponents not exceeding some p and consider

$$(P_n(t))^{p+1} = (P_n(t))^p \cdot P_n(t).$$

Both factors on the right are polynomials. When we multiply them out, the power t^k occurs $k+1$ times—if we multiply the term t^k from $(P_n(t))^p$ by $t^0 = 1$ from $P_n(t)$, or if we multiply t^{k-1} from $(P_n(t))^p$ by t^1 from $P_n(t)$, ..., or if we multiply $t^0 = 1$ from $(P_n(t))^p$ by t^k from $P_n(t)$. Since the coefficient of t^j in $(P_n(t))^p$ is $C(p+j-1,j)$ by the inductive assumption, the coefficient in question is

$$C(p+0-1,0) + \cdots + C(p+j-1,j) + \cdots + C(p+k-1,k) = C(p+k,k)$$

due to equation (1.4.3) and Exercise 1.4.4(1), thus proving the claim. □

Problem 4.3.2. Find the coefficient of t^k in the polynomial $(P_n(t))^p$ for $n < k \le 2n$ and $p \ge 2$, if the polynomial $P_n(t)$ is given by (4.3.8).

To proceed with the method of GF, we introduce some operations on sequences and the corresponding operations on their GF. Linear combinations of sequences, that is, the multiplication of a sequence by a number (a *scalar*) and the addition of sequences, are defined straightforwardly, term-wise.

Example 4.3.2. Consider the sequences

$$\mathbf{a} = \{1, 0, 1, 0, 1, \ldots\}$$

and

$$\mathbf{b} = \{0, 1, 0, 1, 0, \ldots\},$$

that is, $a_n = (1+(-1)^n)/2$ and $b_n = (1-(-1)^n)/2, n \ge 0$, and their GF $f_{\mathbf{a}}(t)$ and $f_{\mathbf{b}}(t)$. The linear combination of **a** and **b** with coefficients α and β is defined as $\alpha\mathbf{a} + \beta\mathbf{b} = \{\alpha, \beta, \alpha, \beta, \alpha, \beta, \ldots\}$, and the corresponding GF is the linear combination of the GF with the same coefficients **a** and **b**, that is, $f_{\alpha\mathbf{a}+\beta\mathbf{b}}(t) = \alpha f_{\mathbf{a}}(t) + \beta f_{\mathbf{b}}(t)$.

Problem 4.3.3. Use formula (4.3.6) to show that in Example 4.3.2

$$\alpha f_{\mathbf{a}}(t) + \beta f_{\mathbf{b}}(t) = \frac{\alpha + \beta t}{1 - t^2}.$$

Problem 4.3.4. What properties (commutativity, associativity, etc.) do these operations on sequences possess? Notice that the sequence $\{0, 0, \ldots\}$ is the neutral element for the term-wise addition of sequences.

To define a “multiplication” of sequences, we consider two sequences $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$ and $\mathbf{b} = \{b_n\}_{n=0}^{\infty}$, and polynomials $P_{\mathbf{a}}(t) = \sum_{n=0}^{p} a_n t^n$ and $Q_{\mathbf{b}}(t) = \sum_{n=0}^{q} b_n t^n$, whose coefficients are initial terms of the sequences **a** and **b**, respectively. Let

$$R(t) = P_{\mathbf{a}}(t) Q_{\mathbf{b}}(t) = \sum_{n=0}^{p+q} c_n t^n.$$

Problem 4.3.5. Show that the coefficients c_n of the polynomial $R(t)$ for $n \le \min\{p, q\}$ are given by

$$c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0.$$

In particular, $c_0 = a_0 b_0, c_1 = a_0 b_1 + a_1 b_0, c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0, \ldots$.

Taking into account the latter equation and looking for an operation on sequences that corresponds to the multiplication of polynomials or power series, we arrive at the following definition, which mimics the *Cauchy rule* of multiplication of power series.

Definition 4.3.1. Given two sequences **a** and **b**, the sequence $\mathbf{c} = \{c_n\}_{n=0}^{\infty}$, where $c_n = a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0, n = 0, 1, 2, \ldots$, is called their *convolution* and is denoted by $\mathbf{c} = \mathbf{a} * \mathbf{b}$.

Now we give the major definition of this section.

Definition 4.3.2. Let a sequence **a** satisfy property (4.3.3). The function $f_{\mathbf{a}}(t)$ in (4.3.2) is called the *Generating Function* (GF) of the sequence **a**.

Example 4.3.3. The GF of the finite sequence $\{1, 0, 1, 0, 1, 0, 1\}$ is

$$f(t) = 1 + 0 \cdot t + 1 \cdot t^2 + 0 \cdot t^3 + 1 \cdot t^4 + 0 \cdot t^5 + 1 \cdot t^6 = 1 + t^2 + t^4 + t^6 = \frac{1 - t^8}{1 - t^2},$$

the GF of the infinite sequence $\{1, 0, 1, 0, 1, 0, \ldots\}$ is

$$f(t) = 1 + t^2 + t^4 + \cdots = \frac{1}{1 - t^2}.$$

When we employ the method of GF and work, instead of sequences, with their GF, at the last step we must return from the derived GF to its sequence and we want to be certain that this sequence is the one we looked for. The method of GF is based on the following statement.

Theorem 4.3.1. *There exists a one-to-one correspondence between the set of sequences satisfying* (4.3.3) *and the set of power series G with a positive radius of convergence. This correspondence preserves algebraic operations, which means that a linear combination of sequences corresponds to a linear combination, with the same coefficients, of their GF, and the convolution of sequences corresponds to the product of their GF.*

Proof. It should be mentioned that we have always considered the largest possible value of the radius of convergence, that is, if $R_{\mathbf{a}}$ is the radius of convergence of series $\mathbf{a}$, then there is no $R > R_{\mathbf{a}}$ such that the series $\sum_{n=0}^{\infty} a_n t^n$ converges in the disk $|t| < R$. The statement on the one-to-one correspondence follows immediately from the uniqueness of the Taylor series [52, pp. 651–652]. The correspondence of linear combinations is obvious. The conclusion regarding convolution follows from Definition 4.3.1 (the *Cauchy rule of multiplication of power series*). □

The set of GF has an algebraic structure of a *ring*; the definition can be found, for example, in [37]. For us that means only that we can add and multiply sequences and their corresponding GF using the standard *commutative, associative and distributive* rules and keeping in mind that by the product of two sequences, we understand their *convolution*.

Proposition 4.3.1. *Prove that the set of sequences, satisfying* (4.3.4)*, is a commutative ring with the unity element* $\mathbf{1} = \{1, 0, \ldots\}$ *with respect to the following operations:*
(1) *the usual term-wise multiplication by real numbers,*
(2) *the usual term-wise addition of sequences as addition, and*
(3) *the convolution of sequences as multiplication.*

Proof. Let $|a_n| \le A_1 + (A_2)^n$ and $|b_n| \le B_1 + (B_2)^n$ for all $n \ge 0$. Since $A_2 > 0, B_2 > 0$,

$$(A_2 + B_2)^n = \sum_{k=0}^{n} C(n,k) A_2^k B_2^{n-k} \ge A_2^n + B_2^n,$$

so that $|a_n + b_n| \le A_1 + A_2^n + B_1 + B_2^n \le (A_1 + B_1) + (A_2 + B_2)^n$. Thus, the sum $\mathbf{a} + \mathbf{b}$ also satisfies (4.3.4).

It remains to prove that the convolution $\mathbf{c} = \mathbf{a} * \mathbf{b}$ satisfies (4.3.4); since the verification of other ring axioms is straightforward we leave it to the reader. Thus, let $\mathbf{a}, \mathbf{b} \in H$ and

$$|a_i b_{n-i}| \le A_1 B_1 + A_1 B_2^{n-i} + A_2^i B_1 + A_2^i B_2^{n-i}.$$

From this

$$|c_n| \le (n+1) A_1 B_1 + A_1 \sum_{i=0}^{n} B_2^{n-i} + B_1 \sum_{i=0}^{n} A_2^i + B_2^n \sum_{i=0}^{n} (A_2/B_2)^i.$$

Set $\delta = \max\{1; A_1; B_1; A_2; B_2\}$, thus, $|c_n| \le 4(n+1)\delta^{n+1}$, and since $n + 1 \le 2^n$, we get $|c_n| \le C_1 + C_2^n$ with constants C_1 and C_2 for all $n = 0, 1, \ldots$. □

Problem 4.3.6. Prove the first part of Proposition 4.3.1 by making use of the inequality $|a_n + b_n| \le 2 \max\{|a_n|; |b_n|\}$.

Problem 4.3.7. Prove that the set of power series with a nonzero radius of convergence is a commutative ring with the usual addition and multiplication. The unity of this ring is a constant function $f(t) = 1 = 1 + 0 \cdot t + 0 \cdot t^2 + \cdots$.

Theorem 4.3.1 explains why the method of GF is useful—instead of performing tedious calculations with sequences, we work with (analytic) functions and have available powerful techniques of algebra and analysis. At the end we return back to the sequence we sought for. At that point we can use the following well-known formulas expressing the Taylor coefficients of a function f through its derivatives [52, p. 654] or through contour integrals [50, p. 174],

$$a_n = \frac{1}{n!} f_{\mathbf{a}}^{(n)}(0) = \frac{1}{2\pi i} \oint_{|z|=\epsilon} z^{-n-1} f_{\mathbf{a}}(z)\, dz, \quad n = 0, 1, \ldots,$$

where ϵ is small so as the circumference $|z| = \epsilon$ lies inside the circle of convergence of $f_{\mathbf{a}}$. If we know $f_{\mathbf{a}}$ exactly or approximately, these formulas allow to find the numbers a_n or to estimate their asymptotic behavior.

Depending upon a particular problem, it may be suitable to use other systems of *linearly independent functions* ϕ instead of the powers t^n. In particular, we will see that in problems, where the ordering of elements must be taken into account, it is useful to employ *exponential generating functions* (EGF) based on the system $\phi = \{t^n/n!\}_{n=0}^{\infty}$. General methods of constructing the GF are discussed in detail in [22].

Definition 4.3.3. A function

$$e_{\mathbf{a}}(t) = \sum_{n=0}^{\infty} a_n \frac{t^n}{n!} \tag{4.3.9}$$

is called the *Exponential Generating Function* (EGF) of a sequence **a**.

Example 4.3.4. For the sequence $\mathbf{s} = \{1, 1, \ldots\}$, the GF is

$$f_{\mathbf{s}}(t) = 1 + 1 \cdot t + 1 \cdot t^2 + \cdots + 1 \cdot t^n + \cdots = \frac{1}{1-t},$$

while its EGF is

$$e_{\mathbf{s}}(t) = 1 + 1 \cdot \frac{t}{1!} + 1 \cdot \frac{t^2}{2!} + \cdots + 1 \cdot \frac{t^n}{n!} + \cdots = e^t.$$

For the sequence $\mathbf{b} = \{1, 1 \cdot 3, 1 \cdot 3 \cdot 5, \ldots\}$, the EGF is $e_{\mathbf{b}}(t) = (1 - 2t)^{-3/2}$.

In the case of EGF, the definition of the convolution must be modified.

Definition 4.3.4. The sequence $\mathbf{d} = \{d_n\}_{n=0}^{\infty}$, where

$$d_n = \sum_{i=0}^{n} C(n, i) a_i b_{n-i}$$

is called the *binomial convolution* (or the Hurwitz composition) of the sequences **a** and **b**.

Problem 4.3.8. State and prove an analogue of Theorem 4.3.1 for EGF.

In addition to algebraic operations considered in Theorem 4.3.1, some other operations on sequences and the corresponding transformations of their GF are useful. For a sequence $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$ and a fixed natural number k, we consider a sequence $\mathbf{b} = \{b_n\}_{n=0}^{\infty}$, where $b_0 = b_1 = \cdots = b_{k-1} = 0$ and $b_n = a_{n-k}$ for $n \ge k$. Then clearly,

$$f_{\mathbf{b}}(t) = t^k f_{\mathbf{a}}(t). \tag{4.3.10}$$

On the other hand, if $b_n = a_{n+k}$ for all $n \ge 0$, then

$$f_{\mathbf{b}}(t) = t^{-k}\big(f_{\mathbf{a}}(t) - a_0 - a_1 t - \cdots - a_{k-1}t^{k-1}\big).$$

The sequences $\{b_n\}_{n=0}^{\infty}$ are called *shifts* of the sequence **a**.

The term-wise differentiation and integration of GF imply the following results.

Theorem 4.3.2. *Let* $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$ *be a given sequence.*
(1) *If* $\mathbf{b} = \{b_n\}_{n=0}^{\infty}$, $b_n = (n+1)a_{n+1}$, *then* $f_{\mathbf{b}}(t) = \frac{d}{dt} f_{\mathbf{a}}(t)$.
(2) *If* $\mathbf{c} = \{c_n\}_{n=0}^{\infty}$, $c_n = na_n$, *then* $f_{\mathbf{c}}(t) = t\frac{d}{dt} f_{\mathbf{a}}(t)$.
(3) *If* $\mathbf{d} = \{d_n\}_{n=0}^{\infty}$, $d_n = \frac{a_{n-1}}{n}$, $n \ge 1$, *and* $d_0 = 0$,*then* $f_{\mathbf{d}}(t) = \int_0^t f_{\mathbf{a}}(x)\,dx$.
(4) *If* $\mathbf{l} = \{l_n\}_{n=0}^{\infty}$, $l_n = \frac{a_n}{n+1}$, *then* $f_{\mathbf{l}}(t) = \frac{1}{t}\int_0^t f_{\mathbf{a}}(x)\,dx$. □

The proof of the following lemma is immediate.

Lemma 4.3.1. *If* $\mathbf{s} = \{1, 1, \ldots\}$, *that is,* $s_n = 1, \forall n \ge 0$, *then for any sequence* $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$

$$\begin{aligned}
&a_0 \cdot s_0 = a_0,\\
&a_0 \cdot s_1 + a_1 \cdot s_0 = a_0 + a_1,\\
&a_0 \cdot s_2 + a_1 \cdot s_1 + a_2 \cdot s_0 = a_0 + a_1 + a_2,
\end{aligned}$$

etc. Therefore, the convolution of sequences

$$\mathbf{a} * \mathbf{s} = \{a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots, a_0 + a_1 + \cdots + a_n, \ldots\}$$

is the sequence of consecutive partial sums of the sequence **a**. □

Problem 4.3.9. Prove that the sequence $\mathbf{I} = \{1, 0, 0, \ldots\}$ is the unit element for the convolution, that is, $\mathbf{a} * \mathbf{I} = \mathbf{I} * \mathbf{a} = \mathbf{a}$ for any sequence **a**.

Due to Lemma 4.3.1, the sequence $\mathbf{s} = \{1, 1, \ldots\}$ is called the *summator*. We know from Theorem 4.3.1 and Example 4.3.4 that the GF for the sequence $\mathbf{a} * \mathbf{s}$ is $(1-t)^{-1} f_{\mathbf{a}}(t)$.

From this observation we can, for instance, immediately conclude that the coefficient of, say, t^{37} in the Taylor series of the rational function

$$(1 - 3t^2 - 4t^7 + 12t^{21} - 5t^{45})(1 - t)^{-1} \tag{4.3.11}$$

is $1 - 3 - 4 + 12 = 6$. We keep the notation for the summing sequence $\mathbf{s} = \{1, 1, \ldots\}$ for the rest of this chapter.

Problem 4.3.10.
(1) Find the coefficients of $t^n, n \geq 0$, in the Taylor series of the rational function (4.3.11).
(2) Why are there different formulas for the coefficients for $n < 45$ and for $n \geq 45$?

The same argument based on the summing property of the sequence $\mathbf{s} = \{1, 1, \ldots\}$ and the identity

$$\begin{aligned}\frac{1}{1-t} &= (1 + t + t^2 + \cdots + t^9) \times (1 + t^{10} + t^{20} + \cdots + t^{90}) \\ &\quad \times (1 + t^{100} + t^{200} + \cdots + t^{900}) \times \cdots, \quad |t| < 1,\end{aligned} \tag{4.3.12}$$

which proof we leave to the reader, immediately implies that any natural number has a base 10 representation, and this representation is unique.

In the rest of this section and in the next one we consider various applications of the method of GF and solve problems.

Problem 4.3.11. Compute the sum $1^2 + 2^2 + \cdots + n^2$ for any natural n.

Solution. There are different ways to approach this problem. If we know the value of this sum, we can carry out a simple inductive proof, as we have done in Problem 1.1.4. However, we are going to apply the method of GF to demonstrate the essential ingredients of the method. We do not even have to know the sum in advance—the method allows us to find the sum explicitly.

Let us then introduce the sequence $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$, where $a_n = 1^2 + 2^2 + \cdots + n^2$. We immediately observe that this is a sequence of partial sums for a simpler sequence $\mathbf{b} = \{b_n\}_{n=0}^{\infty}$ with $b_n = n^2, n \geq 0$. We also know from Lemma 4.3.1 that to find explicitly the sequence $\mathbf{a}$, we can convolve the sequence $\mathbf{b}$ and the summator $\mathbf{s}$. Hence we conclude that $\mathbf{a} = \mathbf{s} * \mathbf{b}$, which is equivalent to $f_{\mathbf{a}}(t) = (1-t)^{-1} f_{\mathbf{b}}(t)$. This equation tells us that, if we find GF $f_{\mathbf{b}}$ explicitly, we will be able to calculate $f_{\mathbf{a}}$ and then solve the problem by computing its coefficients. Let us try to simplify the problem even more and reduce the sequence of squares $\mathbf{b}$ to the sequence of natural numbers themselves. From calculus we remember that $\frac{d}{dx}(x^2) = 2x$. This observation gives us a plan of the solution.

We begin with the *summator* $\mathbf{s} = \{1, 1, 1, \ldots\}$ and its GF

$$f_{\mathbf{s}}(t) = \frac{1}{1-t} = 1 + t + t^2 + \cdots + t^n + t^{n+1} + \cdots;$$

notice that the coefficient of t^n here is 1. Differentiating both sides of this equation (if $|t| < 1$, we can differentiate the series term-wise) we again get (4.3.7) with $p = 2$,

$$\frac{1}{(1-t)^2} = 1 + 2t + 3t^2 + 4t^3 + \cdots + nt^{n-1} + (n+1)t^{n+2} + \cdots; \qquad (4.3.13)$$

hence the function $\frac{1}{(1-t)^2}$ is the GF of the sequence $\mathbf{d} = \{1, 2, 3, \ldots\}$.

We should be careful here, since the indices start at zero, and we have $d_n = n + 1$. Thus, we have to shift this sequence, which in terms of GF corresponds to multiplication by t—see Theorem 4.3.2(2). Therefore, we introduce a shifted sequence $\mathbf{c} = \{c_n\}$ with $c_n = n, n \geq 0$, and derive

$$f_{\mathbf{c}}(t) = tf_{\mathbf{d}}(t) = t\frac{1}{(1-t)^2}.$$

Then, the coefficient of t^n in the series $f_{\mathbf{c}}(t)$ is $c_n = n$. Repeating this step, that is, differentiating $f_{\mathbf{c}}(t)$ and multiplying by t, we get the GF for the sequence of squares $\mathbf{b} = \{n^2\}_{n=0}^{\infty}$,

$$f_{\mathbf{b}}(t) = t\frac{d}{dt}\left\{t\frac{1}{(1-t)^2}\right\};$$

thus we know without any calculation that the coefficient of t^n in the Taylor series of the above function $f_{\mathbf{b}}(t)$ is $b_n = n^2$. Finally, the GF for the sequence we want in this problem is derived if we multiply the function $f_{\mathbf{b}}$ by the GF of the summator, that is, we have to consider

$$f_{\mathbf{a}}(t) = (1-t)^{-1}t\frac{d}{dt}\left\{t\frac{1}{(1-t)^2}\right\} = t(1+t)(1-t)^{-4}.$$

From (4.3.7) with $p = 4$ we get the equation

$$(1-t)^{-4} = \sum_{n\geq 0} \frac{(n+1)(n+2)(n+3)}{6}t^n,$$

thus,

$$\begin{aligned} t(1+t)(1-t)^{-4} &= \sum_{n\geq 0} \frac{(n+1)(n+2)(n+3)}{6}t^{n+1} + \sum_{n\geq 0} \frac{(n+1)(n+2)(n+3)}{6}t^{n+2} \\ &= \sum_{n\geq 0} \frac{1}{6}n(n+1)(2n+1)t^n \end{aligned}$$

and the coefficient of t^n here gives the required formula

$$a_n = 1^2 + 2^2 + \cdots + n^2 = \frac{1}{6}n(n+1)(2n+1).$$

We will solve this problem one more time, now using the *generating polynomials*, that is, truncated power series instead of infinite power series. Indeed, if we want to find a_n, it suffices to consider polynomials of nth degree and truncate all powers greater than n in all computations. This approach can be traced back at least to Niven [43, Chap. 7]. We demonstrate the method in detail here and also in Problem 4.4.5 in Section 4.4.

Consider a polynomial (4.3.8)

$$P_n(t) = 1 + t + t^2 + \cdots + t^n = \frac{1 - t^{n+1}}{1 - t}.$$

Repeating the same steps as above, we compute the functions $\frac{d}{dt}P_n(t)$, $t\frac{d}{dt}P_n(t)$, $t\frac{d}{dt}\{t\frac{d}{dt}P_n(t)\}$, and finally

$$P_n(t)\left(t\frac{d}{dt}\left\{t\frac{d}{dt}P_n(t)\right\}\right)$$
$$= \frac{t(1 - t^{n+1})(1 + t - (n+1)^2t^n + (2n^2 + 2n - 1)t^{n+1} - n^2t^{n+2})}{(1-t)^4}.$$

Since we are interested in the coefficient of t^n, only two terms in the numerator of the latter fraction, namely t and t^2 can contribute to this coefficient. Using Problem 4.3.1 to compute the coefficients of t and t^2, we again find the same expression $a_n = \frac{1}{6}n(n + 1)(2n + 1)$ as in Problem 1.1.4. □

In this problem we have used the operator $t\frac{d}{dt}$ and its square $(t\frac{d}{dt})^2$. The following problem[6] treats arbitrary natural degrees of this operator.

Problem 4.3.12.

(1) For any $n = 1, 2, 3, \ldots$ prove that

$$\left(t\frac{d}{dt}\right)^n \frac{1}{1-t} = 1^nt + 2^nt^2 + 3^nt^3 + \cdots.$$

(2) Prove that

$$1^nt + 2^nt^2 + 3^nt^3 + \cdots = \frac{P_n(t)}{(1-t)^{n+1}}$$

where P_n is a polynomial of degree n with $P_n(0) = 0$ and all the other coefficients positive; moreover, $P_n(1) = n!$.

Using GF we can derive new inversion formulas, distinct from the Möbius inversion in Theorem 4.2.1.

6 See [45, Problem I.45] where this and more general problems are considered.

Theorem 4.3.3. *Consider two sequences* $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$ *and* $\mathbf{b} = \{b_n\}_{n=0}^{\infty}$ *related by the infinite set of equations*

$$a_n = \sum_{k=0}^{n} (-1)^k C(m,k) b_{n-k}, \quad \forall n = 0, 1, 2, \dots, \tag{4.3.14}$$

where a natural number m is fixed. Then

$$b_n = \sum_{k=0}^{n} C(m+k-1,k) a_{n-k}, \quad \forall n = 0, 1, 2, \dots. \tag{4.3.15}$$

Vice versa, equations (4.3.15) *imply* (4.3.14).

Proof. It suffices to notice that by the binomial formula (1.4.4) with $a = 1$ and $b = -t$, $f_{\mathbf{a}}(t) = (1-t)^m f_{\mathbf{b}}(t)$, thus, $f_{\mathbf{b}}(t) = (1-t)^{-m} f_{\mathbf{a}}(t)$, which implies (4.3.15). □

The next problem demonstrates a useful method of construction of GF.

Problem 4.3.13. Find a GF for the sequence

$$\{C(m,0), C(m,1), \dots, C(m,m), 0, 0, \dots\}$$

and use it to calculate again the binomial coefficients $C(m,k)$.

Solution. Consider a polynomial in $m+1$ variables (*indeterminates*) $t, x_1, x_2, \dots, x_m$ and expand it against the powers of t:

$$(1+x_1 t)(1+x_2 t)\cdots(1+x_m t) = 1 + (x_1 + x_2 + \cdots + x_m)t + (x_1 x_2 + \cdots + x_{m-1} x_m)t^2 + \cdots + (x_1 x_2 \cdots x_m)t^m. \tag{4.3.16}$$

The coefficient of t^k here is the sum of all k-element products of the indeterminates $x_1, x_2, \dots, x_m$. There is a one-to-one correspondence between these products and k-element subsets of the set $X = \{x_1, x_2, \dots, x_m\}$. By definition of combinations without repetition, the number of such subsets is $C(m,k)$. Thus, equation (4.3.16) *generates* an explicit roster of all k-combinations of the elements of an m-element set, $0 \le k \le m$.

If we do not need the complete list of combinations, it is convenient to put in (4.3.16) $x_1 = x_2 = \cdots = x_m = 1$, thus reducing (4.3.16) to the binomial formula (1.4.4)

$$(1+t)^m = \sum_{k=0}^{m} C(m,k) t^k. \tag{4.3.17}$$

Differentiating (4.3.17) k times and substituting $t = 1$, we again derive formula (1.4.1), $C(m,k) = m!/((m-k)!k!)$. □

In the same fashion we can construct the GF

$$\sum_{r=0}^{\infty} C_{\mathrm{rep}}(n,r)t^r$$

for the number of combinations with unlimited repetition from elements of n types. Again, we introduce symbols for these elements and list, at least potentially, all of these combinations:

$$\begin{aligned} f(t) &= (1 + x_1 t + x_1^2 t^2 + \cdots) \times (1 + x_2 t + x_2^2 t^2 + \cdots) \times \cdots \\ &\quad \times (1 + x_n t + x_n^2 t^2 + \cdots) \\ &= 1 + (x_1 + x_2 + \cdots)t \\ &\quad + (x_1^2 + \cdots + x_n^2 + x_1 x_2 + \cdots + x_1 x_n + x_2 x_3 + \cdots + x_{n-1} x_n)t^2 + \cdots . \end{aligned} \tag{4.3.18}$$

Setting $x_1 = \cdots = x_n = 1$, (4.3.18) becomes

$$f(t) = (1 + t + t^2 + \cdots)^n = (1 - t)^{-n} = \sum_{r=0}^{\infty} C_{\mathrm{rep}}(n,r)t^r.$$

Differentiating this series r times and substituting $t = 0$, we again derive equation (1.4.6), $C_{\mathrm{rep}}(n,r) = C(n + r - 1, r)$.

Here as well as in the preceding problems, we can avoid operations with infinite series by considering instead of (4.3.18) generating polynomials,

$$f_r(t) = (1 + x_1 t + x_1^2 t^2 + \cdots + x_1^r t^r) \times \cdots \times (1 + x_n t + x_n^2 t^2 + \cdots + x_n^r t^r).$$

Problem 4.3.14. Find the generating polynomials for the sequences

$$\begin{aligned} I &= \{1, 0, 0, \ldots, \}, \\ S &= \{1, 1, 1, \ldots, \}. \end{aligned}$$

Problem 4.3.15. Under which condition is a sequence of polynomials $P_0(t), P_1(t), \ldots$ the sequence of generating polynomials for a numerical sequence $\{a_n\}_{n=0}^{\infty}$?

In the same way we can explicitly construct GF for combinations satisfying arbitrary restrictions on any of its elements. Moreover, since we usually do not need an explicit list of all combinations or other combinatorial objects in terms of the indeterminates $x_1, x_2, \ldots$, there may be no need to employ these parameters. For example, if we want to find the quantity considered in Problem 1.4.14, we can immediately write down its GF as $(1 + t + \cdots + t^{r_1})(1 + t + t^2 + \cdots)^{n-1}$, which by differentiation returns the same formula (1.4.8).

As another example, we compute the number $\widehat{C}(n,r)$ of combinations of elements of n types taken r at a time with unlimited repetition, under the additional restriction that every combination must contain at least one element of each type. This number is also equal to the number of ways to place r identical balls in $n \le r$ different urns so that no urn is empty. Comparing with (4.3.18), we see that to satisfy this requirement, it is necessary to delete the term 1 from each infinite series in the product. Thus, we immediately construct the GF as $f(t) = (t + t^2 + \cdots)^n$. From this we get $f(t) = \sum_{r=n}^{\infty} C(r-1, r-n)t^r$, and finally,

$$\widehat{C}(n,r) = \begin{cases} 0, & r < n, \\ C(r-1, r-n), & r \ge n. \end{cases}$$

Next we consider the simplest ordered combinatorial objects—the arrangements. If we try to find a GF for arrangements as a polynomial similar to (4.3.16), we fail, and the reason for the failure can be easily seen. We cannot derive the complete list of all arrangements, because the multiplication is commutative, $(x_1 \cdot x_2 + x_2 \cdot x_1)t^2 = 2x_1x_2t^2$, etc. One way to overcome this obstacle is to use noncommutative variables to avoid the appearance of like terms as in the example above. However, we can proceed in a more conventional way.

Notice that formula (4.3.17) can be rewritten as

$$(1+t)^m = \sum_{r=0}^{m} A(m,r)\frac{t^r}{r!},$$

which is the EGF for the number of arrangements, so that in the case of ordered totalities the EGF may work better.

Indeed, given p identical objects x, there exists the unique arrangement $(x, x, \dots, x)$ containing all of these objects, therefore, the EGF for this arrangement is

$$e(t) = 1 \cdot \frac{t^p}{p!}.$$

If an arrangement contains $k < p$ objects, the EGF can also be formed in the unique way as $1 \cdot \frac{t^k}{k!}$, since the objects are indistinguishable. Then the EGF for arrangements containing any number $k, 0 \le k \le p$, of these objects is

$$e(t) = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots + \frac{t^p}{p!}$$

and the EGF for arrangements that may contain any finite number of indistinguishable objects, now appears as an infinite series

$$e(t) = 1 + \frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots.$$

We recognize here the power series for the exponential function e^t, which has an infinite radius of convergence.

It is easily seen now that, if there are p indistinguishable (identical) objects of one type and q indistinguishable objects of another type, then the EGF for arrangements with repetition, containing any number of elements of these two types, is

$$e(t) = \left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots + \frac{t^p}{p!}\right)\left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots + \frac{t^q}{q!}\right).$$

It is clear now that the EGF for arrangements with repetition of elements of m types without any restrictions on their repetitions is

$$e(t) = \left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots\right)^m = e^{mt}.$$

The Taylor series for $e^{mt} = 1 + \frac{mt}{1!} + \frac{m^2t^2}{2!} + \cdots + \frac{m^nt^n}{n!} + \cdots$ again recovers formula (1.3.1), $A_{\text{rep}}(m, n) = m^n$.

Problem 4.3.16 (Problem 4.1.5 revisited). Among n-arrangements with repetition from the set $A = \{0, 1, 2, 3\}$, how many contain at least one digit 1, at least one 2, and at least one 3?

Solution. Since there are no restrictions on the symbol 0, the corresponding factor in the EGF is the same series $1 + \frac{t}{1!} + \frac{t^2}{2!} + \cdots = e^t$. However, to ensure the presence of each of the three other symbols 1, 2, 3, we must delete terms $1 = \frac{t^0}{0!}$ from the corresponding series, and these series become $\frac{t}{1!} + \frac{t^2}{2!} + \cdots = e^t - 1$. Thus, the EGF for the quantities in this problem is

$$e^t(e^t - 1)^3 = \sum_{n\geq 0}(4^n - 3^{n+1} + 3 \cdot 2^n - 1)\frac{t^n}{n!}. \qquad \square$$

Problem 4.3.17. Solve again Problems 1.4.18–1.4.20 by making use of EGF. For the reader's convenience we recall these problems.

Problem 1.4.18. Find the number of n-arrangements with repetition from the set $A = \{0, 1\}$, containing an even number of 0s.

Problem 1.4.19. Find the number of n-arrangements with repetition from the set $A = \{0, 1, 2\}$, containing an even number of 0s.

Problem 1.4.20. Find the number of n-arrangements with repetition from the set $A = \{0, 1, 2, 3\}$, containing an even number of 0s and an even number of 1s.

Solution. Since a 0 can appear only an even number of times, the corresponding factor in the EGF is $1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \cdots = (1/2)(e^t + e^{-t})$. All other factors are the same as in the preceding problem, that is, these are e^t. Thus, we get the following EGF.

In Problem 1.4.18: $(1/2)(e^t + e^{-t})e^t = \sum_{n\geq 0} 2^{n-1}\frac{t^n}{n!}$.

In Problem 1.4.19: $(1/2)(e^t + e^{-t})e^{2t} = \sum_{n\geq 0} \frac{1}{2}(3^n + 1)\frac{t^n}{n!}$.

In Problem 1.4.20: $\frac{1}{4}(e^t + e^{-t})^2e^{2t} = 1 + \sum_{n\geq 1}(4^{n-1} + 2^{n-1})\frac{t^n}{n!}$. $\square$

Problem 4.3.18. Compute the sum $F(j,k,n) = \sum_{i=1}^{k} C(n-i,j)$, where j,k,n are given natural numbers and $n \geq j+k$.

Solution. Set $a_i(j) = C(n-i,j)$, hence, $a_i(j) = 0$ for $i > n-j$. Let $f_i(t)$ be the GF of the sequence $\{a_i(j)\}_{j=0}^{\infty}$; we are to find partial sums of this sequence. We will compute them by making use of the binomial formula (4.3.17), namely,

$$f_i(t) = \sum_{j=0}^{\infty} a_i(j)t^j = \sum_{j=0}^{n-i} C(n-i,j)t^j = (1+t)^{n-i}.$$

Let f_F be the GF of the sequence $F(j,k,n)$, where we consider n and k as fixed parameters and j as a variable:

$$\begin{aligned} f_F(t) &= \sum_{j=0}^{\infty} F(j,k,n)t^j = \sum_{j=0}^{\infty} \left\{ \sum_{i=1}^{k} a_i(j) \right\} t^j = \sum_{i=1}^{k} \left\{ \sum_{j=0}^{\infty} a_i(j)t^j \right\} \\ &= \sum_{i=1}^{k} (1+t)^{n-i} = \frac{1}{t}(1+t)^n - \frac{1}{t}(1+t)^{n-k}. \end{aligned}$$

We notice that the series here are actually finite sums and at the very last step we used the formula for the sum of a geometric progression,

$$\sum_{i=1}^{k} q^{n-i} = q^{n-k} \sum_{l=0}^{k-1} q^l = q^{n-k} \frac{q^k - 1}{q-1}$$

with $q = 1+t \neq 1$. Applying (4.3.17) twice, we get

$$F(j,k,n) = C(n,j+1) - C(n-k,j+1),$$

assuming that $C(j,j+1) = 0$. ☐

Exercises 4.3.

Exercise 4.3.1. Find explicitly the GF for the following finite or infinite sequences.
(1) $\mathbf{a_1} = \{1,1,1,\ldots\}$,
(2) $\mathbf{a_2} = \{0,0,0,1,1,1\}$,
(3) $\mathbf{a_3} = \{0,0,0,1,1,1,0,0,0,1,1,1,\ldots\}$,
(4) $\mathbf{a_4} = \{1,1,1,0,0,0,1,1,1,0,0,0,\ldots\}$,
(5) $\mathbf{a_5} = \{a_n\}_{n=0}^{\infty}, a_n = C(5,n)$,
(6) $\mathbf{a_6} = \{a_n\}_{n=0}^{\infty}, a_n = (-1)^n$,
(7) $\mathbf{a_7} = \{a_n\}_{n=0}^{\infty}, a_n = \cos(\alpha n)$,
(8) $\mathbf{a_8} = \{a_n\}_{n=0}^{\infty}, a_n = \sin(\alpha n)$.

Exercise 4.3.2. This problem refers to Exercise 4.1.1. Let $a_n, n = 0,1,2,\ldots$, be the number of elements in the set S possessing exactly n of the properties $P_j, j = 1,2,3,4$. Write down explicitly the GF of the sequence $\{a_n\}_{n=0}^{\infty}$.

Exercise 4.3.3. Using the result of Exercise 1.1.31 prove that the GF $f_{\mathbf{a}}$ and the EGF $e_{\mathbf{a}}$ of any sequence **a** are connected by the equation

$$f_{\mathbf{a}}(t) = \int_0^\infty e^{-x} e_{\mathbf{a}}(xt)\,dx.$$

Exercise 4.3.4. Introduce a sequence of complementary partial sums (tails) of the sequence **a**,

$$c_n = a_{n+1} + a_{n+2} + \cdots, \quad n = 0, 1, 2, \ldots.$$

Prove that, if the GF $f_{\mathbf{a}}(1)$ is convergent, then

$$(1-t) f_{\mathbf{c}}(t) = f_{\mathbf{a}}(1) - f_{\mathbf{a}}(t).$$

Exercise 4.3.5. Prove (4.3.7) by mathematical induction.

Exercise 4.3.6. Prove identity (4.3.12).

Exercise 4.3.7. Find the coefficients of t^7 and t^{11} in the Taylor series of the fraction $\frac{1-2t^2+3t^5-t^8+10t^{10}}{2-3t+t^2}$.

Exercise 4.3.8. Compute the sum $\sum_{k=0}^{m} C(n,k)C(m,k), n \geq m$.

Exercise 4.3.9. Compute the sum $S(m,p) = \sum_{k=0}^{m} (-1)^k C(p,k), p \geq m$.

Exercise 4.3.10. There are 10 married couples living at a townhouse. In how many ways is it possible to select a committee out of these people, consisting of 2 men and 3 women?

Exercise 4.3.11. Solve the preceding problem if no person can serve on the committee together with her/his spouse.

Exercise 4.3.12. In how many ways can one buy 2 different books if a bookstore has 3 bestsellers?

Exercise 4.3.13. Solve the preceding problem if one can also buy two copies of the same title.

Exercise 4.3.14. How many ways are there to buy for gifts 20 copies of these 3 bestsellers? We can buy 20 copies of the same title.

Exercise 4.3.15. How many ways are there to buy for gifts 20 copies of these 3 bestsellers if we want to buy at least one and no more than 10 copies of the same title?

Exercise 4.3.16. How many 10-letter words are there composed of 5 vowels and the letter “z”, which contain each vowel at least once?

Exercise 4.3.17. Use the identity $(1+t)^m \cdot (1+t)^p = (1+t)^{m+p}$ and GF to prove the equation

$$C(m+p,n) = \sum_{k=0}^{m} C(m,k)C(p,n-k).$$

Exercise 4.3.18. Use GF and the identity

$$(1-t)^{-1-n}(1+t)^{-1-n} = \left(1-t^2\right)^{-1-n}$$

to prove the equation

$$\sum_{j=0}^{2m}(-1)^j C(n+j,n)C(n+2m-j,n) = C(n+m,m).$$

Exercise 4.3.19. Find the sum of the third powers and the sum of the fourth powers of the first n natural numbers.

Exercise 4.3.20. A bookstore has four novels, in English, French, German, and Russian, 100 copies of each. In how many ways is it possible to buy 50 books so that you have even numbers of English, French, and German books and an odd number of Russian books? Answer the same question if the number of English and French titles together does not exceed 4. Answer the same question if the quantity of German books is twice or more than that of Russian books.

Exercise 4.3.21. Let f be the GF of a sequence $\{a_0, a_1, \ldots\}$. Which sequence is generated by the function f^2?

Exercise 4.3.22. Find appropriate analogues of (4.3.12) and use them to prove that each natural number has the unique binary (that is, base 2), ternary (base 3), quaternary (base 4), etc., representation.

Exercise 4.3.23. Use GF or EGF to prove, similarly to Theorem 4.3.3, the following pairs of inversion formulas.

(1) The equations $a_n = \sum_{k=0}^{n}(-1)^k C(n,k)b_{n-k}, \forall n = 0,1,2,\ldots$, are equivalent to $b_n = \sum_{k=0}^{n} C(n,k)a_{n-k}, \forall n = 0,1,2,\ldots$.
(2) The equations $a_n = \sum_{k=0}^{n} C(n+p,k+p)b_k, \forall n = 0,1,2,\ldots$, are equivalent to $b_n = \sum_{k=0}^{n}(-1)^{n-k}C(n+p,k+p)a_k, \forall n = 0,1,2,\ldots$.

Exercise 4.3.24. Does Theorem 4.3.3 remain valid if the equation

$$a_{n_0} = \sum_{k=0}^{n_0}(-1)^k C(m,k)b_{n_0-k}$$

fails for only one natural n_0? The same question with regard Exercise 4.3.23.

Exercise 4.3.25. Show that the assertion $u_n = \sum_{m=0}^{n} C(n,m)v_m$ ($\forall n \geq 0$), of Exercise 4.2.5 is equivalent to the equation $f_u(t) = e^t f_v(t)$ between the EGF of the sequences $u = \{u_n\}$ and $v = \{v_n\}$.

Exercise 4.3.26. Restore all computations in Problem 4.3.12.

Exercise 4.3.27. Find again the number of combinations with and without repetition using generating polynomials, that is, truncated GF.

Exercise 4.3.28. To facilitate performance of the participants of a contest, the Combi Club bought 15 identical chocolate bars. In how many ways is it possible to distribute them among five participants of the contest, so that each contestant receives at least two but no more than four chocolates?

Exercise 4.3.29. How many three-term geometric progressions (a, aq, aq^2) like $2, 6, 18$, and four-term geometric progressions (a, aq, aq^2, aq^3) like $3, 9, 27, 81$, are there in the set $\{1, 2, 3, \dots, 99, 100\}$?

Exercise 4.3.30. Prove that, if $f_\mathbf{a}, f_\mathbf{b}, f_\mathbf{c}, f_\mathbf{d}$ are the GF of sequences $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$, respectively, and $f_\mathbf{d} = f_\mathbf{a} \cdot f_\mathbf{b} \cdot f_\mathbf{c}$, then

$$d_n = \sum_{i+j+k=n} a_i b_j c_k$$

where the sum runs over all nonnegative integer solutions of the equation $i + j + k = n$.

Exercise 4.3.31. Use equations (2.1.3)–(2.1.5) and the result of Exercise 2.1.17 to derive the following GF:

(1) $(1+t)^{p(p-1)/2} = \sum_{q\geq 0} C(p(p-1)/2, q)t^q$ for the number of simple labeled graphs of order p and size q.
(2) $(1+t)^{p(p+1)/2} = \sum_{q\geq 0} C(p(p+1)/2, q)t^q$ for the number of labeled graphs of order p and size q with loops but without multiple edges.
(3) $(1-t)^{-p(p-1)/2}$ for the number of labeled graphs of order p and size q without loops but with multiple edges.
(4) $(1-t)^{-p(p+1)/2}$ for the number of labeled graphs of order p and size q with loops and multiple edges.

Exercise 4.3.32.

(1) Use the method of undetermined coefficients and the equation

$$(1+x)^{1/2} \times (1+x)^{1/2} = 1 + x$$

to compute the coefficients of

$$(1+x)^{1/2} = a_0 + a_1 x + a_2 x^2 + \cdots,$$

namely,

$$(1+x)^{1/2} = 1 + \frac{1}{2}x + \cdots + \frac{(1/2)(1/2-1)(1/2-2)\cdots(1/2-n+1)}{n!}x^n + \cdots.$$

(2) Use the same method to derive the equation

$$(1+x)^{-1/2} = 1 - \frac{1}{2^2}C(2,1)x + \frac{1}{2^4}C(4,2)x^2 + \cdots + \frac{(-1)^n}{2^{2n}}C(2n,n)x^n + \cdots.$$

(3) Which property of the binomial coefficients can be deduced from the equation $[(1+x)^{-1/2}]^2 = \frac{1}{1+x}$?

