Number Theory Revealed: A Masterclass

Andrew Granville

AMS American Mathematical Society

Number Theory Revealed: A Masterclass

Andrew Granville

AMS American Mathematical Society
Providence, Rhode Island

Cover design by Marci Babineau.

Front cover image of Srinivasa Ramanujan in the playing card: Oberwolfach Photo Collection, https://opc.mfo.de/; licensed under Creative Commons Attribution Share Alike 2.0 Germany, https://creativecommons.org/licenses/by-sa/2.0/de/deed.en.

Front cover image of Andrew Wiles in playing card, credit: Alain Goriely.

2010 *Mathematics Subject Classification.* Primary 11-01, 11A55, 11B30, 11B39, 11D09, 11D25, 11N05, 11N25.

For additional information and updates on this book, visit
www.ams.org/bookpages/mbk-127

Library of Congress Cataloging-in-Publication Data

Cataloging-in-Publication Data has been applied for by the AMS.
See http:www.loc.gov/publish/cip/.

---

10 9 8 7 6 5 4 3 2 1 24 23 22 21 20 19

Contents

Preface

This is a modern introduction to number theory, aimed at several different audiences: students who have little experience of university level mathematics, students who are completing an undergraduate degree in mathematics, as well as students who are completing a mathematics teaching qualification. Like most introductions to number theory, our contents are largely inspired by Gauss's *Disquisitiones Arithmeticae* (1801), though we also include many modern developments. We have gone back to Gauss to borrow several excellent examples to highlight the theory.

There are many different topics that might be included in an introductory course in number theory, and others, like the law of quadratic reciprocity, that surely must appear in any such course. The first dozen chapters of the book therefore present a "standard" course. In the *masterclass* version of this book we flesh out these topics, in copious appendices, as well as adding five additional chapters on more advanced themes. In the *introductory* version we select an appendix for each chapter that might be most useful as supplementary material.[1] A "minimal" course might focus on the first eight chapters and at least one of chapters 9 and 10.[2]

Much of modern mathematics germinated from number-theoretic seed and one of our goals is to help the student appreciate the connection between the relatively simply defined concepts in number theory and their more abstract generalizations in other courses. For example, our appendices allow us to highlight how modern algebra stems from investigations into number theory and therefore serve as an introduction to algebra (including rings, modules, ideals, Galois theory, p-adic numbers,...). These appendices can be given as additional reading, perhaps as student projects, and we point the reader to further references.

Following Gauss, we often develop examples *before* giving a formal definition and a theorem, firstly to see how the concept arises naturally, secondly to conjecture a theorem that describes an evident pattern, and thirdly to see how a proof of the theorem emerges from understanding some non-trivial examples.

[1] In the main text we occasionally refer to appendices that only appear in the *masterclass* version.
[2] Several sections might be discarded; their headings are in ***bold italics.***

Why study number theory? Questions arise when studying any subject, sometimes fascinating questions that may be difficult to answer precisely. Number theory is the study of the most basic properties of the integers, literally taking integers apart to see how they are built, and there we find an internal beauty and coherence that encourages many of us to seek to understand more. Facts are often revealed by calculations, and then researchers seek proofs. Sometimes the proofs themselves, even more than the theorems they prove, have an elegance that is beguiling and reveal that there is so much more to understand. With good reason, Gauss called number theory the "*Queen of Mathematics*", ever mysterious, but nonetheless graciously sharing with those that find themselves interested. In this first course there is much that is accessible, while at the same time natural, easily framed, questions arise which remain open, stumping the brightest minds.

Once celebrated as one of the more abstract subjects in mathematics, today there are scores of applications of number theory in the real world, particularly to the theory and practice of computer algorithms. Best known is the use of number theory in designing cryptographic protocols (as discussed in chapter 10), hiding our secrets behind the seeming difficulty of factoring large numbers which only have large prime factors.

For some students, studying number theory is a life-changing experience: They find themselves excited to go on to penetrate more deeply, or perhaps to pursue some of the fascinating applications of the subject.

Why give proofs? We give proofs to convince ourselves and others that our reasoning is correct. Starting from agreed upon truths, we try to derive a further truth, being explicit and precise about each step of our reasoning. A proof must be readable by people besides the author. It is a way of communicating ideas and needs to be persuasive, not just to the writer but also to a mathematically literate person who cannot obtain further clarification from the writer on any point that is unclear. It is not enough that the writer believes it; it must be clear to others. The burden of proof lies with the author.