Exercise 4.3.33. Let $\tau(p)$ be the number of rooted labeled trees of order p and

$$\vartheta(t) = \sum_{p=1}^{\infty} \tau(p)\frac{t^p}{p!}$$

be the corresponding EGF. Prove that

$$\vartheta(t) = -W(t),$$

where $W(t)$ is the Lambert W function, that is, the (many-valued) solution of the transcendental equation $We^W = -t$. A rooted tree is a tree with a singled out vertex.

Exercise 4.3.34. Prove that the convolution of sequences is commutative and associative, that is, $\mathbf{a} * \mathbf{b} = \mathbf{b} * \mathbf{a}$ and $\mathbf{a} * (\mathbf{b} * \mathbf{c}) = (\mathbf{a} * \mathbf{b}) * \mathbf{c}$.

Exercise 4.3.35. Prove that the binomial convolution of sequences is commutative and associative.

Exercise 4.3.36. Derive the EGF for the Stirling numbers of the second kind $S_2(k,n)$,

$$\frac{1}{n!}(e^t - 1)^n = \sum_{k\geq n+1} S_2(k,n)\frac{t^k}{k!}.$$

Exercise 4.3.37. Use (4.3.13) to prove the equation

$$\sum_{k=1}^{\infty} \frac{k}{2^k} = 2.$$

The latter implies immediately that $\sum_{k=1}^{\infty} \frac{k-1}{2^k} = 1$.

Exercise 4.3.38. What is the probability to get a heads at the first flip of a fair coin? At the second flip? At the third? ... At the nth flip? Use the previous problem to find the expected value of the number of flips before the first head occurs.

Exercise 4.3.39. Compute the generating function of the sequence $f(n) = \alpha^n/n!$ for even n, and 0 for odd n.

Exercise 4.3.40. Express the generating function of the sequence $g(n) = f(n+1) - f(n)$ through the generating sequence of the sequence $f(n)$.

4.4 Generating functions II. Applications

In this section we employ the method of GF to study partitions and compositions of natural numbers. Then we take up linear difference equations (or recurrence relations) with constant coefficients and solve more problems.

Coffee-time browsing

- en.allexperts.com/e/n/no/norman_macleod_ferrers.htm (N. M. Ferrers)
- http://en.wikipedia.org/wiki/Partition_(number_theory)#Ferrers_diagram (Partitions in Number Theory)
- en.allexperts.com/e/y/yo/Young_tableau.htm (Young Tableaus (diagrams))
- www.gap-system.org/~history/Mathematicians/Fibonacci.html (Fibonacci summary)
- http://images.google.com/images?q=fibonacci+sequence&rls=com.microsoft:en-us:IE-SearchBox&oe=UTF-8&sourceid=ie7&rlz=1I7DKUS&um=1&ie=UTF-8&ei=RhojS7SPOc3jlAeLz9XzCQ&sa=X&oi=image_result_group&ct=title&resnum=4&ved=0CB4QsAQwAw (Fibonacci numbers)
- www.answers.com/topic/godfrey-harold-hardy (Hardy's biography)
- http://en.wikipedia.org/wiki/Fran%C3%A7ois_%C3%89douard_Anatole_Lucas (Lucas' biography)
- www.gap-system.org/~history/Biographies/Jacobi.html (Jacobi's) biography
- http://images.google.com/images?q=jacobian&rls=com.microsoft:en-us:IE-SearchBox&oe=UTF-8&sourceid=ie7&rlz=1I7DKUS&um=1&ie=UTF-8&ei=7qQjS63CLMXHlAfnl6nzCQ&sa=X&oi=image_result_group&ct=title&resnum=4&ved=0CCIQsAQwAw (Jacobians)

Problem 4.4.1. The postage for a letter is 84 cents. In how many ways can one buy stamps to send the letter if the post office has only 42-cent and 1-cent stamps?

Solution. The simplest way is to buy two 42-cent stamps. However, it is also possible to buy one 42-cent stamp and 42 penny stamps, or to buy 84 penny stamps. Therefore, there are three ways to pay the postage, namely,

$$\begin{aligned} 84\text{ (cents)} &= 1\text{ (cent)} \times 84 \\ &= 1\text{ (cent)} \times 42 + 42\text{ (cents)} \times 1 \\ &= 42\text{ (cents)} \times 2. \end{aligned}$$ □

In this problem we represent the integer number 84 as a sum of integer numbers in several ways. It should be also noticed that stamps can be put in any order, while in other problems the order of addends can be essential. Similar problems[7] occur in many applications. In this section we study them using the method of GF. We start with a formal definition.

7 These problems are discussed in detail by G. Andrews [3].

Definition 4.4.1. A set of ordered pairs of natural numbers

$$\Pi(n,k) = \{(n_1,k_1),(n_2,k_2),\dots,(n_l,k_l)\}$$

where $k_1+\dots+k_l = k$ and $n_1k_1+\dots+n_lk_l = n$, is called a *partition of a natural number n* in k terms (or addends) $n_1, n_2, \dots, n_l$. Since the addition is commutative, without loss of generality we will always list the terms of partitions in increasing order of their terms, that is, we always assume that $1 \le n_1 < n_2 < \cdots < n_l$.

For instance, the partition $\Pi(15,8) = \{(1,4),(2,3),(5,1)\}$, where $l = 3$, $n_1 = 1, n_2 = 2, n_3 = 5$, $k_1 = 4, k_2 = 3, k_3 = 1$, $4+3+1 = 8$ and $1\cdot 4+2\cdot 3+5\cdot 1 = 15$, corresponds to the following representation of 15 as a sum of 8 terms: $15 = (1+1+1+1)+(2+2+2)+(5)$. In Problem 4.4.1 we found three partitions of the number 84, $\prod(84,84) = \{(1,84)\}$, that is, $84 = 1\cdot 84$, $\prod(84,42) = \{(1,42),(42,1)\}$, that is, $84 = 1\cdot 42+42\cdot 1$, and $\prod(84,2) = \{(42,2)\}$, that is, $84 = 42\cdot 2$.

Theorem 4.4.1. *The GF for the number of partitions is*

$$\begin{aligned}&(1+t+t^2+t^3+\cdots)\times(1+t^2+t^4+t^6+\cdots)\times\cdots\\&\quad\times(1+t^k+t^{2k}+t^{3k}+\cdots)\times\cdots\\&\quad=\{(1-t)(1-t^2)\times\cdots\times(1-t^k)\times\cdots\}^{-1}.\end{aligned} \tag{4.4.1}$$

Proof. If we multiply out all the series in (4.4.1), the power t^n appears as the product

$$t^n = t^{1\cdot k_1'}\times t^{2\cdot k_2'}\times\cdots\times t^{i\cdot k_i'}\times\cdots.$$

Here the exponent $1\cdot k_1'$ indicates that the number n contains k_1' of units, that is, k_1' of the infinitely many infinite series in (4.4.1) contributed the term $1 = t^0$ as a factor to t^n. Next, the exponent $2\cdot k_2'$ shows that n contains k_2' of 2s, etc.; some k_i' may be equal to zero. Leaving out zero exponents, denoting nonzero terms by k_i, and using the geometrical series (4.3.6) we deduce (4.4.1). □

Problem 4.4.2. Solve again Problem 4.4.1 by making use of GF (4.4.1).

Problem 4.4.3. Show that the GF for the number of partitions with different addends, that is, with $k_1 = k_2 = \cdots = k_l = 1$ (no term repeats), is $(1+t)\times(1+t^2)\times\cdots\times(1+t^k)\times\cdots$. The GF of partitions with all terms not exceeding a given number q, that is, with $n_i \le q$ for every i, is

$$\begin{aligned}&(1+t+t^2+t^3+\cdots)\times(1+t^2+t^4+t^6+\cdots)\times\cdots\\&\quad\times(1+t^q+t^{2q}+t^{3q}+\cdots)=\{(1-t)(1-t^2)\cdots(1-t^q)\}^{-1}.\end{aligned} \tag{4.4.2}$$

The coefficient of t^n in (4.4.1) depends only on n and counts all the partitions of n with any $k = 1,2,\dots$. However, if we consider the number of partitions of n, containing precisely k parts, this quantity depends on two integer parameters n and k. If the quantities we sought depend on several parameters, it may be useful to employ GF of two or more variables. In the next theorem we find a GF for partitions of integers, taking also into account both n and the number k of terms in a partition.

Theorem 4.4.2. *The GF of the partitions containing exactly k terms is*

$$P(t,k) = t^k\{(1-t)(1-t^2)\cdots(1-t^k)\}^{-1}. \tag{4.4.3}$$

Proof. To consider both parameters, n and k of $\Pi(n,k)$, we introduce a function of two variables

$$\begin{aligned} F(t,u) &= (1 + ut + u^2t^2 + u^3t^3 + \cdots) \times (1 + ut^2 + u^2t^4 + u^3t^6 + \cdots) \\ &\quad \times (1 + ut^i + u^2t^{2i} + u^3t^{3i} + \cdots) \times \cdots = \{(1-ut) \times (1-ut^2) \times \cdots\}^{-1}. \end{aligned} \tag{4.4.4}$$

Multiplying out the series in (4.4.4), we get the addends (cf. the proof of Theorem 4.4.3)

$$u^{k_1'} \times t^{k_1'} \times u^{k_2'} \times t^{2k_2'} \times \cdots \times u^{k_i'} \times t^{ik_i'} \times \cdots = u^{k_1'+k_2'+\cdots} t^{k_1'+2k_2'+\cdots}$$

that is the exponent of the factor t contains k_1' of units, k_2' of twos, and so on, totaling to $k_1' + k_2' + \cdots$ addends. Therefore, expanding $F(t,u)$ in powers of u, we derive

$$F(t,u) = \sum_{k=0}^{\infty} P(t,k)u^k \tag{4.4.5}$$

where the coefficient $P(t,k)$ includes only powers t^n such that the corresponding partition of n contains exactly k terms. Hence, if we expand $P(t,k)$ against the powers of t, the coefficient of t^n will give the number of partitions of n consisting exactly of k parts. Thus, $P(t,k)$ is the GF we are looking for, since it lists partitions with precisely $k \geq 1$ parts.

Now we find $P(t,k)$ explicitly. First we remark that by (4.4.4) and (4.4.5), $P(t,0) = F(t,0) = 1$. Moreover, one can directly verify the equation

$$(1-ut)F(t,u) = F(t,ut).$$

Inserting series (4.4.5) in this equation, we get

$$P(t,k) - tP(t,k-1) = t^kP(t,k), \quad k = 1,2,\ldots$$

so that $(1-t^k)P(t,k) = tP(t,k-1), k \geq 1$. If we replace here k with $k-1$, we get $(1 - t^{k-1})P(t,k-1) = tP(t,k-2)$. From these two equations

$$(1-t^k)(1-t^{k-1})P(t,k) = t^2P(t,k-2), \quad k \geq 2.$$

Repeating this process, that is, reducing the latter to $P(t,k-3)$, then to $P(t,k-4),\ldots$, and eventually to $P(t,0)$, we get (4.4.3). □

Problem 4.4.4. Show that the number of partitions of number $2r+k$ in $r+k$ parts does not depend on k.

Solution. By Theorem 4.4.2, this number is the coefficient of t^{2r+k} in $P(t, r+k)$. Since $(2r+k)-(r+k)=r$, this coefficient is equal to the coefficient of t^r in the Taylor series of the function $\{(1-t)(1-t^2)\cdot\dots\cdot(1-t^{r+k})\}^{-1}$. However, the latter coefficient depends only on the first r factors

$$\{(1-t)(1-t^2)\cdots(1-t^r)\}^{-1}.$$

Then, due to (4.4.2) this coefficient is equal to the number of partitions of r into addends not exceeding r. Since a partition of r cannot contain a term bigger than r itself, the latter condition ("addends not exceeding r") imposes no restriction on partitions, thus the quantity we want is the total number of the partitions of r, which does not depend on k. □

This problem and some others can be easily solved by making use of special diagrams.

Definition 4.4.2. Let be $k_1+\dots+k_l = k$ and $n_1k_1+\dots+n_lk_l = n$. The *Ferrers* (or *Young*) *diagram* of a partition

$$\Pi(n,k) = \{(n_1,k_1),(n_2,k_2),\dots,(n_l,k_l)\}$$

is a set of $n = n_1k_1+\dots+n_lk_l$ dots in the plane, situated in $k = k_1+\dots+k_l$ rows such that for any $i = 1,2,\dots,l$ there are k_i rows containing n_i dots each.

We always consider *normalized diagrams*, such that the left-most dots of all rows form a vertical column and the numbers of points in consecutive rows, from top to bottom, do not increase. Thus, diagrams explicitly display the terms of a partition as horizontal rows of dots, from largest to smallest. For example, the Ferrers diagram in Fig. 4.2 depicts the partition $\Pi(13,6) = \{(1,3),(3,2),(4,1)\}$.

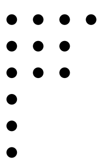


Figure 4.2: The Ferrers diagram of the partition $\Pi(13,6)$.

Another solution of Problem 4.4.4. Since a partition $\Pi(2r+k, r+k)$ contains $r+k$ parts, the left-most column of its normalized Ferrers diagram consists of $r+k$ dots. Hence, the complementary part of the diagram contains $(2r+k)-(r+k)=r$ dots, so that this part corresponds to a partition of the number r. Vice versa, if we append a column, consisting of $r + k$ dots, on the left to the normalized diagram of any partition of a number r, we derive the diagram of some partition $\Pi(2r + k, r + k)$, which proves the statement. □

Next we consider the compositions, or ordered partitions of integer numbers. To formalize ordering, we again use the language of mappings. We recall the notation of a natural segment $\mathbf{N}_m = \{1, 2, \dots, m\}$.

Definition 4.4.3. A *composition* of a natural number $n \in \mathbf{N}$, containing m parts, is a mapping $f : \mathbf{N}_m \to \mathbf{N}$ such that $f(1) + f(2) + \cdots + f(m) = n$.

Example 4.4.1. The partition (Fig. 4.2)

$$\Pi(13, 6) = \{(1, 3), (3, 2), (4, 1)\}$$

that is $13 = (1 + 1 + 1) + (3 + 3) + (4)$, generates a composition $13 = \{1, 1, 1, 3, 3, 4\}$. This composition can be realized as a mapping

$$f_1 : \{1, 2, 3, 4, 5, 6\} \to \mathbf{N}$$

such that $f_1(1) = f_1(2) = f_1(3) = 1, f_1(4) = f_1(5) = 3$, and $f_1(6) = 4$. However, the mapping with the same domain and codomain,

$$f_2 : \{1, 2, 3, 4, 5, 6\} \to \mathbf{N},$$

such that $f_2(1) = f_2(2) = f_2(5) = 1, f_2(4) = f_2(6) = 3$, and $f_2(3) = 4$, corresponds to another composition with the same parameters $n = 13$ and $m = 6$, namely, $13 = \{1, 1, 4, 3, 1, 3\}$.

Using the same reasoning as before, we see for ourselves that the GF for the compositions of a number n, consisting of m parts, is

$$(t + t^2 + \cdots)^m = t^m (1 - t)^{-m}.$$

To find the coefficient of t^n in this series, we have to compute the coefficient of t^{n-m} in the series $(1 - t)^{-m}$—see (4.3.7). Hence, we have proved the next statement.

Theorem 4.4.3. *There are*

$$C(n - 1, m - 1) \tag{4.4.6}$$

compositions of a number n containing m parts. □

Remark 4.4.1. Formula (4.4.6) simultaneously counts combinations with unlimited repetitions from elements of m types containing at least one element of each type.

Corollary 4.4.1. *There are $C(n + m - 1, m - 1)$ compositions if a composition can contain zero terms. To prove this, we can just replace each term x_j of a composition with $x_j - 1$, thus increasing n by m.* □

Problem 4.4.5. In how many ways is it possible to get the sum of n after rolling a fair die several times, if after each roll we add up the numbers it landed on?

Solution. The question, as it is stated, is ambiguous because the answer depends on whether or not we consider the order, in which the outcomes occur. For instance, we can get the sum of 4 as the result of rolling a 3 followed by a 1, or as the result of rolling a 1 followed by a 3—are these two outcomes different or do we consider them the same and identify such results?

These are two different problems. First we assume that the result depends on the order of addends. Thus, we have to compute the number of compositions of a natural number n, such that any part of a composition does not exceed 6, and there is no restriction on the number of parts. If the number m of rollings, that is, the number of parts of a composition, is given, we have a one-to-one correspondence between the compositions we sought and mappings

$$f : \{1, 2, \ldots, m\} \to \{1, 2, 3, 4, 5, 6\}$$

such that $f(1) + f(2) + \cdots + f(m) = n$. Since any composition has at least one part, such mappings are listed by the polynomial

$$(t + t^2 + \cdots + t^6)^m = t^m (1 + t + t^2 + \cdots + t^5)^m$$

and the GF for such mappings is

$$\sum_{m=1}^{\infty} t^m (1 + t + t^2 + \cdots + t^5)^m = \frac{t(1 + t + t^2 + \cdots + t^5)}{1 - t - t^2 - \cdots - t^6}. \tag{4.4.7}$$

If we want to avoid operations with infinite series, we can do that by specifying an n and truncating all series, keeping in mind that m cannot exceed n—even if a die shows a 1 every time, it is sufficient to roll the die n times to accumulate the sum of n.

For example, if $n = 4$, we can consider a polynomial

$$\sum_{m=1}^{4} (t + t^2 + \cdots + t^6)^m.$$

Moreover, if $n = 4$, the faces with 5 and 6 are irrelevant for the problem, and we can even consider a polynomial of a smaller degree,

$$\sum_{m=1}^{4} (t + t^2 + t^3 + t^4)^m.$$

After simple calculation we find the coefficient of t^4 in the latter polynomial to be 8, that is, the sum of 4 can be obtained in 8 ways shown in equation (4.4.8):

$$4 = \begin{cases} 4 \\ 1+3 \\ 3+1 \\ 2+2 \\ 1+1+2 \\ 1+2+1 \\ 2+1+1 \\ 1+1+1+1. \end{cases} \tag{4.4.8}$$

Let us return to an arbitrary n. If we allow m to be equal to 0, assuming by definition that there exists the unique composition consisting of zero parts, then the GF is simpler:

$$\sum_{m=0}^{\infty} t^m (1 + t + t^2 + \cdots + t^5)^m = \frac{1}{1 - t - t^2 - \cdots - t^6}$$

but obviously, the latter has the same coefficient of $t^n, \forall n \geq 1$, as the former one given by (4.4.7).

If we specify the number m of rollings, then the GF is $(t + t^2 + \cdots + t^6)^m$, where by the multinomial theorem Exercise 1.5.19 the coefficient of t^n is

$$\sum_{k_1+k_2+\cdots+k_6=m,\ k_1+2k_2+\cdots+6k_6=n} C(m; k_1, k_1, \ldots, k_6)$$

and $C(m; k_1, k_1, \ldots, k_6)$ are *multinomial coefficients* (1.5.1). Here k_i can be any whole numbers including zero. For example, if $n = 4$ and $m = 3$, the latter sum contains the only addend with $k_1 = 2$ and $k_2 = 1$ and reduces to $C(3; 1, 2) = 3$. Indeed, among the eight terms in (4.4.8) exactly three terms, $1 + 1 + 2, 1 + 2 + 1$, and $2 + 1 + 1$, contain 3 addends.

Now we solve Problem 4.4.5 assuming that the result does not depend on the order of faces. Therefore, we have to find the number of partitions with all parts not exceeding 6, and the GF is given by (4.4.2) with $q = 6$,

$$(1 + t + t^2 + \cdots) \times (1 + t^2 + t^4 + t^6 + \cdots) \times \cdots \times (1 + t^6 + t^{12} + t^{18} + \cdots).$$

In particular, the coefficient of t^4 here is 5, consequently, an unordered sum of 4 can be obtained in five ways, including a 4 itself, namely,

$$4 = \begin{cases} 4, \\ 3+1, \\ 2+2, \\ 2+1+1, \\ 1+1+1+1. \end{cases} \qquad \square$$

In the end of this section we apply the method of GF for solving linear recurrence relations (called also linear difference equations) with constant coefficients.

Definition 4.4.4. A sequence $\{a_n\}_{n=0}^{\infty}$ satisfies a linear (non-homogeneous) *difference equation* of order r if for all $n = 0, 1, 2, \ldots$, the equations

$$a_{n+r} = c_1 \cdot a_{n+r-1} + c_2 \cdot a_{n+r-2} + \cdots + c_r \cdot a_n + d_{n+r} \tag{4.4.9}$$

hold where $c_1, \ldots, c_r$ are given constant coefficients, $c_r \neq 0$, and $\{d_n\}_{n=0}^{\infty}$ is a given sequence. If all $d_n = 0$, then (4.4.9) is called *homogeneous*.

A well-known example of such a sequence is the sequence of the *Fibonacci numbers*

$$\{1, 1, 2, 3, 5, 8, \ldots\}$$

satisfying the difference equation $a_{n+2} = a_{n+1} + a_n$ for all $n \geq 0$. The theory of linear difference equations is in many instances similar to the theory of linear ordinary differential equations. In particular, the reader can readily prove the following *superposition principle* for linear difference equations.

Proposition 4.4.1.
(1) *If two sequences satisfy a linear homogeneous difference equation, then any linear combination of these sequences also satisfies this equation.*
(2) *A linear difference equation of order r has r linearly independent solutions; to specify a solution, one must assign r additional conditions* □

To develop the theory of linear difference equations, we prove in Theorem 4.4.4 that, if a sequence $\{a_n\}_{n=0}^{\infty}$ satisfies a difference equation (4.4.9), then its GF is a rational function. First we give one more definition.

Definition 4.4.5. The polynomial

$$g(t) = t^r - c_1 t^{r-1} - c_2 t^{r-2} - \cdots - c_{r-1} t - c_r$$

is called the *characteristic polynomial* of difference equation (4.4.9).

Let the characteristic polynomial $g(t)$ have the roots $\alpha_1, \alpha_2, \ldots, \alpha_s$ with multiplicities $l_1, l_2, \ldots, l_s$, $l_1 + l_2 + \cdots + l_s = r$. Introduce another polynomial

$$k(t) = t^r g(1/t).$$

It is well known that, if all coefficients c_i in (4.4.9) are real numbers (we consider only this case), then complex roots of g, if there are any, must appear in pairs of complex conjugate numbers, that is, if $\alpha = a + b\imath$ is a root, where a, b are real numbers and $\imath$ is the imaginary unit, $\imath^2 = -1$, then $\overline{\alpha} = a - b\imath$ also must be a root of the same multiplicity [37]. Moreover, g can be factored as

$$g(t) = (t - \alpha_1)^{l_1} \times \cdots \times (t - \alpha_s)^{l_s};$$

therefore

$$k(t) = (1 - \alpha_1 t)^{l_1} \times \cdots \times (1 - \alpha_s t)^{l_s}. \tag{4.4.10}$$

We use these observations to find the general solution of any linear homogeneous difference equation with constant coefficients. As before, the GF of the sequence $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$ is denoted by $f_{\mathbf{a}}(t) = \sum_{n=0}^{\infty} a_n t^n$.

Theorem 4.4.4. *The GF $f_{\mathbf{a}}$ of a sequence $\mathbf{a}$, satisfying a homogeneous equation* (4.4.9), *is a rational function,*

$$f_{\mathbf{a}}(t) = \frac{h(t)}{k(t)} \tag{4.4.11}$$

where the polynomial k of degree r was defined in (4.4.10) *and h is a polynomial of degree at most $r - 1$. Moreover,*

$$f_{\mathbf{a}}(t) = \sum_{n=0}^{\infty} \left\{ \sum_{i=1}^{s} P_i(n)\alpha_i^n \right\} t^n.$$

Thus, the general term of the sequence $\mathbf{a}$ is given by the expression

$$a_n = \sum_{i=1}^{s} P_i(n)\alpha_i^n, \quad n \geq 0, \tag{4.4.12}$$

where each polynomial P_i has degree $l_i - 1$, in particular, $P_i = \text{const}$ whenever α_i is a simple root.

Vice versa, if a_n are given by (4.4.12), *then the sequence $\{a_n\}_{n=0}^{\infty}$ satisfies the homogeneous equation* (4.4.9).

Proof. Multiplying the GF $f_{\mathbf{a}}(t) = \sum_{n=0}^{\infty} a_n t^n$ by the polynomial $k(t) = t^r g(1/t)$, and making use of equation (4.4.9), we readily verify that $h(t) = f_{\mathbf{a}}(t) \cdot k(t)$ is a polynomial of degree at most $r - 1$, which immediately implies (4.4.11).

To prove (4.4.12), we decompose the rational function (4.4.11) in the sum of partial fractions,

$$f_{\mathbf{a}}(t) = \sum_{i=1}^{s} \sum_{j=1}^{l_i} \beta_{ij}(1 - \alpha_i t)^{-j}$$

where β_{ij} are constants. Now (4.3.7) implies the expansion

$$(1 - \alpha t)^{-j} = \sum_{n=0}^{\infty} C(n + j - 1, j - 1)\alpha^n t^n.$$

Note also that

$$\sum_{j=1}^{l_i} \beta_{ij} C(n+j-1, j-1)\alpha_i^n = P_i(n)\alpha_i^n$$

where $P_i(n)$ is a polynomial of degree at most $l_i - 1$; moreover, any such polynomial can be obtained by an appropriate choice of the constants β_{ij}. This proves (4.4.12).

Each polynomial $P_i(n)$ has at most l_i nonzero coefficients, since its degree does not exceed $l_i - 1$. Hence, these polynomials altogether have $l_1 + \cdots + l_s = r$ coefficients and to find them we need r additional conditions; for example, we can assign the first r terms $\{a_0, a_1, \ldots, a_{r-1}\}$ of the sequence **a**. □

Remark 4.4.2. We immediately see that, if a root is simple, then the corresponding polynomial is a constant, hence it follows from (4.4.12) that, if all roots are simple, then every a_n is a linear combination of the nth powers of the s roots of the characteristic polynomial.

Problem 4.4.6. How many n-arrangements are there with repetition from the two-element set $A = \{a, b\}$, such that no two symbols a are situated next to one another?

Solution. Denote the number of these arrangements by $f(n)$. We define $f(0) = 1$, for there is a unique empty arrangement; it is also clear that $f(1) = 2$ since there are two such 1-element arrangements, namely (a) and (b). If $n \geq 2$, we can split all such arrangements into two disjoint subsets—those beginning with an a and beginning with a b. The second subset contains $f(n-1)$ arrangements because the first character b puts no restriction on the second symbol. However, the first subset contains $f(n-2)$ arrangements, since the first a must be followed by a b to avoid two repeating symbols a. By the sum rule

$$f(n) = f(n-1) + f(n-2), \quad n \geq 2. \tag{4.4.13}$$

Consequently, the sequence $f(n)$ satisfies the homogeneous second order difference equation (4.4.13) with the characteristic polynomial $g(t) = t^2 - t - 1$. The quadratic equation $g(t) = 0$ has two different roots $\alpha_1 = \frac{1}{2}(1 + \sqrt{5})$ and $\alpha_2 = \frac{1}{2}(1 - \sqrt{5})$, hence their multiplicities are $l_1 = l_2 = 1$. Thus, $P_i, i = 1, 2$, are zero-degree polynomials, that is, constants. Denoting $P_1 = p$ and $P_2 = q$, we get by (4.4.12)

$$f(n) = p\left(\frac{1+\sqrt{5}}{2}\right)^n + q\left(\frac{1-\sqrt{5}}{2}\right)^n, \quad n \geq 0.$$

The initial conditions $f(0) = 1$ and $f(1) = 2$ give the system of two linear algebraic equations

$$\begin{cases} p + q = 1, \\ \frac{1}{2}(1+\sqrt{5})p + \frac{1}{2}(1-\sqrt{5})q = 2, \end{cases}$$

which results in $p = \frac{\sqrt{5}+3}{2\sqrt{5}}$, $q = \frac{\sqrt{5}-3}{2\sqrt{5}}$, and finally

$$f(n) = \frac{\sqrt{5}+3}{2\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n + \frac{\sqrt{5}-3}{2\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n, \quad n \geq 0. \tag{4.4.14}$$

□

The terms of the sequence $\{1, 2, 3, 5, 8, 13, \ldots\}$ are called the *Fibonacci numbers*; they can also be defined by the same difference equation but with the initial conditions $f(0) = f(1) = 1$, leading to the sequence $\{1, 1, 2, 3, 5, 8, 13, \ldots\}$—see Exercise 4.4.2.

Problem 4.4.7. How many directed paths are there in the directed graph in Fig. 4.3, which start either at the vertex A or at B and arrive at the vertex C_n?

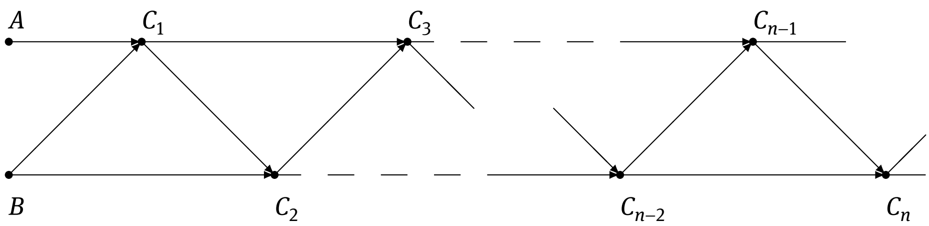


Figure 4.3: A graph in Problem 4.4.7.

Solution. Denote the number of paths from A to C_n by a_n and from B to C_n by b_n. It is clear that whether we start at A or at B, there are two ways to reach the vertex C_n—either through C_{n-1} and then down to C_n, or through C_{n-2} and then directly to C_n by the horizontal edge. Thus, we immediately derive the system of two decoupled linear difference equations

$$\begin{cases} a_n = a_{n-1} + a_{n-2}, \\ b_n = b_{n-1} + b_{n-2}, \end{cases}$$

with the initial conditions $a_1 = a_2 = 1$ and $b_1 = 1, b_2 = 2$. Therefore, a_n are the Fibonacci numbers, $a_n = f(n)$, and b_n are the shifted Fibonacci numbers, $b_n = f(n+1)$, and the total number of paths is (cf. Exercise 4.4.2)

$$a_n + b_n = f(n) + f(n+1) = f(n+2) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+2} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+2}\right].$$

□

One of many applications of the difference equations is the evaluation of determinants. Recall here a few definitions, also needed in subsequent sections. Consider a permutation[8] $g = (x_1, x_2, \ldots, x_m)$ of the first m natural numbers. The permutation g is said to have an *inversion* if there are indices $i < j$ such that $x_i > x_j$. A permutation is called even (odd) if it has an even (odd) number of inversions.

Definition 4.4.6. The *determinant* of an $n \times n$ matrix

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ & & \vdots & \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

is the alternating sum of $n!$ products, such that each product contains one element from every row and from each column of the matrix. Here "alternating" means that the product $a_{1,i_1} a_{2,i_2} \cdots a_{n,i_n}$ has the sign $(-1)^{\sigma(i_1,i_2,\ldots,i_n)}$, where $\sigma(i_1, i_2, \ldots, i_n)$ is the number of inversions in the permutation $(i_1, i_2, \ldots, i_n)$. The determinant of the matrix M is denoted by

$$\det(M) = \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ & & \vdots & \\ a_{n1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix}.$$

For instance, if $M = (a)$ is a 1×1 matrix, its determinant is the unique entry of the matrix M, $\det(M) = a$; if $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is a 2×2 matrix, then $\det(M) = ad - bc$. The determinant of an $n \times n$ matrix can be computed by using the expansion across the first row,

$$\begin{aligned} \det(M) = a_{1,1} \det(M_{1,1}) - a_{1,2} \det(M_{1,2}) \\ + a_{1,3} \det(M_{1,3}) - \cdots + (-1)^{n-1} a_{1,n} \det(M_{1,n}) \end{aligned}$$

where $M_{1,i}$, $1 \le i \le n$, are $(n-1) \times (n-1)$ matrices obtained from M by deleting its first row and ith column. The determinant $\det(M_{1,i})$ is called the *minor* of a matrix element $a_{1,i}$. Quite similarly, one can expand a determinant along any row or column. This is a *recursive procedure*: we reduce, step-by-step, the order of a determinant to be computed, until we reach 2×2 determinants that can be calculated straightforwardly, and then we work backward, computing minors of third, fourth, etc., orders. The procedure described is lengthy and there are different ways to speed up the computations; see, e. g. [42]. In the following problem we show that some determinants can be efficiently computed by making use of difference equations.

8 Permutations are studied in more detail in Section 4.5.

Problem 4.4.8. Compute a three-diagonal determinant (a special case of the Jacobi determinant) of order n,

$$d_n = \begin{vmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & 1 & \dots & 0 \\ & & \vdots & & & \\ 0 & \dots & 0 & 0 & 1 & 1 \end{vmatrix}.$$

Solution. Expanding d_n along the first row (or the first column), we immediately derive a recurrence relation $d_n = d_{n-1} - d_{n-2}$. Its characteristic polynomial $g(t) = t^2 - t + 1$ has simple complex roots $\exp\{\pm\frac{\pi i}{3}\}$, so $d_n = pe^{n\frac{\pi i}{3}} + qe^{-n\frac{\pi i}{3}}$, where p and q are constants, that is, polynomials of zero degree. The values of the determinants $d_1 = |1| = 1$ and $d_2 = \left|\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right| = 0$ are immediate, and we use them to set up a system of algebraic linear equations

$$\begin{cases} pe^{i\frac{\pi}{3}} + qe^{-i\frac{\pi}{3}} = 1 \\ pe^{i\frac{2\pi}{3}} + qe^{-i\frac{2\pi}{3}} = 0. \end{cases}$$

Solving this system, we find $p = (e^{i\frac{\pi}{3}} + 1)^{-1}$, $q = (e^{-i\frac{\pi}{3}} + 1)^{-1}$, and finally $d_n = \frac{2}{\sqrt{3}} \sin \frac{(n+1)\pi}{3}$.

Similarly, a determinant of order n,

$$d_n = \begin{vmatrix} 2 & 1 & 0 & 0 & \dots & 0 \\ 1 & 2 & 1 & 0 & \dots & 0 \\ 0 & 1 & 2 & 1 & \dots & 0 \\ & & \vdots & & & \\ 0 & \dots & 0 & 0 & 1 & 2 \end{vmatrix},$$

satisfies the equation $d_n = 2d_{n-1} - d_{n-2}$. However, in this case the characteristic equation $t^2 - 2t + 1 = (t-1)^2 = 0$ has a multiple root $\alpha = 1$ of multiplicity $l = 2$, therefore, the general solution of the equation $d_n = 2d_{n-1} - d_{n-2}$ should be looked at as $d_n = (p \cdot n + q)\alpha^n = p \cdot n + q$. Again, we immediately compute the determinants of orders 1 and 2, $d_1 = |2| = 2$ and $d_2 = \left|\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right| = 3$, and use these values to set up an algebraic system of two linear equations

$$\begin{cases} p + q = 2, \\ 2p + q = 3. \end{cases}$$

From this $p = q = 1$, and we find $d_n = n + 1$. □

Problem 4.4.9. In how many parts do n convex closed curves divide the plane if any two curves have two common points, but no three have a common point?

Solution. Denote by a_n the number of parts generated by n curves. If we add one more, $(n+1)$st curve to the existing n curves, it has two intersection points with each of the initial n curves, totaling to $2n$ points. These $2n$ points split the new, $(n+1)$st curve into $2n$ pieces, and each of these pieces divides exactly one of the initial a_n parts of the plane into two pieces. Therefore, the new curve increases the number of parts, the plane was decomposed to by n curves, by $2n$ resulting in the difference equation

$$a_{n+1} - a_n = 2n. \tag{4.4.15}$$

Unlike the preceding ones, this difference equation is non-homogeneous, therefore, its general solution is the sum of the general solution of the corresponding homogeneous equation and a particular solution of non-homogeneous equation (4.4.15). The homogeneous equation $a_{n+1} - a_n = 0$ has a linear characteristic polynomial $g(t) = t - 1$ with one simple root $\alpha = 1$, thus the general solution of the homogeneous equation is $a_n^{\text{hom}} = p\alpha^n = p$, where p is an arbitrary constant.

The right-hand side of (4.4.15) is $2n$—this is a polynomial in n, therefore we look for a particular solution of the non-homogeneous equation as a polynomial as well. However, $2n$ is a first-degree polynomial in n, and its degree, which is 1, is a root of the characteristic polynomial $g(t)$, so that we must look for a particular solution as a second-degree polynomial $a_n^{nonhom} = qn^2 + rn + s$. Inserting this into (4.4.15) we get $q = 1, r = -1$; s may be any number, we choose $s = 0$. Thus, $a_n^{non\text{-hom}} = n^2 - n$ and $a_n = a_n^{non\text{-hom}} + a_n^{\text{hom}} = n^2 - n + p$. Next, $a_1 = 2$ because one convex closed curve divides the plane in two parts, cf. Theorem 2.5.1. Using this initial condition, we find $p = 2$ and $a_n = n^2 - n + 2$. □

The next problem deals with a sequence $\{c_n\}_{n=0}^{\infty}$ satisfying a nonlinear difference equation. Such equations, like their differential counterparts, can be explicitly solved only in rare occasions. In this problem we are able to solve a nonlinear equation by using the GF of the sequence sought.

Problem 4.4.10. In how many ways can one compute the product of $n+1$ quantities $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ taken in this fixed order, if the multiplication is non-associative?

Solution. Associativity means that $(ab)c = a(bc)$, thus, the question is, in how many ways can we insert parentheses among $n+1$ factors $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ taken in this order? Denote the number of possible products by c_n. For example, if $n = 1$, there is only one way to multiply two elements α_1, α_2, namely, $\alpha_1 \times \alpha_2$, hence $c_1 = 1$. However, for $n = 2$ there are two possibilities, $(\alpha_1 \times \alpha_2) \times \alpha_3$ and $\alpha_1 \times (\alpha_2 \times \alpha_3)$, thus, $c_2 = 2$. We leave it to the reader to verify that $c_3 = 5$. It is convenient to define $c_0 = 1$.

To derive the GF

$$f_{\mathbf{c}}(t) = c_0 + c_1 t + \cdots + c_n t^n + \cdots$$

let us notice that it is always possible to determine the position of the very last multiplication, that is, we can find an element α_r such that in order to compute the entire product, we multiply the left-most r elements, and independently multiply the other rightmost $n+1-r$ elements, and only after that multiply the two partial products; in the example with two factors above $r=1$, and in the example with three factors either $r=2$ or $r=1$. Now let us notice that there are c_{r-1} ways to multiply the left-most r elements and c_{n-r} ways to multiply the other $n+1-r$ elements. By the product rule, there are $c_{r-1}\cdot c_{n-r}$ ways to calculate this product with the last multiplication after the α_r. Since r runs from 1 through n, we have by the sum rule

$$c_n = c_0\cdot c_{n-1} + c_1\cdot c_{n-2} + \cdots + c_{r-1}\cdot c_{n-r} + \cdots + c_{n-1}\cdot c_0, \quad n\ge 1. \tag{4.4.16}$$

Comparing (4.4.16) with the convolution $c * c$, we see that the right-hand side of (4.4.16) is the coefficient of t^n in the power series of $(f_{\mathbf{c}}(t))^2$. Computing $f_{\mathbf{c}}^2(t)$ and using the condition $c_0 = 1$, we derive a quadratic equation for $f_{\mathbf{c}}(t)$,

$$tf_{\mathbf{c}}^2(t) - f_{\mathbf{c}}(t) + 1 = 0,$$

which has roots $\frac{1}{2t}(1\pm\sqrt{1-4t})$. Since $f_{\mathbf{c}}(0) = c_0 = 1$, we see that

$$\frac{1+\sqrt{1-4t}}{2t}$$

is an extraneous root, so that we must set

$$f_{\mathbf{c}}(t) = \frac{1-\sqrt{1-4t}}{2t}.$$

By making use of the result of Exercise 4.3.32

$$(1+x)^{1/2} = 1 + \frac{1}{2}x + \cdots + \frac{\frac{1}{2}(\frac{1}{2}-1)(\frac{1}{2}-2)\cdots(\frac{1}{2}-n+1)}{n!}x^n + \cdots$$

with $x=-4t$, we expand $(1-4t)^{1/2}$ in the power series and get again the Catalan numbers (Remark 1.4.2) $\mathrm{Cat}_n = \frac{1}{n+1}C(2n,n)$. □

Recurrence relations allow us to find certain sums explicitly, in closed form. Let us compute again the sum found in Problems 1.1.4 and 4.3.11.

Problem 4.4.11. Evaluate anew the sum $s(n)=\sum_{k=1}^{n}k^2$, now by making use of a recurrence relation.

Solution. The equation $s(n+1)-s(n)=(n+1)^2$ is obvious. This is a first-order linear non-homogeneous recurrence equation with the characteristic polynomial $t-1=0$ and with a quadratic polynomial on the right. Hence, we look for a particular solution of the non-homogeneous equation as a polynomial of third degree, $s(n)=an^3+bn^2+cn+d$. Inserting this polynomial in the equation and equating the coefficients of n^3, n^2, and n, we find $a=1/3, b=1/2, c=1/6, d$ remains undetermined. We set $d=0$, since after all we have to add the general solution of the corresponding homogeneous equation, which is $p(1)^n = p$, hence, it is a constant p as well. Satisfying the obvious initial condition $s(1) = 1$, we find that $p = 0$ and finally we again derive the formula

$$s(n) = \frac{1}{3}n^3 + \frac{1}{2} + \frac{1}{6}n = \frac{1}{6}n(n+1)(2n+1). \qquad \square$$

The last problem of this section employs some elementary complex analysis. GF, used in this text, have converging Taylor series, therefore their sums are analytic functions within the circle of convergence and the powerful techniques of complex analysis may be used in applications of this method. We consider one typical example, referring the reader to [15] for a detailed treatment of the topic.

Problem 4.4.12. (Hardy) Compute the sum

$$H(m) = \sum_{k=0}^{[m/2]} \frac{(-1)^k}{m-k} C(m-k, k), \quad m = 1, 2, \ldots$$

where $[m/2]$ is the integer part of $m/2$.

Solution. Applying (1.4.2) we readily verify the identity

$$\frac{1}{m-k} C(m-k, k) = \frac{1}{m}\{C(m-k, k) + C(m-k-1, k-1)\}$$

or

$$\frac{1}{m-k} C(m-k, k) = \frac{1}{m}\{C(m-k, m-2k) + C(m-k-1, m-2k)\}. \tag{4.4.17}$$

Consider the contour integral

$$I(m, n) = \frac{1}{2\pi i} \oint_{|w|=\frac{1}{2}} w^{-n-1}(1+w)^m \, dw.$$

The integrand has two singular points, $w = 0$ and $w = -1$, but only the first one lies inside the contour $|w| = \frac{1}{2}$. Computing the residue at the $(n+1)$-fold pole $w = 0$, we deduce the formula $I(m, n) = C(m, n)$. From this and (4.4.17),

$$\frac{1}{m-k} C(m-k, k) = \frac{1}{2\pi i m} \oint_{|w|=\frac{1}{2}} w^{-m+2k-1}(1+w)^{m-k-1} \, dw.$$

Therefore, for $k > [m/2]$ the integrand is a holomorphic function in the disk $|w| < 1$ and

$$\begin{aligned} H(m) &= \sum_{k=0}^{[m/2]} \frac{(-1)^k}{2\pi i m} \oint_{|w|=\frac{1}{2}} w^{-m+2k-1}(1+w)^{m-k-1} \, dw \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2\pi i m} \oint_{|w|=\frac{1}{2}} w^{-m+2k-1}(1+w)^{m-k-1} \, dw. \end{aligned}$$

The geometric series $\sum_{k=0}^{\infty}(\frac{-w^2}{1+w})^k$ converges uniformly for $|w| \leq q < 1$; hence, the order of summation and integration can be interchanged, yielding

$$H(m) = \frac{1}{2\pi i m} \oint_{|w|=\frac{1}{2}} w^{-m-1}(2+w)(1+w)^{m-1}\left\{\sum_{k=0}^{\infty}(-w^2/(1+w))^k\right\} dw$$
$$= \frac{1}{2\pi i m} \oint_{|w|=\frac{1}{2}} w^{-m-1}(2+w)(1+w)^{m} \frac{dw}{1+w+w^2}$$

where we have used the formula for the sum of geometric series. Instead of calculating the residue at a multiple pole $w = 0$, it is more convenient to change the direction we traverse the contour and consider this integral over the boundary of the exterior domain $|w| > \frac{1}{2}$, where the integrand has only two *simple* poles at $w = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$, since the residue at infinity is equal to zero. Computing residues at these points, we get

$$H(m) = \frac{2(-1)^m}{m}\cos\left(\frac{2}{3}m\pi\right) = \begin{cases} (-1)^m \frac{2}{m} & \text{if } m \equiv 0 \pmod 3, \\ (-1)^{m-1}\frac{1}{m} & \text{if } m \equiv \pm 1 \pmod 3). \end{cases}$$ □

Exercises 4.4.