The word "proof" can mean different things in different disciplines. In some disciplines a "proof" can be several different examples that justify a stated hypothesis, but this is inadequate in mathematics: One can have a thousand examples that work as predicted by the hypothesis, but the thousand and first might contradict it. Therefore to "prove" a theorem, one must build an incontrovertible argument up from first principles, so that the statement must be true in every case, assuming that those first principles are true.

Occasionally we give more than one proof of an important theorem, to highlight how inevitably the subject develops, as well as to give the instructor different options for how to present the material. (Few students will benefit from seeing *all* of the proofs on their first time encountering this material.)

Motivation. Challenging mathematics courses, such as point-set topology, algebraic topology, measure theory, differential geometry, and so on, tend to be dominated at first by formal language and requirements. Little is given by way of motivation. Sometimes these courses are presented as a prerequisite for topics that will come later. There is little or no attempt to explain what all this theory is good

for or why it was developed in the first place. Students are expected to subject themselves to the course, motivated primarily by trust.

How boring! Mathematics surely should not be developed only for those few who already know that they wish to specialize and have a high tolerance for boredom. We should help our students to appreciate and cherish the beauty of mathematics. Surely courses should be motivated by a series of interesting questions. The right questions will highlight the benefits of an abstract framework, so that the student will wish to explore even the most rarified paths herself, as the benefits become obvious. Number theory does not require much in the way of formal prerequisites, and there are easy ways to justify most of its abstraction.

In this book, we hope to capture the attention and enthusiasm of the reader with the right questions, guiding her as she embarks for the first time on this fascinating journey.

Student expectations. For some students, number theory is their first course that formulates abstract statements of theorems, which can take them outside of their "comfort zone". This can be quite a challenge, especially as high school pedagogy moves increasingly to training students to learn and use sophisticated techniques, rather than appreciate how those techniques arose. We believe that one can best use (and adapt) methods if one fully appreciates their genesis, so we make no apologies for this feature of the elementary number theory course. However this means that some students will be forced to adjust their personal expectations. Future teachers sometimes ask why they need to learn material, and take a perspective, so far beyond what they will be expected to teach in high school. There are many answers to this question; one is that, in the long term, the material in high school will be more fulfilling if one can see its long-term purpose. A second response is that every teacher will be confronted by students who are bored with their high school course and desperately seeking harder intellectual challenges (whether they realize it themselves or not); the first few chapters of this book should provide the kind of intellectual stimulation those students need.

Exercises. Throughout the book, there are a lot of problems to be solved. Easy questions, moderate questions, hard questions, exceptionally difficult questions. No one should do them all. The idea of having so many problems is to give the teacher options that are suitable for the students' backgrounds:

An unusual feature of the book is that exercises appear embedded in the text.[3] This is done to enable the student to complete the proofs of theorems as one goes along.[4] This does not require the students to come up with new ideas but rather to follow the arguments given so as to fill in the gaps. For less experienced students it helps to write out the solutions to these exercises; more experienced students might just satisfy themselves that they can provide an appropriate proof.

[3]Though they can be downloaded, as a separate list, from `www.ams.org/granville-number-theory`.

[4]Often students have little experience with proofs and struggle with the level of sophistication required, at least without adequate guidance.

Other questions work through examples. There are more challenging exercises throughout, indicated by the symbol † next to the question numbers, in which the student will need to independently bring together several of the ideas that have been discussed. Then there are some really tough questions, indicated by the symbol ‡, in which the student will need to be creative, perhaps even providing ideas not given, or hinted at, in the text.

A few questions in this book are open-ended, some even phrased a little misleadingly. The student who tries to develop those themes her- or himself, might embark upon a rewarding voyage of discovery. Once, after I had set the exercises in section 9.2 for homework, some students complained how unfair they felt these questions were but were silenced by another student who announced that it was so much fun for him to work out the answers that he now knew what he wanted to do with his life!

At the end of the book we give hints for many of the exercises, especially those that form part of a proof.

Special features of our syllabus. Number theory sometimes serves as an introduction to "proof techniques". We give many exercises to practice those techniques, but to make it less boring, we do so while developing certain themes as the book progresses, for examples, the theory of recurrence sequences, and properties of binomial coefficients. We dedicate a preliminary chapter to induction and use it to develop the theory of sums of powers. Here is a list of the main supplementary themes which appear in the book:

Special numbers: Bernoulli numbers; binomial coefficients and Pascal's triangle; Fermat and Mersenne numbers; and the Fibonacci sequence and general second-order linear recurrences.

Subjects in their own right: Algebraic numbers, integers, and units; computation and running times: Continued fractions; dynamics; groups, especially of matrices; factoring methods and primality testing; ideals; irrationals and transcendentals; and rings and fields.

Formulas for cyclotomic polynomials, Dirichlet L-functions, the Riemann zeta-function, and sums of powers of integers.

Interesting issues: Lifting solutions; polynomial properties; resultants and discriminants; roots of polynomials, constructibility and pre-Galois theory; square roots (mod n); and tests for divisibility.

Fun and famous problems like the *abc*-conjecture, Catalan's conjecture, Egyptian fractions, Fermat's Last Theorem, the Frobenius postage stamp problem, magic squares, primes in arithmetic progressions, tiling with rectangles and with circles.

Our most unconventional choice is to give a version of Rousseau's proof of the law of quadratic reciprocity, which is directly motivated by Gauss's proof of Wilson's Theorem. This proof avoids Gauss's Lemma so is a lot easier for a beginning student than Eisenstein's elegant proof (which we give in section 8.10 of appendix 8A). Gauss's original proof of quadratic reciprocity is more motivated by the introductory material, although a bit more complicated than these other two proofs.

We include Gauss's original proof in section 8.14 of appendix 8C, and we also understand $(2/n)$ in his way, in the basic course, to interest the reader. We present several other proofs, including a particularly elegant proof using Gauss sums in section 14.7.

Further exploration of number theory. There is a tremendous leap in the level of mathematical knowledge required to take graduate courses in number theory, because curricula expect the student to have taken (and appreciated) several other relevant courses. This is a shame since there is so much beautiful advanced material that is easily accessible after finishing an introductory course. Moreover, it can be easier to study other courses, if one already understands their importance, rather than taking it on trust. Thus this book, *Number Theory Revealed*, is designed to lead to two subsequent books, which develop the two main thrusts of number theory research:

In *The distribution of primes: Analytic number theory revealed*, we will discuss how number theorists have sought to develop the themes of chapter 5 (as well as chapters 4 and 13). In particular we prove the prime number theorem, based on the extraordinary ideas of Riemann. This proof rests heavily on certain ideas from complex analysis, which we will outline in a way that is relevant for a good understanding of the proofs.

In *Rational points on curves: Arithmetic geometry revealed*, we look at solutions to Diophantine equations, especially those of degree two and three, extending the ideas of chapter 12 (as well as chapters 14 and 17). In particular we will prove Mordell's Theorem (developed here in special cases in chapter 17) and gain a basic understanding of modular forms, outlining some of the main steps in Wiles's proof of Fermat's Last Theorem. We avoid a deep understanding of algebraic geometry, instead proceeding by more elementary techniques and a little complex analysis (which we explain).

References. There is a list of great number theory books at the end of our book and references that are recommended for further reading at the end of many chapters and appendices. Unlike most textbooks, I have chosen to not include a reference to every result stated, nor necessarily to most relevant articles, but rather focus on a smaller number that might be accessible to the reader. Moreover, many readers are used to searching online for keywords; this works well for many themes in mathematics.[5] However the student researching online should be warned that Wikipedia articles are often out of date, sometimes misleading, and too often poorly written. It is best to try to find relevant articles published in expository research journals, such as the *American Mathematical Monthly*,[6] or posted at arxiv.org which is "open access", to supplement the course material.

The cover (designed by Marci Babineau and the author)**.**

In 1675, Isaac Newton explained his extraordinary breakthroughs in physics and mathematics by claiming, "*If I have seen further it is by standing on the shoulders*

[5] Though getting just the phrasing to find the right level of article can be challenging.

[6] Although this is behind a paywall, it can be accessed, like many journals, by logging on from most universities, which have paid subscriptions for their students and faculty.

of Giants." Science has always developed this way, no more so than in the theory of numbers. Our cover represents five giants of number theory, in a fan of cards, each of whose work built upon the previous luminaries.

Modern number theory was born from Pierre de Fermat's readings of the ancient Greek texts (as discussed in section 6.1) in the mid-17th century, and his enunciation of various results including his tantalizingly difficult to prove "Last Theorem." His "Little Theorem" (chapter 7) and his understanding of sums of two squares (chapter 9) are part of the basis of the subject.