Exercise 4.4.1. Prove Proposition 4.4.1.

Exercise 4.4.2. Solve again equation (4.4.13) for the Fibonacci numbers,

$$f(n) = f(n-1) + f(n-2),$$

now with the initial conditions $f(0) = f(1) = 1$, and derive the formula

$$f(n) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

Calculate the first six Fibonacci numbers by making use of this formula.

The number $\frac{1}{2}(1+\sqrt{5})$ is called the *golden ratio*.

Exercise 4.4.3. Prove that the Fibonacci numbers f_n satisfy the equations

(1) $f(m+n) = f(m)f(n) + f(m-1)f(n-1)$,
(2) $f(1) + f(3) + \cdots + f(2n+1) = f(2n+2)$,
(3) $1 + f(2) + f(4) + \cdots + f(2n) = f(2n+1)$,
(4) $f(n+1) = C(n,0) + C(n-1,1) + \cdots + C(n-k,k), k = [n/2]$,
(5) $f(n+2) - 1 = f(0) + f(1) + \cdots + f(n)$;
(6) the latter equation is a special instance of the identity

$$\sum_{k=0}^{m} C(m+k-1,k)f(n-k) + \sum_{k=1}^{m} C(n+k-1,n)f(2m+1-2k) = f(n+2m).$$

Exercise 4.4.4. Prove that the sum of any 8 consecutive Fibonacci numbers is not a Fibonacci number.

Exercise 4.4.5. Let a sequence $\{g(n)\}_{n=0}^{\infty}$ satisfy the difference equation

$$g(n) = ag(n-1)$$

where a is a constant number and $g(0) = \alpha$. Prove that the EGF of this sequence, $g(t)$, satisfies the functional equation

$$e^{-at}g(t) - e^{at}g(-t) = 0.$$

Exercise 4.4.6. Prove that, for any solution of the difference equation

$$a_{n+2} = a_{n+1} + a_n$$

independently from the initial data, both absolute values $|a_{n+1}a_{n-1} - a_n^2|$ and $|a_{n+2}a_{n-1} - a_{n+1}a_n|$ do not depend on n. For the Fibonacci numbers, each of these values is 1.

Exercise 4.4.7. Prove that the nth Fibonacci number $f(n)$ is equal to the continuant (a special three-diagonal determinant)

$$f(n) = \begin{vmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & 1 & \dots & 0 \\ & & \vdots & & & \\ 0 & \dots & 0 & 0 & -1 & 1 \end{vmatrix}.$$

Exercise 4.4.8. Prove that for every natural n the number

$$a_n = \left(\frac{3+\sqrt{5}}{2}\right)^n + \left(\frac{3-\sqrt{5}}{2}\right)^n - 2$$

is equal to m^2 for some natural m if n is odd, or else it is equal to $5m^2$ for some natural m if n is even.

Exercise 4.4.9. A frog sits initially at the point of the number line marked by 1. From any point $k, k = 1, 2, \dots$, it can jump for one or two steps to the right, either to $k+1$ or to $k+2$. In how many ways can the frog reach the point n from its initial location? Two ways are identical if the frog visits the same points.

Exercise 4.4.10. How many 12-digit natural numbers are there, containing only digits 3 and 9, such that no two digits 3 come together?

Exercise 4.4.11. Consider all bit strings of length n, that is, n-permutations

$$\beta = (\beta_1, \beta_2, \dots, \beta_n)$$

where each β_i is either a 0 or a 1. How many of these permutations are there such that $\beta_i = \beta_{i \,(\mathrm{mod}\, n)+1} = 0$ for every $i, 1 \le i \le n$?

Exercise 4.4.12. Prove that the GF for the Fibonacci numbers $f(n)$, is

$$f(1)t + f(2)t^2 + \cdots + f(n)t^n + \cdots = \frac{t}{1 - t - t^2}.$$

Exercise 4.4.13. Prove that the nth Fibonacci number $f(n)$ is the closest integer to the power $\frac{1}{\sqrt{5}}(\frac{1+\sqrt{5}}{2})^n$.

Exercise 4.4.14. The *Lucas numbers* L_n satisfy the same difference equation (4.4.13) as the Fibonacci numbers $f(n)$, $L_n = L_{n-1} + L_{n-2}$; however, with the initial conditions $L_0 = 2$ and $L_1 = 1$. Find an explicit formula for the Lucas numbers. Compare the first six Lucas and Fibonacci numbers.

Exercise 4.4.15. Prove that $L_n = f(n-1) + f(n+1)$.

Exercise 4.4.16. Try to find a particular solution of the non-homogeneous difference equation (4.4.15) using a first-degree polynomial $pn + q$ and see for yourself why this approach does not work and we had to use a second-degree polynomial.

Exercise 4.4.17. Find the general solutions of difference equations
(1) $x_{n+2} + x_{n+1} - 2x_n = 0$,
(2) $x_{n+2} + 4x_n = 0$,
(3) $x_{n+3} - 3x_{n+2} + 3x_{n+1} - x_n = 0$,
(4) $x_{n+2} + 2x_{n+1} - 3x_n = 5 \cdot 2^n, x_0 = 0, x_1 = 1$,
(5) $x_{n+2} + 2x_{n+1} - 3x_n = 5, x_0 = 0, x_1 = 1$.

Exercise 4.4.18. Prove that, if $a_n = b_n + b_{n-1}, \forall n \geq 1$, and $a_0 = b_0$, then $f_{\mathbf{a}}(t) = (1 + t)f_{\mathbf{b}}(t)$.

Exercise 4.4.19. An *arithmetic progression* can be defined as a solution of the recurrence relation $a_{n+1} - a_n = d$. Find the general term a_n of the arithmetic progression as a function of d and a_1 by solving this recurrence relation.

Exercise 4.4.20. A *geometric progression* can be defined as a solution of the recurrence relation $a_{n+1} = q \cdot a_n$. Find the general term a_n of the geometric progression by solving this recurrence relation.

Exercise 4.4.21. Kate invested $1 200 at 6 % interest rate compounded monthly. If after every month she withdraws $50, find the balance in her account after one year.

Exercise 4.4.22. Prove that the Catalan numbers Cat_n satisfy the recurrence relation $\mathrm{Cat}_{n+1} = \frac{4(4n^2-1)}{(n+1)(n+2)}\mathrm{Cat}_{n-1}, n \geq 1$.

Exercise 4.4.23. Find a sequence **a** such that its GF is
(1) $f_{\mathbf{a}}(t) = \sqrt{2 - t}$,
(2) $f_{\mathbf{a}}(t) = \log(1 + t)$.

Exercise 4.4.24. The characteristic equation of a linear homogeneous recurrence relation has roots 0, 1, −1, 3. Find the general solution of this recurrence relation and write the relation explicitly.

Exercise 4.4.25. The ratio of a geometric progression is $\frac{1+\sqrt{5}}{2}$. Prove that each term of the sequence, starting from the second one, is the difference of the two its neighbors.

Exercise 4.4.26. Solve the following systems of difference equations
(1)

$$\begin{cases} x_{n+1} = y_n - 2, \\ y_{n+1} = x_n + 3, \end{cases}$$

(2)

$$\begin{cases} x_n = y_{n-1} - y_{n-2} + 4, \\ y_n = y_{n-1} + x_{n-1}, \end{cases} \qquad x_1 = 3, x_2 = 5, y_1 = 1.$$

Exercise 4.4.27. Find the GF for the number a_n of whole (nonnegative integer) solutions $\{x, y, z, t\}$ of the equation

$$x + 2y + 5z + 7t = n.$$

Exercise 4.4.28. In how many ways can a natural number n be written as a sum of three natural addends?

Exercise 4.4.29. In how many ways can a natural number be represented as a sum of certain natural addends?

Exercise 4.4.30. Use formula (1.2.2) in Exercise 1.2.10 to prove that the number of partitions of n with k terms is the same as the number of partitions of n with each term not exceeding k.

Exercise 4.4.31. Denote by $\mathrm{Comp}(m, n; k)$ the number of compositions of a natural number m with n parts not exceeding k. Prove the recurrence relations

$$\mathrm{Comp}(m, n; k) = \mathrm{Comp}(m - 1, n; k) + \mathrm{Comp}(m - 1, n - 1; k) - \mathrm{Comp}(m - k - 1, n - 1; k)$$

and

$$\mathrm{Comp}(m, n; k) = \sum_{j=0}^{n} C(n, j)\mathrm{Comp}(m - jk, n - k; k - 1).$$

Exercise 4.4.32. How many ways are there to pay 90 cents using quarters, dimes, and nickels? First set up the GF for this quantity.

Exercise 4.4.33. Is it possible to change a silver dollar using exactly 50 coins? If yes, in how many ways?

Exercise 4.4.34. The postage for a letter is 97 cents. In how many ways is it possible to buy stamps if the post office has 42-cent, 20-cent, 3-cent, and 1-cent stamps? Consider two cases, when the order of stamps does or does not matter.

Exercise 4.4.35. Together, 30 members of the Combi Club have composed 40 problems for the Club contest. Among the members there are freshmen, sophomores, juniors, seniors, and graduate students. Any two students of the same rank composed the same number of problems, while any two students of different ranks composed different numbers of problems. How many students composed one problem?

Exercise 4.4.36. In how many ways can a number 1 000 000 be represented as a product of three natural numbers, if the order of factors does not count?

Exercise 4.4.37. Find the number of terms in an expansion $(x+y+z)^n$ after combining like terms. For example, the expansion $(x+y)^2 = x^2 + 2xy + y^2$ contains three terms.

Exercise 4.4.38. Which is larger, the number of all partitions of a natural number n or the number of partitions of $2n$ in n parts?

Exercise 4.4.39. In how many parts is a sphere divided by n planes containing the center of the spheres, if no three planes contain the same diameter of the sphere?

Exercise 4.4.40. At a hot dog eating contest, everyone of n participants ate no more than m hot dogs. Denote by $c_i, 1 \le i \le n$, the number of frankfurters consumed by the ith contestant, and by $d_k, 0 \le d_k \le m$, the number of contestants consumed at least k hot dogs. Prove that

$$c_1 + c_2 + \cdots + c_n = d_1 + d_2 + \cdots + d_m.$$

Exercise 4.4.41. Find closed-form formulas for the following sums
(1) $\sum_{k=1}^{n} k2^k$,
(2) $\sum_{k=1}^{n} k^2 2^k$,
(3) $\sum_{k=1}^{n} k^2 2^{-k}$.

Exercise 4.4.42. Compute a three-diagonal determinant of order n

$$d_n = \begin{vmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & -1 & 0 & \dots & 0 \\ & & & \vdots & & & \\ 0 & 0 & \dots & 0 & 0 & 1 & -1 \\ 0 & 0 & \dots & 0 & 0 & 1 & 1 \end{vmatrix}.$$

Exercise 4.4.43. For every integer $n \geq 0$, compute the determinant of order $k+1$

$$d_{n,k} = \begin{vmatrix} C(n,0) & C(n,1) & \dots & C(n,k) \\ C(n+1,0) & C(n+1,1) & \dots & C(n+1,k) \\ \vdots & \vdots & \vdots & \vdots \\ C(n+k,0) & C(n+k,1) & \dots & C(n+k,k) \end{vmatrix}.$$

Definition 4.4.7. If we drop the sign $(-1)^\sigma$ of each term in Definition 4.4.6 of the determinant of a matrix, the resulting number is called the *permanent* of this matrix.

Exercise 4.4.44. Compute the permanents of the square matrices leading to the determinants in Exercises 4.4.7, 4.4.8, 4.4.42, 4.4.43.

Exercise 4.4.45. Prove that, if $a_n = c_n a_{n-1} + d_n$, where $\{c_n\}$ and $\{d_n\}$ are given sequences, $c_0 = 0$, and $a_0 = d_0$, then

$$a_n = \sum_{k=0}^{n} \left(\prod_{j=k+1}^{n} c_j \right) d_k.$$

Assume that $\prod_{j=n+1}^{n} = 1$.

Exercise 4.4.46. In how many parts do n lines split a plane, if no two lines are parallel and no three of them intersect at a point?

Exercise 4.4.47. In how many ways can a convex n-gon be split in $n-2$ triangles by $n-3$ nonintersecting diagonals? Derive the GF and compare it with that for the Catalan numbers Cat_n.

Exercise 4.4.48. Recall that the Bell numbers B_n (Definition 1.1.7) count the number of partitions of an n-element set. Derive a difference equation for the Bell numbers, $B_n = \sum_{k=1}^{n} C(n-1,k-1)B_{n-k}$. This is a linear difference equation of variable order.

Exercise 4.4.49. Let $\mathrm{Cat}(t)$ be the GF for the Catalan numbers Cat_n and a be a whole number. Prove that $\frac{(\mathrm{Cat}(t))^a}{\sqrt{1-4t}}$ is the GF for the sequence of binomial coefficients $\{C(2k+a,k)\}$ and $(\mathrm{Cat}(t))^a$ for $\{\frac{a}{a+2k}C(2k+a,k)\}$ [40].

Exercise 4.4.50. Use EGF to prove the inversion formulas

$$a_n = \sum_k C(n,k)(x+k)^{n-k} b_k, \quad \forall n,$$

$$\Updownarrow$$

$$b_n = \sum_k C(n,k)(-1)^{n-k}(x+n)^{n-k-1}(x+k)a_k, \quad \forall n.$$

Exercise 4.4.51. Prove that the permanent of the permutation matrix of order n is equal to 1.

Exercise 4.4.52. Compute the permanent of the matrix $\left(\begin{smallmatrix} 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{smallmatrix}\right)$.

4.5 Enumeration of equivalence classes

This section is devoted to the Pólya–Redfield enumeration theory, which gives a general method of deriving GF in various problems, in particular, in the problems where one has to find the number of equivalence classes. We have already solved such problems when all the equivalence classes had the same cardinality and it was enough to apply Lemma 1.1.4 or some equivalent statements. However, simple examples like Problem 4.5.1 show that equivalence classes can have different cardinalities.

Moreover, not only can equivalence classes have different cardinalities, but the elements of sets under consideration may have various weights. Pólya's theory applies to such problems as well. The subsequent exposition of the main Pólya theorem [44] follows N. G. De Bruijn [5, pp. 144–184]. Other approaches to this theory can be found, for example, in [47] or in [7, Appendix by J. Riget].

Coffee-time browsing

- http://en.wikipedia.org/wiki/Nicolaas_Govert_de_Bruijn (De Bruijn's biography and work)
- http://en.wikipedia.org/wiki/Plato (Plato)
- www.georgehart.com/virtual-polyhedra/platonic-info.html (Platonic Solids)
- http://en.wikipedia.org/wiki/William_Burnside (Burnside's biography)
- http://en.wikipedia.org/wiki/Ferdinand_Georg_Frobenius (Frobenius' biography)
- chemistry.about.com/od/chemistryglossary/a/valencedef.htm (Valence)

The next problem illustrates some basic ideas of the theory.

Problem 4.5.1. Consider ten Hindu–Arabic numerals $0, 1, \dots, , 9$. Some of them, like 7, after rotating upside down through 180° in their plane become meaningless symbols. However, some others after this rotation interchange with another digit, for example, 6 becomes 9 and vice versa; moreover, the digits $0, 1, 8$ do not change at all.

Denote by D the set of the whole numbers with five digits, if a number consists of less than five digits, we add a few zeros in front of such a number, like 00236; thus $|D| = 10^5$. Two such five-digit numbers are said to be equivalent if one of them can be transformed into another by rotating through the angle of 180° or 0° without removing off the plane. How many non-equivalent numbers are there?

Solution. To solve the problem, we consider two mappings $g_i : D \to D, i = 0, 1$, where $g_0 : D \to D$ is the identical mapping of D, which clearly does not change any number, while the mapping $g_1 : D \to D$ rotates a number upside down if the result is a number in D, and leaves the number unchanged if the number cannot be rotated. For example, $g_1(19\,806) = 90\,861$ and $g_1(12\,880) = 12\,880$. The reader can readily verify that the binary relation on the set of all whole five-digit numbers described in this problem is an equivalence relation. Thus, Problem 4.5.1 can be stated as follows.

Consider the following binary relation ϱ on the set D, which is easily seen to be an equivalence relation.

Two numbers $d_1, d_2 \in D$ are in the binary relation ϱ if and only if either $d_1 = d_2$, that is, $d_2 = g_0(d_1)$, or $d_2 = g_1(d_1)$. How many equivalence classes exist in D with respect to this equivalence relation?

We finish the solution of Problem 4.5.1 after the proof of Lemma 4.5.1.

Hereafter, bijections of a set D onto itself are called *substitutions* or *permutations* of the elements of D. The discussion in Problem 4.5.1 suggests that two elements d_1 and d_2 of a given set D should be considered indistinguishable (identical, equivalent) if there exists a *substitution* g of the elements of D such that $g(d_1) = d_2$. Any set G of substitutions generates the following binary relation ρ on D.

Two elements $d_1, d_2 \in D$ are ϱ-equivalent if and only if there exists a substitution $g \in G$ such that $g(d_1) = d_2$.

We want to find conditions that guaranty that this binary relation ϱ is an equivalence relation on D, that is, it is reflexive, symmetric, and transitive. It is clear that in order for the binary relation ϱ to be reflexive it is sufficient if G contains the identical substitution. For ϱ to be symmetric, it is enough if along with each substitution g, its inverse g^{-1} also belongs to G. To guarantee the transitivity of the binary relation ϱ, the set of substitutions G must contain the superposition $g_1 \circ g_2$ of any two of its elements $g_1, g_2 \in G$. Thus, to generate an equivalence relation on D, it suffices for G to have a special algebraic structure, namely, to be a *group* of substitutions with the superposition of substitutions as the group operation. We recall the definition. □

Definition 4.5.1. A set X with a binary operation $\circ$ defined on the Cartesian product $X \times X$ is called a *group* if

(1) X has the *neutral element* e such that $e \circ x = x \circ e = x, \forall x \in X$,
(2) the operation $\circ$ is *associative*, that is, $x \circ (y \circ z) = (x \circ y) \circ z, \forall x, y, z \in X$,
(3) each element $x \in X$ is *invertible*, that is, for any $x \in X$ there exists an element $x^{-1} \in X$ such that $x \circ x^{-1} = x^{-1} \circ x = e$.

The cardinality of a group is called the *order of the group*.

Example 4.5.1. For instance, all $n!$ substitutions of an n-element set D make a group, called the symmetric group Sym_n; it follows from Lemma 1.1.7 that the order of this group is $n!$.

Thus, in this section we consider an m-element set $D = \{d_1, d_2, \dots, d_m\}$ and a group of substitutions G acting on D. The group operation is the superposition of substitutions. This group may be the entire *symmetric group* $\mathrm{Sym} = \mathrm{Sym}_m$ of all substitutions of D, in which case $|G| = m!$, or it may be any subgroup of Sym, for example, a trivial group $\{e\}$ of order 1 consisting of the only neutral element (substitution) $e \in G$.

We study certain properties of *substitutions*. Any substitution $g \in G$ splits D into disjoint subsets called *cycles* or *orbits*. Namely, fix an element $d \in D$ and consider a sequence of elements

$$g(d),\ g(g(d)) = g^2(d), \quad g(g(g(d))) = g^3(d), \dots$$

where we have used a standard notation $g(g^k) = g^{k+1}, g^1 = g$. Since the set D is finite, after a several steps some element in this sequence must repeat. Suppose that for a given $d \in D$ all k elements $d, g(d), g^2(d), \dots, g^{k-1}(d)$ are different, but $g^k(d) = d$. Then these k elements

$$D_1 = \{d, g(d), g^2(d), \dots, g^{k-1}(d)\}$$

are said to form a *cycle of length k*.

If a cycle $D_1 \neq D$, that is $k < m$, then there exists an element $d_1 \in D \setminus D_1$, which generates another cycle starting with d_1 and disjoint with D_1, and so on. After a finite number of steps each element of D will get into one and only one cycle. It is possible that each element forms its own 1-element cycle; this is the case for the neutral (identical) substitution e. As another extreme, it is possible that all elements of D belong to one cycle.

Definition 4.5.2. If a substitution g splits D in b_1 cycles of length 1 (1-element cycles), b_2 cycles of length 2, b_3 cycles of length 3, etc., g is said to have the *cycle type* $(b_1, b_2, b_3, \dots)$.

Since D is finite, all but a finite number of entries of the cycle type sequence are zeros; more specifically, it is obvious that $b_{m+1} = b_{m+2} = \cdots = 0$. It is also clear that $b_1 + b_2 + b_3 + \cdots = |D|$. Thus we write cycle type sequences as m-element sequences $(b_1, b_2, b_3, \dots, b_m)$. Consequently, any substitution gives rise to a partition of the integer number $|D|$, however, *different substitutions can yield the same partition*.

Substitutions acting on an m-element set $D = \{x_1, \dots, x_m\}$, can be conveniently represented by $2 \times m$ matrices (we use the same symbol for a substitution and for its matrix)

$$g = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ x_{j_1} & x_{j_2} & \dots & x_{j_m} \end{pmatrix}$$

where $x_{j_i} = g(x_i)$. For example, the identical substitution e is given by the matrix

$$g_0 = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ x_1 & x_2 & \dots & x_m \end{pmatrix}$$

clearly exhibiting its structure—every element remains unmoved and makes its own cycle, the cycle type of e is $(m, 0, 0, \dots, 0)$. On the other hand, a substitution g_c given by the matrix

$$g_c = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ x_2 & x_3 & \dots & x_1 \end{pmatrix} \tag{4.5.1}$$

consists of the only cycle of length m, because g_c moves x_1 to x_2, then x_2 to $x_3, \dots, x_n$ to x_1; hence it has the cycle type $(0, \dots, 0, 1)$, where all elements, but a 1 at the mth place, are zeros.

Definition 4.5.3. Let a group of substitutions G act on an m-element set D and a substitution $g \in G$ have the cycle type $(b_1, b_2, b_3, \dots, b_m)$. A monomial

$$p_g(x_1, x_2, \dots, x_m) = x_1^{b_1} x_2^{b_2} \dots x_m^{b_m}$$

where $x_1, x_2, \dots, x_m$ are indeterminates, is called the *cycle index of a substitution* g.

A polynomial (the group average of the cycle indices of substitutions)

$$P_G(x_1, x_2, \dots, x_m) = \frac{1}{|G|} \sum_{g \in G} p_g(x_1, x_2, \dots, x_m) = \frac{1}{|G|} \sum_{g \in G} x_1^{b_1} x_2^{b_2} \cdots x_m^{b_m} \tag{4.5.2}$$

is called the *cycle index of the group* G acting on the set D.

Example 4.5.2. The identical (neutral) substitution e has the cycle type $(m, 0, 0, \dots, 0)$, hence

$$p_e(x_1, x_2, \dots, x_m) = x_1^m.$$

If $G_e = \{e\}, |G_e| = 1$, then

$$P_{G_e}(x_1, x_2, \dots, x_m) = x_1^m. \tag{4.5.3}$$

Remark 4.5.1. In polynomial (4.5.3) only x_1 is an essential variable, all others are fictitious.

Problem 4.5.2. Consider three sets of substitutions $G_e = \{e\}, G_c = \{g_c\}$, and $G_2 = \{e, g_c\}$, where the substitution g_c was defined by matrix (4.5.1). Does any of G_e, G_c, and G_2 make up a group?

Problem 4.5.3. Find the cycle index of the group of rotations of a square[9] in the plane, when this group acts on the set D of the vertices of the square.

Solution. Here $|D| = 4$ and $G = \{g_0, g_1, g_2, g_3\}$, where g_k is a rotation of the square over the angle of $k\pi/2, k = 0, 1, 2, 3$; the identical substitution is $g_e = g_0$. By (4.5.3), the cycle index of the identical rotation g_0 is x_1^4. With regard to the substitutions g_1

9 See Exercise 4.5.10.

and g_3, all the vertices make one cycle of length 4, hence their cycle type is $(0, 0, 0, 1)$ and corresponding monomials are x_4. In particular,

$$g_1 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_3 & x_4 & x_1 \end{pmatrix} \quad \text{and} \quad g_3 = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ x_4 & x_1 & x_2 & x_3 \end{pmatrix}.$$

This is an example of two different substitutions with the same cycle structure. Under the action of g_2 the set of vertices breaks up into two cycles, each consisting of two nonadjacent vertices, hence $g_2 = \left(\begin{smallmatrix} x_1 & x_2 & x_3 & x_4 \\ x_3 & x_4 & x_1 & x_2 \end{smallmatrix}\right)$ has the type $(0, 2, 0, 0)$ and its index is x_2^2. Averaging these four monomials, we have by (4.5.2)

$$P_G(x_1, x_2, x_3, x_4) = \frac{1}{4}(x_1^4 + x_2^2 + 2x_4). \tag{4.5.4}$$

□

Problem 4.5.4. Find the cycle index of the group of rotations of a right tetrahedron[10] with an equilateral base, when this group acts on the set D of faces of the tetrahedron.

Solution. Here $|D| = 4$ and the group G contains only three substitutions: the identical substitution e and two rotations over the angles $\pm 120^\circ$ about the height of the pyramid, that is, about the axis perpendicular to the base. Thus,

$$P_G(x_1, x_2, x_3, x_4) = \frac{1}{3}(x_1^4 + 2x_1x_3). \tag{4.5.5}$$

□

Problem 4.5.5. A tetrahedron is called regular if all its faces are equilateral triangles. Find the cycle index of the group of rotations of a regular tetrahedron,[11] when this group acts on the set D of faces (or vertices, which is equivalent) of the tetrahedron.

Solution. Now we can rotate the tetrahedron about the axes perpendicular to each face, thus we have eight rotations over $\pm 120^\circ$ angles about these axes. However, for a regular tetrahedron there is another kind of rotations—after rotating such a tetrahedron through the angle of 180° about the axis connecting the midpoints of two skew edges, the tetrahedron coincides with itself. Since a pair of edges can be chosen in $C(3, 2) = 3$ ways, there are three such substitutions. In total,

$$P_G(x_1, x_2, x_3, x_4) = \frac{1}{12}(x_1^4 + 8x_1x_3 + 3x_2^2). \tag{4.5.6}$$

□

10 See Exercise 4.5.10.

11 See Exercise 4.5.10.

Problem 4.5.6. Find the cycle index of the group of rotations of a cube[12] acting on the set of
(1) vertices,
(2) edges,
(3) faces of the cube.

Solution. It is readily seen that there are 24 different rotations of a cube splitting in the following five types.
(A) Identical rotation (≡ neutral substitution) e.
(B) Three 180° rotations about the lines connecting the centers of opposite parallel faces.
(C) Six ±90° rotations about the same lines as in **(B)**.
(D) Six 180° rotations about the lines connecting the midpoints of opposite parallel edges.
(E) Eight ±120° rotations about the lines connecting the opposite vertices.
(1) A cube has 8 vertices, thus in this case $|D| = 8$ and the **(A)**-type substitutions have the cycle type $(8, 0, 0, \ldots)$, **(B)**-type substitutions have the cycle type $(0, 4, 0, \ldots)$, **(C)**-type substitutions have the cycle type $(0, 0, 0, 2, 0, \ldots)$, **(D)**-type substitutions have the same cycle type $(0, 4, 0, \ldots)$ as the **(B)**-type ones, and **(E)**-type substitutions have the cycle type $(2, 0, 2, 0, \ldots)$. Therefore,

$$P_G(x_1, x_2, \ldots, x_8) = \frac{1}{24}(x_1^8 + 9x_2^4 + 6x_4^2 + 8x_1^2x_3^2). \tag{4.5.7}$$

(2) A cube has 12 edges, therefore now $|D| = 12$ and the substitutions of the same five kinds have cycle types, respectively, $(12, 0, 0, \ldots)$, $(0, 6, 0, \ldots)$, $(0, 0, 0, 3, 0, \ldots)$, $(2, 5, 0, \ldots)$, and $(0, 0, 4, 0, \ldots)$. Hence,

$$P_G(x_1, x_2, \ldots, x_{12}) = \frac{1}{24}(x_1^{12} + 3x_2^6 + 6x_4^3 + 6x_1^2x_2^5 + 8x_3^4).$$

(3) Now $|D| = 6$ and we deduce, in the same way as before,

$$P_G(x_1, x_2, \ldots, x_6) = \frac{1}{24}(x_1^6 + 3x_1^2x_2^2 + 6x_1^2x_4 + 6x_2^3 + 8x_3^2). \tag{4.5.8}$$

□

Consider again a finite set D and a group G of substitutions, acting on the elements of D. This group generates the following equivalence relation on D—two elements $d_1, d_2 \in D$ are said to be equivalent if there exists a substitution $g \in G$ such that $g(d_1) = d_2$. As we have already noted, group axioms of G imply that this binary relation is an equivalence relation. Next we derive an important formula for the number

12 See Exercise 4.5.10.

of classes of equivalence, which is traditionally called the Burnside lemma. To state it, we need a definition.

Definition 4.5.4. By $\psi(g)$ we denote the number of *fixed elements* of the substitution $g \in G$, that is, the number of elements $d \in D$ such that $g(d) = d$.

Lemma 4.5.1. (Burnside or Cauchy–Frobenius lemma [9, p. 278]) *If a group of substitutions G generates an equivalence relation on a finite set D, then the number of the equivalence classes is*

$$n = \frac{1}{|G|} \sum_{g \in G} \psi(g). \tag{4.5.9}$$

Remark 4.5.2. Therefore, the number of the equivalence classes is the average over the group G of the numbers of fixed elements of the substitution $g \in G$.

Proof. We calculate twice the cardinal number of the set of ordered pairs of substitutions and their fixed elements,

$$X = \{(g, d) \mid g \in G, d \in D, g(d) = d\}.$$

On the one hand, if a substitution $g \in G$ is fixed, then the number of these ordered pairs is $\psi(g)$; summing up over all $g \in G$, we have

$$|X| = \sum_{g \in G} \psi(g).$$

On the other hand, let $\eta(d)$ be the number of substitutions $g \in G$ such that $g(d) = d$ for a particular element d. Thus,

$$|X| = \sum_{d \in D} \eta(d)$$

and we derive the equation

$$\sum_{d \in D} \eta(d) = \sum_{g \in G} \psi(g).$$

For a fixed $d \in D$ let us consider a set $G_d = \{g \in G \mid g(d) = d\}$. This is a subgroup[13] of G of order $|G_d| = \eta(d)$.

Let $d_1 \in D$ be equivalent to d in the above sense, hence, there exists a substitution $h \in G$ such that $h(d_1) = d$. If $g(d) = d_1$, then $h(g(d)) = h(d_1) = d$ and $h \circ g \in G_d$. Therefore, to each substitution $g \in G$ with the property $g(d_1) = d$ there corresponds a

13 Exercise 4.5.11.

substitution $g_1 \in G_d$. Vice versa, if $g_1 \in G_d$, then $g(d) = d_1$, where $g = h^{-1} \circ g_1$. Whence the number of elements g such that $g(d) = d_1$ is equal to $|G_d|$.

Now, let $K(d)$ denote the equivalence class containing an element d. Any substitution $g \in G$ shifts d to an element of the same equivalence class. We have also shown that for every element d_1 equivalent to d, $d_1 \sim d$, the number of substitutions g such that $g(d) = d_1$, is the same and is equal to $|G_d|$. Thus,

$$\eta(d) = |G_d| = |G|/|K(d)|.$$

Summing up these equations over all $d' \in K(d)$ yields the equation

$$\sum_{d' \in K(d)} \eta(d') = |G|.$$

To complete the proof of the Burnside lemma, we have to add up all these equations over all the equivalence classes. □

The Burnside lemma will be essentially used in the proof of Pólya's theorem, however, we can immediately apply it to solve Problem 4.5.1.

Solution of Problem 4.5.1. Since the identity substitution does not move any element of D, $\psi(g_0) = 10^5$. To compute $\psi(g_1)$, we notice that $10^5 - 5^5$ numbers contain a digit, which does not turn over, these digits being $2, 3, 4, 5, 7$. Moreover, there are 3×5^2 "symmetric" numbers, which do not change after the rotation. Indeed, to determine such a number, one has to select the middle digit, which is either 0, or 1, or 8, and then to choose the first two digits of the number from the set $0, 1, 6, 8, 9$; these three digits determine the number completely. Therefore, $\psi(g_1) = 10^5 - 5^5 + 3 \times 5^2$. By formula (4.5.9) there are

$$\frac{1}{2}(10^5 + 10^5 - 5^5 + 3 \times 5^2) = 98\,475$$

non-equivalent numbers. □

Using the Burnside lemma, we can straightforwardly solve more problems on bracelets and similar things.

Problem 4.5.7. A bracelet consists of five beads of the same size and shape, but of three different colors. Two bracelets are considered to be identical (equivalent) if we cannot distinguish them after rotating about the wrist without flipping (not taking them off the wrist). How many different bracelets are there?

Solution. Let D be the set of all possible placements of these five beads in the vertices of a regular pentagon, and $G = \{g_0, g_1, g_2, g_3, g_4\}$ be the group of rotations of this pentagon; here g_k is the rotation of the pentagon through the angle of $\frac{2}{5}k\pi$ radians, $k = 0, 1, 2, 3, 4$, $g_0 = e$ being the identical rotation. Assuming all five beads to be different physical entities, we identify the elements of D with arrangements with repetition, that is, $|D| = 3^5 = 243$. Thus, $\psi(g_0) = 243$ and $\psi(g_1) = \psi(g_2) = \psi(g_3) = \psi(g_4) = 3$, since a bracelet is a fixed element with respect to a nontrivial rotation only if all beads are of the same color. By (4.5.9), the number of the equivalence classes, that is, the number of different bracelets, is $\frac{1}{5}(243 + 3 \times 4) = 51$.

If a bracelet consists of n distinguishable beads, the answer is $\frac{1}{n}(n!+0+\cdots+0) = (n-1)!$. Similarly, n people can be arranged in a dancing circle in $\frac{1}{n}(n!+0+\cdots+0) = (n-1)!$ ways. We assume that a bracelet (or a dancing circle) is located in a plane and rotate it in *this* plane about the axis perpendicular to the plane. However, if we can put the bracelet off the wrist, turn it over and then put it back on the wrist—imagine all dancers standing upside down, then there are only $\frac{1}{2}(n-1)!$ indistinguishable bracelets. □

Remark 4.5.3. Compare this problem with Problem 4.2.1.

Problem 4.5.8. Solve Problem 4.5.7 for other numbers of beads and colors, for example, if there are 6 beads and 3 or 4 colors.

To develop a higher power theory, we introduce another set R and consider the power set R^D (Definition 1.1.14). The group G generates a certain equivalence relation on the power set R^D.

Definition 4.5.5. Two mappings, $f_1 : D \to R$ and $f_2 : D \to R$ are called *equivalent* if there exists a substitution $g \in G$ such that $f_1 \circ g = f_2$; the equivalence of mappings is denoted by $f_1 \sim f_2$. Since G is a group, the group axioms induce the fact that this is an equivalence relation.

Problem 4.5.9. Verify that the binary relation described in this definition is an equivalence relation.

Problem 4.5.10.
(1) Given six different colors and assuming that not all of them must be used, in how many geometrically distinct ways is it possible to paint faces of a cube?
Two colorings are called geometrically distinct if it is impossible to transfer one coloring to another by rotating a cube.
(2) In how many geometrically distinct ways is it possible to paint faces of a cube in two colors, blue and green.

Solution. (1) In general, there are $P(6) = 6! = 720$ colorings, but many of them should be identified. We know (Problem 4.5.6) that there are 24 rotations of the cube, thus any equivalence class in this problem consists of 24 elements, and the number of nonequivalent colorings is $720 \div 24 = 30$.

(2) Now the equivalence classes have different cardinalities. First we solve the second part of the problem by a direct enumeration. Later on we solve it by making use of Theorem 4.5.1. Let D be the set of faces, $|D| = 6$, and $R = \{\text{blue}, \text{green}\}$, $|R| = 2$. When we paint a cube, we assign a color to each face of the cube, thus each coloring c can

be viewed as a mapping $c : D \to R$, that is an element of the power set R^D. By Theorem 1.1.6, $|R^D| = 2^6 = 64$, but some of these functions are equivalent and must be identified.

Namely, there is the unique coloring if all faces are blue. If there are five blue faces and one green face, then any of six faces can be chosen as this unique green face and all of these six colorings are equivalent. In the case of four blue faces and two green ones, these green faces can be opposite to one another, thus generating three different colorings, or these two green faces can be adjacent, that is, incident to the same edge, so that giving 12 equivalent colorings. If there are three blue and three green faces, then there are eight equivalent colorings, when the three faces of the same color are incident to the same vertex. In addition, there are 12 more colorings, when two faces of the same color are parallel (opposite to one another) and the third face of the same color is adjacent to both of them—indeed, there are three ways to pick an axis perpendicular to two parallel faces, and after that there are four ways to place a connecting face, thus, $3 \cdot 4 = 12$. Similarly, one can count colorings containing two blue faces, one blue face, and no blue face. All in all, we have $1+6+3+12+8+12+12+3+6+1 = 64$ colorings splitting into 10 different equivalence classes. □

It is useful to endow the elements of R with *weights* by considering one more set W and a mapping $w : R \to W$. We will have to multiply the weights, and for this purpose we assume that W is a commutative ring as discussed before Proposition 4.3.1.

Definition 4.5.6. The image $w(r)$ of an element $r \in R$ is called the *weight* of r and the product

$$w(f) = \prod_{d \in D} w(f(d))$$

is called the *weight of a function* $f \in R^D$.

Lemma 4.5.2. *Equivalent functions have the same weight.*

Proof. If $f_1 \sim f_2$, then there exists a substitution $g \in G$ such that $f_1 \circ g = f_2$. Therefore,

$$w(f_2) = \prod_{d \in D} w(f_2(d)) = \prod_{d \in D} w(f_1(g(d))) = \prod_{d' \in D} w(f_1(d')) = w(f_1)$$

since $\prod_{d \in D} = \prod_{d'=g(d) \in D}$; this equation holds because the substitution g is bijective. □

Whence the following definition is well-posed.

Definition 4.5.7. The *weight* $W(F)$ *of an equivalence class* $F \subset R^D$ is the weight $w(f)$ of any mapping f in the equivalence class F.

In particular, if for any $r \in R$ its weight is $w(r) = 1$, then also $w(f) = 1$ for any function $f \in R^D$ and so that $W(F) = 1$ for each equivalence class F.

Definition 4.5.8. The set R here is called a *reserve*. The sum of weights of all elements of the reserve R is called the *inventory* and is denoted by

$$\operatorname{Inv}(R) = \sum_{r\in R} w(r).$$

Depending on our choice of weights, the inventory gives more or less detailed description of a reserve. For instance, if a student has three books and the weight of each of them includes its title and the author's name, the inventory is a formal sum, nothing but a list of three entries, giving a certain description of student's books. If she has two books in mathematics and one in physics and we assign weights M and P, respectively, to these books, the inventory becomes $2M + P$—from this inventory we see only subjects but not the titles of the books. If these books have 342, 229, and 400 pages, respectively, totaling to 971 pages, and we use these numbers as weights, then the inventory becomes 971, giving us only the idea of the total thickness of these books. Next, if we use the prices of books, say, \$99.95, \$125, and \$129, then the inventory is \$353.95, representing only the total price of these books and nothing else.

Lemma 4.5.3. *Let D and R be finite sets and $D = D_1 \cup \cdots \cup D_k$ be a partition of D. If a subset $S \subset R^D$ consists of all mappings that are constant on each set $D_j, 1 \le j \le k$, then*

$$\operatorname{Inv}(S) = \prod_{j=1}^{k}\left\{\sum_{r\in R}(w(r))^{|D_j|}\right\}. \tag{4.5.10}$$

Proof. Define a function $\psi : D \to \{1,2,\dots,k\}$ by the equation $\psi(d) = j$ whenever $d \in D_j \subset D$, that is, $\psi(d)$ is the index of the subset D_j in the partition, such that the element $d \in D_j$; obviously, $d \in D_{\psi(d)}$. Any mapping $f \in S$ can be represented as a composite function $f = \varphi \circ \psi$, where a function $\varphi : \{1,2,\dots,k\} \to R$. The mapping $\varphi = \varphi_f$ is uniquely defined by f because the latter is constant on all subsets $D_j, 1 \le j \le k$. Vice versa, given any function $\varphi : \{1,2,\dots,k\} \to R$, the superposition $f = \varphi\circ\psi$ is a piece-wise constant function, therefore, this function $f = f_\varphi \in S$. Hence, there exists a one-to-one correspondence between S and the power set $R^{\{1,2,\dots,k\}}$, which proves the equation $|S| = |R|^k$.

Expand now the right-hand side of (4.5.10),

$$\left\{\sum_{r\in R}(w(r))^{|D_1|}\right\} \times \left\{\sum_{r\in R}(w(r))^{|D_2|}\right\} \times\cdots\times \left\{\sum_{r\in R}(w(r))^{|D_k|}\right\}.$$

Multiplying these sums out, we derive a set of products each containing one addend from every sum. There are $|R|^k = |S|$ such products, a generic one being

$$(w(r_{i_1}))^{|D_1|} \times (w(r_{i_2}))^{|D_2|} \times\cdots\times (w(r_{i_k}))^{|D_k|}. \tag{4.5.11}$$

Considering product (4.5.11), it is natural to introduce a function

$$\widetilde{\varphi} : \{1,2,\dots,k\} \to R$$

as follows: $\widetilde{\varphi}(j) = r_{i_j}, 1 \le j \le k$. This shows that product (4.5.11) is equal to the weight $w(\widetilde{f})$, where the function $\widetilde{f} = \widetilde{\varphi} \circ \psi$. Indeed,

$$w(\widetilde{f}) = \prod_{d\in D} w(\widetilde{f}(d)) = \prod_{j=1}^{k} \prod_{d\in D_j} w(\widetilde{f}(d)) = \prod_{j=1}^{k} \{w(\widetilde{f}(d))\}^{|D_j|}$$

because $\widetilde{f}$ is constant on any D_j. It is easily seen that starting at different products (4.5.11), this construction leads to different functions $\widetilde{\varphi}$, and thus to different functions $\widetilde{f}$. Since there are as many products (4.5.11) as functions $\widetilde{\varphi}$, we conclude that by multiplying out all terms in (4.5.10) we get the sum of the weights of all functions in S. However, this sum is precisely the inventory $\mathrm{Inv}(S)$, which completes the proof of Lemma 4.5.3. □

Corollary 4.5.1. *If a partition of D contains only 1-element sets, then S is the power set R^D and* (4.5.10) *simplifies to*

$$\mathrm{Inv}(R^D) = (\mathrm{Inv}(S))^{|D|}. \tag{4.5.12}$$

Problem 4.5.11. In how many ways can three people distribute m tokens among themselves, so that the first and the second persons get an equal number of tokens?