The first modern number theory book, Gauss's *Disquisitiones Arithmeticae*, on which this book is based, was written by Carl Friedrich Gauss at the beginning of the 19th century. As a teenager, Gauss rethought many of the key ideas in number theory, especially the law of quadratic reciprocity (chapter 8) and the theory of binary quadratic forms (chapter 12), as well as inspiring our understanding of the distribution of primes (chapter 5).

Gauss's contemporary Sophie Germain made perhaps the first great effort to attack Fermat's Last Theorem (her effort is discussed in appendix 7F). Developing her work inspired my own first research efforts.

Srinivasa Ramanujan, born in poverty in India at the end of the 19th century, was the most talented untrained mathematician in history, producing some extraordinary results before dying at the age of 32. He was unable to satisfactorily explain many of his extraordinary insights which penetrated difficult subjects far beyond the more conventional approaches. (See appendix 12F and chapters 13, 15, and 17.) Some of his identities are still inspiring major developments today in both mathematics and physics.

Andrew Wiles sits atop our deck. His 1994 proof of Fermat's Last Theorem built on the ideas of the previous four mentioned mathematicians and very many other "giants" besides. His great achievement is a testament to the success of science building on solid grounds.

Thanks. I would like to thank the many inspiring mathematicians who have helped me shape my view of elementary number theory, most particularly Bela Bollobas, Paul Erdős, D. H. Lehmer, James Maynard, Ken Ono, Paulo Ribenboim, Carl Pomerance, John Selfridge, Dan Shanks, and Hugh C. Williams as well as those people who have participated in developing the relatively new subject of "additive combinatorics" (see sections 15.3, 15.4, 15.5, and 15.6). Several people have shared insights or new works that have made their way into this book: Stephanie Chan, Leo Goldmakher, Richard Hill, Alex Kontorovich, Jennifer Park, and Richard Pinch. The six anonymous reviewers added some missing perspectives and Olga Balkanova, Stephanie Chan, Patrick Da Silva, Tristan Freiberg, Ben Green, Mariah Hamel, Jorge Jimenez, Nikoleta Kalaydzhieva, Dimitris Koukoulopoulos, Youness Lamzouri, Jennifer Park, Sam Porritt, Ethan Smith, Anitha Srinivasan, Paul Voutier, and Max Wenqiang Xu kindly read subsections of the near-final draft, making valuable comments.

Gauss's *Disquisitiones Arithmeticae*

In July 1801, Carl Friedrich Gauss published *Disquisitiones Arithmeticae*, a book on number theory, written in Latin. It had taken five years to write but was immediately recognized as a great work, both for the new ideas and its accessible presentation. Gauss was then widely considered to be the world's leading mathematician, and today we rate him as one of the three greatest in history, alongside Archimedes and Sir Isaac Newton.

The first four chapters of *Disquisitiones Arithmeticae* consist of essentially the same topics as our course today (with suitable modifications for advances made in the last two hundred years). His presentation of ideas is largely the model upon which modern mathematical writing is based. There follow several chapters on quadratic forms and then on the rudiments of what we would call Galois theory today, most importantly the constructibility of regular polygons. Finally, the publisher felt that the book was long enough, and several further chapters did not appear in the book (though Dedekind published Gauss's disorganized notes, in German, after Gauss's death).

One cannot overestimate the importance of *Disquisitiones* to the development of 19th-century mathematics. It led, besides many other things, to Dirichlet's formulation of ideals (see sections 3.19, 3.20 of appendix 3D, 12.8 of appendix 12A, and 12.10 of appendix 12B), and the exploration of the geometry of the upper half-plane (see Theorem 1.2 and the subsequent discussion).

As a young man, Dirichlet took his copy of *Disquisitiones* with him wherever he went. He even slept with it under his pillow. As an old man, it was his most prized possession even though it was in tatters. It was translated into French in 1807, German in 1889, Russian in 1959, English only in 1965, Spanish and Japanese in 1995, and Catalan in 1996!

Disquisitiones is no longer read by many people. The notation is difficult. The assumptions about what the reader knows do not fit today's reader (for example, neither linear algebra nor group theory had been formulated by the time Gauss wrote his book, although *Disquisitiones* would provide some of the motivation for developing those subjects). Yet, many of Gauss's proofs are inspiring, and some have been lost to today's literature. Moreover, although the more advanced two-thirds of *Disquisitiones* focus on binary quadratic forms and have led to many of today's developments, there are several themes there that are not central to today's research. In the fourth book in our trilogy (!), *Gauss's* Disquisitiones Arithmeticae *revealed*, we present a reworking of Gauss's classic, rewriting it in modern notation, in a style more accessible to the modern reader. We also give the first English version of the missing chapters, which include several surprises.