Solution. We give two solutions of this problem, one straightforward and elementary, and another based on Lemma 4.5.3.

First solution of Problem 4.5.11. We have to find in how many ways it is possible to split the number m into two whole addends, such that any addend or even two of them may be 0, because one or two of these people may get nothing. If m is even, this can be done in $\frac{m}{2} + 1$ ways, and if m is odd, then there are $\frac{m+1}{2}$ ways to split the number.

Second solution of Problem 4.5.11. We use the problem to demonstrate the machinery of applying Lemma 4.5.3. In this particular problem the following solution is obviously longer than the first one, however, we introduce in this solution an important technique, which has much broader applications.

Consider two sets, $D = \{p_1, p_2, p_3\}$ and $R = \{0, 1, 2, \dots, m\}$, and a partition $D = D_1 \cup D_2$, where $D_1 = \{p_1, p_2\}$ and $D_2 = \{p_3\}$. Let S have the same sense as in Lemma 4.5.3, that is, a function $f \in S$ if and only if the images of p_1 and p_2 are the same, $f \in S \subset R^D \iff f(p_1) = f(p_2)$. To solve the problem, we have to find the number of functions $f \in S$ satisfying an additional restriction

$$f(p_1) + f(p_2) + f(p_3) = m.$$

Assign the weights $w(i) = x^i, i = 0, 1, \dots, m$, where x is indeterminate, to the elements of R. If $f \in R^D$, then

$$w(f) = \prod_{d\in D} w(f(d)) = w(f(p_1)) \cdot w(f(p_2)) \cdot w(f(p_3)) = x^{f(p_1)+f(p_2)+f(p_3)}.$$

Therefore, the functions we sought after, have the weight x^m, and the number of such functions is the coefficient of x^m in $\mathrm{Inv}(S)$. It should be noted that in this case the inventory is the GF of the reserve. By (4.5.10) with $k = 2$,

$$\begin{aligned}\mathrm{Inv}(S) &= (1 + x^2 + x^4 + \cdots + x^{2m})(1 + x + x^2 + \cdots + x^m)\\ &= \frac{1 - x^{2m+2}}{1 - x^2} \cdot \frac{1 - x^{m+1}}{1 - x} = (1 - x^2)^{-1}(1 - x)^{-1} + g(x)\\ &= \frac{1}{4}(1 + x)^{-1} + \frac{1}{2}(1 - x)^{-2} + \frac{1}{4}(1 - x)^{-1} + g(x).\end{aligned}$$

Here we separated the function g, because in the problem we are interested in the coefficient of x^m, and the function g has a power series, which begins with the term x^{m+1}, therefore, g contributes nothing in the coefficient of x^m. Using equation (4.3.7), we straightforwardly verify that the latter expression leads to the same answer as the first solution, namely, $\frac{m+1}{2} + \frac{1+(-1)^m}{4}$. □

Remark 4.5.4. We have simultaneously found in this problem the number of compositions consisting of two parts, such that one part is even.

The proof of Lemma 4.5.3 carries out without any change if $|R| = \infty$, so long as all occurring series are convergent. In this case (4.5.10) and (4.5.12) read that the corresponding series are equal. This remark often allows one to simplify computations. For instance, in Problem 4.5.11 it is convenient to use $R = \{0, 1, 2, \ldots\}$; then

$$\mathrm{Inv}(S) = (1 + x^2 + x^4 + \cdots)(1 + x + x^2 + \cdots) = \frac{1}{1 - x^2} \cdot \frac{1}{1 - x},$$

and the latter expression has, certainly, the same coefficient of x^m as the preceding one.

Now we state the main result of this section.

Theorem 4.5.1. *Let D and R be finite sets, $|D| < \infty$, $|R| < \infty$. Let a group of permutations G act on D and $P_G(x_1, x_2, \ldots)$ be the cycle index of G. Then the inventory* $\mathrm{Inv}(R^D)$, *that is, the complete list of the weights of all equivalence classes generated in the power set R^D by the group G, is*

$$\sum_F W(F) = P_G\Big(\sum_{r\in R} w(r), \sum_{r\in R} (w(r))^2, \sum_{r\in R} (w(r))^3, \ldots\Big). \tag{4.5.13}$$

Corollary 4.5.2. *In particular, if $w(r) = 1, \forall r \in R$, then as it was mentioned, $W(F) = 1, \forall F \in R^D$, and the left-hand side of* (4.5.13) *is equal to the number of the equivalence classes, which is, therefore,*

$$n = P_G(|R|, |R|, \ldots). \tag{4.5.14}$$

□

Before proving Theorem 4.5.1, we apply it to solve again Problem 4.5.10.

Another solution of Problem 4.5.10. In this problem D is the set of faces of a cube, $|D| = 6$, G is the group of rotations of a cube, whose cycle index was found in Problem 4.5.6(3), and $R = \{\text{blue}, \text{green}\}$, $|R| = 2$. Inserting $x_1 = x_2 = \cdots = x_6 = 2$ into formula (4.5.8), we find again

$$n = \frac{1}{24}(2^6 + 3 \cdot 2^4 + 6 \cdot 2^3 + 6 \cdot 2^3 + 8 \cdot 2^2) = 10. \qquad \square$$

Proof of Theorem 4.5.1. Let ω be an arbitrary possible value of the weight and $S_\omega = \{f \in R^D : w(f) = \omega\}$. Lemma 4.5.2 implies that, if an equivalence class intersects with S_ω, then this entire class is a subset of S_ω. Thus, we can restrict the action of the group G on S_ω; let n_ω denote the number of equivalence classes belonging to S_ω. By Lemma 4.5.1 being applied to S_ω, we conclude

$$n_\omega = \frac{1}{|G|} \sum_{g \in G} \psi_\omega(g) \tag{4.5.15}$$

where $\psi_\omega(g)$ is the number of functions $f \in R^D$ such that $f = f \circ g$ and $w(f) = \omega$.

The quantity

$$\omega \cdot n_\omega = \underbrace{\omega + \omega + \cdots + \omega}_{n_\omega \text{ addends}}$$

is the inventory of all equivalence classes having the weight ω. Therefore, if we multiply (4.5.15) by ω and sum up these equations over all possible values of ω, we get the inventory of all the equivalence classes,

$$\sum_F W(F) = \frac{1}{|G|} \sum_\omega \sum_{g \in G} \omega \psi_\omega(g)$$

where the order of summation can be interchanged since all sums are finite. However, $\sum_\omega \omega \psi_\omega(g) = \sum_{f \in R^D : f = f \circ g} w(f)$; hence

$$\sum_F W(F) = \frac{1}{|G|} \sum_{g \in G} \left(\sum_{f \in R^D : f = f \circ g} w(f) \right).$$

Considering now the definition of the cycle index (4.5.2), we see that to prove the theorem, it remains to prove the following claim. $\square$

Proposition 4.5.1. *Given a permutation g with the cycle type* $(b_1, b_2, \ldots, b_m)$*, the expression*

$$\sum_{f \in R^D : f = f \circ g} w(f)$$

comes out of the monomial

$$x_1^{b_1} \times x_2^{b_2} \times \cdots \times x_m^{b_m}$$

after replacing the indeterminate x_1 *with* $\sum_{r\in R} w(r)$*, then replacing* x_2 *with* $\sum_{r\in R}(w(r))^2$*, then* x_3 *with* $\sum_{r\in R}(w(r))^3$*, ...,* k_k *with* $\sum_{r\in R}(w(r))^k$*, and so forth.*

Proof. We observe that when g acts on D, the latter breaks down into disjoint cycles $D_1, D_2, \dots, D_k$, and the condition $f = f \circ g$ implies

$$f(d) = f(g(d)) = f(g^2(d)) = \cdots$$

therefore, f is constant on any cycle $D_j, 1 \le j \le k$.

Vice versa, if f is constant on every cycle contained in g, then $f \circ g = f$, which means $g(d)$ always belongs to the same cycle as d. Thus, one can apply Lemma 4.5.3, yielding

$$\sum_{f\in R^D: f=f\circ g} w(f) = \prod_{j=1}^{k} \sum_{r\in R} (w(r))^{|D_j|}. \tag{4.5.16}$$

Given a substitution g with the cycle type $(b_1, b_2, \dots, b_m)$, a 1 occurs b_1 times among the cardinalities $|D_1|, |D_2|, \dots, |D_k|$, a 2 appears b_2 times, etc. Hence (4.5.16) can be written as

$$\begin{aligned}\sum_{f\in R^D: f=f\circ g} w(f) &= \left\{\sum_{r\in R} w(r)\right\}^{b_1} \times \left\{\sum_{r\in R} (w(r))^2\right\}^{b_2} \times \cdots \\ &= x_1^{b_1} \times x_2^{b_2} \times \cdots \times x_m^{b_m}\big|_{x_i=\sum_{r\in R}(w(r))^i, 1\le i\le m},\end{aligned}$$

which proves Proposition 4.5.1 and consequently, Theorem 4.5.1. □

In the rest of the section we consider various applications of the Pólya–Redfield enumeration theory.

Problem 4.5.12. In how many geometrically distinct ways is it possible to paint the vertices of a square in blue and green colors? Two colorings are said to be geometrically identical if they can be made indistinguishable by rotating the square in the plane about its center.

Solution. Given a vertex, we choose a color for that vertex, therefore, a coloring is a mapping from the set $D = \{v_1, v_2, v_3, v_4\}$ of the vertices of the square to the set $R = \{\text{blue}, \text{green}\}$. The equivalence relation on the set of colorings is generated by the group of rotations of a square acting on vertices D of the square. The cycle index of this group is given by equation (4.5.4),

$$P_G(x_1, x_2, x_3, x_4) = \frac{1}{4}(x_1^4 + x_2^2 + 2x_4).$$

Choosing all weights to be 1 and setting in (4.5.14) $x_1 = x_2 = x_4 = |R| = 2$, we get $n = P_G(2, 2, 2, 2) = 6$. □

Problem 4.5.13. In how many geometrically distinct ways can one paint the faces of a regular tetrahedron in two colors? Two colorings are said to be geometrically identical (equivalent) if they can be made indistinguishable by rotating the tetrahedron in space about its center.

Solution. The cycle index of the group of rotations is given by equation (4.5.6). Combining it with (4.5.14) and substituting there $x_i = |R| = 2$, we compute $n = 5$, that is, there are five different ways to color the regular tetrahedron, which can be easily verified by inspection. □

Problem 4.5.14. For any natural n, $\frac{1}{24}(n^8 + 17n^4 + 6n^2)$ is an integer number.

Solution. Formulas (4.5.7) and (4.5.14) imply that this is the number of geometrically different colorings of the vertices of a cube in n colors. □

The next problem was solved by Pólya in his original article [44].

Problem 4.5.15. In how many geometrically distinct ways can we place three blue balls, two green balls, and a pink ball at the vertices of a regular octahedron?

Solution. The group of rotations of a regular octahedron acts on the set D of its vertices, $|D| = 6$. It is clear that the cycle index of this group is the same as that of the group of rotations of a cube acting on the set of the faces of the latter. The cycle index of the latter group of substitutions is given by (4.5.8). Since we have three different colors, we select a reserve $R = \{\text{blue}, \text{green}, \text{pink}\}$.

To distinguish different colors, we introduce weights on R as indeterminates $w(\text{blue}) = b$, $w(\text{green}) = g$ and $w(\text{pink}) = p$, hence

$$\sum_{r\in R} w(r) = b + g + p$$
$$\sum_{r\in R} w^2(r) = b^2 + g^2 + p^2$$
$$\sum_{r\in R} w^3(r) = b^3 + g^3 + p^3$$

etc. Then, by (4.5.13) and (4.5.8), we derive the complete list of all possible colorings,

$$\begin{aligned}\sum_F W(F) &= P_G\Big(\sum_{r\in R} w(r), \sum_{r\in R} (w(r))^2, \sum_{r\in R} (w(r))^3, \dots\Big)\\ &= \frac{1}{24}\{(b+g+p)^6 + 3(b+g+p)^2(b^2+g^2+p^2)^2\\ &\quad + 6(b+g+p)^2(b^4+g^4+p^4) + 6(b^2+g^2+p^2)^3 + 8(b^3+g^3+p^3)^2\}.\end{aligned}$$

The equivalence classes we are looking for, have weight b^3g^2p. Multiplying out all factors in the expression for P_G, we see that this monomial appears in the sum three times, so that there are three geometrically different colorings in this problem. □

Problem 4.5.16. How many different molecules are there (see Fig. 4.4), which contain a four-valent atom of the carbon C in the center with four endings X, where X may be either the hydrogen H, or the chlorine Cl, or the methyl group CH_3, or the ethyl group C_2H_5?

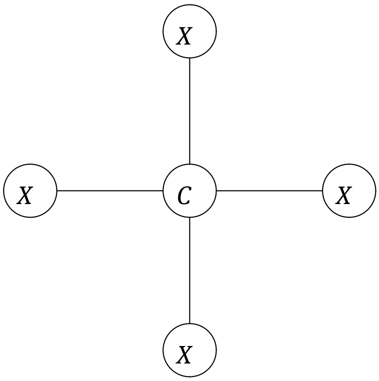


Figure 4.4: Molecules in Problem 4.5.16.

Solution. Put into a correspondence to each molecule a regular tetrahedron, whose vertices are labeled by the symbols X. We have to calculate the number of equivalence classes in the power set R^D, where D is the set of vertices of the tetrahedron, $|D| = 4$, and $R = \{H, Cl, CH_3, C_2H_5\}$. The cycle index of this group of rotations is given by (4.5.6). Thus, the number of molecules is equal to $P_G(4, 4, 4, 4) = 36$.

If we want to list these molecules with regard to the number of the hydrogen atoms contained, it is convenient to choose weights $w(H) = h$ and

$$w(Cl) = w(CH_3) = w(C_2H_5) = 1.$$

Then the sum of the squares of the weights is $h^2 + 1^2 + 1^2 + 1^2 = h^2 + 3$, etc., and we deduce from (4.5.13)

$$P_G(h + 3, h^2 + 3, h^3 + 3, h^4 + 3) = h^4 + 3h^3 + 6h^2 + 11h + 15.$$

The coefficients in the latter tell us that there is the unique molecule containing four atoms of hydrogen, there are three molecules with three hydrogen atoms, six molecules with two hydrogen atoms, 11 molecules with one hydrogen atom, and 15 molecules do not contain hydrogen at all. We note that $15 = P_G(3, 3, 3, 3)$, since the latter molecules can have only three possible endings. □

In the following problems we apply the techniques of this section to compute once again the number of permutations and combinations.

Problem 4.5.17. Find the number of n-permutations with repetition from an m-3element set A.

Solution. Since the permutations were defined as special mappings, we set $D = \{1, 2, \ldots, n\}$ and $R = A$. We do not have to identify any mappings, therefore, we can use the trivial 1-element group of substitutions $G = \{g_0\}$, where g_0 is the identical substitution, and all weights $w = 1$. Combining (4.5.3) and (4.5.14), we arrive again at (1.3.1),

$$A_{\text{rep}}(m, n) = P_G(|R|, |R|, \ldots) = m^n. \qquad \square$$

Problem 4.5.18. Find the number of r-combinations with repetition from elements of n types.

Solution. These combinations can be put in a one-to-one correspondence[14] with mappings $f : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, r\}$, such that

$$f(1) + f(2) + \cdots + f(n) = r,$$

where $f(i)$, $1 \le i \le n$, stands for the number of elements of the ith type in this combination. Introduce weights on R as in the second solution of Problem 4.5.11, by $w(i) = x^i$, and set $D = \{1, 2, \ldots, n\}$. The mappings we look after are listed by the term x^r in the inventory $\text{Inv}(R^D)$, hence by means of (4.5.12) we get

$$\text{Inv}(R^D) = \left\{\sum_{i \in R} w(i)\right\}^{|D|} = (1 + x + x^2 + \cdots + x^r)^n$$

which yields again formula (1.4.6), $C_{\text{rep}}(n, r) = \frac{(n+r-1)!}{(n-1)!r!}$.

Calculations are simpler if we take the infinite reserve $R = \{0, 1, 2, \ldots\}$, leading to

$$\text{Inv}(R^D) = (1 + x + x^2 + \cdots)^n.$$

Similarly, by making use of (4.5.12) one can find the formula for the number of compositions of integer numbers (cf. Section 4.4). The distinction between the latter and the combinations with repetition is that a combination may omit elements of certain types, but a composition cannot contain zero elements. Thus, in the case of compositions we must use the reserve $R = \{1, 2, 3, \ldots\}$. $\square$

The last problem in this section deals with colorings of binary trees.

14 This means, in particular, that the combinations with repetition can be defined as such mappings.

Problem 4.5.19. Consider a binary tree with seven vertices (Fig. 4.5), which have to be painted in two colors. Two colorings are called equivalent if one can be derived from the other by rotating either the entire tree through 180° about the vertical symmetry axis or any of its subtrees through 180° about the horizontal symmetry axis. For example, the coloring in Fig. 4.4(a) is equivalent to that in Fig. 4.4(b), but is not equivalent to the one in Fig. 4.4(c). How many non-equivalent colorings are there?

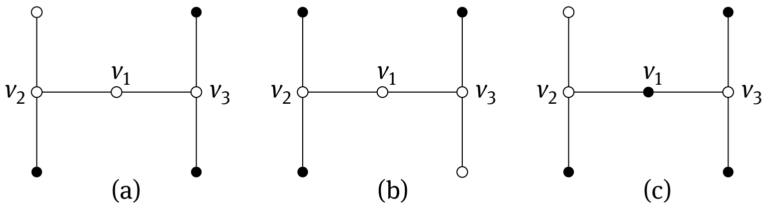


Figure 4.5: Three binary trees in Problem 4.5.19.

Solution. Let D be the set of the vertices of the tree, $|D| = 7$, and R the set of available colors, $|R| = 2$. To each coloring there corresponds a mapping $f \in R^D$, and this is easily seen to be a one-to-one correspondence. Let $G = \{g_0, g_1, \ldots, g_7\}$ be the group of substitutions acting on D, where g_0 is the identical rotation, g_1 is rotation about the vertex v_1, g_2 is rotation about the vertex v_2, g_3 is rotation about the vertex v_3, and $g_4 = g_2 \circ g_3$, $g_5 = g_1 \circ g_2$, $g_6 = g_1 \circ g_3$, $g_7 = g_1 \circ g_2 \circ g_3$. It is readily verified that the cycle indices of permutations g_0–g_7 are, respectively,

$$x_1^7,\ x_1x_2^3,\ x_1^5x_2,\ x_1^5x_2,\ x_1^3x_2^2,\ x_1x_2x_4,\ x_1x_2x_4,\ x_1x_2^3;$$

thus the cycle index of the group G is

$$P_G(x_1, x_2, \ldots, x_7) = \frac{1}{8}(x_1^7 + 2x_1x_2^3 + 2x_1^5x_2 + x_1^3x_2^2 + 2x_1x_2x_4).$$

By (4.5.14), the number of colorings is $P_G(2, 2, 2, 2, 2, 2, 2) = 42$. □

Exercises 4.5.

Exercise 4.5.1. Verify that the binary relation in Problem 4.5.1 satisfies the axioms of the binary relation.

Exercise 4.5.2.
(1) Extend Problem 4.5.1, computing now the number of non-equivalent 6-digit and 7-digit numbers.
(2) What is the sum of all these numbers?
(3) How many of them are multiple of 4?
(4) Solve Problem 4.5.1 assuming that digit 1 cannot be rotated upside down.

Exercise 4.5.3. Find the cycle type and cycle index of the substitution

$$g = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} \\ x_2 & x_3 & x_1 & x_4 & x_6 & x_5 & x_7 & x_8 & x_{10} & x_9 \end{pmatrix}.$$

Exercise 4.5.4. Find the coefficient of t^5 in the expansion of $\prod_{k=1}^{8}(t + a_k)$.

Exercise 4.5.5. A disk is divided in p equal sectors by p radii, where p is a prime number. In how many different ways is it possible to paint the disk in n colors, if two colors cannot be used on the same sector? Two colorings are to be identified if they can be made indistinguishable by a rotation of the disk in the plane about its center.

Exercise 4.5.6. A family of substitutions $F = \{g_0, g_1, g_2\}$, where

$$g_0 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad g_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}, \quad g_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix},$$

acts on a set $X = \{1, 2, 3, 4\}$. Can we apply Lemma 4.5.1 to the family F?

Exercise 4.5.7. Prove that the number of fixed elements of a substitution acting on an n-element set is $n! \sum_{k=0}^{n}(-1)^k/k!$. Compare with Problem 4.1.9.

Exercise 4.5.8. Find the number of equivalence classes induced on a set $X = \{1, 2, 3, 4\}$ by the group of substitutions $\{g_0, g_1, g_2, g_3\}$, where

$$g_0 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad g_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix},$$

$$g_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}, \quad g_4 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}.$$

Exercise 4.5.9.
(1) A 2×2 square is divided in four 1×1 squares. Using 6 colors, in how many geometrically distinct ways is it possible to paint the big square so that neighboring (having a common side) small squares have different colors?
(2) Solve the same problem for a 3×3 square split in nine 1×1 squares.

Exercise 4.5.10. Verify that rotations of a cube, acting on its vertices or on its sides, or faces, form groups with the superposition of rotations as the group operation. Answer the same question regarding the rotations of a square or a regular tetrahedron. The cycle indices of these groups were found in Problems 4.5.3–4.5.6.

Exercise 4.5.11. Prove that the set G_d in the proof of Lemma 4.5.1 is a subgroup of order $\eta(d)$ of the original group of permutations G.

Exercise 4.5.12. Draw explicitly all six different colorings of the vertices of a square in two colors—see Problem 4.5.12.

Exercise 4.5.13. In how many geometrically distinct ways can we paint the 2×2 checker-board in two colors? The same question for the 3×3 checker-board.

Exercise 4.5.14.
(1) In how many geometrically distinct ways can we paint the faces of a cube in three colors? In no more than six colors?
(2) Solve the same problems for edge coloring. Answer the same question if there are four or five colors available.
(3) Solve the same problems for a right tetrahedron.

Exercise 4.5.15. Using the six digits $1, 2, \ldots, 6$, the faces of a die can be marked in geometrically different ways. For instance, 1 and 2 can be on two opposite or on two adjoint faces. How many differently labeled dice are there?

Exercise 4.5.16. In how many geometrically distinct ways can we paint the edges or faces of a regular tetrahedron in two colors? The same question if there are three colors available.

Exercise 4.5.17. In how many geometrically distinct ways can 12 friends, 6 girls and 6 boys, ride a carousel with 12 seats, if all boys are considered to be indistinguishable and all girls are indistinguishable either?

Exercise 4.5.18. There are 999 students in the Small College and each of them has recently passed four tests with scores 7, 8, 9, or 10. What is the largest possible number of students such that any two of them have different sets of scores and the sum of the four grades is odd? Is an even number?

Exercise 4.5.19. Represent the substitutions defined in Exercises 4.5.6 and 4.5.8 as products of transpositions; which of them are odd and which of them are even?

Exercise 4.5.20. Compute the cycle index of the symmetric group Sym_5 (see Example 4.5.1) acting on 2-element subsets of the set $\{1, 2, 3, 4, 5\}$.

Exercise 4.5.21. Compare the parity (odd/even), that is, the number of cycles of a substitution (permutation) with the parity of the number of inversions in the substitution (Definition 4.4.6 and before)—is it the same?

Exercise 4.5.22. In how many ways a party of n friends can be sitting around the round table, if two ways are considered identical, if one can be derived from another by a shift of the whole party around the table clock-wise the same for every person?

5 Existence theorems in combinatorics

Three topics considered in this chapter have one essential feature in common—the existence of combinatorial configurations in question is not obvious at all and must be proved. In Section 5.1 we prove Ramsey's theorem—a far-reaching extension of the Dirichlet or pigeonhole principle. Section 5.2 is devoted to the famous Philip Hall's marriage theorem. Its quantitative version on a lower estimate of the number of systems of distinct representatives is given as well as a few equivalent statements, in particular, Denés König's and Dilworth's theorems. Section 5.3 gives an introduction to the theory of combinatorial block designs. Finally in Section 5.4 we consider in more detail the systems of triples including the proof, due to Hilton [30], of the necessary and sufficient conditions of the existence of the Steiner triple systems.

5.1 Ramsey's theorem

Coffee-time browsing

– http://en.wikipedia.org/wiki/Frank_P._Ramsey (Ramsey life and work)
– www.cs.umd.edu/~gasarch/ramsey/ramsey.html (Applications of Ramsey theory)
– www-history.mcs.st-andrews.ac.uk/Biographies/Dirichlet.html (Dirichlet life and work)
– www-history.mcs.st-and.ac.uk/Biographies/Schur.html (Schur's biography)
– http://en.wikipedia.org/wiki/Paul_Erd%C5%91s#Biography (Erdos' biography)
– http://www.oakland.edu/enp/ (What is Erdos number?)
– http://www-history.mcs.st-and.ac.uk/Biographies/Szekeres.html (George and Esther Szekeres summary)
– http://en.wikipedia.org/wiki/Erd%C5%91s%E2%80%93Szekeres_theorem (Erdos-Szekeres theorem)

Problem 5.1.1. There are six pairs of socks of six different colors in a drawer. Which is the smallest number of separate socks that must be drawn at random to ensure that the owner gets at least one complete pair?

Solution. Let us consider the worst-case scenario. This case occurs if one gets six socks of six different colors. After that, any seventh sock makes a complete pair of socks. It suffices, therefore, to draw seven socks, and moreover, the number seven cannot be reduced to six. □

The problem can be stated in set-theory terms as follows.

A set A is partitioned in six subsets, $A = A_1 \cup \cdots \cup A_6$, with every $|A_i| \geq 2, 1 \leq i \leq 6$. Which is the smallest cardinality of a subset $B \subset A$ such that the intersection of B with at least one of the subsets A_i contains two or more elements?

It is said that these two or more elements represent the sets A_i in B.

https://doi.org/10.1515/9783110751185-005

The existence of a solution to Problem 5.1.1 is clear. Ramsey's theorem treats significantly more general situations, when even the existence of a solution is far from obvious, whereas the cardinality of the solution in most cases is unknown yet. To state the theorem, we have to formalize a concept of a collection of elements containing several identical copies of the same element, for such a collection[1] is *not* a set in the standard set-theory meaning. In particular, we must consider families of subsets containing several copies of the same subset.

Definition 5.1.1. For a set S, a mapping

$$U : \{1, 2, \ldots, t\} \to 2^S$$

is called a *family of subsets* of S containing t terms, or just a t-family of subsets. The family U is denoted by $U = (S_1, S_2, \ldots, S_t)$, where $S_i = U(i), 1 \le i \le t$. The mapping U does not have to be injective, thus some of the sets S_i can coincide with one another.

We again use mappings to distinguish certain objects. Even if two terms, S_i and S_j, of a family are equal as sets, we consider them as different terms of the family, because they have different subscripts, that is, different preimages with respect to the mapping U.

Example 5.1.1. Let $t = 3$, $S = \{1, 2, 3\}$, $S_1 = S_2 = \{1, 2\}$, and $S_3 = \{2, 3\}$. Then $U = (S_1, S_2, S_3)$ is an example of a 3-family.

Definition 5.1.2. If $S = S_1 \cup S_2 \cup \cdots \cup S_t$, where $S_i \cap S_j = \emptyset, i \ne j, 1 \le i, j, \le t$, then the t-family $U = (S_1, S_2, \ldots, S_t)$ is called an *(improper) ordered partition of the set S in t parts*, or a t-partition of S. It is called improper because $\emptyset \subset S$ and so that U can contain empty terms.

Example 5.1.2. Let again $t = 3$, $S = \{1, 2, 3\}$, but now $S_1 = \emptyset$, $S_2 = \{1, 2\}$, and $S_3 = \{3\}$. Then $U = (S_1, S_2, S_3)$ is an example of an improper 3-family.

Theorem 5.1.1. (Ramsey) *Consider a set X, natural numbers p and t, and any ordered t-partition of the set 2^X_p of all p-subsets of X,*

$$2^X_p = A_1 \cup A_2 \cup \cdots \cup A_t \tag{5.1.1}$$

where t subsets A_i are the parts of this partition. Then, for arbitrary natural numbers $p, q_1, q_2, \ldots, q_t$ such that $1 \le p \le q_i$ for all $1 \le i \le t$, there exists the smallest natural number $R = R_p(q_1, q_2, \ldots, q_t)$ with the following property:

If $|X| \ge R$, then there exists an index $i, 1 \le i \le t$, and a subset $B \subset X$ such that $|B| = q_i$ and $2^B_p \subset A_i$.

The numbers $R_p(q_1, q_2, \ldots, q_t)$ are called Ramsey numbers.

It is useful to restate Theorem 5.1.1 in terms of colorings of graphs.

1 "Sets" with repeating objects are sometimes called *multisets*.

Theorem 5.1.2. *For arbitrary natural numbers* $p, q_1, q_2, \dots, q_t$ *such that* $1 \le p \le q_i$ *for* $1 \le i \le t$, *there exists the smallest natural number* $R = R_p(q_1, q_2, \dots, q_t)$ *with the following property.*

Let G *be a graph of size* $q = q_G \ge R$, *that is, with* q *edges. Consider all subgraphs of* G *of size* p *and color each of their edges in one of the given* t *colors. Then, for some* $i, 1 \le i \le t$, G *has a monochromatic subgraph* G' *of size* q_i, *that is all subgraphs of* G' *of size* p *have the same color.*

Ramsey's theorem gives a precise meaning to the following intuitively clear statement:

For an arbitrary subdivision of a set in a prescribed number of parts, all these parts cannot simultaneously be small if the set is sufficiently large.

First we consider the special case $p = 1$ of Theorem 5.1.1. In this case, the set $2^X_p = 2^X_1$ of the 1-element subsets of X can be identified with the set X itself and (5.1.1) can be thought of as a t-partition of X. Now the theorem claims that, for any t-partition $X = A_1 \cup \dots \cup A_t$ of the set X and for any integers $q_1 \ge 1, \dots, q_t \ge 1$, there is an index $i, 1 \le i \le t$, and a subset $B \subset X$ such that $|B| = q_i$ and $B \subset A_i$, whenever $|X|$ is large enough. Similarly to Problem 5.1.1, in the case $p = 1$ we have

$$R_1(q_1, q_2, \dots, q_t) = (q_1 - 1) + \dots + (q_t - 1) + 1 = q_1 + \dots + q_t - t + 1. \tag{5.1.2}$$

In Problem 5.1.1 $t = 6, q_1 = \dots = q_6 = 2, p = 1$ and $R_1(2, 2, 2, 2, 2, 2) = 7$.

If $p = 1$ and $q_i \ge 2, 1 \le 2 \le t$, we derive the following *Dirichlet* or *pigeonhole principle*.

Proposition 5.1.1. *If* R *objects (for example, pigeons) are placed in* t *boxes (cages) and* $R \ge t+1$, *then at least one box (cage) must contain two or more of these objects (pigeons). Moreover, if* R *objects are placed in* t *boxes, then there is a box containing at least* $[\frac{R-1}{t}]+1$ *objects.*

Before taking up the proof, we solve a few problems by making use of the Dirichlet principle to show a variety of its applications.

Problem 5.1.2. How many numbers should be chosen from the set $\{1, 2, 3, 4\}$ to ensure that at least one pair of these numbers adds up to 5?

Solution. Among all $6 = C(4, 2)$ 2-element subsets of the given set, only two pairs, $\{1, 4\}$ and $\{2, 3\}$ satisfy the condition. Thus, if we select only two numbers, a and b, it may happen that $a \in \{1, 4\}$ and $b \in \{2, 3\}$, therefore, $a + b \ne 5$. However, any third number chosen completes either the pair $\{1, 4\}$ or the pair $\{2, 3\}$. Consequently, it is enough to choose three numbers. In a more formal language of Theorem 5.1.1, we set $t = 2$ (two pairs-cages), $q_1 = q_2 = 2$, and by (5.1.2) we have $R_1(2, 2) = 3$. □

Problem 5.1.3. How many numbers are to be chosen from the set $\{1, 2, 3, 4, 5\}$ to ensure that at least one pair of these numbers adds up to 5?

Solution. We still have two favorable pairs $\{1,4\}$ and $\{2,3\}$. However, if we select two numbers, representing these two pairs, then the third number may happen to be 5. Consequently, now we have to select at least 4 numbers. In the formal language, in this problem $t = 3$ (two pairs and a singleton 5), $q_1 = q_2 = q_3 = 2$, so that by (5.1.2), $R_1(2,2,2) = 4$. □

Problem 5.1.4. Which is the smallest number of integers that one has to choose from the set $T = \{1, 2, \ldots, 30\}$ in order to ascertain that there are three numbers among the chosen, whose sum is multiple of 3?

Solution. The sum of three integers is a multiple of 3 if and only if 3 divides the sum of the remainders after dividing these numbers by 3. Consider three subsets of the set T,

$$T_0 = \{3,6,9,\ldots,30\}, \quad T_1 = \{1,4,7,\ldots,28\}, \quad T_2 = \{2,5,8,\ldots,29\}.$$

Obviously we can select three numbers, say one number in T_1 and two in T_2, whose sum is not divisible by 3. Moreover, we can choose two integers in T_0 and two in T_1 (or in T_2), and this quadruple also does not satisfy the problem. However, any other number from T being combined with these four numbers, solves the problem. Thus, we have to choose at least five numbers. □

Remark 5.1.1. It is instructive to translate this solution to the language of the Dirichlet principle, similarly to Problems 5.1.2–5.1.3.

Problem 5.1.5. A box has the shape of a cube with the side of one meter. There are 2 001 flies in the box. Prove that at least three of them are in a ball of radius $5\sqrt{3}$ cm.

Solution. Divide the box into $10^3 = 1\,000$ small cubes of side 10 cm. A diagonal of each such cube is $10\sqrt{3}$ cm. Since $2 \times 1\,000 < 2\,001$, there is at least one small cube $\mathcal{C}$ with three flies inside—in the problem we suppose that a fly is a mathematical dimensionless point and exclude surfaces of the small cubes from consideration. The cube $\mathcal{C}$ together with its three flies lies completely in the ball of radius $5\sqrt{3}$ cm circumscribed about $\mathcal{C}$. □

In the following lemma we consider another special case of Theorem 5.1.1 with $p = t = 2$.

Lemma 5.1.1. *For any integer $q \geq 2$,*

$$R_2(q,2) = R_2(2,q) = q.$$

Proof. The equation $R_2(q,p) = R_2(p,q)$ is obvious due to symmetry, therefore, it is enough to prove that $R_2(2,q) = q$. This equation says that, for any set X with $|X| \geq q$ and for any partition of its 2-element subsets into two groups, $\mathcal{A}_1$ and $\mathcal{A}_2$, one of which may be empty, either there exists a 2-element subset $A \subset X, |A| = 2$, such that $A \in \mathcal{A}_1$, or there exists a q-element subset $A \subset X, |A| = q$, such that $2_2^A \subset \mathcal{A}_2$. Moreover, q is the smallest cardinal number with this property. It is important to keep in mind that we consider *ordered* partitions of 2_2^X.

Thus, to proceed with the proof, let X be an arbitrary q-element set and $2_2^X = \mathcal{A}_1 \cup \mathcal{A}_2$ be any 2-partition of the set of the 2-element subsets of X. If $\mathcal{A}_1 \neq \emptyset$, which means that $\mathcal{A}_1$ contains at least one pair $\{a, b\}$, then we can set $A = \{a, b\}$—this 2-element set $A \subset X$ solves the problem. Otherwise, that is, if $\mathcal{A}_1 = \emptyset$, we set $A = X$. That establishes the inequality $R_2(2, q) \leq q$.

To prove that the strict inequality $R_2(2, q) < q$ cannot hold, we consider a set X with $|X| \leq q - 1$. Then for the ordered 2-partition $2_2^X = \emptyset \cup \mathcal{A}_2$ there is no subset $A \subset X$ with the properties we sought. Indeed, here $\mathcal{A}_1 = \emptyset$, hence there is no 2-element subset of A belonging to $\mathcal{A}_1$, and also A cannot contain a q-element subset since $|X| \leq q - 1$. This yields the equation $R_2(2, q) = q$. □

We shall prove Theorem 5.1.1 only in the case $p = t = 2$ following [16]. A proof of the general case can be found, for example, in [24, p. 73–74]. Restate the theorem in the case $p = t = 2$.

Theorem 5.1.3. (Erdös–Szekeres) *For arbitrary integer numbers $k, l \geq 2$ there exists the smallest number $R = R_2(k, l) = R_2(l, k)$ such that for any set X with $|X| \geq R$ and for any ordered 2-partition $2_2^X = \mathcal{A}_1 \cup \mathcal{A}_2$ of all 2-element subsets of X either there exists a subset $T \subset X$, such that $|T| = k$ and $2_2^T \subset \mathcal{A}_1$, or there exists a subset $U \subset X$ such that $|U| = l$ and $2_2^U \subset \mathcal{A}_2$.*

Proof. We will prove the theorem by mathematical induction on both k and l, using Lemma 5.1.1 as the basis of induction. Let $k, l > 2$. By the inductive assumption, there exist numbers $R_2(k - 1, l)$ and $R_2(k, l - 1)$ defined in the statement of Theorem 5.1.3. We will prove that there exists a number $R_2(k, l)$ with the required property and this number satisfies the inequality

$$R_2(k, l) \leq R_2(k - 1, l) + R_2(k, l - 1).$$

Denote

$$p = R_2(k - 1, l) + R_2(k, l - 1)$$

and consider a set X with $|X| \geq p$ and any partition $2_2^X = \mathcal{A}_1 \cup \mathcal{A}_2$. Choose an arbitrary $x \in X$ and introduce two sets, $A = \{y \in X \mid \{x, y\} \in \mathcal{A}_1\}$ and $B = \{y \in X \mid \{x, y\} \in \mathcal{A}_2\}$.

Since $|A| + |B| \geq p - 1$, then either $|A| \geq R_2(k - 1, l)$ or $|B| \geq R_2(k, l - 1)$, otherwise we would have had $|A| + |B| \leq p - 2$. These two cases are symmetric and it is sufficient to consider only one of them; suppose that $|A| \geq R_2(k - 1, l)$. Assuming this inequality, we construct a 2-partition $2_2^A = \widehat{\mathcal{A}_1} \cup \widehat{\mathcal{A}_2}$, where $\widehat{\mathcal{A}_i} = \mathcal{A}_i \cap 2_2^A, i = 1, 2$. By the inductive assumption, either there exists $\widehat{T} \in 2_{k-1}^A$ such that $2_2^{\widehat{T}} \subset \widehat{\mathcal{A}_1}$, or else there exists $\widehat{U} \in 2_l^A$ such that $2_2^U \subset \widehat{\mathcal{A}_2}$. However, $A \subset X \setminus \{x\}$ and $\widehat{\mathcal{A}_i} \subset \mathcal{A}_i, i = 1, 2$. In the first case, since the set $\widehat{T}$ exists, we let $T = \widehat{T} \cup \{x\} \in 2_k^X$, therefore $2_2^T \subset \widehat{\mathcal{A}_1} \cup \mathcal{A}_1 = \mathcal{A}_1$. In the second case, when a nonempty set $\widehat{U}$ exists, it is immediately clear that $U \in 2_l^X$ and $2_2^U \subset \mathcal{A}_2$. □

Corollary 5.1.1. *The numbers $R_2(k,l)$ are finite for all $k,l \geq 2$.*

Corollary 5.1.2.

$$R_2(k,l) \leq \binom{k+l-2}{k-1}.$$

Consider two applications of Ramsey's theorem.

Theorem 5.1.4. *Let k,l be two arbitrary natural numbers and $\{x_1,\dots,x_n\}$ be any set of distinct real numbers. There exists a number $R = R(k,l)$ such that if $n \geq R$ then in the sequence $\{x_1,\dots,x_n\}$ there is either an increasing subsequence of length k, or a decreasing subsequence of length l.*

Proof. Introduce the set of indices $X = \{1,2,\dots,n\}$ and split all ordered pairs (i,j) of indices from X with $i < j$ as $2^X_2 = \mathcal{A}_{\text{inc}} \cup \mathcal{A}_{\text{dec}}$, where $(i,j) \in \mathcal{A}_{\text{inc}}$ if $x_i < x_j$ and $(i,j) \in \mathcal{A}_{\text{dec}}$ if $x_i > x_j$. It is clear now that the statement follows from Theorem 5.1.1. □

The next statement can be proved quite similarly.

Corollary 5.1.3. (Schur) *For every positive integer k there is a (large enough) natural number $n = n(k)$ such that for any partition of the set $\{1,2,\dots,n\}$ in k subsets at least one of these subsets contains three numbers x,y,z, such that $z = x + y$.*

Problem 5.1.6. Assuming that any two people either are familiar with one another or are not, prove that among any six people either there are three who are familiar with one another or there are three who do not know each other.

Proof. Let X be any 6-element set. Consider a 2-partition $2^X_2 = \mathcal{A}_1 \cup \mathcal{A}_2$, assigning to $\mathcal{A}_1$ all pairs of people familiar with one another and to $\mathcal{A}_2$ all remaining pairs. If T is a set of pairwise familiar people, then $2^T_2 \subset \mathcal{A}_1$; on the other hand, if U is a set of pairwise unfamiliar people, then $2^U_2 \subset \mathcal{A}_2$. We have to prove that we can find either a set T such that $|T| = 3$, or a set U such that $|U| = 3$.

Hence, to solve the problem it is sufficient to prove that $R_2(3,3) \leq 6$. We actually prove a stronger statement $R_2(3,3) = 6$, which in particular solves Exercise 5.1.20. □

Lemma 5.1.2. $R_2(3,3) = 6$.

Proof. The following well-known proof uses graph-theory language. Consider a complete graph K_6 modeling this party of six. If two people are familiar with one another, then the edge, connecting the vertices corresponding to these two people, is marked by Y; all other edges are marked by N—in the example in Fig. 5.1 some irrelevant for the proof labels Y and N are omitted.

Consider any vertex of the graph, say, the vertex F in Fig. 5.1. Among five edges incident to F, at least three edges must have the same labels—in Fig. 5.1 three edges are labeled by Y. The three end-vertices of these three edges, which are different from the F, make a triangle—in Fig. 5.1 this is ΔABC. If all the three sides of this triangle

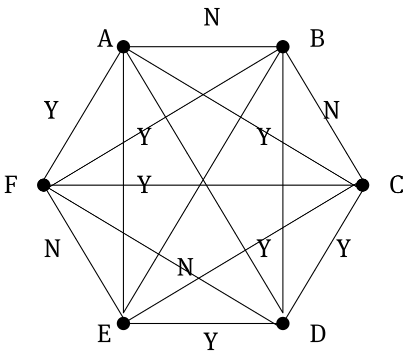


Figure 5.1: Proof of Lemma 5.1.2.

bear the same label, then these vertices form a triple of elements we look for. If these sides have different labels, then there exists a side s of this triangle with the same label as that marking three initial edges (incident to F)—the side s and the two edges connecting its end-vertices with the initial vertex form the triangle we seek. In Fig. 5.1, the edge AC is labeled by Y, as well as AF and CF, hence, the triangle ΔACF has the same marks on all the three of its sides.

We have proved that $R_2(3,3) \le 6$. The following example shows that the right-hand side 6 in this inequality cannot be decreased to 5, that is, the conclusion cannot be claimed for all 5-element sets. Indeed, let $X = \{a,b,c,d,e\}$, thus,

$$2_2^X = \{\{a,b\},\{a,c\},\{a,d\},\{a,e\},\{b,c\},\{b,d\},\{b,e\},\{c,d\},\{c,e\},\{d,e\}\}.$$

Consider two sets of pairs,

$$\mathcal{A}_1 = \{\{a,c\},\{a,e\},\{b,d\},\{b,e\},\{c,d\}\}$$

and

$$\mathcal{A}_2 = \{\{a,b\},\{a,d\},\{b,c\},\{c,e\},\{d,e\}\}.$$

It is obvious that there is no 3-element set $Y \subset X$ such that either $2_2^Y \subset \mathcal{A}_1$ or $2_2^Y \subset \mathcal{A}_2$. □

Problem 5.1.7. Where in the proof was the pigeonhole principle used with $t = 2$?

Remark 5.1.2. Lemma 5.1.2 can be stated as follows: If edges of the complete graph K_6 are colored in two colors, then the graph contains a triangle with edges of the same color. Exercise 5.1.19 can be restated similarly if one considers a three-coloring of the edges of K_{17}.

Problem 5.1.8. Is it true or false that among any six natural numbers either there are three pairwise mutually prime or there are three numbers whose common divisor is greater than one?

Problem 5.1.9 (Cf. Exercise 1.1.28). Consider a simple graph of order $k \geq 2$. Show that this graph has at least two vertices of the same degree.

Solution. If the graph has two isolated vertices, they have the same (zero) degree. Otherwise, let X be the set of non-isolated vertices and $|X| = t+1$, $t \geq 0$. Simple graphs have no parallel edges nor loops, thus, a vertex in X can have any degree from 1 through t; we denote by A_i the subset of vertices of degree i, $1 \leq i \leq t$. Since $|X| > t$, by (5.1.2) with $q_1 = \cdots = q_t = 2$ we have $R_1(2, \ldots, 2) = t + 1$. Therefore, there exists an i such that $1 \leq i \leq t$ and $|A_i| \geq 2$. □

Problem 5.1.10. Consider the complete graph K_n and an arbitrary coloring of its edges in two colors A and B. Show that for any natural numbers p and q there is a natural number $n_0(p, q)$ such that for any $n \geq n_0(p, q)$ the graph either contains an A-colored subgraph of order p or a B-colored subgraph of order q.

Solution. It suffices to apply Theorem 5.1.2 to the set of vertices of the graph, decomposing all pairs of vertices in two subsets depending upon the color of the edge connecting these two vertices. □

The following, almost obvious, statement is equivalent to the Dirichlet principle.

Problem 5.1.11. Consider two finite sets X and Y and a function $f : X \to Y$. If for any $y \in Y$ its preimage $f^{-1}(\{y\})$ contains at most k elements, then $|X| \leq k|Y|$. □

Exercises 5.1.

Exercise 5.1.1. A family has eight siblings. Prove that at least two of them were born the same day of week.

Exercise 5.1.2. There are 12 red, 10 blue, 10 green, and 8 yellow pencils. Which is the smallest number of pencils that we must pick at random if we need
(1) at least 6 pencils of the same color?
(2) at least 6 green pencils?
(3) at least 1 pencil of each color?
(4) at least 4 pencils of the same color?

Exercise 5.1.3.
(1) In a class of 37 students, are there 4 of them who celebrate their birthday the same month?
(2) Answer the same question for a class with 36 students.

Exercise 5.1.4. A student claims that at least 4 people in her class were born in the same month. Which is the smallest size of the class?

Exercise 5.1.5. Among given 11 lines in a plane no two are parallel. Prove that we can find two lines among them, such that the angle between them is less than 17^o.

Exercise 5.1.6. How many are there key chains with 7 identical apartment keys and 3 identical lobby keys?

Exercise 5.1.7. Prove that among any 6 integers there are two numbers such that 5 divides their difference.

Exercise 5.1.8. Which is the largest cardinality of a set of natural numbers not exceeding 10, if no number among them is twice another number?

Exercise 5.1.9. There are three identical pairs of black socks and three pairs of white socks in a drawer. Which is the smallest number of separate socks that must be drawn at random to ensure that one gets at least one complete pair of black socks?

Exercise 5.1.10. There are 70 balls of the same size but of different colors in a box, among them 20 red, 20 green and 20 yellow balls; the others are black and white balls. Which is the smallest number of balls to be chosen at random from the box to ensure that at least 10 same-color balls are selected?

Exercise 5.1.11. Consider n-digit natural numbers with $n \geq 3$, whose decimal representations contain only three digits 1,2,3. How many of such numbers contain each of these digits at least once?

Exercise 5.1.12. How many are there natural numbers less than 84 900 000 and mutually prime with that number?

Exercise 5.1.13. A high school rented 11 buses for the senior prom. The maximal load of every bus is 40 students. Which are the smallest and the largest number of seniors in the school this year, if at least three buses carry the same number of students?

Exercise 5.1.14.
(1) Prove that any 6-element sequence of natural numbers contains either three numbers going in increasing order or three numbers going in decreasing order.
(2) Is this conclusion true for 5-element sequences?
(3) Is this conclusion true for 9-element sequences and 4-element subsequences?

Exercise 5.1.15.
(1) Arrange the integers from 1 through 100 inclusive, so that this ordering does not contain an increasing subsequence of length 11, nor a decreasing subsequence of length 11.
(2) Prove that no such arrangement is possible for the first 101 natural numbers, that is, prove that any permutation of the integers from 1 through 101 inclusive either contains an increasing subsequence of length 11, or a decreasing subsequence of length 11.

Exercise 5.1.16. A test consists of five problems. Five students took the test and each of them solved at least two problems. Prove that at least two students solved the same number of problems.

Exercise 5.1.17. The conclusion of Problem 5.1.8 is false. How to change its statement to make it true?

Exercise 5.1.18. Prove that among any nine people there are either three pairwise familiar with one another or four pairwise unfamiliar, that is, prove that $R_2(3,4) \le 9$. Moreover, in a party of eight this property fails. In other words, prove that the Ramsey number $R_2(3,4) = 9$.

Exercise 5.1.19. Among 17 students there are people collecting stamps, postcards, and coins. Each pair of students has one and only one common hobby. Prove that there are at least three students with a mutual hobby. Is it always possible to find four people with a mutual hobby among these 17 students? Rephrase this problem in terms of the Ramsey numbers and in terms of colorings of the complete graph K_{17}.

Exercise 5.1.20. In Theorem 5.1.3 we have proved that

$$R_2(k,l) \le R_2(k-1,l) + R_2(k,l-1).$$

Show that if both $R_2(k-1,l)$ and $R_2(k,l-1)$ are even, then this inequality is strict, that is, the equality case cannot occur here.

Exercise 5.1.21. Prove that $R_2(3,5) = 14$ and $R_2(4,4) = R_2(3,6) = 18$.

Exercise 5.1.22. Prove that $R_2(k,l) \le C(k+l-2, k-1), k,l \ge 2$.

Exercise 5.1.23. Given 20 pairwise distinct natural numbers less than 65, prove that among the pairwise differences of these numbers there are at least four equal numbers.

Exercise 5.1.24. There are white, black, and brawn gloves in a drawer, at least two pairs of each color. Which is the smallest number of gloves that one has to pick at random in order to get two pairs (four gloves) of the same color?

Exercise 5.1.25. Which is the smallest number of integers to be chosen from the set $T = \{1, 2, \dots, 15\}$ so that the difference of two of the numbers chosen is 6?

Exercise 5.1.26. Prove that among any 101 integer numbers there are at least two numbers such that 100 divides their difference.

Exercise 5.1.27. In a small town there are 10 000 cars, whose license plates are numbered by 4-digit numbers. If a number has less than 4 digits, we append in front of it a few zeros like 0012. More than half of the cars are registered in the central district of the town. Prove that there is a car in the central district, whose number is the sum of numbers of two other cars from this district.

Exercise 5.1.28. Show that if 6 points are selected at random inside a square of side 1 cm., then at least two of them are less than 0.5 cm apart.

Exercise 5.1.29. All edges of a complete graph with 17 vertices are colored in three colors. Prove that there is a triangle in the graph whose edges have the same color.

Exercise 5.1.30. The 6-element set $T = \{1, 12, 23, 34, 45, 56\}$ possesses the following property: for any two numbers in T the last digits of their sum and of their difference are not 0. Prove that 6 is the biggest number with this property, that is, prove that any 7-element set of integers contains a pair of numbers such that 10 divides either their sum or their difference.

Exercise 5.1.31. If six different numbers are chosen from the set $T = \{1, 2, \dots, 10\}$, then there are at least two consecutive numbers among the six.

Exercise 5.1.32. Eight numbers are chosen from the set $T = \{1, 2, \dots, 10\}$. Show that there are at least three pairs of these numbers with the sum 11. Is this conclusion true if only seven numbers are chosen?

Exercise 5.1.33. 22 people gathered at the alumni reunion at a Small College, among them engineers, chemists, and business people. Show that at least one major was represented by eight or more alumni.

Exercise 5.1.34. A set of integers contains at least two numbers congruent modulo 11 (Definition 1.1.18). Which is the smallest cardinality of such a set of integers?

Exercise 5.1.35. Find a coloring of the edges of the complete graph K_{13} in two colors, blue and green, so that no subgraph of order 3 has only blue edges and no subgraph of order 5 has only green edges.

Exercise 5.1.36. Find a coloring of the edges of the complete graph K_{17} in two colors, so that no subgraph of order 4 is monochromatic.

Exercise 5.1.37. Prove that for any two mutually prime integers m and n there is a natural number k such that n divides $m^k - 1$.

Exercise 5.1.38. The sum of all entries of a 10×10 zero-one matrix is 81. Prove that the matrix contains a row and a column such that the sum of elements in these two lines is at least 17.

Exercise 5.1.39. A township of 51 houses occupies a square of 1 mile side. Prove that at least three houses are inside a circle of radius $\frac{1}{7}$ mi.

Exercise 5.1.40. Prove Corollaries 5.1.2 and 5.1.3.

Exercise 5.1.41. Prove that for every integer $l \geq 1$ there is the smallest natural number $n = n(l)$ such that for any partition of the set $\{1, 2, \dots, n\}$ in two subsets at least one of these subsets contains $l+1$ numbers $x_1, \dots, x_l, x_{l+1}$ satisfying the equation $x_1 + \cdots + x_l = x_{l+1}$. Prove in particular that $n(2) = 5$.

Exercise 5.1.42. Prove that a decimal expansion of any non-zero rational number is a periodic (repeating) decimal; it can start with a finite pre-period. For example, $2/15 = 0.133\ldots$—here the period is 3 and the pre-period is 1.

Exercise 5.1.43. How many integers are there between 0 and 10^n which do not contain the same two digits going together?

Exercise 5.1.44. Prove that $R_2(k,k) \geq \text{const}\, k \cdot 2k/2$.

5.2 Systems of distinct representatives

Several statements considered in this section are equivalent to each other as well as to some other results such as Menger's theorem on disjoint chains in graphs or the maximal flow theorem—see, for instance, [19, p. 11 and p. 55]. They have numerous applications. Hereafter we systematically use ordered families of sets $U = (S_1, \ldots, S_n)$ in the sense of Definitions 5.1.1 and 5.1.2. In the end of the section we consider matchings in bipartite graphs.

Coffee-time browsing

- http://www-history.mcs.st-and.ac.uk/Biographies/Halmos.html (Halmos' biography)
- http://www.youtube.com/watch?v=ONvYPldXoZs (I Want To Be A Mathematician)
- http://www-history.mcs.st-and.ac.uk/Mathematicians/Menger.html
- http://www.iit.edu/csl/am/about/menger/about.shtml (Carl Menger)
- www.gap-system.org/~history/Mathematicians/Hall.html (Ph. Hall's life and work)
- http://www.gap-system.org/~history/Biographies/Hall_Marshall.html (M. Hall's life and work)
- http://robertborgersen.info/Presentations/GS-05R-1.pdf (Equivalence of Seven Major Combinatorial Theorems)
- http://www.viswiki.com/en/D%C3%A9nes_K%C5%91nig (Konig's biography)
- http://www.viswiki.com/en/Robert_P._Dilworth (Dilworth's biography)
- http://myweb.lsbu.ac.uk/~whittyr/MathSci/TheoremOfTheDay/CombinatorialTheory/Dilworth/TotDDilworth.pdf (Dilworth's Theorem)
- http://russell.lums.edu.pk/~cs211bs08/slides/proofsfrombookthreethms.pdf (More on the Marriage Theorem)

Definition 5.2.1. Given a set S and a family $U = (S_1, \ldots, S_n)$ of its subsets, a set $D = \{a_1, \ldots, a_n\} \subset S$ is called a *system of distinct representatives* (SDR) or a *transversal* of the family U, if there exists a permutation $(j_1, j_2, \ldots, j_n)$ of indices $1, 2, \ldots, n$ such that $a_{j_i} \in S_i, 1 \leq i \leq n$.

This definition becomes transparent from the following example.

Example 5.2.1. Let $S = \{1, 2\}$. The 3-family $(\{1\}, \{2\}, \{1, 2\})$ cannot have a SDR, since the two available elements, 1 and 2, cannot represent the three subsets $\{1\}, \{2\}$, and $\{1, 2\}$ of the family. However, the 2-family $(\{1\}, \{1, 2\})$ has one SDR $D = \{1, 2\}$.

It is clear from this example, that in order to have a SDR, the family of subsets cannot contain more terms than the cardinality of the union of these subsets. It turns out that this simple necessary condition is also sufficient for the existence of a SDR.

Definition 5.2.2. A family $U = (S_1, \ldots, S_n)$ of subsets of a set S satisfies *Hall's condition* **(H)** if the inequality

$$|S_{i_1} \cup S_{i_2} \cup \cdots \cup S_{i_k}| \geq k \tag{5.2.1}$$

holds for each $k, 1 \leq k \leq n$, and for any set of indices $(i_1, i_2, \ldots, i_k)$.

Theorem 5.2.1. *An ordered family $U = (S_1, \ldots, S_n)$ of subsets of a finite set S has a SDR if and only if U satisfies the condition* ***(H)***.

Proof. The necessity of the condition is clear. Indeed, if the family U has a SDR $\{a_1, \ldots, a_n\}$, then for any set of indices $(i_1, i_2, \ldots, i_k)$ the k-element set $\{a_{i_1}, \ldots, a_{i_k}\}$ is a SDR for a sub-family $(S_{i_1}, \ldots, S_{i_k})$, thus the union $S_{i_1} \cup \cdots \cup S_{i_k}$ contains at least these k elements $a_{i_1}, \ldots, a_{i_k}$, and the condition **(H)** is valid. □

The sufficiency of the condition **(H)** follows from the next theorem of M. Hall, which gives also a lower bound of the number of SDR. It is convenient to introduce the following notation.

For a family $U = (S_1, \ldots, S_n)$ of subsets of a finite set S denote

$$\mu = \mu_U = \min_{1 \leq i \leq n} |S_i|.$$

Theorem 5.2.2. *If the family $U = (S_1, \ldots, S_n)$ of subsets of a finite set S satisfies the condition* ***(H)*** *then U has at least $\mu!$ SDR if $\mu \leq n$ and at least $\frac{\mu!}{(\mu-n)!}$ SDR if $\mu \geq n$.*

Proof. We will prove the statement by mathematical induction on the number of terms n of the family. If $n = 1$, then the family U consists of one set S_1, and the inequality $|S_1| = \mu \geq n = 1$ is true for any natural μ. Since in this case any element of S_1 makes a SDR, there are exactly

$$\frac{\mu!}{(\mu - n)!} = \frac{\mu!}{(\mu - 1)!} = \mu$$

SDR, and this establishes the basis of induction.

To make an inductive step, we pick a natural n and assume that the conclusion of the theorem is valid for all families with less than n terms, that is, under the condition

(H) any family of less than n terms has at least the above-mentioned number of SDR. To prove the same conclusion for any n-family of subsets, we proceed in two steps.

First, introduce a strengthened condition $\widetilde{\mathbf{(H)}}$. Say that a family U satisfies the condition $\widetilde{\mathbf{(H)}}$ if the inequality

$$|S_{i_1} \cup S_{i_2} \cup \cdots \cup S_{i_k}| \geq k+1$$

holds for all k and for all sets of indices $(i_1, i_2, \ldots, i_k)$ with $1 \leq k \leq n-1$, while for $k = n$ we still assume (5.2.1) as in **(H)**.

Consider a family U satisfying the strengthened condition $\widetilde{\mathbf{(H)}}$. Choose an element $a \in S_1$ and consider sets $S_i' = S_i \backslash \{a\}, 2 \leq i \leq n$. A new family $U' = (S_2', \ldots, S_n')$ satisfies the original condition **(H)**, since it consists of less than n subsets, and by the strengthened condition $\widetilde{\mathbf{(H)}}$

$$\left|S_{i_1}' \cup S_{i_2}' \cup \cdots \cup S_{i_k}'\right| \geq |S_{i_1} \cup S_{i_2} \cup \cdots \cup S_{i_k}| - 1 \geq k.$$

It is obvious that, for any i, $|S_i'| \geq \mu - 1$. If $\mu \leq n$, then $\mu - 1 \leq n - 1$. However, the size of the new family U' is $n-1$, and by the inductive assumption, the new family U' has at least $(\mu-1)!$ of SDR. If $\mu > n$, then $\mu - 1 > n - 1$ and U' has at least $\frac{(\mu-1)!}{(\mu-n)!}$ of SDR. Since the element $a \in S_1$ was excluded from all of the sets in the new family U', this element, being appended to any SDR for U', makes up a SDR for the original family U. Now, the element a can be chosen at least in μ ways, thus, multiplying the assumed number of SDR for the new family U' by μ, we arrive at the conclusion of Theorem 5.2.2 under the strengthened condition $\widetilde{\mathbf{(H)}}$.

Suppose now that the condition $\widetilde{\mathbf{(H)}}$ fails but the original condition **(H)** holds true. Hence, for some $k, 1 \leq k \leq n-1$, and for some set of indices $(i_1, i_2, \ldots, i_k)$ the equality

$$|S_{i_1} \cup S_{i_2} \cup \cdots \cup S_{i_k}| = k$$

holds good. Without loss of generality we assume that $i_j = j$, that is,

$$|S_1 \cup S_2 \cup \cdots \cup S_k| = k.$$

We notice that now the parameter μ satisfies

$$\mu = \min |S_i| \leq |S_1| \leq |S_1 \cup S_2 \cup \cdots \cup S_k| = k,$$

where k is the size of the family $U_k = (S_1, S_2, \ldots, S_k)$. It follows by the inductive assumption that this family U_k has at least $\mu!$ of SDR; let $D = \{a_1, \ldots, a_k\}$ be any of them. To complete the proof, we now show that the shortened family $U'' = (S_{k+1}'', \ldots, S_n'')$, where $S_j'' = S_j \setminus D$, also satisfies the condition **(H)**.

Suppose on the contrary, that the family U'' does not satisfy the condition **(H)**. Then one can find a sub-family $(S_{j_1}'', \ldots, S_{j_l}''), k+1 \leq j_i \leq n$, such that

$$\left|S_{j_1}'' \cup \cdots \cup S_{j_l}''\right| = l'' < l.$$

However, the families U_k and U'' were constructed mutually disjoint, so that

$$\begin{aligned}&\left|S_1 \cup S_2 \cup \cdots \cup S_k \cup S''_{j_1} \cup \cdots \cup S''_{j_l}\right| \\ &\quad = |S_1 \cup S_2 \cup \cdots \cup S_k \cup S_{j_1} \cup \cdots \cup S_{j_l}| = k + l'' < k + l.\end{aligned}$$

Thus, the condition (5.2.1) fails also for the original family U, which contradicts the premise. Hence, the family U'' satisfies the condition **(H)** and by the inductive assumption has at least one SDR. Combining this SDR with any of $\mu!$ SDR for the family U_k, as we did with the element a above, we derive $\mu!$ SDR for the original family U. The proof of Theorem 5.2.2 is now complete. Simultaneously we proved Theorem 5.2.1. □

A constructive proof of the existence of SDR can be found, for instance, in [24, Section 5.1]. Theorem 5.2.1 is sometimes called the *marriage theorem* or the *theorem on village weddings* due to the following its reformulation.

Problem 5.2.1. Among young people attending a party, each boy is familiar with at least m of the attending girls, however, every girl knows no more than m boys. Demonstrate that every boy can marry a girl he has been familiar with.

Solution. Denote the number of boys by p and let G_i be the set of girls familiar with the ith boy. We prove that the family $(G_1, \dots, G_p)$ satisfies the condition **(H)**. Otherwise, there would existed a set of indices $i_1, i_2, \dots, i_k, k \le p$, such that $|G_{i_1} \cup \cdots \cup G_{i_k}| \le k - 1$, which means that k boys, say, $b_{i_1}, b_{i_2}, \dots b_{i_k}$, together have at most $k - 1$ familiar girls; let us denote these girls by $g_{j_1}, g_{j_2}, \dots g_{j_l}$, where $l \le k - 1$.

Consider all pairs $(b_\alpha, g_\beta), \alpha \in \{i_1, \dots, i_k\}, \beta \in \{j_1, \dots, j_l\}$, such that the boy b_α is familiar with the girl g_β—obviously, this is a symmetric binary relation. Denote by Υ the total number of such pairs. By assumption, for a fixed β there are no more than m such pairs, and since $l \le k - 1$, all in all we have $\Upsilon \le m(k - 1)$. On the other hand, for a fixed α there are no less than m such pairs, thus, $\Upsilon \ge mk$ implying that $mk \le m(k - 1)$. This contradiction proves that the family $(G_1, \dots, G_p)$ satisfies the Hall condition, consequently, it has a SDR, say, $\{g_{i_1}, g_{i_2}, \dots g_{i_p}\}$. The latter exactly means that a boy b_j is familiar with the girl $g_{i_j}, 1 \le j \le p$. □

Analyzing this solution, we immediately derive the following sufficient condition for the existence of a SDR.

Problem 5.2.2. Let $U = (S_1, \dots, S_n)$ be a family of subsets of a finite set S. Prove that if all the subsets S_i have the same cardinality k, $|S_i| = k, 1 \le i \le n$, and each element of the set S belongs to exactly k of the subsets S_i, then the family U has a SDR.

SDR have many applications. Let (A) and (B) denote two m-partitions, in the sense of Definition 1.1.7, of a finite set T,

$$T = A_1 \cup A_2 \cup \cdots \cup A_m = B_1 \cup B_2 \cup \cdots \cup B_m.$$

Consider an m-element subset $E \subset T, |E| = m$, such that $A_i \cap E \neq \emptyset$ and $B_i \cap E \neq \emptyset$ for each $i, 1 \leq i \leq m$. Then obviously

$$|A_i \cap E| = |B_i \cap E| = 1, \quad 1 \leq i \leq m.$$

Definition 5.2.3. The set E is called a *system of mutual representatives* of the partitions (A) and (B).

Clearly, a system of mutual representatives exists if and only if one can renumber sets in the partitions so that $|A_i \cap B_i| \neq \emptyset, 1 \leq i \leq m$. We prove a criterion for the existence of a system of mutual representatives analogous to the condition **(H)**.

Theorem 5.2.3. *Two m-partitions (A) and (B) have a system of mutual representatives if and only if for every $k, 1 \leq k \leq m$, and for any set of indices $i_1, i_2, \ldots, i_k$ the union $A_{i_1} \cup \cdots \cup A_{i_k}$ contains no more than k of the sets $B_1, B_2, \ldots, B_m$.*

Proof. The necessity is obvious, as in Theorem 5.2.1. For, if

$$A_1 \cup \cdots \cup A_k \supset B_1 \cup B_2 \cup \cdots \cup B_{k+1}$$

then k elements $a_1, \ldots, a_k$, $a_i \in A_i, 1 \leq i \leq k$, cannot represent $k+1$ sets $B_1, B_2, \ldots, B_{k+1}$.

To establish the sufficiency, we consider a set $S = \{A_1, \ldots, A_m\}$ and introduce the family $U = (S_1, \ldots, S_m)$, where S_i is the totality of sets A_j such that $A_j \cap B_i \neq \emptyset, 1 \leq i \leq m$. We prove that the family U satisfies the Hall condition **(H)**. On the contrary, if for some k the union $S_1 \cup \cdots \cup S_{k+1}$ would contain at most k elements (sets) $A_{i_1}, \ldots, A_{i_k}$, then it were

$$A_{i_1} \cup \cdots \cup A_{i_k} \supset B_1 \cup B_2 \cup \cdots \cup B_{k+1}$$

notwithstanding the assumption. Hence due to Theorem 5.2.1, the family U has a SDR. Now we can renumber the components of the partition (A) so that this SDR becomes $D = \{A_1, \ldots, A_m\}$ and arrive at the conclusion by making use of the remark before the theorem. □

Theorem 5.2.3 can be reformulated as follows.

Theorem 5.2.4. *Two m-partitions (A) and (B) have a system of mutual representatives if and only if for any $k = 1, 2, \ldots, m$ no k of the sets A_i are contained in the union of less than k of the sets B_j.*

Problem 5.2.3. $m \times p$ couples attend a party. The gentlemen belong to m professions, p men in each trade. The attending ladies belong to m clubs, p women in each club. Show that it is possible to select m pairs for a dance representing all clubs and all professions.

Solution. Introduce a set C, whose elements are $m \times p$ couples at the party, and consider two of its partitions:

$$C = A_1 \cup A_2 \cup \cdots \cup A_m = B_1 \cup B_2 \cup \cdots \cup B_m.$$

Here each $A_i, i = 1, 2, \ldots, m$, stands for the set of all pairs where gentlemen have the same profession, and every B_j stands for the set of all pairs where the ladies belong to the same club. Since

$$|A_1| = \cdots = |A_m| = |B_1| = \cdots = |B_m| = p.$$

Theorem 5.2.4 certainly applies, and the partitions (A) and (B) have a system of mutual representatives—they form a set of m pairs we need. □

As another demonstration of the power of Hall's theorem, we consider its application to counting bases in finite-dimensional vector spaces. Recall that a *basis* in a *vector space* V is a linearly-independent set of vectors, which spans the whole space, meaning that any vector of the space can be expanded through the basis vectors.

Theorem 5.2.5. *Any two bases of a finite-dimensional vector space consist of the same number of vectors.*

Proof. Let $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_m\}$ be two bases of a vector space V. Expand the vectors x_i against y_j,

$$x_i = \sum_{k=1}^{m_i} f_{i_k} y_{j_k}, \text{ where } m_i \le m \text{ and all } f_{i_k} \ne 0,$$

and consider the family $U = (S_1, \ldots, S_n)$ of sets $S_i = \{y_{j_1}, \ldots, y_{j_{m_i}}\} \subset V$. Thus, for each $i, 1 \le i \le n$, the set S_i consists of the basis vectors y_{j_k} spanning the vector x_i.

We claim that this family satisfies the condition **(H)**. Otherwise, we would be able to express certain k basis vectors x_i through less than k vectors y_j, which means that these *basis* vectors x_i are linearly dependent. This contradiction implies that the n-family U has a SDR. Since these n distinct representatives belong to the set $S = \{y_1, y_2, \ldots, y_m\}$, we must have $n \le m$. The reversed inequality follows the same lines due to symmetry, so that $m = n$. □

Problem 5.2.4. For the latter theorem to hold true, it is sufficient that the vectors $\{y_j\}$ span the entire space and the vectors $\{x_i\}$ are linearly independent. Then the invariance of the dimension of the vector space follows.

The next application of Hall's theorem deals with extremal combinatorial problems. For detailed exposition of this topic see, for example, [23], here we consider only the *assignment problem*. Suppose that there are n jobs that must be assigned to n employees on a one-to-one basis. The *utility* (usefulness or uselessness) of the ith worker at the jth position is measured by the entry $t_{i,j}$ of the *utility matrix* T. Any assignment is given by a permutation

$$\Pi = \begin{pmatrix} 1 & 2 & \dots & n \\ j_1 & j_2 & \dots & j_n \end{pmatrix}$$

where the first row lists the employees and $j_i = \Pi(i)$ denotes a job the ith worker is assigned to. To improve performance, we have to maximize the sum

$$\sum_{i=1}^{n} t_{i,\Pi(i)}$$

over all permutations Π. A direct solution of the problem by the brute force enumeration of all $n!$ permutations is unfeasible even for moderate values of n. However, the problem has an effective algorithmic solution.

Theorem 5.2.6. *Let $T = (t_{i,j})$ be an $n \times n$ matrix with real entries. Then*

$$\max_{\Pi} \sum_{i=1}^{n} t_{i,\Pi(i)} = \min\left(\sum_{i=1}^{n} w_i + \sum_{j=1}^{n} v_j\right) \tag{5.2.2}$$

where the maximum on the left is taken over all n-permutations of the set $S = \{1, 2, \dots, n\}$ and the minimum on the right is taken over all numbers w_i and v_j, such that

$$w_i + v_j \ge t_{i,j} \text{ for all } 1 \le i, j \le n.$$

This common extreme value of (5.2.2) *is attained for certain values $j_i = \Pi^*(i)$ such that*

$$w_i + v_{\Pi^*(i)} = t_{i,\Pi^*(i)}, \quad i = 1, \dots, n,$$

and the permutation Π^* solves the assignment problem.

Proof. We prove the theorem only for the integer-valued utilities $t_{i,j}$. The general case can be found, for example, in [24, Sect. 7.1].

For a given integer-valued matrix T we can always find the integer numbers w_i, v_j, such that $w_i + v_j \ge t_{i,j}, 1 \le i, j \le n$. It suffices, for example, to set all $v_j = 0$ and $w_i = \max_{1 \le j \le n} t_{i,j}$. Then $w_i + v_{\Pi(i)} \ge t_{i,\Pi(i)}$ for any permutation Π, and summing up over $i = 1, 2, \dots, n$, we deduce

$$\sum_{i=1}^{n} w_i + \sum_{j=1}^{n} v_j \ge \sum_{i=1}^{n} t_{i,\Pi(i)}.$$

This readily yields the inequality $m \ge M$, where m and M are, respectively, the minimum and the maximum appearing in (5.2.2). We use Hall's theorem to prove that actually $m = M$. The subsequent proof is *constructive*, that is, it gives an algorithmic solution of the assignment problem.

For the given entries $t_{i,j}$, we have already shown the existence of integers w_i and v_j such that $w_i + v_j \geq t_{i,j}$. Fix a subscript i and a number w_i, and try, if it is possible, to increase v_j as long as the latter inequality holds true. This way, for each i we find at least one j such that $w_i + v_j = t_{i,j}$. Denote by $S_i, 1 \leq i \leq n$, the set of subscripts j such that $w_i + v_j = t_{i,j}$, where w_i and v_j were chosen as described above. Introduce also the n-family $U = (S_1, \dots, S_n)$. If the family U has a SDR $\{j_1, \dots, j_n\}$, then $w_i + v_{j_i} = t_{i,j_i}$ and the permutation Π^*, such that $\Pi^*(i) = j_i$, solves the problem. Therefore, to complete the proof of Theorem 5.2.6, we have to construct a SDR for U.

Suppose that the family U does not have a SDR. Then, by Theorem 5.2.1, the condition **(H)** fails for U. In turn, this means that there are subscripts $i_1, \dots, i_k$, where $1 \leq k \leq n$, such that the union $S_{i_1} \cup \dots \cup S_{i_k}$ contains at most $k-1$ different subscripts j. Denote $K = \{i_1, \dots, i_k\}$ and $S_K = S_{i_1} \cup \dots \cup S_{i_k}$; by the assumption, $|S_K| = l \leq k-1$. Denote also

$$\widehat{w}_i = \begin{cases} w_i - 1 & \text{if } i \in K, \\ w_i & \text{if } i \notin K, \end{cases}$$

$$\widehat{v}_j = \begin{cases} v_j + 1 & \text{if } j \in K, \\ v_j & \text{if } j \notin K. \end{cases} \tag{5.2.3}$$

Clearly, if $i \notin K$, then $\widehat{w}_i + \widehat{v}_j \geq t_{i,j}$. If $i \in K$ and $j \in S_K$, then $\widehat{w}_i + \widehat{v}_j = (w_i - 1) + (v_j + 1) = w_i + v_j \geq t_{i,j}$ as well. Finally, if $i \in K$ and $j \notin S_K$, then due to the definition of S_i, $w_i + v_j \neq t_{i,j}$, that is, $w_i + v_j \geq t_{i,j} + 1$; hence $\widehat{w}_i + \widehat{v}_j = w_i + v_j - 1 \geq t_{i,j}$. Thus, we have proved that $\widehat{w}_i + \widehat{v}_j \geq t_{i,j}, 1 \leq i, j \leq n$, in all possible cases. The following equation is also obvious:

$$\sum_{i=1}^{n} \widehat{w}_i + \sum_{j=1}^{n} \widehat{v}_j = \sum_{i=1}^{n} w_i + \sum_{j=1}^{n} v_j - k + l$$

that is, when we replace w_i and v_j with, respectively, $\widehat{w}_i$ and $\widehat{v}_j$, this sum decreases by an integer $k - l > 0$.

However, this sum is bounded from below by M. Therefore, after finitely many such steps the sum $\sum_{i=1}^{n} w_i + \sum_{j=1}^{n} v_j$ cannot be decreased any more, which means that after a finite number of steps the modified family U satisfies the condition **(H)**, and so that it has a SDR. The corresponding permutation, as at the beginning of the proof, solves the assignment problem. □

Problem 5.2.5. Solve the assignment problem for the 4×4 utility matrix

$$T = \begin{pmatrix} 1 & 2 & 4 & 5 \\ 3 & 3 & 2 & 6 \\ 4 & 1 & 2 & 3 \\ 2 & 1 & 3 & 4 \end{pmatrix}.$$

Solution. We illustrate here the algorithm of Theorem 5.2.6. Set initially $v_1 = v_2 = v_3 = v_4 = 0$. Since we need

$$w_1 + 0 = \max\{t_{1,j} \mid 1 \le j \le 4\}$$

we choose $w_1 = 5$. Similarly we find $w_2 = 6, w_3 = 4, w_4 = 4$. Now we observe that these w_i and the corresponding v_j have the pairs of indices $(1,4), (2,4), (3,1), (4,4)$, leading to the first set $K = \{1,2,3,4\}$ and the corresponding set $S_K = \{1,4\}$—let us recall that S_K comprises all the second indices from the index pairs above. The family of three sets $(\{4\}, \{4\}, \{1,4\})$ contains only two distinct elements and obviously does not have a SDR, therefore, we have to change K and S_K. For each index $i \in K$ we decrease w_i by 1 and for each index $j \in S_K$ we increase v_j by 1. The new values are $w_1 = 4, w_2 = 5, w_3 = 3, w_4 = 3$ and $v_1 = v_4 = 1, v_2 = v_3 = 0$.

We are to recalculate K and S_K with these new w_i and v_j, so that $w_i + v_j = t_{i,j}$. Now we have pairs of indices $(1,3), (1,4), (2,4), (3,1), (4,3), (4,4)$ and the corresponding family of four sets $(\{3,4\}, \{4\}, \{1\}, \{3,4\})$, comprising only three elements. This family also does not have a SDR and we have to repeat the basic step of the algorithm. It is clear that $i = 3$ can only be represented by $j = 1$, thus, we define the second set $K = \{1,2,4\}$ and the corresponding $S_K = \{3,4\}$. These K and S_K tell us to decrease w_1, w_2, w_4 by 1 and to increase v_3, v_4 by 1, leading to new values $w_1 = 3, w_2 = 4, w_3 = 3, w_4 = 2$ and $v_1 = 1, v_2 = 0, v_3 = 1, v_4 = 2$. These values result in the same $K = \{1,2,4\}$ and $S_K = \{3,4\}$, thus, we have to repeat the basic step one more time. After this step, the family of sets (of subscripts) derived is (S_1, S_2, S_3, S_4), where

$$S_1 = \{2,3,4\}, \quad S_2 = \{2,4\}, \quad S_3 = \{1\}, \quad S_4 = \{1,2,3,4\},$$

which obviously has SDR, for example, $(3,2,1,4)$. Thus, the permutation

$$\Pi^* = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

solves the problem and gives the maximum value $t_{1,3} + t_{2,2} + t_{3,1} + t_{4,4} = 15$. To compute the extreme value of the utility, we also can, due to the duality relation (5.2.2), use the last values of w_i and v_j; these values are (check that!) $(2+3+3+1)+(1+0+2+3) = 15$. □

The next application of Hall's theorem is concerned with *zero-one matrices*, that is, the matrices consisting of 0s or 1s. In what follows *a line* means either a row or a column of a matrix. Evidently, we can consider not only zero-one matrices but those with elements of arbitrary nature and separate all the elements into two disjoint classes. Moreover, zeros and ones in this theory are symmetric as well as rows and columns.

Definition 5.2.4. A set of rows, containing all 1s in a zero-one matrix is called a *covering* of the matrix. A collection of 1s in a zero-one matrix, such that no two 1s among them belong to the same line, is called an *independent* set of 1s.

Theorem 5.2.7. (König) *The minimum number of lines in any covering of a zero-one matrix is equal to the maximum number of independent* 1s *in the matrix.*

Proof. Let m be the cardinality of a minimal covering and M be the maximum number of independent 1s in a zero-one matrix $A = (a_{ij})$. To cover all 1s, one should at the very least cover these M independent 1s. By virtue of the independence condition, no pair of 1s among these M 1s can be covered by a line, hence at least M lines are needed and $m \geq M$.

We use Theorem 5.2.1 to prove the opposite inequality. Let the minimum covering consist of r rows and c columns, $r + c = m$. The numbers m and M certainly do not change when we rearrange rows or columns. Therefore, by changing the order of rows and columns, we can assume that the r rows appearing in the covering, are r uppermost rows of the matrix, and similarly the c columns appearing in the covering are c leftmost columns.

Introduce sets $S_i = \{j \mid u_{i,j} = 1, j > s\}, i = 1, \ldots, r$, that is, S_i contains the numbers of columns to the right of the sth column, such that the element at the intersection of such a column and the ith row is a 1. We show that a family $U = (S_1, \ldots, S_r)$ satisfies the condition **(H)**. Otherwise, it would be possible to find among these S_i sets k sets, such that their union contains at most $k - 1$ elements. Considering the construction of these sets, this would mean that in the corresponding k rows, in the columns to the right of the sth column there are altogether only $k - 1$ 1s.

But the 1s, located at the intersection of these k rows with leftmost s columns are covered by these columns—recall that the first s columns belong to the covering. Hence, if we remove these k rows from the covering, we can *uncover* at most $k - 1$ of 1s in these rows to the right of the sth column. Now, to cover these 1s, we need at most $k - 1$ columns, and by adding such $k - 1$ columns to the covering instead of the k rows removed, we derive a new covering consisting of at most $m - 1$ lines, which contradicts the minimality of m. By Theorem 5.2.1, the family U has a SDR—namely, r 1s in the upper r rows, such that no two of them are in the same row and all these 1s are in the columns with numbers greater than s.

In exactly the same way, we can choose s 1s in the leftmost s columns, such that no two 1s among them are covered by one column and all are in the rows with indices greater than r. Obviously, no two 1s among the chosen $r + s = m$ 1s are in the same line, thus, $m \leq M$. □

Many other applications of Hall's theorem, such as for instance, the calculation of permanents or construction of the Latin squares, can be found in [24] or [49]. We only prove the following beautiful theorem of Frobenius dealing with determinants. The latter were introduced in Section 4.4, Definition 4.4.6. We recall here that the determinant of an $n \times n$ matrix is an alternating sum of $n!$ products, called here the *terms* of the determinant.

Theorem 5.2.8. *In order for each of the $n!$ terms of the determinant of an $n \times n$ matrix A to be equal to zero it is necessary and sufficient that there are k rows and $n - k + 1$*

columns in A, with $1 \leq k \leq n$, not necessarily in succession, such that all entries $a_{i,j}$ at the intersection of these rows and columns are zeros.

Proof. Since a product (in our case a term of the determinant) vanishes if and only if it contains at least one vanishing factor, we can consider only zero-one matrices, replacing all non-zero elements of A with 1s. Let m and M have the same meaning as in the preceding theorem. If $M \geq n$, then A contains at least n independent (in the sense of Definition 5.2.4) non-zero elements, that is, independent 1s. Thus they form a non-zero term of the determinant $\det(A)$, and if all terms of $\det(A)$ vanish, then $M < n$. By Theorem 5.2.7, $m = M < n$, and if this m-covering of all 1s in A consists of r rows and s columns, then all other elements, which are 0s, are situated in complementary $n - r$ rows and $n - s$ columns. The intersection of these lines is a zero $(n - r) \times (n - s)$ matrix, and since $s = m - r$, we have $n - r + n - s = 2n - m > n$.

To complete the proof, it suffices now to notice that this reasoning is word-by-word reversible. □

The next statement, Dilworth's theorem, is concerned with properties of partially ordered sets (posets)—see Definition 1.1.13. Notice that if a poset is not a chain, that is, not all of its elements are pairwise comparable, then it can be decomposed in the union of disjoint chains, and this decomposition generally is not unique.

Example 5.2.2. Consider a poset $X = \{a, b, c, d, f\}$, where a binary relation of partial order is given by

$$\varrho = \{(a, c), (a, d), (b, d), (b, f)\},$$

that is, $a \prec c, a \prec d, b \prec d$, and $b \prec f$. Then we can represent X as the union of chains in several ways, for instance, as

$$X = \{a, c\} \cup \{b, d\} \cup \{f\}$$

or as

$$X = \{a, d\} \cup \{b, f\} \cup \{c\}$$

or else as

$$X = \{a, c\} \cup \{b\} \cup \{d\} \cup \{f\}.$$

Theorem 5.2.9. (Dilworth) *Let $X = \{x_1, x_2, \dots, x_n\}$ be a finite poset. The minimum number of disjoint chains containing all elements of X is equal to the maximum number of pairwise noncomparable elements of X.*

Proof. Let m be the minimal number of chains covering the set X and M be the maximal number of pairwise noncomparable elements in X. Since a chain cannot contain two noncomparable elements, it is obvious that $m \geq M$. We use König's theorem, Theorem 5.2.7, to prove the opposite inequality. Consider a matrix $A = (a_{i,j})$, where $a_{i,j} = 1$ if and only if $x_i \prec x_j$ and $i \neq j$; otherwise $a_{i,j} = 0$. First we prove two lemmas. □

Lemma 5.2.1. *For any independent set F of* 1s *in the matrix A there exists a partition Δ of the n-element set X into disjoint chains such that*

$$|F| + |\Delta| = n.$$

Proof. Let

$$F = \{a_{i_1,i_2}, a_{i_3,i_4}, \ldots, a_{i_{2k-1},i_{2k}}\}$$

which means that $x_{i_1} \prec x_{i_2}, \ldots, x_{i_{2k-1}} \prec x_{i_{2k}}$. Thus, the elements $x_{i_1}, \ldots, x_{i_{2k}}$ can be grouped in chains containing two or more elements each. These chains are mutually disjoint due to the independence of F. If in addition to these chains, we consider all other elements of X as 1-element chains, we derive a partition of X into disjoint chains; call this partition Δ. Denote by l_j the number of elements in the jth chain. Since these chains contain all elements of X and are disjoint, we have

$$n = \sum_{j=1}^{|\Delta|} l_j = \sum_{j=1}^{|\Delta|} (l_j - 1) + |\Delta| = |F| + |\Delta|$$

because the $l_j - 1$ 1s in F correspond to l_j elements of X, which make the jth chain in Δ. □

Definition 5.2.5. A covering of a matrix is called *irreducible* if it fails to be a covering after removal of any line from it.

Lemma 5.2.2. *Let a zero-one matrix A correspond to a poset X, $|X| = n$, and T be an irreducible covering of* 1s *in A. Then there exists a subset $U \subset X$ such that*

$$|U| + |T| = n$$

and U consists of pairwise incomparable elements.

Proof. Let the covering T consist of rows $i_1, \ldots, i_k$ and columns $j_1, \ldots, j_m$. First we prove that all these indices are different.

On the contrary, if $i_1 = j_1$, then due to the irreducibility of the covering T there is an element $a_{r,j_1} = 1$ such that the rth row does not belong to T, and also there is an element $a_{i_1,s} = 1$ such that the sth column does not belong to T. The transitivity of a partial order and the equation $i_1 = j_1$ imply $x_r \prec x_s$. Suppose that $r = s$. Then we have $x_{i_1} \prec x_s = x_r \prec x_{j_1} = x_{i_1}$ and the antisymmetry of a partial order leads to the equation $x_{i_1} = x_s$, meaning that the element $a_{i_1,s} = 1$ is located on the principal diagonal, contrary to the definition of A. Thus, $r \neq s$ and $x_r \prec x_s$, implying $a_{r,s} = 1$.

We have noticed, however, that no line of the covering T can cover this unity. This contradiction shows that all indices in T are distinct.

We denote $U = X \setminus \{x_{i_1}, \dots, x_{i_k}, x_{j_1}, \dots, x_{j_m}\}$. Since T is a covering, the elements of U are pairwise incomparable and $n = |U| + |T|$. □

Completion of the proof of the Dilworth theorem. We have to establish the inequality $m \le M$. Let $\widehat{F}$ be a maximal set of independent 1s in A. By Lemma 5.2.1, there exists the corresponding partition $\widehat{\Delta}$; clearly, $m \le |\widehat{\Delta}|$. On the other hand, let $\widehat{T}$ be a minimal covering of 1s in A. By Lemma 5.2.2, there is a subset $\widehat{U} \subset X$ corresponding to $\widehat{T}$; clearly, $|\widehat{U}| \le M$. Theorem 5.2.7 implies the equation $|\widehat{F}| = |\widehat{T}|$. Thus, $|\widehat{\Delta}| = |\widehat{U}|$ and finally $|\widehat{U}| \le M \le m \le |\widehat{\Delta}| = |\widehat{U}|$, that is, $m = M$. □

Now we establish the equivalence of the three major results of this section.

Theorem 5.2.10. *The theorems of Hall, König, and Dilworth are equivalent.*

Proof. We only have to deduce Hall's Theorem 5.2.1 from Dilworth's Theorem 5.2.9. Given a set $S = \{x_1, \dots, x_m\}$, let us consider a family $U = (S_1, \dots, S_n)$ of its subsets. Hall's theorem asserts that **(H)** is a necessary and sufficient condition for the existence of a SDR. Since the necessity is immediate—see the proof of Theorem 5.2.1, we assume that the family U satisfies the condition **(H)** and shall prove that U has a SDR.

Introduce a poset $X = \{R_1, \dots, R_m, C_1, \dots, C_n\}$, where the partial ordering is defined by the following three conditions:

(1) $R_i \prec C_j$ if and only if $x_i \in S_j$,
(2) $C_j \prec C_j, \forall j, 1 \le j \le n$,
(3) $R_i \prec R_i, \forall i, 1 \le i \le m$.

If all elements of a subset $\{R_{i_1}, \dots, R_{i_l}, C_{j_1}, \dots, C_{j_k}\}$ are pairwise incomparable, this signifies that no element among $x_{i_1}, \dots, x_{i_l}$ belongs to any of the subsets $S_{j_1}, \dots, S_{j_k}$. Hence, the condition **(H)** yields the inequality $k + l \le m$. Therefore, no subset in X, consisting of pairwise incomparable elements, can have the cardinality greater than m. This can be rephrased as follows: The maximum number of pairwise incomparable elements in X does not exceed m. Moreover, there is a subset in X, namely $\{R_1, \dots, R_m\}$, consisting of exactly m pairwise incomparable elements, thus this maximal number is m. By Theorem 5.2.9, the set X can be decomposed in m disjoint chains. There may be three kinds of chains: 2-element chains $\{R_{i_p}, C_{j_q}\}$, 1-element chains $\{R_{i_p}\}$, and 1-element chains $\{C_{j_q}\}$. After some renumbering, this decomposition can be written as

$$\{R_1, C_1\}, \dots, \{R_t, C_t\}, \{R_{t+1}\}, \dots, \{R_m\}, \{C_{t+1}\}, \dots, \{C_n\}.$$

Since any chain contains at most one element R_i, every R_i must belong to a chain, and there are exactly as many chains as there are elements $R_i, 1 \le i \le m$. Thus, no 1-element chain $\{C_j\}$ can exist and we must have $t = n \le m$. Therefore, the above chain decomposition actually is

$$\{R_1, C_1\}, \dots, \{R_t, C_t\}, \{R_{t+1}\}, \dots, \{R_m\}.$$

This shows that $x_1 \in S_1, \dots, x_n \in S_n$, and we constructed a SDR $\{x_1, \dots, x_n\}$ for the family U. □

In the end of this section we apply Hall's theorem to study matchings in bipartite graphs.

Definition 5.2.6. Consider a bipartite graph $G = (V_1 \cup V_2, E)$. Any set of its edges is called a *matching* from V_1 to V_2. A matching $\mathcal{M}$ in a bipartite graph G is called complete if there is a one-to-one correspondence between V_1 and a subset of V_2 such that the corresponding vertices are connected by the $\mathcal{M}$-edges. A matching $\mathcal{M}$ in a bipartite graph G is called maximal if no other matching in G contains more edges than $\mathcal{M}$.

A matching can be described as a (nonsymmetric) binary relation between the sets V_1 and V_2. If $X \subset V$ is a set of vertices, we denote by $\Gamma(X)$ the set of all vertices adjacent with some vertex in X. We again immediately observe a necessary condition of the existence of a complete matching in a bipartite graph $G = (V_1 \cup V_2, E)$, namely, the inequality $|\Gamma(X)| \geq |X|$ for any subset $X \subset V_1$. Similarly to other results in this section, the following theorem asserts that this natural necessary condition is also sufficient.

Theorem 5.2.11. *A complete matching in a bipartite graph* $G = (V_1 \cup V_2, E)$ *exists if and only if* $|\Gamma(X)| \geq |X|$ *for every subset* $X \subset V_1$.

Proof. We prove that the statement is equivalent to Hall's theorem, Theorem 5.2.1. Let S be a finite set. To any family of sets $U = (S_1, \dots, S_n), S_j \subset S, 1 \leq j \leq n$, there corresponds its bipartite *incidence graph* $G = (V_1 \cup V_2, E)$, where V_1 and V_2 are arbitrary sets with $|V_1| = n, |V_2| = |S|$, and a point in V_1 is connected with a point in V_2 if and only if the corresponding subset in U contains the corresponding element in S. Vice versa, to any bipartite graph we can quite similarly put in a correspondence a family of sets U. It is obvious that a complete matching in G exists if and only if the family U has a SDR. □

Exercises 5.2.

Exercise 5.2.1.
(1) A clothing store has suits of two designs and two colors. Is it possible to choose two suits for display representing both designs and both colors?
(2) If there are suits of three designs and three colors, can the display show all the three designs and three colors using only two suits?

Exercise 5.2.2. Prove Theorem 5.2.6 for non-integer utilities $t_{i,j}$.

Exercise 5.2.3. Revisit Problem 5.2.5 and find all other possible SDR and all other solutions of the assignment problem. Make sure that they return the same maximum value as in the solution above.

Exercise 5.2.4. Solve the assignment problem with the utility matrix

$$T = \begin{pmatrix} 3 & 2 & 2 & 5 & 4 \\ 5 & 6 & 2 & 6 & 1 \\ 3 & 1 & 2 & 4 & 6 \\ 2 & 3 & 3 & 5 & 4 \\ 4 & 2 & 1 & 1 & 3 \end{pmatrix}.$$

Exercise 5.2.5. In addition to the three chain decompositions presented in Example 5.2.2, find all other possible decompositions of the set X in the example.

Exercise 5.2.6. Consider all 3-element families of subsets of the set $X = \{a, b, c, d\}$ without repeating subsets. Which of them have SDR? Find them.

Exercise 5.2.7. Prove that if a family $U = (S_1, \ldots, S_n)$ of subsets of a finite set $S = \{x_1, \ldots, x_m\}$ satisfies Hall's condition **(H)**, then U has the unique SDR if and only if $|S_1 \cup \cdots \cup S_n| = n$.

Exercise 5.2.8. How many are there $n \times n$ zero-one matrices that have exactly one 1 in each line (row or column)?

Exercise 5.2.9. How many are there $m \times n$ zero-one matrices such that the sum of all its elements is k?

Exercise 5.2.10. Prove that if a $2n \times 2n$ zero-one matrix contains $3n$ 1s, then it is possible to find n rows and n columns, which cover all 1s in the matrix. However, there exists a zero-one $2n \times 2n$ matrix with $3n + 1$ 1s, such that no set of n rows and n columns covers all of the 1s.

Exercise 5.2.11. An *edge-cover* in a graph is a set S of vertices such that every edge is incident to a vertex in S. Prove that König's theorem can be stated as follows: *The maximum size of a matching in a bipartite graph is equal to the minimum size of an edge-cover in the graph.*

Exercise 5.2.12. The next problem represents another kind of problems on systems of (not necessarily distinct) representatives.

In Really Fraternal College there are 2 006 fraternities and sororities each of which comprises more than half of all college students. Many of the students belong to several sororities or fraternities. Prove that it is possible to find at most 10 students at the college that represent every sorority and fraternity, that is, for each sorority and fraternity there is someone among these 10 students who belongs to this sorority or fraternity.

Exercise 5.2.13. Let a square matrix A of order n contains a zero sub-matrix of order $p \times q$. Prove that if $p + q > n$, then both the determinant and the permanent of A are zero.

Exercise 5.2.14. Consider the complete graph K_{13}. Paint its edges in two colors, so that the graph does not have 3-subgraphs of only one color and does not have 5-subgraphs of another color.

5.3 Block designs

In this section we are concerned with methods of selecting subsets in a given set subject to various restrictions on the elements of these subsets. Such methods are important in scheduling, planning of experiments, and many other problems.

Coffee-time browsing

- http://en.wikipedia.org/wiki/Bruck%E2%80%93Chowla%E2%80%93Ryser_theorem (Bruck–Ryser–Chowla Theorem)
- http://en.wikipedia.org/wiki/Sarvadaman_Chowla (Chowla's biography)
- http://en.wikipedia.org/wiki/Herbert_John_Ryser (Ryser's biography)
- www.gap-system.org/~history/Biographies/Diophantus.html (Diophantus' biography)
- http://www-history.mcs.st-and.ac.uk/Biographies/Diophantus.html (Diophantus' biography)
- http://www.gap-system.org/~history/Biographies/Lagrange.html (Lagrange's biography)

Problem 5.3.1. Organizers of a football tournament invited nine teams and rented three stadiums. Each team must play any other exactly once. How should the organizers schedule the games to finish the tournament as soon as possible?

Solution. All in all, $C(9,2) = 36$ games are necessary. Therefore, each stadium should host $36 \div 3 = 12$ games, because if one field hosts less than 12 games, then another field must do more than 12, which would make the tournament longer. The organizers can split all nine teams into groups, called hereafter *blocks*, of three, assign each block to a stadium and schedule mini-series within each group. Any such mini-tournament consists of $C(3,2) = 3$ games. When the first series of the three simultaneous mini-tournaments within blocks is over, the organizers reshuffle all teams in new blocks of three, making sure that no pair of the teams meets again, and repeat this procedure until each team plays all others. Since it is necessary to have 12 games at each field and any mini-series consists of three games, we expect at least $12 \div 3 = 4$ shuffles of three blocks with each block consisting of three teams.

It is not at all clear that this procedure works, and to finish the solution, we have to present all blocks explicitly. Denoting the participating teams by $t_1, \ldots, t_9$, we arrange them in blocks as follows. The first shuffle is

$$B_1 = \{t_1, t_2, t_3\}, \quad B_2 = \{t_4, t_5, t_6\}, \quad B_3 = \{t_7, t_8, t_9\}.$$

The second shuffle is

$$B_4 = \{t_1, t_4, t_7\}, \quad B_5 = \{t_2, t_5, t_8\}, \quad B_6 = \{t_3, t_6, t_9\}.$$

The third shuffle is

$$B_7 = \{t_1, t_5, t_9\}, \quad B_8 = \{t_3, t_4, t_8\}, \quad B_9 = \{t_2, t_6, t_7\}.$$

And the last, fourth shuffle is

$$B_{10} = \{t_3, t_5, t_7\}, \quad B_{11} = \{t_1, t_6, t_8\}, \quad B_{12} = \{t_2, t_4, t_9\}.$$

We observe that, for each pair $\{t_i, t_j\}, 1 \leq i, j \leq 9, i \neq j$, the teams t_i and t_j play each other exactly once. □

Solving this problem, we selected 12 subsets-blocks of the given set. In the example, these 12 blocks, $B_1, \dots, B_{12}$, satisfy the following obvious conditions:
- Each block consists of three elements.
- Each element of the given set appears in exactly four blocks.
- Each pair of elements meets precisely in one block.

Such configurations are called (combinatorial) *block designs*. They are useful in many problems, like scheduling, experiment planning and many others. Formalizing the conditions above, we arrive at the following definition.

Definition 5.3.1. Let $X = \{x_1, \dots, x_v\}$ be a finite set, whose subsets are hereafter called blocks. A family (in the sense of Definition 5.1.1) of blocks $B = (B_1, \dots, B_b)$ is called a *balanced incomplete block design* (BIBD) with parameters (v, b, k, r, λ), denoted by $S(v, b, k, r, \lambda)$, if
- Each block contains k elements, $|B_1| = \cdots = |B_b| = k$.
- Each element of the set X belongs to exactly r blocks.
- Each pair of elements of X appears in precisely λ blocks.

These configurations are called *balanced*, because of the uniformity of the preceding conditions. They are called *incomplete*, because a block design does not necessarily contain all the 2^v_k k-element subsets of X. It is worth recalling that by Definition 5.1.1 of a family of subsets, some or even all blocks B_i can coincide as sets. The solution of Problem 5.3.1 gives an example of the BIBD $S(9, 12, 3, 4, 1)$. Here are two more examples of BIBD:

BIBD $S(7, 7, 3, 3, 1)$:

$$B_1 = \{1, 3, 7\}, \quad B_2 = \{1, 2, 4\}, \quad B_3 = \{2, 3, 5\}, \quad B_4 = \{3, 4, 6\},$$
$$B_5 = \{4, 5, 7\}, \quad B_6 = \{1, 5, 6\}, \quad B_7 = \{2, 6, 7\},$$

BIBD $S(13, 26, 3, 6, 1)$:

$$B_1 = \{1,2,5\}, \quad B_2 = \{2,3,6\}, \quad B_3 = \{3,4,7\}, \quad B_4 = \{4,5,8\},$$
$$B_5 = \{5,6,9\}, \quad B_6 = \{6,7,10\}, \quad B_7 = \{7,8,11\}, \quad B_8 = \{8,9,12\},$$
$$B_9 = \{9,10,13\}, \quad B_{10} = \{1,10,11\}, \quad B_{11} = \{2,11,12\}, \quad B_{12} = \{3,12,13\},$$
$$B_{13} = \{1,4,13\}, \quad B_{14} = \{1,3,8\}, \quad B_{15} = \{2,4,9\}, \quad B_{16} = \{3,5,10\},$$
$$B_{17} = \{4,6,11\}, \quad B_{18} = \{5,7,12\}, \quad B_{19} = \{6,8,13\}, \quad B_{20} = \{1,7,9\},$$
$$B_{21} = \{2,8,10\}, \quad B_{22} = \{3,9,11\}, \quad B_{23} = \{4,10,12\}, \quad B_{24} = \{5,11,13\},$$
$$B_{25} = \{1,6,12\}, \quad B_{26} = \{2,7,13\}.$$

As we will see, the existence of a BIBD $S(v, b, k, r, \lambda)$ imposes certain restrictions on the parameters v, b, k, r, λ. To this end it is convenient to introduce the *incidence matrix* of a block design.

Definition 5.3.2. Consider a BIBD $S(v, b, k, r, \lambda)$ built on a v-element set X and consisting of b blocks. The *incidence matrix of* $S(v, b, k, r, \lambda)$ is a zero-one $b \times v$ matrix $M = (m_{i,j})$ with the entries

$$m_{i,j} = \begin{cases} 1 & \text{if } x_i \in B_j, \\ 0 & \text{if } x_i \notin B_j. \end{cases} \tag{5.3.1}$$

To find necessary conditions which the parameters of a BIBD $S(v, b, k, r, \lambda)$ must satisfy, we calculate the number of 1s in M in two different ways. On the one hand, each element belongs to r blocks, hence, each row in the matrix contains r 1s, thus, the matrix contains in total $r{\cdot}v$ 1s. On the other hand, each block consists of k elements, therefore, each one of b columns, representing b blocks, contains k 1s, totaling to $b \cdot k$. Thus, we get a necessary condition for a BIBD $S(v, b, k, r, \lambda)$ to exist,

$$b \cdot k = r \cdot v. \tag{5.3.2}$$

To derive another necessary condition of the existence of a BIBD $S(v, b, k, r, \lambda)$, we choose an element, say x_1, and compute how many times all the ordered pairs $(x_1, x_i), 1 \neq i$, appear in all blocks. The element x_1 appears in r blocks and in each of the blocks it makes up pairs with $k - 1$ other elements, altogether generating $r(k - 1)$ such pairs. On the other hand, since $|X| = v$, there are $v - 1$ different pairs (x_1, x_i) with $i \neq 1$ and each of them appears λ times, adding up to $\lambda(v - 1)$. Thus we get another necessary condition,

$$r(k - 1) = \lambda(v - 1). \tag{5.3.3}$$

Conditions (5.3.2) and (5.3.3) are necessary but as we will see, are not sufficient for the existence of BIBD $S(v, b, k, r, \lambda)$.

To study BIBD, we need a few simple properties of their incidence matrices. Let M be the incidence matrix of the BIBD $S(v,b,k,r,\lambda)$. Introduce a $v \times v$ matrix $N = M \cdot M^T$, where M^T is the transpose of M, that is, the matrix M flipped about its main diagonal. We compute the matrix N in the next lemma. Hereafter, subscripts indicate the dimensions of a matrix or a vector.

Lemma 5.3.1.

$$N = \begin{pmatrix} r & \lambda & \dots & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ \dots & \dots & \dots & \dots & \dots \\ \lambda & \dots & \dots & \lambda & r \end{pmatrix} = (r-\lambda)I_v + \lambda J_v. \tag{5.3.4}$$

Moreover, $w_v M = k w_b$, which is equivalent to

$$JM = kJ \tag{5.3.5}$$

where I is the unit matrix, that is, all of its diagonal elements are 1s and all off-diagonal elements are 0s, J is a $v \times v$ matrix and w is a vector all of whose elements are 1s.

Proof. An element $n_{i,j}$ of the matrix N is the dot product of the ith and jth rows of the incidence matrix M. Therefore, $n_{i,i}$ is equal to the number of 1s in the ith row of M, which is r. If $i \neq j$, then $n_{i,j} = m_{i,1}m_{j,1} + \dots + m_{i,b}m_{j,b}$, and the addend $m_{i,s}m_{j,s} = 1$ if and only if $m_{i,s} = m_{j,s} = 1$, which means that $a_i \in B_s$ and $a_j \in B_s$. However, each pair a_i, a_j meets in λ blocks, thus each pair contributes λ unities to $n_{i,j}$. This proves (5.3.4).

To prove the second statement of the lemma, it suffices to notice that each column of M contains exactly k 1s, which is expressed by (5.3.5). □

The converse of Lemma 5.3.1 is also true—see Exercise 5.3.5.

We will also use equation (5.3.4) rewritten in other terms. Introduce linear forms

$$L_j(x_1, \dots, x_v) = \sum_{i=1}^{v} m_{i,j} x_i, \quad 1 \le j \le b, \tag{5.3.6}$$

where $m_{i,j}$ are elements of the matrix M. Then (5.3.4) can be written as

$$L_1^2 + \dots + L_b^2 = (r-\lambda)(x_1^2 + \dots + x_v^2) + \lambda(x_1 + \dots + x_v)^2. \tag{5.3.7}$$

To compute the determinant of the matrix $N = M \cdot M^T$, we subtract its first column from all the subsequent columns and then add the 2nd, 3rd, $\dots$, vth rows to the first one. The resulting matrix is triangular, hence its determinant is the product of the diagonal elements,

$$\det(N) = (r-\lambda)^{v-1}(v\lambda - \lambda + r). \tag{5.3.8}$$

If $r = \lambda$, then $\lambda(k-1) = \lambda(v-1)$ by (5.3.3), thus, $v = k$ and the block design is trivial in the sense that it consists of several identical repeating blocks—copies of the basic set X. The strict inequality $r < \lambda$ is impossible, for it would imply $r(k-1) < \lambda(k-1)$, hence $\lambda(v-1) < \lambda(k-1)$ by (5.3.3), therefore, $v < k$. The latter would mean that a block contains more elements than the entire basic set.

Thus, hereafter we assume that $r > \lambda$. Hence

$$v\lambda - \lambda + r = v\lambda + (r - \lambda) > 0$$

and (5.3.8) implies the inequality $\det(N) > 0$. We have proved that N is a non-singular matrix and since N is a $v \times v$ matrix, its rank is v.

The rank of a product of matrices cannot exceed the rank of any factor in the product. In addition, the rank of M cannot exceed the number of its columns, which is b. Hence, we deduce the *Fisher inequality*,

$$v \leq b \tag{5.3.9}$$

valid for any BIBD $S(v, b, k, r, \lambda)$. Moreover, (5.3.9) and (5.3.2) imply an inequality $k \leq r$ for any such BIBD.

Definition 5.3.3. A BIBD $S(v, b, k, r, \lambda)$ is called *symmetric* if $v = b$; in this case equation (5.3.2) implies also that for symmetric BIBD $k = r$. Therefore, the symmetric block designs have only three independent parameters and will be denoted by $S(v, k, \lambda)$.

For example, the BIBD $S(7, 7, 3, 3, 1) = S(7, 3, 1)$ presented above is symmetric. Symmetric block designs are dealt with in the following statement.

Theorem 5.3.1. *The incidence matrix of a symmetric block design $S(v, k, \lambda)$ satisfies the relations*

$$MM^T = (k - \lambda)I + \lambda J = M^T M \tag{5.3.10}$$

and

$$JM = kJ = MJ. \tag{5.3.11}$$

Proof. Only the right equations require proofs, since the left ones are, respectively, (5.3.4) and (5.3.5) rewritten for the symmetric case $v = b$ and $k = r$. The right equation in (5.3.11) follows immediately from symmetry, since it tells that every row in M contains $k = r$ 1s, that is, each element belongs to $r = k$ blocks. Thus, we have to prove only the right equation in (5.3.10).

To prove it, we multiply (5.3.4) on the left by the inverse matrix M^{-1}, which exists due to the non-singularity of $N = MM^T$, and get the equation

$$M^T = (k - \lambda)M^{-1} + \lambda M^{-1} J. \tag{5.3.12}$$

Similarly, the equation $kJ = MJ$ implies $kM^{-1}J = J$, which together with (5.3.12) gives $M^T = (k - \lambda)M^{-1} + (\lambda/k)J$. Multiplying the latter on the left by M and using (5.3.11) we complete the proof. □

Remark 5.3.1. This theorem is a particular case of a more general theorem by Ryser [24, p. 130].

Studying the combinatorial block designs we are concerned with two problems.
- Does there exist a design $S(v, b, k, r, \lambda)$ with the given parameters v, b, k, r, λ?
- If a design $S(v, b, k, r, \lambda)$ does exist, how may one construct it explicitly?

The following theorem gives necessary conditions for the existence of symmetric block designs.

Theorem 5.3.2. (Bruck–Ryser–Chowla) *Let there exist a symmetric BIBD $S(v, k, \lambda)$.*
(1) *If the number of elements v is even, then the difference $\alpha = k - \lambda$ is a perfect square.*
(2) *If v is odd, then the Diophantine equation*

$$z^2 = (k - \lambda)x^2 + (-1)^{\frac{v-1}{2}} \lambda y^2$$

has a non-trivial solution in integer numbers x, y, z. The triple $x = y = z = 0$ obviously satisfies this equation; non-trivial *means that at least one of the numbers x, y, z is non-zero, in other words $x^2 + y^2 + z^2 > 0$.*

Proof. (1) It follows from (5.3.3) that in the symmetric case $k(k - 1) = \lambda(v - 1)$, hence $v\lambda - \lambda + k = k^2$. Therefore, we deduce from (5.3.8) the equation

$$(\det(M))^2 = \det(N) = (k - \lambda)^{v-1}(v\lambda - \lambda + k) = (k - \lambda)^{v-1}k^2.$$

The latter implies that $(k - \lambda)^{v-1}$ must be a square, which is impossible for an odd number $v - 1$ unless the base $k - \lambda$ is a square.

(2) To prove the theorem in the case of odd v, we need the following lemma. We leave it to the reader to verify this claim by direct calculation. □

Lemma 5.3.2. *If*

$$\begin{cases} y_1 = b_1x_1 - b_2x_2 - b_3x_3 - b_4x_4, \\ y_2 = b_2x_1 + b_1x_2 - b_4x_3 + b_3x_4, \\ y_3 = b_3x_1 + b_4x_2 + b_1x_3 - b_2x_4, \\ y_4 = b_4x_1 - b_3x_2 + b_2x_3 + b_1x_4, \end{cases} \tag{5.3.13}$$

then

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 = \alpha(x_1^2 + x_2^2 + x_3^2 + x_4^2) \tag{5.3.14}$$

where $\alpha = b_1^2 + b_2^2 + b_3^2 + b_4^2$.

Let us notice that the determinant of system (5.3.13) is equal to α^2. Thus, if b_1, b_2, b_3, b_4 are integers, then the solutions $x_1 - x_4$ of (5.3.13) can be expressed as linear forms with rational coefficients through $y_1 - y_4$, and the common denominator of all these coefficients is α^2. Moreover, in the symmetric case equation (5.3.7) becomes

$$L_1^2 + \cdots + L_v^2 = \alpha(x_1^2 + \cdots + x_v^2) + \lambda(x_1 + \cdots + x_v)^2. \tag{5.3.15}$$

We will use the following theorem (see, for example, [27, p. 302]).

Lagrange's Theorem on Four Squares. *If zero addends are allowed, then every natural number can be written as a sum of four squares.*

For example, $9 = 3^2 + 0^2 + 0^2 + 0^2$, $10 = 3^2 + 1^2 + 0^2 + 0^2$, $11 = 3^2 + 1^2 + 1^2 + 0^2$, $12 = 3^2 + 1^2 + 1^2 + 1^2$.

Proof of Theorem 5.3.2 *when v is odd.* Let $v \equiv 1 \pmod 4$, that is, $v - 1$ is a multiple of 4. Applying Lagrange's theorem to number $\alpha = k - \lambda$, we can write

$$\alpha = b_1^2 + b_2^2 + b_3^2 + b_4^2. \tag{5.3.16}$$

Next, we split the variables $x_1, \ldots, x_{v-1}$ into quadruples

$$(x_1, x_2, x_3, x_4), \ldots, (x_{v-4}, x_{v-3}, x_{v-2}, x_{v-1}).$$

Considering (5.3.16), we apply formula (5.3.14) to each quadruple,

$$\alpha(x_i^2 + x_{i+1}^2 + x_{i+2}^2 + x_{i+3}^2) = y_i^2 + y_{i+1}^2 + y_{i+2}^2 + y_{i+3}^2;$$

thus (5.3.15) becomes

$$L_1^2 + \cdots + L_v^2 = y_1^2 + \cdots + y_{v-1}^2 + \alpha x_v^2 + \lambda(x_1 + \cdots + x_v)^2. \tag{5.3.17}$$

The rest of the proof is based on the observation that (5.3.17) is an identity with rational coefficients in indeterminates $x_1, \ldots, x_v$, or which is equivalent, in $y_1, \ldots, y_v$, and we use these indeterminates to derive the Diophantine equation we sought after.

First, we set $y_v = x_v$ and eliminate all $x_i, 1 \le i \le v$, from (5.3.17) by considering system (5.3.13) for each quadruple $x_i, x_{i+1}, x_{i+2}, x_{i+3}$ with the same coefficients b_1, b_2, b_3, b_4, and solving all these systems for $x_i, 1 \le i \le v$. After this elimination (5.3.17) becomes

$$L_1^2 + \cdots + L_v^2 = y_1^2 + \cdots + y_{v-1}^2 + \alpha y_v^2 + \lambda w^2 \tag{5.3.18}$$

where $L_1, \ldots, L_v$ and $w = x_1 + \cdots + x_v$ are linear forms of indeterminates $y_1, \ldots, y_v$ with rational coefficients.

Let $L_1 = c_1 y_1 + \cdots + c_v y_v$. If $c_1 \ne 1$, then the equation $L_1 = y_1$ allows us to express y_1 through $y_2, \ldots, y_v$ as a linear form with rational coefficients. Otherwise, that is, if $c_1 = 1$, we consider the equation $L_1 = -y_1$. In either case $L_1^2 = y_1^2$, thus (5.3.18) becomes

$$L_2^2 + \cdots + L_v^2 = y_2^2 + \cdots + y_{v-1}^2 + \alpha y_v^2 + \lambda w^2.$$

The latter identity depends only on $y_2, \dots, y_v$. Continuing in the same fashion and eliminating, one by one, the indeterminates $y_2, \dots, y_{v-1}$, we end up with the equation

$$L_v^2 = \alpha y_v^2 + \lambda w^2$$

where L_v and w are rational multiples of the last remaining variable y_v. Setting here y_v to be equal to an integer multiple $x \neq 0$ of the denominators of L_v and w, we derive a relation between three integer numbers $x \neq 0, y, z$:

$$z^2 = \alpha x^2 + y^2. \tag{5.3.19}$$

Therefore, we have proved the theorem in the case $v \equiv 1 \pmod 4$. The case $v \equiv 3 \pmod 4$ is treated similarly, but to apply Lagrange's theorem in this case, we have to introduce a new variable x_{v+1} and add the term αx_{v+1}^2 to both sides of (5.3.15). Similar calculations lead to equations $\alpha x^2 = y_{v+1}^2 + \lambda w^2$ and

$$z^2 = \alpha x^2 - \lambda y^2. \tag{5.3.20}$$

Diophantine equations (5.3.19) and (5.3.20) together complete the proof of Theorem 5.3.2. □

Theorem 5.3.2 implies, in particular, that the necessary conditions (5.3.2)–(5.3.3), which in the symmetric case reduce to one equation

$$k(k-1) = \lambda(v-1)$$

are not sufficient. For example, it is readily verified that the values $v = 43, k = 7, \lambda = 1$ satisfy the latter equation. However, Theorem 5.3.2 gives in this case the equation $z^2 = 6x^2 - y^2$, which has no non-trivial solution.

Exercises 5.3.

Exercise 5.3.1. A teacher arranges her 4 first-graders in a 2×2 square. For how many days can she make these arrangements so that every child has a new neighbor in her row?

Exercise 5.3.2. Solve the previous problem if 40 kids must be arranged in a 10×4 rectangle.

Exercise 5.3.3. There are 20 students in a class. In the classroom, there are 10 desks with two seats each. On the first day of each week a teacher rearranges the students, so that any two students seat at the same desk if and only if they have never seated together before. For how many weeks can the teacher do that?

Exercise 5.3.4. Arrange several pennies, nickels, dimes, quarters, and half-dollars in a 4×4 square, so that each row, each column, and each of two diagonals consist of different coins and the total sum of all 16 coins is the largest.

Exercise 5.3.5. Prove the converse of Lemma 5.3.1, that is, prove that, given a zero-one matrix M, whose elements satisfy (5.3.4)–(5.3.5), there exists a BIBD $S(v, b, k, r, \lambda)$ with the incidence matrix M.

Exercise 5.3.6. Deduce equation (5.3.7) from (5.3.4).

Exercise 5.3.7. Restore details of the calculation of the determinant leading to equation (5.3.8).

Exercise 5.3.8. Compute the determinant of system (5.3.13).

Exercise 5.3.9. Prove Lemma 5.3.2.

Exercise 5.3.10. Derive in detail equation (5.3.20).

Exercise 5.3.11. Prove that the equation $z^2 = 6x^2 - y^2$ has no non-trivial solution.

Exercise 5.3.12. Do there exist BIBD $S(43, 43, 7, 7, 1)$ and $S(15, 21, 5, 7, 2)$?

5.4 Systems of triples

In the last section we study systems of triples, that is, the block designs with 3-element blocks. In particular, we find for what values of v the systems of triples exist.

Coffee-time browsing

- http://en.wikipedia.org/wiki/Kirkman%27s_schoolgirl_problem (Kirkman's schoolgirl problem)
- www-history.mcs.st-andrews.ac.uk/Biographies/Kirkman.html (Kirkman's biography)
- http://www-history.mcs.st-and.ac.uk/Mathematicians/Steiner.html (Steiner's biography)
- http://en.wikipedia.org/wiki/E._H._Moore (E. H. Moore)
- http://images.google.com/images?q=Fano+plane&rls=com.microsoft:en-us:IE-SearchBox&oe=UTF-8&sourceid=ie7&rlz=1I7DKUS&um=1&ie=UTF-8&ei=j90jS8SGO8rUlAfBzMWLCg&sa=X&oi=image_result_group&ct=title&resnum=4&ved=0CCIQsAQwAw (Many Fano Planes)

If $k = 3$, equations (5.3.2) and (5.3.3) read

$$3b = rv, \quad 2r = \lambda(v - 1),$$

leading to

$$r = \frac{1}{2}\lambda(v - 1), \quad b = \frac{1}{6}\lambda v(v - 1). \tag{5.4.1}$$

Since r and b are integer numbers, equations (5.4.1) give the following necessary conditions for a system of triples to exist.

Proposition 5.4.1. *If a system of triples $S(v, b, 3, r, \lambda)$ exists, then the product $\lambda(v-1)$ is even and the product $\lambda v(v-1)$ is divisible by 6, that is,*

$$\lambda(v-1) \equiv 0 \pmod 2, \quad \lambda v(v-1) \equiv 0 \pmod 6 \tag{5.4.2}$$

It turns out that these necessary conditions are also sufficient for the existence of systems of triples. Moreover, similar conditions, which follow from (5.3.2)–(5.3.3), are also necessary and sufficient for the existence of block designs with $k = 4$ and any λ, but are not sufficient if $k \geq 5$ [24, Chap. 15]. It is, however, known [59] that for any given k and λ there exists a number v_0 such that for all $v \geq v_0$ conditions (5.3.2)–(5.3.3) are not only necessary but also sufficient for the existence of a block design $S(v, b, k, r, \lambda)$.

We consider in more detail the systems of triples with $\lambda = 1$. They are called *Steiner triple systems*. If $k = 3$ and $\lambda = 1$, then for a given v two other parameters, b and r, are uniquely determined from (5.4.1), thus, we denote the Steiner triple systems by $S(v)$ and call v the *order* of the system.

When $\lambda = 1$, (5.4.2) implies $v - 1 \equiv 0 \pmod 2$ and $v(v-1) \equiv 0 \pmod 6$. Therefore, v is odd and $v(v-1) = 6t$, where t is an integer; thus, there are only three possible cases,

$$v = 6t + 1, \quad v = 6t + 3, \quad v = 6t + 5.$$

However, if $v = 6t + 5$, then $v(v-1) = (6t+5)(6t+4) = 6(6t^2 + 9t) + 20$, which is not divisible by 6, hence for $v \equiv 5 \pmod 6$ systems $S(v)$ do not exist, and we have only two possibilities left, $v \equiv 1 \pmod 6$ and $v \equiv 3 \pmod 6$. Such values of v are called *admissible*. It turns out that for each admissible v Steiner triple systems $S(v)$ do exist. The proof below follows Hilton [30] and is recursive. First we prove two theorems of Moore, which give algorithms for constructing a system $S(v)$ from given systems with smaller values of the parameter v, and then we prove that each admissible value $v = 6t + 1$ or $v = 6t + 3$ can be expressed through smaller admissible values of v, such that those algorithms can be applied.

Definition 5.4.1. Let two block designs, S' and S'', be built on the sets X' and X'', respectively, and $\mathcal{B}', \mathcal{B}''$ stand for the families of their blocks. The designs S' and S'' are called *isomorphic* if there exist two one-to-one correspondences

$$\varphi : X' \to X''$$

and

$$\psi : \mathcal{B}' \to \mathcal{B}''$$

compatible with the incidence relations in these designs. The latter means that, for any blocks $B_1 \in \mathcal{B}'$ and $B_2 \in \mathcal{B}''$, and for any elements $x_1 \in B_1$ and $x_2 \in B_2$, the equality

$x_2 = \varphi(x_1)$ holds if and only if $B_2 = \psi(B_1)$. If $S' = S'' = S$, then an isomorphism is called an automorphism of S.

Since superposition of mappings is associative, it is (almost) obvious that all the automorphisms of a block design make a multiplicative group with respect to the superposition.

Problem 5.4.1. Prove this statement. □

Theorem 5.4.1. (Moore) *If there are Steiner triple systems $S(v_1)$ and $S(v_2)$, then there exists also a Steiner triple system $S(v)$ with $v = v_1 \cdot v_2$, containing a subsystem isomorphic to $S(v_1)$ and a subsystem isomorphic to $S(v_2)$.*

Proof. Consider any two sets $X = \{x_1, \dots, x_{v_1}\}$ and $Y = \{y_1, \dots, y_{v_2}\}$, such that $|X| = v_1$ and $|Y| = v_2$. To simplify notation, we consider, without loss of generality, the sets $X = \mathbf{N}_{v_1} = \{1, \dots, v_1\}$ and $Y = \mathbf{N}_{v_2} = \{1, \dots, v_2\}$. By the assumption, there exist Steiner triple systems $S(v_1)$ and $S(v_2)$ on these sets, respectively. We construct a system $S(v_1 \cdot v_2)$ on the Cartesian product $Z = X \times Y$, $|Z| = v_1 \cdot v_2$. The set Z consists of ordered pairs of natural numbers, and the following algorithm determines which triples of the elements of Z, that is, which triples of these ordered pairs make blocks in $S(v)$.

Let $z_{i,j} = (x_i, y_j) \in Z$. A triple $\{z_{i,r}, z_{j,s}, z_{k,t}\}$ is a block in $S(v)$ if and only if one of the following three mutually-exclusive conditions holds true.

(1) The triple $\{x_i, x_j, x_k\}$ is a block in $S(v_1)$ and $r = s = t$.
(2) The triple $\{y_r, y_s, y_t\}$ is a block in $S(v_2)$ and $i = j = k$.
(3) The triple $\{x_i, x_j, x_k\}$ is a block in $S(v_1)$ and the triple $\{y_r, y_s, y_t\}$ is a block in $S(v_2)$.

We have to verify that this system of triples satisfies the definition of Steiner triple system $S(v)$ with $v = v_1 \cdot v_2$. First we check that each pair of elements meets in exactly one block. Let $\{z_{i,r}, z_{j,s}\}$ be an arbitrary pair in Z. If $i = j$, then the pair $\{y_r, y_s\}$ meets in the unique block $\{y_r, y_s, y_u\}$ of $S(v_2)$, since $S(v_2)$ is a Steiner triple system. Thus, the pair $\{z_{i,r}, z_{j,s}\}$ uniquely determines the block $\{z_{i,r}, z_{i,s}, z_{i,u}\}$. Moreover, the pair $\{z_{i,r}, z_{j,s}\}$ does not appear in two other parts of the algorithm and cannot generate any more triples. The same argument works if $r = s$.

Next, if $i \neq j$ and $r \neq s$, then the pair $\{x_i, x_j\}$ uniquely determines the block $\{x_i, x_j, x_k\}$ in $S(v_1)$, that is, we found the index k. Similarly, the pair $\{y_r, y_s\}$ uniquely determines the block $\{y_r, y_s, y_u\}$ in $S(v_2)$, thus, we found the index u. If a pair $\{z_{i,r}, z_{j,s}\}$ were to meet in two blocks $\{z_{i,r}, z_{j,s}, z_{k,u}\}$ and $\{z_{i,r}, z_{j,s}, z_{l,t}\}$, this would mean that the system $S(v_1)$ contained two different blocks $\{x_i, x_j, x_p\}$ and $\{x_i, x_j, x_q\}$ with $p \neq q$, contrary to the definition.

We still have to verify that any element $z_{i,r}$ of Z belongs to $r = \frac{v-1}{2}$ blocks. The element $z_{i,r}$ enters $r_2 = \frac{v_2-1}{2}$ blocks together with x_i, and it appears in $r_1 = \frac{v_1-1}{2}$ blocks together with y_r. We compute now how many blocks in $S(v)$ contain $z_{i,r}$ and do not contain x_i or y_r. We have just shown that two elements $z_{i,r}$ and $z_{j,s}$ determine the block uniquely, thus it suffices to calculate in how many ways it is possible to find a pair $z_{j,s}$ for a given element $z_{i,r}$ with $j \neq i$ and $s \neq r$. To this end we compute the total number of elements $v_1 \cdot v_2$ in Z less the number of elements containing x_i or y_r save the $z_{i,r}$ itself, and take a half of that amount, since the order of elements in a pair does not matter. This calculation yields $\frac{1}{2}(v_1v_2 - v_1 - v_2 + 1)$. Thus, the total number of blocks containing $z_{i,r}$ is

$$r = \frac{v_1 - 1}{2} + \frac{v_2 - 1}{2} + \frac{1}{2}(v_1v_2 - v_1 - v_2 + 1) = \frac{1}{2}(v_1v_2 - 1) = \frac{1}{2}(v - 1).$$

We have proved that the algorithm returns the Steiner triple system we sought. The triples with $r = s = t = 1$ form a subsystem isomorphic to $S(v_1)$, and the triples with $i = j = k = 1$ form a subsystem isomorphic to $S(v_2)$. □

Theorem 5.4.2 (Moore). *Given three natural numbers v_1, v_2, v_3. If there exist systems $S(v_1)$ and $S(v_2)$, and either $v_3 = 1$ or there exists a system $S(v_3)$ such that the system $S(v_2)$ contains a subsystem isomorphic to $S(v_3)$, then there exists a system $S(v)$ of order $v = v_3 + v_1(v_2 - v_3)$, containing a subsystem of order v_1, a subsystem of order v_3, and v_1 subsystems of order v_2.*

Proof. If $v_2 = v_3$, then $v = v_2$ and there is nothing to prove, for the given system $S(v_2)$ contains a subsystem isomorphic to $S(v_3)$. Thus, we assume $v_2 - v_3 \geq 1$, set $s = v_2 - v_3$, and use the union of the following $v_1 + 1$ sets:

$$\begin{aligned} X &= \{x_1, x_2, \ldots, x_{v_3}\}, \\ Y_1 &= \{y_{1,1}, y_{1,2}, \ldots, y_{1,s}\}, \\ &\ldots \\ Y_{v_1} &= \{y_{v_1,1}, y_{v_1,2}, \ldots, y_{v_1,s}\}, \end{aligned}$$

as the set of elements of the system $S(v)$ under construction. We describe now an algorithm generating the system $S(v)$. A triple of elements from the union $X \cup Y_1 \cup \cdots \cup Y_{v_1}$ is a block in $S(v)$ if and only if one of the following three mutually exclusive conditions holds true.

(1) A triple $\{x_i, x_j, x_k\}$ is a block in $S(v)$ if this triple is a block in a system $S(v_3)$ derived from the base set X. If $v_3 = 1$, then this case is vacuous.
(2) For each $i, 1 \leq i \leq v_1$, we construct a system $S(v_2)$ from the elements of the set $X \cup Y_i$; this system exists by the assumption and contains a system $S(v_3)$ built from X. The system $S(v)$ will include all blocks of $S(v_2)$ except for those that belong to $S(v_3)$ and are listed in step (1) of the algorithm. These blocks contain no more than one element $x_j \in X$ and are either $\{x_j, y_{i,k}, y_{i,l}\}$ or $\{y_{i,k}, y_{i,l}, y_{i,m}\}$.
(3) Finally, we construct a system $S(v_1)$ on the set of numbers $\{1, 2, \ldots, v_1\}$. If $\{i, j, k\}$ is a block of this system, we include in $S(v)$ all triples $\{y_{i,x}, y_{j,y}, y_{k,t}\}$, such that the second subscripts satisfy the congruence $x + y + t \equiv 0 \pmod{s}$.

To prove that this system is a Steiner triple system $S(v)$, the reader can carry over the argument similar to the proof of Theorem 5.4.1. □

Now we can prove a criterion of the existence of the Steiner triple systems.

Theorem 5.4.3. *A Steiner triple system $S(v), v \in \mathbf{N}$, exists if and only if $v \geq 3$ is admissible, that is, $v = 3$, or $v = 6t + 1$, or $v = 6t + 3$ with any $t \in \mathbf{N}$.*

Proof. The necessity of these conditions has been already proven. To prove their sufficiency, we follow the recurrent argument of A. Hilton [30] and start by constructing $S(v)$ with all admissible $v \leq 36$, that is, with $v = 3, 7, 9, 13, 15, 19, 21, 25, 27, 31, 33$. Indeed, $S(3)$ is a trivial system with one block, the systems $S(7)$, $S(9)$, and $S(13)$ were presented in Section 5.3, the existence of systems $S(15)$, $S(19)$, $S(21)$, $S(25)$, $S(27)$, $S(31)$, and $S(33)$ follows from Theorems 5.4.1–5.4.2 by virtue of the equations

$$\begin{aligned}
15 &= 1 + 7(3 - 1),\\
19 &= 1 + 9(3 - 1),\\
21 &= 7 \cdot 3,\\
25 &= 1 + 3(9 - 1),\\
27 &= 3 \cdot 9,\\
31 &= 1 + 15(3 - 1),\\
33 &= 3 + 3(13 - 3),
\end{aligned}$$

if one takes into consideration the demonstrated existence of $S(3)$, $S(7)$, $S(9)$, $S(13)$.

Next we present formulas, expressing all bigger admissible values of v, that is, $v = 6t + 1$ and $v = 6t + 3$ with $v > 36$ through smaller admissible values of v, hence one can straightforwardly apply Theorem 5.4.2. If an admissible $v \neq 36t + 13$, then we have the following 11 cases:

$$\begin{aligned}
v &= 36t + 1 = 1 + 3\big((12t + 1) - 1\big),\\
v &= 36t + 3 = 1 + (18t + 1)(3 - 1),\\
v &= 36t + 7 = 1 + (6t + 1)(7 - 1),\\
v &= 36t + 9 = 3 + (6t + 1)(9 - 3),\\
v &= 36t + 15 = 1 + (18t + 7)(3 - 1),\\
v &= 36t + 19 = 1 + (6t + 3)(7 - 1),\\
v &= 36t + 21 = 3 + (6t + 3)(9 - 3),\\
v &= 36t + 25 = 1 + 3\big((12t + 9) - 1\big),\\
v &= 36t + 27 = 1 + (18t + 13)(3 - 1),\\
v &= 36t + 31 = 1 + (18t + 15)(3 - 1),\\
v &= 36t + 33 = 3 + 3\big((12t + 13) - 3\big).
\end{aligned}$$

If $v = 36t+13$, a tempting simple approach would be to write $v = 7+(6t+1)(13-7)$; however, this does not work, since by Exercise 5.4.8 the system $S(13)$ cannot contain a subsystem of order 7, thus we cannot apply Theorem 5.4.2. Therefore, the case $v = 36t + 13$ must be split into several subcases. If t is even, say $t = 2k$, then

$$v = 36t + 13 = 1 + (6k + 1)(13 - 1)$$

and Theorem 5.4.2 is again applicable. Suppose now that t is odd and there exists $r \geq 1$ such that

$$t = 2^{2r-2} + 2^{2r-4} + \cdots + 2^2 + 2^0.$$

If here $r = 1$, then $v = 49 = 7 \cdot 7$ and Theorem 5.4.1 works. If $r > 1$ is odd, that is, $r = 2s + 3, s \geq 0$, then

$$v = 36t + 13 = 9 + \left(18(2^{4s} + \cdots + 2^0\right) + 1)(49 - 9),$$

while if r is even, that is, $r = 2s + 2, s \geq 0$, then

$$v = 36t + 13 = 3 + \left(18(2^{4s} + \cdots + 2^0\right) + 1)(13 - 3).$$

Therefore, in all these cases v can be expressed through smaller admissible values of v, so that Moore's theorems can be applied.

Finally we have to consider the case of an odd t with a representation

$$t = 2^{2r+\alpha_s} + \cdots + 2^{2r+\alpha_0} + 2^{2r-2} + 2^{2r-4} + \cdots + 2^2 + 1$$

where $r \geq 1$, $0 < \alpha_0 < \alpha_1 < \cdots < \alpha_s$; among the terms $2^{2r+\alpha_1}, \ldots, 2^{2r+\alpha_s}$ there also may be powers of 4, but $2^{2r+\alpha_0}$ is not such a power. Thus,

$$v = 36t + 13 = 1 + (3t + 1) \cdot 3 \cdot 2^2 = 1 + (3x + 1) \cdot 3 \cdot 2^{2r+2}$$

where $x = 2^{\alpha_0} + \cdots + 2^{\alpha_s}$. Since $\alpha_0 > 0$, x must be an even number, hence the remainder after dividing x by 4 is either 0 or 2. Therefore, if $t = 4n + 1$, then

$$v = 36t + 13 = 1 + (6n + 1)\left((3 \cdot 2^{2r+2} + 1\right) - 1),$$

and if $t = 4n + 3$, then

$$v = 36t + 13 = 1 + (18n + 15)\left((2^{2r+3} + 1\right) - 1).$$

To complete the proof, it only remains to notice that in the latter case

$$2^{2r+3} + 1 = 6 \cdot 4^r + 2(3 + 1)^r + 1 = 6l + 3.$$

Thus, any admissible v can be expressed through smaller admissible values of v and Theorems 5.4.1–5.4.2 apply. □

Problem 5.4.2. Construct a system $S(7)$ by making use of Moore's theorems.

Solution. If we divide the order $v = 7$ by 36, the remainder is 7, so that by the general algorithm of Theorem 5.4.3 we can use the representation

$$v = 36t + 7 = 1 + (6t + 1)(7 - 1)$$

with $t = 0$. However, it is more instructive here to write $7 = 1 + 3(3 - 1)$ and use the algorithm of Theorem 5.4.2 with $v_1 = v_2 = 3$ and $v_3 = 1$.

As in the proof of this theorem, we introduce the four sets

$$\begin{aligned} X &= \{x\}, \\ Y_1 &= \{y_{1,1}, y_{1,2}\}, \\ Y_2 &= \{y_{2,1}, y_{2,2}\}, \\ Y_3 &= \{y_{3,1}, y_{3,2}\}. \end{aligned}$$

The set X itself does not make a 3-element block; however, the unions $X \cup Y_1, X \cup Y_2$ and $X \cup Y_3$ generate three blocks $\{x, y_{1,1}, y_{1,2}\}, \{x, y_{2,1}, y_{2,2}\}$ and $\{x, y_{3,1}, y_{3,2}\}$. To work out the third part of the algorithm of Theorem 5.4.2, one must consider the set $\{1, 2, \dots, v_1\} = \{1, 2, 3\}$. This set generates the only 3-element set $\{1, 2, 3\}$. Thus, we must consider all possible triples $\{y_{1,k}, y_{2,l}, y_{3,m}\}$ and solve the congruence $k + l + m \equiv 0 \pmod 2$. This is equivalent to solving a Diophantine equation $k + l + m = 2s$ in integers k, l, m, where s is any integer number and $1 \le k, l, m \le 2$. Subject to the latter restrictions, the equation can be readily solved; it has four solutions, $(1, 1, 2), (1, 2, 1), (2, 1, 1), (2, 2, 2)$. These four triples generate four more blocks, in addition to the initial blocks, namely,

$$\{y_{1,1}, y_{2,1}, y_{3,2}\},\ \{y_{1,1}, y_{2,2}, y_{3,1}\},\ \{y_{1,2}, y_{2,1}, y_{3,1}\},\ \{y_{1,2}, y_{2,2}, y_{3,2}\}.$$

Setting here $x = 1, y_{1,1} = 2, y_{1,2} = 4, y_{2,1} = 3, y_{2,2} = 7, y_{3,1} = 6$, and $y_{3,2} = 5$, we arrive at the system $S(7, 7, 3, 3, 1)$ considered above. Other choices of the parameters give other, but isomorphic block designs. □

Problem 5.4.3. Construct Steiner system $S(631)$.

Solution. Since $v = 631 = 6t + 1$ with $t = 105$, the value $v = 631$ is admissible and system $S(631)$ exists. First we find what smaller admissible values are required by the algorithm. The algorithm depends upon the divisibility by 36; therefore, we divide $v = 631$ by 36, $631 = 17 \cdot 36 + 19$ with the remainder equal to 19, and use an appropriate representation from the chart, $v = 36t + 19 = 1 + (6t + 3)(7 - 1)$, with $t = 17$; thus, $v = 1 + 105 \cdot (7 - 1)$. We want to apply now Theorem 5.4.2 with $v = 631, v_3 = 1, v_2 = 7$, and $v_1 = 105$. The case $v_3 = 1$ is trivial, the system $S(7)$ was presented above; thus, to apply Theorem 5.4.2, we need a system $S(105)$. By the same token, $105 = 36 \cdot 2 + 33$ with the remainder 33, that is we have to use the representation $105 = 3 + 3(37 - 3)$. Keeping in mind the existence of the trivial system $S(3)$, to apply Theorem 5.4.2 we need to construct a system $S(37)$. However, $37 = 36 \cdot 1 + 1$ with the remainder 1, hence to get a system $S(37)$, we use the representation $37 = 1+3(13-1)$, where the components $S(3)$ and $S(13)$ have been already proven to exist.

To construct $S(631)$, we now work backward. First we use the systems $S(3)$ and $S(13)$ to construct $S(37)$ by Theorem 5.4.2—cf. the preceding problem. Using $S(3)$ and $S(37)$, we construct $S(105)$, and finally from $S(7)$ and $S(105)$ we derive $S(631)$. □

Exercises 5.4.

Exercise 5.4.1.

(1) Write down all ordered triples from eight digits $1, 2, \ldots, 7, 8$.
(2) Find the number of triples consisting of three different digits (without repetition).
(3) Find the number of triples consisting of two different digits, like $(1, 2, 1)$.
(4) Find the number of triples consisting of the same digit, similar to $(3, 3, 3)$.
(5) Which is the largest set of triples such that any pair of digits belongs to at most one triple?
(6) Which is the smallest set of triples such that any pair of digits belongs to at least one triple?

Exercise 5.4.2. Write down all triples of the elements of the set $X = \{1, 2, 3, 4, 5\}$

(1) Such that each pair of elements of X belongs to at most one triple and the number of triples is the largest.
(2) Such that each pair of elements of X belongs to at least one triple and the number of triples is the smallest.

Exercise 5.4.3. Consider a system of triples

$$S = \{\{1, 2, 3\}, \{1, 4, 5\}, \{1, 6, 7\}, \{2, 4, 6\}, \{2, 5, 7\}\}.$$

Which triples from the list

$$\{2, 3, 4\}, \{5, 6, 7\}, \{3, 4, 6\}, \{3, 5, 7\}, \{3, 4, 7\}, \{3, 6, 7\}, \{4, 5, 6\}, \{1, 2, 3\}$$

should be added to S to make it a Steiner system $S(7)$?

Exercise 5.4.4. Complete the solution of Problem 5.4.3 by constructing explicitly all blocks of $S(631)$. How many blocks does this BIBD contain?

Exercise 5.4.5. Figure 5.2 called the Fano plane, has occurred in various areas of mathematics. How does it represent the Steiner triple system $S(7, 7, 1)$?

Exercise 5.4.6. Show that the Steiner triple systems can be generated by decompositions of a complete graph K_n in triangles without common edges.

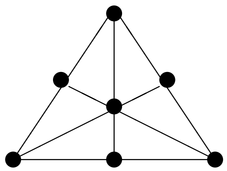


Figure 5.2: The Fano plane.

Exercise 5.4.7. There are 100 professors at a college. Every day three of them have a lunch together at the college cafeteria. Is it possible to schedule their visits during some period of time so that every two of them lunch together exactly once?

Exercise 5.4.8. Let $S(w)$ be a Steiner subsystem of the Steiner system of triples $S(v)$. Prove that $w \le \frac{1}{2}(v-1)$.

Exercise 5.4.9. Prove that if two Steiner subsystems of a Steiner system of triples have a nonempty intersection, then the intersection also is a Steiner system of triples.

Exercise 5.4.10. Find the necessary conditions, similar to (5.4.2), of the existence of systems of quadruples $S(v, b, 4, r, \lambda)$. Specialize these conditions when $\lambda = 1$.

Exercise 5.4.11. Do BIBD $S(7, 7, 4, 4, 1)$ and $S(13, 13, 4, 4, 1)$ exist? If either of them exists, construct it.

Exercise 5.4.12. Nine professors must proctor 12 exams in 4 days, so that each test must be observed by a committee of 3 professors. Compose a schedule of the exams such that every pair and every triple of the professors do not meet more than once during the exams.

Exercise 5.4.13. At the Test College students have four exams every day during seven days in a row. Is it possible to arrange eight professors to proctor these exams in pairs so that the same pair of professors does not proctor two exams?

Exercise 5.4.14. An ice hockey team has nine forwards. The team plays four games in a row. Prove that it is possible to set up the triples of field players, so that no two forwards play twice in the same triple.

Exercise 5.4.15. Every year the Combi Club holds a meeting where each member of the Club must present his or her results for the past year to every other member of the club. However, the Club has a very small classroom, where only 3 people can be at a time. If every such small meeting of 3 members lasts 30 minutes without a break between 3-party meetings, and this year there are 15 club members, then how many these meetings are necessary and for how long will the room be occupied?

Exercise 5.4.16. Solve the previous problem if the club has (a) 14, (b) 16 members.

6 Secondary structures of the RNA

6.1 RNAs, graphs, and the Cauchy–Hadamard formula

Graph theory has endless applications, which sometimes lead to new kinds of graphs. In this chapter we show a simple biological application of the graph theory. Since M. Waterman [56] defined the Ribonucleic Acid (RNA) secondary structure in graph-theoretical terms, the derivation of the upper bounds or of the precise asymptotic formulas for the various structures became an important problem; the number of relevant papers is growing; see, e. g., [12, 20] and the references therein. Derivation of these estimates in the current literature is invariably based on the deep result of complex analysis that can be traced back to G. Darboux. However, the Darboux theorem and its modern analogs are well beyond the current undergraduate curriculum.

The goal of this chapter is to show that precise upper bounds for the number of secondary structures in many cases can be derived quite elementary, well within the power of an undergraduate student taking an introductory complex analysis class. The method is based on the well-known Cauchy–Hadamard formula for the radius of convergence of Taylor series, or even on its real-valued relative—the root-test for convergence of the power series.[1]

For the reader's convenience, we review a few simple biological notions, fundamental to the RNA and relevant to our topic here. We expound them in terms of the elementary concepts of graph theory as in Chapter 2. Then we explain our method, and in Section 6.5 we consider examples of its application.

All living organisms consist of *cells*. A cell is a hull (membrane) filled with the chemical "bullion", containing different molecules. Some cells can split, giving rise to other cells. During this process, the biological molecules can also split and replicate themselves. However, our world is not perfect, and this process can introduce changes (mutations), errors, thus leading to new cells, maybe with different features. This process can lead to the biological evolution, but also to certain negative changes, like illnesses.

Some cells contain important (macro)molecules, called ribonucleic acids (RNA). These acids contain essential genetic information, for example, about viruses, thus, it is important to know their structure. Biologists distinguish primary, secondary, and tertiary structures of RNA; below we consider only the first two of them. Unlike the double helix of the DNA, each RNA is a *linearly ordered strand*, or just a *string*, consisting of other molecules, called *ribonucleotides*. This string is the backbone of any

---

1 These issues were, in particular, discussed at the workshop *Teaching Discrete and Algebraic Mathematical Biology to Undergraduates* at the Mathematical Biosciences Institute at Ohio State University, Columbus, OH, 7/29/2013–8/02/2013.

https://doi.org/10.1515/9783110751185-006

RNA. Traditionally, it is represented by a horizontal straight segment with nodes occupied by the nucleotides. If the RNA contains n ribonucleotides, we select n points of the segment and number them consecutively from the left to the right by the natural numbers $1, 2, \ldots, n$; thus, in Fig. 6.1 $n = 6$. These dots represent the ribonucleotides in the RNA. For more information the reader can consult, for example, [46] or [54].

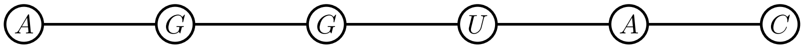


Figure 6.1: A string (primary structure) $\mathcal{A}$–$\mathcal{G}$–$\mathcal{G}$–$\mathcal{U}$–$\mathcal{A}$–$\mathcal{C}$.

When we start studying new objects, it is often necessary to know their quantity. In particular, it is important to know the number of the primary and secondary structures of RNA. It is not always possible to find a precise formula for the number of the secondary structures subject to various restrictions. And even if such a formula is derived, it can be very cumbersome, and therefore useless. That is why a lot of work has been done to derive different *asymptotic* formulas for the numbers of various secondary structures; see, e. g., [46, 54] and the references therein. A formula is called asymptotic, if it gives better and better *relative approximation* of a quantity under consideration, when an important parameter (like time or size) is approaching a crucial threshold; for example, if time tends to infinity.

There are four different ribonucleotides, called *adenine* (denoted hereafter $\mathcal{A}$), *cytosine* ($\mathcal{C}$), *guanine* ($\mathcal{G}$), and *uracil* ($\mathcal{U}$). The linear ordering of these four nucleotides in either order, where each of them can repeat indefinitely, is called the *primary structure* of the RNA. Thus, the primary structure of an RNA can be depicted by drawings like this:

Pictures like this are called *graphs*; we studied them in Chapter 2. The graph in Fig. 6.1 is *labeled*—its vertices are labeled by the symbols of the nucleotides. Graph theory is a mathematical theory, and even though mathematics by itself cannot solve biological problems, it can give useful insights and help people to solve biological problems [57].

An RNA molecule is not rigid like a metal bar; it is flexible and can be conveniently thought of as a smooth flexible string, which can be crumpled, and then stretched again without any noticeable change.

Imagine now that we attached small pieces of velcro tape at some places of this string. If we now fold it over, then these pieces of velcro tape can hook one another, and we cannot easily stretch the tape in a linear structure as before. In real molecules instead of velcro tape there are certain pairs of nucleotides. If they happen to be close enough one to another, they are capable of forming chemical bonds. According to Watson–Crick and their non-named assistants, three pairs of the nucleotides, namely, A–U, G–C, and U–G, *in either order*, can make these bonds. For example, the string in Fig. 6.1 can fold over, like the one in Fig. 6.2, where the new ties are pairs $C–G$ and $G–C$.

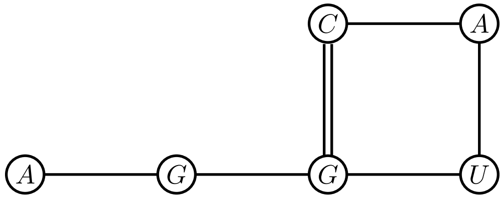


Figure 6.2: A secondary structure built on the primary string $A–G–G–U–A–C$.

This folding is called the *secondary structure* of the RNA molecule. Since different nucleotides can come close to each other, a primary structure can generate many secondary structures, which drastically complicates the analysis of the RNA. It is supposed that the secondary structure is a two-dimensional object, thus it can be drawn in the plane.

The *secondary structure* of the RNA describes the ordering and location of these base pairs of the nucleotides. The secondary structure is responsible for many crucial biological phenomena, and the graph theory is helpful in discovering these structures. To this end, special graphs, called *diagrams* are useful [46]. The diagrams visualize the primary structure and possible ways to generate the corresponding secondary structures. In the next section we consider the primary and secondary structures of the RNA and these diagrams in more detail.

6.2 Counting the primary structures

It is often useful to know the number of the objects studied. We begin by solving an easy problem of the calculating the number of the RNA primary structures. If an RNA can contain infinitely long strains of the nucleotides, then there clearly are infinitely many primary structures. Therefore, we consider the RNA of some specified length, say n. We denote as $\widetilde{R}(n)$ the number of different linear strings, containing n nucleotides, without any restrictions on the neighboring ones. Since every string starts with one of the four nucleotides, either $\mathcal{A}$, or $\mathcal{C}$, or $\mathcal{G}$, or $\mathcal{U}$, followed by a string of length $n-1$, we can immediately produce the basic equation,

$$\widetilde{R}(n) = 4 \times \widetilde{R}(n-1).$$

Such equations are called *recurrence relations* or *difference equations*; see in particular Sections 4.3–4.4. In the same fashion,

$$\widetilde{R}(n-1) = 4 \times \widetilde{R}(n-2),$$

and we can *iterate* this equation, getting the equation

$$\widetilde{R}(n) = 4\widetilde{R}(n-1) = 4^2\widetilde{R}(n-2) = 4^3\widetilde{R}(n-3) = \cdots = 4^{n-1}\widetilde{R}(1).$$

Since we have an obvious initial condition $\widetilde{R}(1) = 4$, the total number of primary structures without any restriction is

$$\widetilde{R}(n) = 4^n.$$

We can notice that this is just the number of permutations (or arrangements) with repetitions of n elements of four different kinds of elements.[2]

Thus, the number of the RNA grows exponentially, as 4^n. Therefore, the number of the secondary structures in the literature is usually compared with the exponential function $b^n, b \leq 4$.

Exercise 6.2.1. List and sketch all RNA with $n = 1, 2, 3$.

Now we take up the RNA with restrictions on the neighboring nucleotides. Of course, in real molecules there are always small deviations from the basic rules, that is, certain "forbidden" pairs can occur, even though very infrequently, with a small probability. We neglect these "outliers" and consider only RNA, where *all the pairs* are only of these three types, allowed by Nature:

$$A\text{–}U;\ C\text{–}G;\ G\text{–}C. \tag{6.2.1}$$

The other pairs are forbidden by certain biological considerations. We compute the number of the primary structures satisfying these restrictions.

Let us denote the number of strings of length n and starting with A, as $R_A(n)$, and similarly, $R_C(n), R_G(n), R_U(n)$. If the very first nucleotide is a $\mathcal{U}$, this puts no restriction on the second element, thus,

$$R_U(n) = R(n-1).$$

However, if the first nucleotide is a $\mathcal{A}$, then the second element must be $\mathcal{U}$, thus

$$R_A(n) = R_U(n-1).$$

Similarly,

$$R_C(n) = R_G(n-1) \quad \text{and} \quad R_G(n) = R_C(n-1),$$

and since an RNA must start with a nucleotide,

$$R(n) = R_A(n) + R_C(n) + R_G(n) + R_U(n).$$

2 For all the basic information from combinatorics and graph theory see for example Chapters 1 and 2.

Collecting all these equations, we deduce the equation

$$R(n) = R_U(n-1) + R_G(n-1) + R_C(n-1) + R(n-1). \tag{6.2.2}$$

This is a fundamental equation; it tells that if the first element is A, then the second one must be U, but its length is $n-1$; if the first element is C, then the second one must be G, and its length is $n-1$ also; if the first element is G, then the second one must be C, and again its length is $n-1$; finally, if the first element is U, then there is no restriction on the second position.

The *initial conditions* are the same,

$$R_A(1) = R_C(1) = R_G(1) = R_U(1) = 1; \quad R(1) = 4. \tag{6.2.3}$$

Using equations (6.2.2)–(6.2.3), we can easily compute the number $R(n)$ for any given n; for large n we should probably use computers. For example, if $n = 2$, we get $R(2) = 1 + 1 + 1 + 4 = 7$. Indeed, we can list these strands explicitly,

$$A–U;\ C–G;\ G–C;\ U–A;\ U–C;\ U–G;\ U–U.$$

Exercise 6.2.2. Compute the number of $R(3)$ and $R(4)$ strands and draw them explicitly.

Expositions of the graph theory can be found in many books. However, applications of the graph theory are so abundant that researchers often introduce new special kinds of graphs. For example, a new kind of graphs, called *dendrograms*, was introduced, to study clustering algorithms; see, e. g., [34, p. 152] or Chapter 3 above.

6.3 Diagrams

A *diagram* (more precisely, n-diagram) is a graph with n vertices, which are drawn as n equidistant points on a line L. This line divides the plane into two half-planes, and the diagrams are situated in one of these half-planes, including the boundary line L, say, in the upper half-plane. The vertices are labeled by the natural numbers $\{1, 2, \dots, n\}$. Some of the vertices are connected by smooth arcs, each of which is situated in the upper half-plane. The degree of every vertex, i. e., the number of arcs incident to this vertex, is one.

The arcs can intersect one another or can go without an intersection. A set of k distinct arcs, (i_1, j_1), (i_2, j_2), $\dots$, (i_k, j_k), is called a *k-crossing*, if

$$i_1 < i_2 < \cdots < i_k < j_1 < j_2 < \cdots < j_k \tag{6.3.1}$$

The definition, clearly, makes sense only if $k \geq 2$; the 2-crossing and the 3-crossing are shown in Fig. 6.3 and Fig. 6.4, respectively. We see that a k-crossing diagram has exactly k pairwise intersecting arcs. Moreover, a k-crossing diagram has l-crossing for any l, $2 \leq l \leq k$.

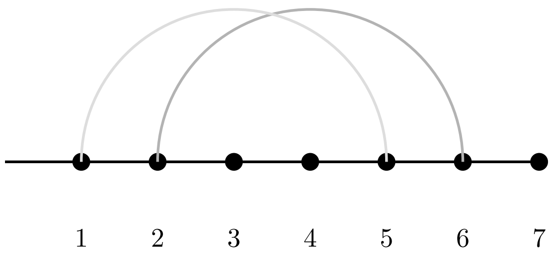


Figure 6.3: 2-crossing.

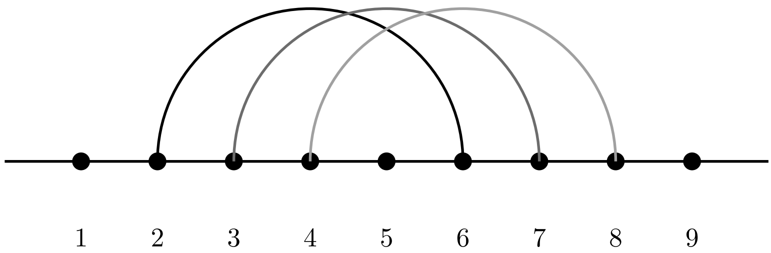


Figure 6.4: 3-crossing.

Exercise 6.3.1. Sketch examples of 4- and 5-crossings.

Exercise 6.3.2. Can three or more arcs intersect at the same point?

Exercise 6.3.3. If no 3 or more arcs intersect at the same point, how many intersections does a k-diagram have?

Exercise 6.3.4. Let a k-diagram have l points where exactly 3 arcs intersect, and no point where more than 3 arcs intersect. How many intersections does this diagram have?

Exercise 6.3.5. Extend the previous problem to diagrams with intersections of any multiplicity.

The point, where two arcs intersect, is stable in the sense that if we slightly move any end-point of the arcs, the arcs are still intersecting. However, it is not the case for the intersection of three or more arcs. One can easily see that such a triple intersection point splits into three double intersections.

It is also useful to consider diagrams without certain crossings. More specifically, a diagram is said to be *k-noncrossing*, or *k-noncrossing partial matching*. A diagram is called *k-noncrossing matching*, if it has no isolated nodes. A set of k distinct arcs is called a k-nesting, if their vertices satisfy the condition (compare with (3))

$$i_1 < i_2 < \cdots < i_k < j_k < j_{k-1} < \cdots < j_1 \tag{6.3.2}$$

A k-nonnesting diagram, obviously, is one without any k-nestings, see Fig. 6.5.

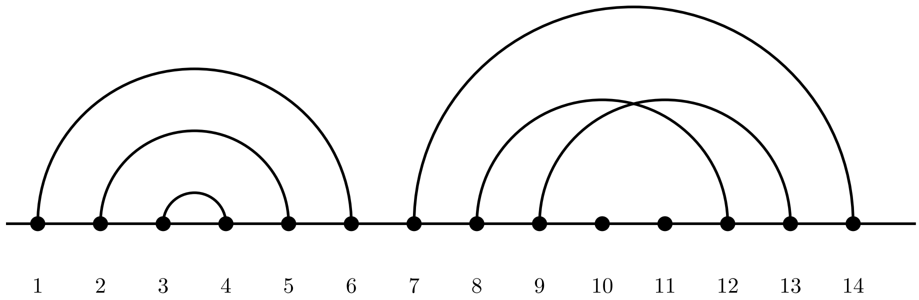


Figure 6.5: A 2-noncrossing case, which is also a 3-nesting structure on the left, and a 3-noncrossing structure on the right.

It is clear that a k-noncrossing is also an l-noncrossing for every $l > k$, and a k-nesting contains l-nestings for all l, $2 \le l < k$.

6.4 Secondary structures

In what follows, the secondary structure ia also a graph, defined by M. Waterman [56]. We consider hereafter only *simple* graphs, i. e., graphs without loops or parallel edges. To a simple graph, there corresponds a zero-one $n \times n$ square matrix, such that its element $ai, j = 1$ if and only if the vertices v_i and v_j are connected with an edge; this matrix is called the *incidence matrix*. The secondary structure is a horizontal backbone of length n, i. e., just a primary structure, enriched by several arcs in the upper half-plane, whose end-points are the nodes of the backbone. In notation we follow [46]. The arcs of the secondary structure are subject to certain restrictions. The arc with ends at the nodes i and $j > i$ is denoted as (i, j), the *length* of the arc is $j - i \ge 1$. Not every family of arcs corresponds to a secondary structure.

A secondary structure is a diagram, i. e., a simple graph on the backbone of the length n, such that the adjacency matrix $A = (a_{i,j})$ possesses the following three properties.

(1) For the basic string to be a backbone, we require that $a_{1,2} = a_{2,3} = \cdots = a_{n-1,n} = 1$.
(2) Next, not counting the neighbors, every point can be adjacent to at most one other point of the backbone. In terms of the incidence matrix, this means that, for any $i, 1 \le i \le n$, there exists at most one j with $j \ne i \pm 1$, such that $a_{i,j} = 1$.

(3) Finally, it is assumed that if the vertices a_i and a_j, $i < j$, are adjacent and $i < k < j$, then the vertex a_k *cannot be adjacent* with any vertex to the left of a_i or to the right of a_j. In terms of the adjacency matrix this means that if $a_{k,l} = 1$ and $i < k < j$, then also $i < l < j$.

It follows that the arcs of a secondary structure *do not intersect*, it is a *noncrossing* structure.

Let us consider the secondary structures with $n = 5$. We start with the basic primary structure with 5 nods, Fig. 6.6.

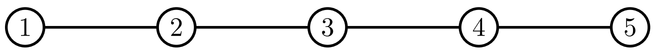


Figure 6.6: The simplest, with no arc, secondary structure; here $n = 5$ and the allowable connections 1–2–3–4–5.

The basic rules (1)–(3) forbid the immediate connection of the neighboring nodes 1–2, but allow connections, by arcs, 1–3, 1–4, and 1–5. In the cases 1–3 and 1–4 there is no room for another arc, but in the case 1–5 the connection 2–4 is possible, see Figures 6.7, 6.8, 6.9, 6.10, 6.11, 6.12, and 6.13.

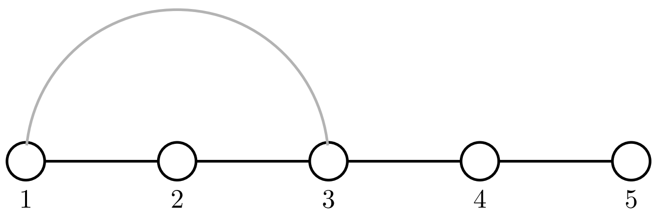


Figure 6.7: The secondary structure with $n = 5$ and the arc 1–3.

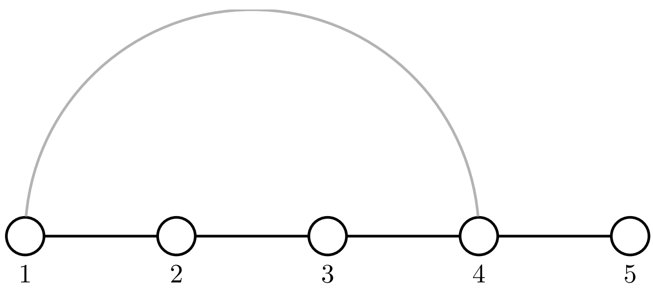


Figure 6.8: The secondary structure with $n = 5$ and the arc 1–4.

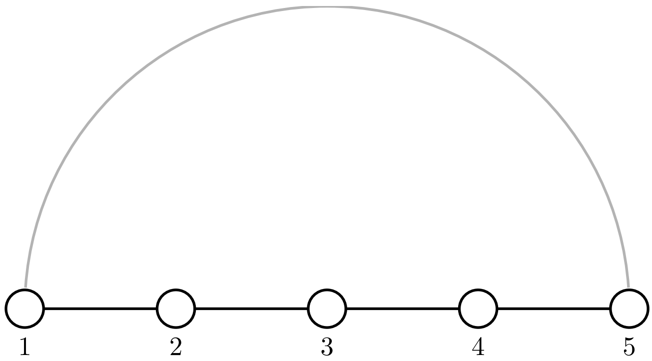


Figure 6.9: The secondary structure with $n = 5$ and the arc 1–5.

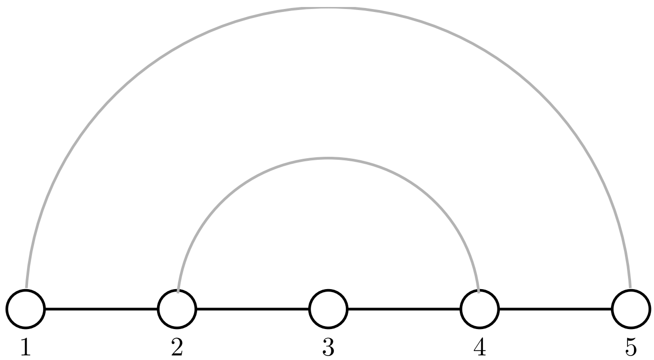


Figure 6.10: The secondary structure with $n = 5$ and the arcs 1–5 and 2–4.

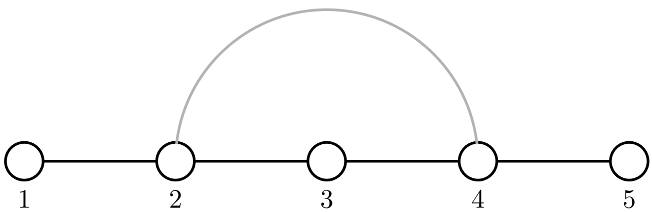


Figure 6.11: The secondary structure with $n = 5$ and the arc 2–4.

Thus, we listed all possible secondary structures, when node (2) is involved. Next, we consider the structures, not involving this nod. If the external arc starts at (2), it can arrive at (4) or (5), generating the arcs 2–4 and 2–5.
Moving further to the right of (2), we have only the possibility 3–5. Therefore, there are 8 secondary structures with $n = 5$.

The 17 possible secondary structures with $n = 6$ are shown in [56].

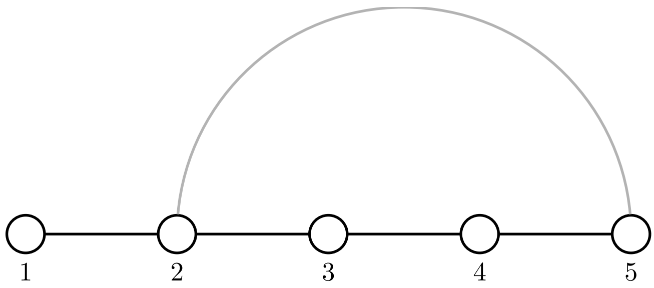


Figure 6.12: The secondary structure with $n = 5$ and the arc 2–5.

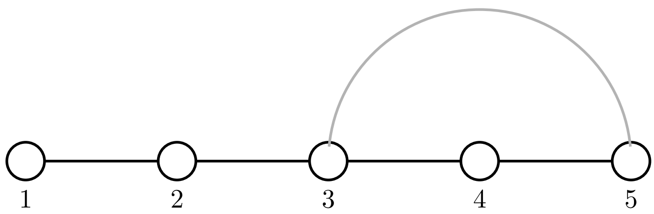


Figure 6.13: The secondary structure with $n = 5$ and the arc 3–5.

Exercise 6.4.1. Sketch the secondary structures with $n = 4$ and $n = 7$; there are 4 and 37 of them, respectively.

6.5 Asymptotic enumeration of the secondary structures. Examples

Even if n is about 10, the total listing of all the secondary structures is unfeasible, so that we want to estimate their quantity. A convenient device for this is their *generating function*—see Sections 4.3–4.4, that is, the power series

$$\varphi(x) = \sum_{0}^{\infty} S(n)x^n.$$

To proceed, we give an equivalent representation of the secondary structures through piecewise linear graphs, called Motzkin paths, lying in the first quadrant $x \geq 0, y \geq 0$ of the coordinate plane and connecting the origin $(0, 0)$ with the point $(n, 0)$, where n is the length of the secondary structure. This subsection is necessary here only to explain the derivation of the major recursive relation for the secondary structures and can be omitted at the first reading. It is important that the secondary structures are noncrossing. We introduce the Motzkin paths on the example of the secondary structure in Fig. 6.10, i. e., the secondary structure with $n = 5$ and arcs 1–5 and 2–4.

Place the backbone on the x-axis, so that its nodes occupy the points with integer coordinates $1, 2, 3, 4, 5$. The node 1 of the backbone is the initial point of an arc. Hence, we connect the origin and the point $(1, 1)$ with a straight segment; its slope is $\tan 45^\circ = 1$. Since the node 2 is also an initial point of an arc, we draw another segment with the same slope from the point$(1, 1)$ to the point $(2, 2)$. The next node, 3, is isolated (*unpaired nucleotide*), therefore, we draw a horizontal segment (the slope is 1) from $(2, 2)$ to the point $(3, 2)$. The length of this horizontal segment is equal to the number of unpaired nucleotides, in the example the length is 1. The next node, 4, is the final point of an arc. Because of that, we draw a segment with the slope of -1 from the point $(3, 2)$ to $(4, 1)$. Since the node 5 is also the terminal point of an arc, the next and last segment goes from $(4, 1)$ to the last node $(5, 0)$ (Fig. 6.14).

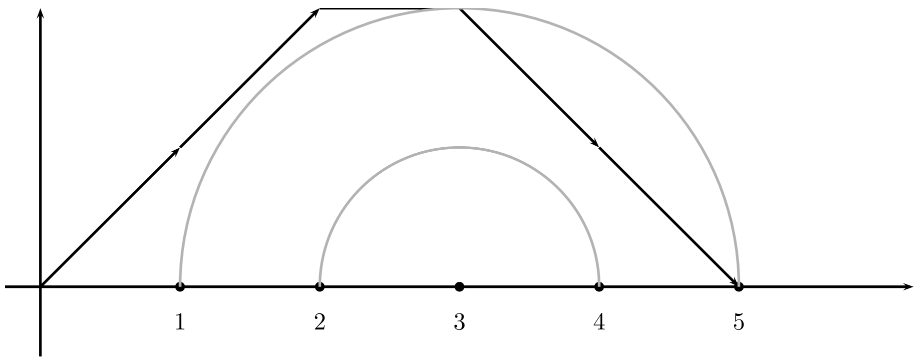


Figure 6.14: The secondary structure with $n = 5$ and the arcs 1–5 and 2–4, and the corresponding Motzkin path.

The description in the previous paragraph exhibits a well-defined algorithm for constructing the Motzkin path, corresponding to a given (noncrossing!) secondary structure.

Exercise 6.5.1. Trace the algorithm and explain, why the condition of noncrossing is important.

Exercise 6.5.2. Draw the Motzkin paths for all other secondary structures in Fig. 6.14.

Exercise 6.5.3. Prove that a Motzkin path cannot have picks,[3] i. e., vertices where merge a side with the positive slope and a side with the negative slope.

Exercise 6.5.4. Prove that the algorithm above is invertible, that is, given a Motzkin path, we can construct the unique (noncrossing) secondary structure. Therefore, there

3 The paths with picks but without horizontal segments are called Dyck paths; see, e. g., [34, p. 57].

is a one-to-one correspondence between the secondary structures and the Motzkin paths.

Exercise 6.5.5. Find the secondary structure corresponding to the Motzkin path in Fig. 6.14.

Motzkin paths can be *concatenated*, i. e., combined together, so that one path, of length, k, say, starts at the origin as before, and then the second path, of the length, l, say, starts at the last nod of the first path. Thus, the new path, the *concatenation* of the two original paths, has the length $n = k + l$; see Fig. 6.14.

Denote the number of secondary structures on n nods with the arc lengths at least λ as $S^\lambda(n)$. Exercise 6.5.4 implies that instead of counting the number of these structures we can count the number of corresponding Motzkin paths.

If we consider secondary structures with arc lengths at least $\lambda \geq 2$, then the concatenation nod can be at any of the points $j = 0, 1, 2, \dots, n - (\lambda + 1)$. Moreover, concatenating any Motzkin path between 0 and j with any Motzkin path between j and n, we get a Motzkin path between 0 and n. Adding the $S^\lambda(n-1)$ 'direct' paths between 1 and n, we derive the nonlinear *recurrent relation* (or *finite-dimensional equation*)

$$S^\lambda(n) = S^\lambda(n-1) + \sum_{j=0}^{n-1-\lambda} S^\lambda(n-2-\lambda)S^\lambda(j). \tag{6.5.1}$$

If $\lambda = 2$, this recurrent relation was derived by Waterman [56]; we follow Reidys [46]. To solve it, we must supply $\lambda + 1$ initial conditions. We assume

$$S^\lambda(0) = S^\lambda(1) = \cdots = S^\lambda(\lambda) = 1.$$

The crucial observation is that recurrent relation (6.5.1) contains a sum of pairwise products quite similar to equation in Problem 4.3.5 for the convolutions. First, consider the case $\lambda = 2$. Multiplying that equation by x^n, after some simple algebra we derive the following equation for the generating function:

$$x^2 s^2(x) + (x - 1 - x^2)s(x) + 1 = 0. \tag{6.5.2}$$

Another important observation is that the latter is a *quadratic equation*[4] for the generating function we sought for. Solving it by the quadratic formula, we find

$$s(x) = \frac{1}{2x^2}(x^2 - x + 1 \pm \sqrt{1 - 2x - x^2 - 2x^3 + x^4}).$$

From the previous equation we see that $s(0) = 1$, therefore, we have to choose the sign "–" above, and finally get

$$s(x) = \frac{1}{2x^2}(x^2 - x + 1 - \sqrt{1 - 2x - x^2 - 2x^3 + x^4}).$$

4 However, the method has the broader scope. In the end we consider an example, where the method is successfully applied to the generating function satisfying a cubic equation.

It is worth noting that we are able to derive equation (6.5.2) because the Motzkin paths can be *concatenated*, and many other important quantities possess this property.

We need the following simple facts from elementary complex analysis. If a power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ has a positive radius of convergence R, then its sum $f(z)$ is an analytic function in the open disc $|z| < R$, R is the distance from the $z = 0$ to the nearest singular point, which must be located on the boundary $|z| = R$, and the Cauchy–Hadamard formula

$$\frac{1}{R} = \limsup_{n\to\infty} \sqrt[n]{|a_n|} \tag{6.5.3}$$

is valid. We can even (with some reservations!) refer to the root test, which is more-or-less real-valued version of the latter.

Example 6.5.1. Consider the generating function $s(x)$ above. The origin $x = 0$ is a *removable singularity*, as can be straightforwardly seen by rationalizing the numerator; indeed,

$$s(x) = \frac{2}{x^2 - x + 1 + \sqrt{1 - 2x - x^2 - 2x^3 + x^4}}.$$

The radicand $1 - 2x - x^2 - 2x^3 + x^4$ is a symmetric polynomial of fourth degree, which can be explicitly solved by dividing over x^4 and substituting $t = x + 1/x$; it has two real roots, $\frac{3\pm\sqrt{5}}{2}$, and two complex roots, $\frac{-1\pm\imath\sqrt{3}}{2}$. The root, closest to the origin, is $\frac{3-\sqrt{5}}{2}$, whence, $s(x)$ is analytic in the open disc $|x| < R = \frac{3-\sqrt{5}}{2}$. Finally, from the previous equations we get the estimate

$$|S_n| \le \left(\frac{1}{3-\sqrt{5}}\right)^n = \left(\frac{3+\sqrt{5}}{2}\right)^n,$$

which is, up to the pre-exponential factor, the estimate derived in [46, p, 65] by making use of certain sophisticated tools of complex analysis.

Example 6.5.2. In the general case, that is, for the secondary structures with any arc length $\lambda \ge 2$, the difference equation is ([46])

$$S^\lambda(n) = S^\lambda(n-1) + \sum_{j=0}^{n-1-\lambda} S^\lambda(n-2-j)S^\lambda(j), \tag{6.5.4}$$

leading again to a quadratic equation for the generating function,

$$x^2(s^{[\lambda]}(x))^2 - (1 - x + x^2 + \cdots + x^\lambda)s^{[\lambda]}(x) + 1 = 0, \tag{6.5.5}$$

which can be explicitly solved for $s^{[\lambda]}$ by the quadratic formula. For example, if $\lambda = 3$, the radicand is the polynomial $1 - 2x - x^2 - x^4 + 2x^5 + x^6$, which can be factored,

$$1 - 2x - x^2 - x^4 + 2x^5 + x^6 = (1 - 2x - x^2)(1 - x^4),$$

with the smallest root $\sqrt{2}-1$. Repeating the same reasoning as above, we get the estimate

$$S^3(n) \le (\sqrt{2}+1)^n.$$

If $\lambda = 4$, the radicand is a polynomial of 8th degree, whose roots are to be evaluated numerically; the smallest one is 0.436911, leading to the estimate

$$S^4(n) \le 2.28879^n.$$

Example 6.5.3. The secondary structures were considered under many various restrictions; see, for example, [54, 55, 36] and the references therein. The approach above works for these results, even if the equation for the generating function is not quadratic, as in the next example. First, we consider so-called *saturated secondary structures*. In this case the generating function $S = S(z) = \sum_{n\ge 0} S - nz^n$ satisfies the system of two nonlinear equations (see, e. g., [36] and the references therein)

$$S(z) = z + z^2 + zT(z) + z^2T + z^2S + z^2S^2$$

and

$$T(z) = z^2S + z^2TS.$$

Eliminating T, we derive the cubic equation for S,

$$z^4S^3 + z^2(z^2-2)S^2 + (1-z^2)S - z(1+z) = 0.$$

To apply our method, we solve the cubic equation by the classical Cardano formula. When the parameter z is within the range of interest, the equation has one real root,

$$S_r(z) = \frac{2-z^2}{3z^3} + \frac{1}{3\sqrt[3]{2}z^3}\left\{A^{1/3} + \frac{\sqrt[3]{2}(z^4-z^2+1)}{3A^{1/3}}\right\},$$

where

$$\begin{aligned} A = & -2z^6 + 30z^4 + 27z^3 + 3z^2 - 2 \\ & + (3z)^{3/2}\sqrt{-4z^7 - 4z^6 + 32z^5 + 60z^4 + 35z^3 + 6z^2 - 5z - 4}. \end{aligned}$$

The equation $A = 0$ simplifies to $(z^4 - z^2 + 1)^3 = 0$ with all the roots on the unit circle. The radicand in the A has the smallest root at $z_0 \approx 0.424687310$, which is the radius of convergence, R, of $S(z)$. It should be mentioned that we choose the branch of the radical which is positive in a right neighborhood of z_0. Hence, we get the bound

$$S_n \le \cdot (1/R)^n = \text{const} \cdot 2.354673^n$$

in complete agreement with [36]).

Example 6.5.4. In the case of *saturated secondary structures*, its generating function S also satisfies a system of the two nonlinear equations

$$S(z) = z + zS(z) + z^2Q(z) + z^2S(z)Q(z)$$

and

$$Q(z) = z^3 + z^2Q(z) + z^4S(z)Q(z) + z^3S(z).$$

Eliminating Q, we get the quadratic equation for S with the discriminant

$$\Delta(z) = z^1 0 - 4z^9 - 2z^8 + 6z^7 + 3z^6 - 8z^5 - z^4 + 4z^3 - z^2 - 2z + 1 = 0.$$

Its smallest root in absolute value is $\approx$ 0.5081360362, which is closest to the origin singular point of $S(z)$, and the upper bound for the number of the saturated secondary structures on n nucleotides is its reciprocal 1.967977, again in perfect agreement with [36]. It should be mentioned that we have used only tools from first-year calculus.

Example 6.5.5. Finally, we consider an example from [20, Theor. 2, p. 352], where the authors derived the generating function for some class of secondary structures,

$$S_0(z) = \frac{1}{2z^2(1+z)^2}\left(1 - z - z^2(1+z)^2 - \sqrt{P(z)}\right),$$

where

$$P(z) = \left(z^4 + 2z^3 + z^2 + z - 1\right)^2 - 4z^3(1+z)^2.$$

Here $z = 0$ is a removal singularity and $z = -1$ is beyond the disc of convergence, since the closest to the origin singular point is the branching point of the radical at the smallest root of the polynomial P. This root is $z_0 \approx 0.32471796$, thus the radius of convergence is $R \approx 3.0795963$ and the upper bound of the number of the secondary structures at question is ·3.0795963, again in agreement with [20].

Acknowledgment

Numerical computations in the examples above were made by making use of the Wolfram Mathematica free online root finder. The author wants to acknowledge the excellent work of this widget.

Answers/solutions to selected problems

I.J.K means problem **K** within a Section **I.J** or exercise **K** in Exercises and Problems **I.J**.

Section 1.1

1.1.1. f is neither, g injective, h surjective, k bijective.
1.1.9. Hint: $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$.
1.1.10. Hint: There are fewer digits than English characters.
1.1.11. $\mathbf{N}_1 \cup \mathbf{N}_3 = \mathbf{N}_3$, $\mathbf{N}_1 \cap \mathbf{N}_3 = \mathbf{N}_1$, $\mathbf{N}_1 \setminus \mathbf{N}_3 = \varnothing$, $\mathbf{N}_3 \setminus \mathbf{N}_1 = \{2, 3\}$.
1.1.13. Any set $\mathbf{Z}'$ containing zero, satisfies $\mathbf{Z}'_e \cup \mathbf{Z}_0 \cup \mathbf{Z}' = \mathbf{Z}$, but only set $\mathbf{Z}'' = \{0\}$ makes $\mathbf{Z}'_e \cup \mathbf{Z}_0 \cup \mathbf{Z}''$ a partition of $\mathbf{Z}$.
1.1.17. 8.
1.1.21. (1) 39 916 800, (2) 40 200.

Exercises 1.1

1.1.1. (1) $2\frac{57}{70}$, (2) $2\frac{17}{60}$, (4) 1.
1.1.11. One zero; 12 zeros; 23 zeros.
1.1.27. $\mathbf{N}_0$.
1.1.30. 2^{2^n}.
1.1.31. Hint: Integrate by parts.

Section 1.2

1.2.1. 20.

Exercises 1.2

1.2.2. $3 \times t + 1 \times 4 = 19$.
1.2.3. $(n!)^2$—Assuming that circular shifts do not generate a new sitting.
1.2.6. $4 \times (3! + 3 + 3 + 1) = 52$.
1.2.7. (a) $T_0 \cup T_1 \cup T_2$; (b) $N_1 \cup P \cup P^c$.
1.2.11. The number of divisors is $4 \times 5 \times 6 \times 7 = 840$, their sum is $(2^4 - 1)(3^5 - 1)(5^6 - 1)(7^7 - 1)$.
1.2.15. $[(9\,999 - 1\,000)/7] + 1 = 1\,286$.
1.2.16. $2^{\frac{n(n-5)}{2}} + 2^{\frac{n(n-6)}{2}} = 2(2n - 11)$.

Exercises 1.3

1.3.3. $A(6, 4) = 360$.
1.3.5. 9!.
1.3.6. $10! - 9! = 9 \cdot 9!$.
1.3.7. Hint: Such numbers must contain either one 4, or one 3 and one 1, or two 2s, or one 2 and two 1, or four 1s, and a complementary number of zeros; the answer is 220.

https://doi.org/10.1515/9783110751185-007

1.3.8. Thus, the numbers are 24 permutations of these four digits, and each digit appears at every position 6 times; the answer is 66660.

1.3.10. s parallel streets divide the town into $s+1$ infinite strips. The first slanted street adds $s+1$ blocks, the second one adds $s+2$, etc. In general, we have $(t+1)(s+1) + \frac{1}{2}t(t-1)$ blocks.

1.3.14. $2 \cdot 8! - 6!$.

1.3.15. $2 \cdot 8! - 6!$.

1.3.19. Hint: In how many ways can you place n different objects, without ordering, in either 2 or three different boxes?

1.3.20. (2) After canceling by $P(k)$, the equation becomes $n \cdots (k+1) = 5$, thus $k = 4$ and $n = 5$.

Exercises 1.4

1.4.1. (1) $A(12,8)$; (2) $\bar{A}(12,8)$; (3) 1; (4) 1.

1.4.3. Hint: The ratio $\frac{n(n+1)(n+2)\cdots(n+k-1)}{k!} = C(n+k-1,k)$.

1.4.5. Hint: Set d_k to be the largest integer such that $C(d_k,k) \le n$.

1.4.8. Hint: Find the largest integer i such that $i! \le n$.

1.4.10. $\frac{C(6,4)}{2^6} = \frac{15}{64}$.

1.4.12. Let 2 005 have k_1 preimages, 2 006—k_2, and 2 007—k_3, then $k_1 + k_2 + k_3 = 2\,006$, and the number $2\,005k_1 + 2\,006k_2 + 2\,007k_3$ must be even. The latter implies that k_1 and k_3 must have the same parity, while the parity of k_2 is immaterial. Therefore, we have to find how many ways there exist to partition the difference $2\,006 - k_2, k_2 = 0, 1, \ldots, 2\,006$, into two addends, which are, necessarily, of the same parity. The answer is $1\,004 \cdot 1\,003$. The number of such functions with an odd sum is $3^{2\,006} - 1\,004 \cdot 1\,003$.

1.4.14. $C(n,2)$.

1.4.23. $C(10,3)$.

1.4.28. $P!C(P+1,S)$, thus $P+1 \ge S$; or $(P-1)!C(P,S)$, thus $P \ge S$.

1.4.43. (1) $C(12,4)$; (2) $\bar{C}(12,4)$.

1.4.52. $C(n,2)$.

1.4.74. $\frac{1}{2}30 \cdot 27 = 405$.

1.4.75. The equation $\frac{1}{2}n(n-3) = 35$ gives $n = 10$.

1.4.78. 6.

Exercises 1.5

1.5.1.
(1) Place a ball in every urn—there is only one way to do that—and then put the two remaining balls in any way, thus $C(4.2) + C(4,1) = 10$.
(2) $5C(4,2) = 30$.
(3) $4 + 5C(4,2) + 4 \cdot 10 + 10 = 84$.
(4) 4.

1.5.2. $35!/(5!)^7$.

1.5.7. (1) $5!/3!$; (2) $2 \cdot 5!/3!$.

1.5.8. $\frac{1}{2}\frac{11!}{1!(4!)^2 2!}$.

1.5.11. $\frac{30!}{(3!)^{10}10!}$; $\frac{30!}{(10!)^3 3!}$.

1.5.21. $\frac{8!}{2! \cdot 1! \cdot 5!}$.

1.5.24. The last two digits must be 12, 24, 32, 44, 52, 64, 72, 84, 92; for each of these pairs the first two digits can be chosen in $P(3,2) = 6$ ways. Thus, there are $9 \cdot 6 = 54$ numbers.

1.5.26. (1) k^n; (2) $C(n+k-1, n)$; (3) $C(k, n)$ (if $k \geq n$).

Section 1.6

1.6.1. (2) $S = \{-\$1(\text{Loss}), 0, \ldots, \$9\,999\}$.

1.6.3.
(1) These events are not disjoint.
(2) The largest probability is 1 if P(Movie AND Restaurant) = 0.3, the smallest probability is 0.7 if P(Movie AND Restaurant) = 0.6.
(3) We must assign the probability P(Movie AND Restaurant).

Exercises 1.6

1.6.4. $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$.

1.6.6. $\frac{1}{26^2 \cdot 10^4}$.

1.6.7. $P(9) + 8 \cdot A(9,8) = 9 \cdot 9!$.

1.6.30. $2 \cdot \frac{1}{3} \cdot \frac{2}{3} + (\frac{1}{3})^2 = \frac{5}{9}$.

1.6.42. To make the average 85, the fifth score must be 87. Thus, the first probability is 1/21, the second one is 9/21, and the third one is 0.

Exercises 2.1

2.1.3.
(1) Yes, it is possible to partition the graph in 10 disjoint subgraphs isomorphic to K_4.
(2) No, due to Lemma 2.1.2.

2.1.4. (1) Yes. (2) No, again by Lemma 2.1.2.

2.1.5. Yes.

2.1.10. Hint: In how many ways is it possible to split $2n$ vertices into n unordered pairs?

2.1.11. Hint: Use again Lemma 2.1.2.

2.1.12. $d \leq v - 1$ and $d \cdot v$ must be an even number.

2.1.16. If two Triplanians exchange a handshake, there must be a third Triplanian exchanging handshakes with both of them, thus we have a triangle of vertices with the sum of degrees equal 6. Therefore, the handshaking lemma says that the sum of degrees of all the vertices for any graph on Triplan is multiple of 6.

Section 2.2

2.2.2. (1)

$$A^* = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 2 & 0 \end{pmatrix}.$$

(2)

$$A \times C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The product $C \times A$ is undefined.

Exercises 2.2

2.2.1. Hint: The five-element set of vertices of g has $2^5 - 1 = 31$ non-empty subsets.

2.2.2. The first and the second diagrams are isomorphic.

2.2.3. G' consists of two connected components, G_1 and G_2.

2.2.8. (1) $C(n,3) + C(n,4) + \cdots + C(n,n) = 2^n - 1 - n - n(n-1)/2$.
(2) $C(5,3) + C(5,4) = 15$.

2.2.12. (1) $n - 2$.
(2) Three walks if the vertices are in the opposite components of the graph, 0 otherwise.

2.2.21. Hint: Split the graph in two or more connected components.

Section 2.3

2.3.2. Hint: You can argue by contradiction and use Lemma 2.2.3.

2.3.4. Since the order of G_0 is 5, its spanning tree must have 4 edges. We start with any edge of weight 1 but cannot include all four such edges for three of them make a cycle. Hence we select any three edges of weight 1 and append any of the two edges of weight 2. There are $(C(4,3) - 1) \cdot 2 = 6$ minimum spanning tree, each of weight 5.

Exercises 2.3

2.3.2. The trees.

2.3.4. Hint: Apply Theorem 2.3.1.

2.3.5. $\{e_1, e_4, e_5\}$, $\{e_1, e_2, e_5\}$, $\{e_1, e_3, e_5\}$, $\{e_2, e_5, e_4\}$, $\{e_3, e_5, e_4\}$.

2.3.8. 1.

2.3.12. $p - 1$.

2.3.16. By EP **2.3.15**, $67 - 35 = 32$ trees.

2.3.23. $3\,000 - (300 - 1) = 2\,701$.

Section 2.4

2.4.1. Such a graph does not exist by Corollary 2.1.1.

2.4.2. The graph on the right ("an open envelope") is semi-Eulerian but not Eulerian, the left graph is neither.

Exercises 2.4

2.4.2. Hint: Count passes through a vertex.

2.4.7. If we number the vertices of K_5 consecutively in either order by v_1, v_2, v_3, v_4, v_5, then a possible Hamiltonian circuit goes consecutively through vertices $v_1, v_3, v_5, v_2, v_4, v_1$.

Exercises 2.5

2.5.1. The answer to all three questions is positive.

2.5.3. Here $p = 6$, thus $q = \frac{1}{2}(5 \cdot 3 + 1) = 8$. By (2.5.1), $f = 2 + 8 - 6 = 4$ including the unbounded component.

2.5.4. No, the road map for this area must be isomorphic to the complete graph K_5, which is not planar.

2.5.10. Any connected component of the graph must have at least $(p-1)/2+1 = (p+1)/2$ edges, thus if the graph has at least two components, then the complement cannot have more than $p - (p + 1)/2 = (p - 1)/2$ edges—a contradiction.

Exercises 3.1

3.1.3. Four clusters: DE, FL, LA, MD, MI, SC, AL, GA, KY, MO, NC, TN, VA, WV, TX.
Five clusters: DE, FL, LA, AL, MD, MI, SC, GA, KY, MO, NC, TN, VA, WV, TX.

3.1.4. This dissimilarity value is 3.

Exercises 3.2

3.2.6. The problem contains the ties starting from the dissimilarity of 2. Hence, the first-level clustering is unique: $\{\{x_2, x_5\}, \{x_1\}, \{x_3\}, \{x_4\}, \{x_6\}\}$; however, the second-level clustering depends upon what edge with the dissimilarity of 2 is chosen first. This clustering may be $\{\{x_2, x_3, x_5\}, \{x_1\}, \{x_4\}, \{x_6\}\}$ or $\{\{x_2, x_5\}, \{x_4, x_6\}, \{x_1\}, \{x_3\}\}$. In turn, these clusterings lead to different clusterings of the next level. After that, the coming conjoint clustering is, of course, unique.

Exercises 4.1

4.1.3. (1) The smallest number is 83.
(2) The longest such sequence of positive numbers consists of 107 numbers, and is infinite if negative numbers are allowed.

4.1.4. By Theorem 4.1.1, the union contains

$$17 + 23 + 41 + 45 + 56 - 6 \cdot C(5, 2) + 4 \cdot (5, 3) - 0 \cdot C(5, 4) = 162 \text{ elements.}$$

4.1.24. (1) The power-set contains $4^3 = 64$ maps, among them there are $A(4,3) = 24$ injective and (since $3 < 4$) no surjective or bijective maps; thus $64 - 24 = 40$ are neither injective nor surjective.

Exercises 4.2

4.2.2. (4) Hint: Insert the latter formula in (4.2.5) and change the order of summation.

4.2.4. Hint: straightforward substitution.

4.2.5. Hint: Choose in EP **4.2.4** $Q_m(t) = t^m, m = 0, 1, \ldots$, compute the corresponding polynomials $P_m(t)$, and use the binomial theorem.

Section 4.3

4.3.2.

$$p\{C(p+k-n-2, p-2) + C(p+k-n-1, p-2) + \cdots + C(p+2n-k-2, p-2)\}$$

4.3.8. Hint: Use the binomial convolution (Definition 4.3.3) and replace $\sqrt[n]{|a_n|}$ in (4.3.3) by $\sqrt[n]{\frac{|a_n|}{n!}}$.

4.3.9. $a_0 = a_1 = 1$, $a_2 = \cdots = a_6 = -2$, $a_7 = \cdots = a_{20} = -6$, $a_{21} = \cdots = a_{44} = 6$, $a_{45} = \cdots = 1$.

4.3.14. $P_0(I,t) = P_1(I,t) = \cdots = 1$; $P_k(S,t) = 1 + t + t^2 + \cdots + t^k = (1 - t^{k+1})/(t-1)$, $k = 0, 1, \ldots$

4.3.16. We have to compute the coefficient of t^{10} in the series

$$(t + t^2 + \cdots)^5 (1 + t + t^2 + \cdots) = t^5 (1 + t + t^2 \cdots)^6,$$

that is, the coefficient of t^5 in $(1 + t + t^2 \cdots)^6$, which is $C(10, 5)$.

Exercises 4.3

4.3.1. (1) $\frac{1}{1-t}$; (2) $t^3(1+t+t^2)$; (3) $\frac{t^3(1+t+t^2)}{1-t^6}$; (4) $\frac{1+t+t^2}{1-t^6}$; (5) $1 + 5t + 10t^2 + 10t^3 + 5t^4 + t^5$; (6) $\frac{1}{1+t}$; (7) $\frac{1-(\cos\alpha)t}{1-2(\cos\alpha)t+t^2}$; (8) $\frac{(\sin\alpha)t}{1-2(\cos\alpha)t+t^2}$.

Section 4.4

4.4.2. $(1-t)^{-1}(1-t)^{-1} = 1 + \cdots + 3t^{78} + \cdots$.

Exercises 4.4

4.4.2. The solution of equation (4.4.13) with the initial values $f(0) = f(1) = 1$ is $f(2) = 1, f(3) = 2, f(4) = 3, f(5) = 5, f(6) = 8, \ldots$

4.4.3. These equations follow either from equation (4.4.13) or from the preceding problem. For example, to derive the equation in (5), we can rewrite its left-hand side as $f(n) + 2f(n-1) + f(n-2) - 1$. Comparing the latter with the right-hand side, we get the equation

$$f(n) - 1 = f(n-1) + f(n-2) - 1 = f(0) + \cdots + f(n-2),$$

which is the equation in (5) with n instead of $n+2$.

4.4.5. By induction, $g(n) = \alpha a^n, n = 0, 1, 2, \ldots$, thus $g(t) = \alpha e^{at}$, and the conclusion follows.

4.4.7. Hint: Expand the determinant over the first column.

4.4.9. $f(n)$.

4.4.10. Hint: Compose a recurrent equation for the number in question.

4.4.13. Hint: Estimate the number $\frac{\sqrt{5}-1}{2\sqrt{5}}$.

4.4.14. $L_n = (\frac{1+\text{sqrt}\,5}{2})^n + (\frac{1-\text{sqrt}\,5}{2})^n$, $L_0 = 2, L_1 = 1, L_2 = 3, L_3 = 4, L_4 = 7, L_5 = 11$.

4.4.15. Follows from EPs **4.4.2** and **4.4.14**.

4.4.17. (4) $x(n) = a + b(-3)^n + \frac{5}{3}2^n$.

4.4.25. Hint: Notice that the sequence satisfies equation (4.4.13).

4.4.41. (1) $n2^{n+2} - (n+1)2^{n+1} + 2$.

Section 4.5

4.5.2. G_2 if $m = 2$ and G_e.

Exercises 4.5

4.5.2. $(3, 2, 1, 0, 0, 0, 0, 0, 0, 0)$, $x_1^3 x_2^2 x_3$.

4.5.7. No, since F is not a group.

4.5.16. The ride is unique.

4.5.17. In both cases, $2 \cdot 4 \cdot 2 \cdot 2^3 = 128$.

4.5.21. $g_1 = (1,2)(3,4)$, $g_2 = (1,2)(1,3)(1,4)$, $g_3 = (1,2)(3,3)(4,4)$, $g_4 = (1,1)(2,2)(3,4)$.

Section 5.1

5.1.7. Since $\deg F = 5$, at least three incident vertices must have the same label.

5.1.11. $|X| = \sum_{y \in Y} |f^{-1}(\{y\})| \le k|Y|$, since for different y the preimages $f^{-1}(\{y\})$ are disjoint.

Exercises 5.1

5.1.1. A week consists of 7 days.

5.1.2. (1) 21; (2) 36; (3) 33; (4) 13.

5.1.3. (1) Yes, since $37 = 3 \cdot 12 + 1$. (2) No.

5.1.5. Parallel translation does not change the angles between the lines, hence we can assume that these lines intersect at a point. Thus, 11 lines make 22 angles,and if each angle is at leat 17°, then the total angle would be at least $22 \cdot 17^\circ = 374^\circ > 360^\circ$.

5.1.11. $6 \cdot 3^{n-3}$.

5.1.16. There are only four different outcomes of the test.

5.1.25. There are 14 differences, $1, 2, \ldots, 14$, and the 9-element set 1,2,3,4,5,6,13,14,15 generates all the differences but the 6. Thus we need to select at least 10 numbers.

5.1.34. There are only 11 different remainders after dividing any integer by 11.

Exercises 5.2

5.2.1. (1) Yes, by Theorem 5.2.3 with $m = 2$.

5.2.8. $n!$.

5.2.9. $C(m \cdot n, k)$ if $m \cdot n \geq k$ and 0 otherwise.

5.2.10. The simplest example is the case $n = 1$ and the 2×2 matrix containing only 1s; these four 1s cannot be covered by one row and one column.

Exercises 5.3

5.3.3. We must build the block design $S(20, b, 2, 19, 1)$, which clearly exists by (5.3.2) with $b = 190$ blocks.

5.3.5. Hint: Reconstruct a BIBD from M.

5.3.8. Hint: Multiply the ith row of the determinant by b_i and add the rows.

5.3.9. Hint: Prove that y and z must have the same parity, that is, either both are odd or both are even, and consider these cases separately.

5.3.10. $S(43, 43, 7, 7, 1)$ does not exist due to Theorem 5.3.2 and the previous problem; for $S(15, 21, 5, 7, 2)$, condition (5.3.2) fails.

Exercises 5.4

5.4.3. Triples $\{3, 4, 7\}$ and $\{3, 5, 6\}$.

5.4.4. The six segments and the circumference represent seven blocks of $S(7)$.

5.4.7. 100 is not an admissible value for $S(v)$.

5.4.10.

$$\begin{cases} \lambda(v - 1) \equiv 0 \pmod 3, \\ \lambda v(v - 1) \equiv 0 \pmod{12}. \end{cases}$$

If $\lambda = 1$, then $v \equiv 1 \pmod{12}$ or $v \equiv 4 \pmod{12}$.

5.4.11. $S(13, 13, 4, 4, 1)$ exists, however, $S(7, 7, 4, 4, 1)$ does not, since for these values of parameters the necessary condition (5.3.3) fails.

5.4.12. Hint: Consider $S(9, 12, 3, 4, 1)$.

5.4.17. Hint: Fix a point outside $S(v)$ and count all the triples where this point meets elements of $S(v)$ and also other elements.

Bibliography

[1] M. Aigner, *Combinatorial Theory*. Springer, Berlin, New York, 1979.
[2] M. Aigner, Catalan and other numbers: a recurrent theme. In *Algebraic Combinatorics and Computer Science*, 347–390. Springer, Milan, 2001.
[3] G.E. Andrews, *The Theory of Partitions*. Cambridge Univ. Press, Cambridge, 1998.
[4] J.G. Michaels, K.H. Rosen, *Applications of Discrete Mathematics*. McGraw-Hill, Inc., New York, 1991.
[5] N.G. De Bruijn, Pólya’s theory of counting. In E. Beckenbach, Ed. *Applied Combinatorial Mathematics*, 144–184. John Wiley and Sons, New York, 1964.
[6] V.V. Belov, E.M. Vorob’ev, V.E. Shatalov, *Graph Theory* (In Russian), Moscow, 1976.
[7] C. Berge, *The Theory of Graphs and its Applications*. Wiley, New York, 1964.
[8] C. Berge, *Principles of Combinatorics*. Academic Press, New York, 1971.
[9] B. Bollobás, *Modern Graph Theory*. Springer Verlag, 1998.
[10] P.J. Cameron, *Combinatorics: Topics, Techniques, Algorithms*. Cambridge Univ. Press, Cambridge, 1994.
[11] C.C. Chen, K.M. Koh, *Principles and Techniques in Combinatorics*. World Sci. Publ. Co., Inc., Singapore, 1992.
[12] P. Clote, E. Kranakis, D. Krizanc, Asymptotic Number of Hairpins of Saturated RNA Secondary Structures. *Bull. Math. Biol.* 75, 2410–2430 (2013).
[13] K.A. Rybnikov, *Combinatorial Analysis. Problems and Exercises*. Nauka, Moscow, 1982 (Russian).
[14] C. Cox, P. Hansen, B. Julesh, *Partitioning Data Sets*. AMS, 1995.
[15] G.P. Egorychev, *Integral Representation and the Computation of Combinatorial Sums*. AMS, Providence, RI, 1984.
[16] P. Erdõs, J. Spenser, *Probabilistic Methods in Combinatorics*. Acad. Press, New York, 1974.
[17] B. Everitt, *Cluster Analysis*. Heinemann Educational Books, Portsmouth, NH, 1974.
[18] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1. John Wiley & Sons, Inc., New York, 1967, vol. 2, 1971.
[19] L.R. Ford Jr., D.R. Fulkerson, *Flows in Networks*. Princeton Univ. Press, Princeton, NJ, 1962.
[20] É. Fusy, P. Clote, Combinatorics of locally optimal RNA secondary structures. *J. Math. Biol.* 68, 341–375 (2014).
[21] A.D. Gordon, *Classification*. Chapman and Hall, 1981.
[22] I.P. Goulden, D.M. Jackson, *Combinatorial Enumeration*. John Wiley & Sons, New York, 1983.
[23] R.L. Graham, M. Grötschel, L. Lovász, *Handbook of Combinatorics*, vol. 1–2. Elsevier, Amsterdam, 1995.
[24] M. Hall Jr., *Combinatorial Theory*, 2nd edn. John Wiley & Sons, Inc., New York, 1986.
[25] P.R. Halmos, *Naive Set Theory*. Van Nostrand, Princeton, N.J., 1960.
[26] F. Harary, E.M. Palmer, *Graphical Enumeration*. Academic Press, New York, 1973.
[27] G.H. Hardy, E.M. Wright, *An Introduction to the Theory of Numbers*. 4th edn. Oxford Clarendon Press, Oxford, 1960.
[28] J.A. Hartigan, *Clustering Algorithms*. John Wiley & Sons, Inc., New York, 1975.
[29] J. Herman, R. Kučera, J. Šimša, *Problems in Combinatorics, Arithmetic, and Geometry*. Springer, New York, 2003.
[30] A.J.W. Hilton, A simplification of Moore’s proof of the existence of Steiner triple systems. *J. Comb. Theory, Ser. A* 13, 422–425 (1972).
[31] L.J. Hubert, Some applications of graph theory to clustering. *Psychometrika* 39, 283–309 (1974).
[32] A.K. Jain, R.C. Dubes, *Algorithms for Clustering Data*. Prentice Hall, 1988.

https://doi.org/10.1515/9783110751185-008

[33] L. Kaufman, P.J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. John Wiley & Sons, Inc., New York, 1990.

[34] A. Kheyfits, *A Primer in Combinatorics*. De Gruyter, Berlin, 2010.

[35] D.E. Knuth, *The Art of Computer Programming*, vol. 1. Addison-Wesley Publ. Company, Reading, MA, 1973.

[36] E. Kranakis, Combinatorics of Canonical RNA Secondary Structures. Preprint, 2011. http://cgi.csc.liv.ac.uk/~ctag/seminars/evangelos20110309.pdf.

[37] S. Lang, *Undergraduate Algebra*, 2nd edn. Springer, 1990.

[38] L. Lovász, *Combinatorial Problems and Exercises*. North Holland, Amsterdam, 1993.

[39] G.E. Martin, *Counting: The Art of Enumerative Combinatorics*. Springer, New York, 2001.

[40] D. Merlini, R. Sprugnoli, M.C. Verri, Lagrange inversion: when and how. *Acta Appl. Math.* 94, 233–249 (2006).

[41] B. Mirkin, *Mathematical Classification and Clustering*. Kluwer, Dordrecht, 1996.

[42] W.K. Nicholson, *Elementary Linear Algebra*, 2nd edn. Raerson, Toronto, 2004.

[43] I. Niven, *Mathematics of Choice*. MAA, Washington, DC, 1965.

[44] G. Pólya, Kombinatorische anzahlbestimmungen für gruppen, graphen und chemische verbindungen. *Acta Math.* 68, 145–254 (1937).

[45] G. Pólya, G. Szegö, *Problems and Theorems from Analysis*, vol. 1, 3rd edn. Springer, Berlin, 1964.

[46] C. Reidys, *Combinatorial Computational Biology of RNA*. Springer, 2011.

[47] J. Riordan, *An Introduction to Combinatorial Analysis*. John Wiley & Sons, Inc., New York, 1958.

[48] G.-C. Rota, *Finite Operator Calculus*. Academic Press, New York, 1975.

[49] H.J. Ryser, *Combinatorial Mathematics*. John Wiley & Sons, Inc., New York, 1963.

[50] R.A. Silverman, *Introductory Complex Analysis*. Dover, NY, 1972.

[51] R.P. Stanley, *Enumerative Combinatorics*, vol. 1–2, 2nd edn. Cambridge Univ. Press, Cambridge, MA, 1997.

[52] S.K. Stein, A. Barcellos, *Calculus and Analytic Geometry*, 5th edn. McGraw-Hill, New York, 1992.

[53] N.Ya. Vilenkin, *Combinatorics*. Acad. Press, New York-London, 1971.

[54] Z. Wang, K. Zhang, RNA Secondary Structure Prediction, Ch. 14. In *Current Topics in Computational Molecular Biology*, 345–363. Tsinghua Univ. Press and MIT, 2002.

[55] W. Wang, M. Zhang, T. Wang, Asymptotic enumeration of RNA secondary structure. *J. Math. Anal. Appl.* 342, 514–523 (2008).

[56] M. Waterman, Secondary Structure of Single-Stranded Nucleic Acids. In *Studies in Foundations and Combinatorics. Advances in Mathematics. Supplementary Studies*, vol. 1, 167–212. Acad. Press, 1978.

[57] E. Wigner, The unreasonable effectiveness of mathematics in the natural sciences. *Commun. Pure Appl. Math.* 13, 1–14 (1960).

[58] R.J. Wilson, *Introduction to Graph Theory*. Acad. Press, New York, 1972.

[59] R.M. Wilson, An existence theory for pairwise balanced designs, I, II, III. *J. Comb. Theory, Ser. A* 13, 220–273 (1972). 18, 71–79 (1975).

Index

Abel transformation 25
antisymmetric *see* binary relations
arrangements
– with repetition 37
– without repetition 37
assignment operator 145
assignment problem 273

backbone 311
balls in urns model 39, 55, 68, 69, 84, 178, 186, 207
Bayes's formula 81
Bell numbers 14
– difference equation 235
Bernoulli's trials 80
bijective *see* mappings
binary relations 15
– antisymmetric 16
– equivalence relations 16
  – equivalence classes 16
  – factor-set 16
  – number of equivalence classes 17
– partial order 16
  – chain 17
– reflexive 15
– symmetric 15
– transitive 16
binomial coefficients 42, 55, 57, 59, 185, 213, 230, 235
binomial convolution (Hurwitz composition) 201
binomial distribution 80
binomial formula 44
binomial theorem 44
birthday problem 81
bit string 231
block designs 284
– incidence matrix 285
– incomplete, balanced (BIBD) 284
– isomorphism 292
  – automorphism 293
– necessary conditions of existence 285
– symmetric 287
Bonferroni inequalities 180
Boolean *see* sets
Boolean functions 64, 65
Bose–Einstein statistics 70
bracelet 243
Bruck–Ryser–Chowla theorem 288
Burnside (Cauchy–Frobenius) lemma 242

cardinality of unions *see* sets
Cartesian (direct) product *see* sets
Cartesian product, cardinality 18
Catalan numbers 55, 57, 65, 228, 235
– recurrence relation 232
Cauchy rule 198
Cauchy–Hadamard criterion 195
Cauchy–Hadamard formula 301
Cayley
– first formula 111
– second formula 117
chain (linearly or totally ordered set) *see* poset
characteristic function *see* sets
chromatic number *see* graphs
circuit *see* graphs
clustering 131
– algorithms
  – agglomerative 133
  – agglomerative single-link 136, 146
  – divisive 134
  – hierarchical 133
  – Hubert's complete-link 155, 156
  – Hubert's single-link 148
  – single-link 134
– amalgamated 133
– clumps 131
– clusters 131
– completely disjoint 133
– dendrogram 154
– (dis)similarity 132
– dissimilarity matrix (table) 132
– link 136
– threshold 132
– threshold graph 137
coloring problems 179, 244, 250, 251, 253, 255, 256, 259, 264
combinations
– with repetition 45, 46, 253
– without repetition 42
complement *see* sets
compositions 218, 233, 248, 253
– generating function 218
conditional probability 78
congruence modulo p *see* natural numbers

connected graph, component *see* graphs
cycle *see* graphs
cyclic sequences 189

de Morgan laws 13
dendrogram *see* clustering
derangement 181
determinant 225
– continuant 231
– expansion across a line 225
– Jacobi determinant 226
– minor 225
Diagram 305
difference *see* sets
difference equations 221
– characteristic polynomial 221
– generating function 222
– superposition principle 221
digraph *see* graphs
Dilworth theorem 278
Diophantine equations 49, 63, 212, 233, 288, 290
Dirichlet principle 259, 264
dot product 102
Dyck path *see* trajectory method

EGF *see* exponential generating function
equivalence relations *see* binary relations
Erdös-Szekeres theorem 261
Euler (totient) function 185
Eulerian circuit (trail) 118
Eulerian graph 118
Euler's theorem 121
events 71
– complementary 74
– disjoint (mutually exclusive) 74
– elementary 71
– exhaustive 74
– independent 77, 79
expected value *see* mathematical expectation
exponential generating function (EGF) *see* method of generating functions

factorial 22
– Stirling asymptotic formula 22
family of subsets 258
Fano plane 298
favorable outcome 72
Fermat little theorem 191
Fermi–Dirac statistics 70
Ferrers diagram 217
– normalized 217
Fibonacci numbers 224, 230
– generating function 232
Fisher's inequality 287
Fleury's algorithm 120
floor function *see* integer part
forest *see* graphs
Frobenius theorem 277
function *see* mappings
Fundamental Theorem of Arithmetic 26

Gauss formula 191
generating polynomials *see* method of generating functions
geometrically identical colorings 244
GF *see* method of generating functions
golden ratio 230
graphs 89
– acyclic 107
– adjacency matrix 102
– bipartite 93
– bipartite graphs
  – matching 281
– chromatic number 182, 186
– coloring problems 106, 179, 182, 212, 258, 263, 267
– complete graph K_p 93
– connected 100
– connected component 100
– contour 99
– cut-edge (bridge) 100
– degree sequence 93
– diagram 90
– directed 90
– edge 89
  – initial (end) vertex 91
  – loop 89
– edge-cover 282
– embedding 91
  – regular 91
– embedding in R^n 90
– enumeration problems 94
– Eulerian characteristic 121
– forest 107
– incidence function 89
– incidence of edges and vertices 89
– isomorphism 94

– labelled 94, 212
– order 89
– planar 91
– plane 121
– regular coloring 182
– simple 90
– size 89
– spanning subgraph (factor) 98
– subgraph 98
– Thomsen graph 122
– tree 107
  – rooted 107
  – spanning tree 109
– vertex 89
  – isolated 89
  – odd (even) 93
  – pendant (leaf) 89
– vertex degree 93
– walks, trails, paths, circuits, cycles 98
– weighted 109
group 237
– cycle index 239
– order 237
– symmetric 237

Hall's condition 269
– strengthened 270
Hall's theorem 269
Hamiltonian graph 120
Hamiltonian path 120
handshaking lemma 93
harmonic numbers 56
hypothesis of equally likely probabilities 74

image *see* mappings
Inclusion–Exclusion Principle 176
injective *see* mappings
integer part 51
intersection *see* sets
inventory 246
inverse image, total preimage *see* preimage
inversion formulas 187, 191, 192, 205, 211, 235
– Möbius inversion 187, 188
inversions in permutations 225
– even (odd) 225

Jordan's theorem 121

Kaplansky lemma 61
König's theorem 277
Kronecker delta 56
Kruskal's algorithm 109, 146

Lagrange's theorem on four squares 289
Lambert *W* function 213
loop *see* graphs
Lucas numbers 232
– generating function 232

mappings
– bijective 5
– codomain 5
– domain 5
– equal 5
– equivalent 244
– image 5
– injective 5
– number of
  – arbitrary mappings 24
  – bijective mappings 24
  – injective mappings 24
– preimage 5
– range 5
– restriction 19
– surjective
  – cardinality 178
– surjective (onto) 5
– weight 245
mathematical expectation 79
matrix 101
– product 102
– symmetric 101
– transpose 101
– zero-one matrix 276
  – covering 276
  – independent set of entries 276
  – irreducible covering 279
Maxwell–Boltzmann statistics 70
method of generating functions 194, 198
– convolution of sequences 198
– examples 193
– exponential generating function 200
– generating polynomials 195, 196, 204, 206, 212, 219
– problems 202, 205, 207–209, 231
– shifts of sequences 201
Möbius function 188
Moore theorems 293, 294

Motzkin path 310
multigraph 90
multinomial
– coefficients 67
– theorem 70
multinomial coefficients 220
multiset 258

natural numbers **N** 4
– combinatorial representation 56
– congruence 24
– factorial representation 57
– natural segment $\mathbf{N}_n$ 4
number of equivalence classes *see* binary relations

ordered pairs 14

partial order *see* binary relations
partitions
– of integers 215, 233
  – generating function 215
– of sets 14
  – number of 178
  – ordered 258
  – Schur's lemma 262
Pascal's triangle 43
path *see* graphs
permanent 235
permutations 37
– with identified elements (with repetition) 67
pigeonhole principle *see* Dirichlet principle
planar, plane graph *see* graphs
Pólya–Redfield theorem 248
Pontryagin–Kuratowski theorem 123
poset *see* partial order
power set *see* sets
preimage *see* mappings
Principle of Mathematical Induction 6
probability axioms 73
probability distribution 72
probability experimental (frequency) 73
product notation 9
Product Rule 30
progression
– arithmetic 25
– geometric 26
Prüfer code 112
pseudograph 90

quantifier
– existential 4
– universal 4

Ramsey
– numbers 258
– theorem 258
random experiment 71
random variables (functions) 77
range *see* mappings
recurrence relations *see* difference equations
reflexive *see* binary relations
reserve 246
RNA secondary structure 301

sample space 71
secondary structure 303
semi-Eulerian graph 118
sets
– Boolean 19
– Cartesian (direct) product 15
– characteristic function 21
– complement 12
– countable 6
– difference 12
– disjoint 12
– empty set 3
– finite 6
– intersection 12
– power set 19
– subset 4
– union 12
  – cardinality 17
– universal set 12
Sieve Formula *see* Inclusion-Exclusion Principle
sigma (summation) notation 8
spanning tree *see* graphs
– minimum 109
– weighted 109
Stirling numbers 191
– of the first kind 186
– of the second kind 178, 186, 213
subfactorial 181
substitutions 237
– cycle index 239
– cycle (orbit) 238
– cycle type 238
– fixed elements 242
– matrix representation 238

sum rule 29
– modified 35
summing sequence (summator) 201
surjective *see* mappings
symmetric *see* binary relations
system of mutual representatives 272
systems of distinct representatives (SDR) 269
systems of quadruples 299
systems of triples 292
– admissible values 292
– Steiner systems 292

telescoping sums 11
totient function *see* Euler (totient) function
trails *see* graphs
trajectory method 54
transitive *see* binary relations
transversal *see* systems of distinct representatives
tree *see* graphs
tree of alternatives 33
triangular numbers 59
tuple (vector) 15

union *see* sets

Vandermonde's identity 56
village weddings (marriage) theorem *see* Hall's theorem

walks *see* graphs
weight
– of a function 245
– of an element 245
– of an equivalence class 245
whole numbers **W** 4