Notation

$\mathbb{N}$ – The *natural numbers*, $1, 2, 3, \ldots$.

$\mathbb{Z}$ – The *integers*, $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$.

Throughout, all variables are taken to be integers, unless otherwise specified.

Usually p, and sometimes q, will denote prime numbers.

$\mathbb{Q}$ – The *rational numbers*, that is, the fractions a/b with $a \in \mathbb{Z}$ and $b \in \mathbb{N}$.

$\mathbb{R}$ – The *real numbers*.

$\mathbb{C}$ – The *complex numbers*.

$$\sum_{\substack{\text{Some variables:}\\ \text{Certain conditions hold}}} \text{summand} \qquad \text{and} \qquad \prod_{\substack{\text{Some variables:}\\ \text{Certain conditions hold}}} \text{summand}$$

mean that we sum, or product, the summand over the integer values of some variable, satisfying certain conditions.

Brackets and parentheses: There are all sorts of brackets and parentheses in mathematics. It is helpful to have protocols with them that take on meaning, so we do not have to repeat ourselves too often, as we will see in the notation below. But we also use them in equations; usually we surround an expression with "(" and ")" to be clear where the expression begins and ends. If too many of these are used in one line, then we might use different sizes or even "{" and "}" instead. If the brackets have a particular meaning, then the reader will be expected to discern that from the context.

$A[x]$ — The set of *polynomials* with coefficients from the set A, that is, $f(x) = \sum_{i=0}^{d} a_i x^i$ where each $a_i \in A$. Mostly we work with $A = \mathbb{Z}$.

$A(x)$ —The set of *rational functions* with coefficients from the set A, in other words, functions $f(x)/g(x)$ where $f(x), g(x) \in A[x]$ and $g(x) \neq 0$.

$[t]$ — The *integer part of* t, that is, the largest integer $\leq t$.

$\{t\}$ — The *fractional part of* (real number) t, that is, $\{t\} = t - [t]$. Notice that $0 \le \{t\} < 1$.

(a, b) — The greatest common divisor of a and b.

$[a, b]$ — The least common multiple of a and b.

$b|a$ — Means b divides a.

$p^k \| a$ — Means p^k divides a, but not p^{k+1} (where p is prime). In other words, k is the "exact power" of p dividing a.

$I(a, b)$ — The set $\{am + bn : \ m, n \in \mathbb{Z}\}$, which is called the *ideal* generated by a and b over $\mathbb{Z}$.

log — The logarithm in base e, the natural logarithm, which is often denoted by "ln" in earlier courses.

Parity – The *parity* of an integer is either even (if it is divisible by 2) or odd (if it is not divisible by 2).

The language of mathematics

"By a *conjecture* we mean a proposition that has not yet been proven but which is favored by some serious evidence. It may be a significant amount of computational evidence, or a body of theory and technique that has arisen in the attempt to settle the conjecture.

An *open question* is a problem where the evidence is not very convincing one way or the other.

A *theorem*, of course, is something that has been proved. There are important theorems, and there are unimportant (but perhaps curious) theorems.

The distinction between open question and conjecture is, it is true, somewhat subjective, and different mathematicians may form different judgements concerning a particular problem. We trust that there will be no similar ambiguity concerning the theorems."

—— Dan Shanks [**Sha85**, p. 2]

Today we might add to this a *heuristic* argument, in which we explore an open question with techniques that help give us a good idea of what to conjecture, even if those techniques are unlikely to lead to a formal proof.

Prerequisites

The reader should be familiar with the commonly used sets of numbers $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$, as well as polynomials with integer coefficients, denoted by $\mathbb{Z}[x]$. Proofs will often use the *principle of induction*; that is, if $S(n)$ is a given mathematical assertion, dependent on the integer n, then to prove that it is true for all $n \in \mathbb{N}$, we need only prove the following:

- $S(1)$ is true.
- $S(k)$ is true implies that $S(k+1)$ is true, for all integers $k \geq 1$.

The example that is usually given to highlight the principle of induction is the statement "$1+2+3+\cdots+n = \frac{n(n+1)}{2}$" which we denote by $S(n)$.[1] For $n = 1$ we check that $1 = \frac{1\cdot 2}{2}$ and so $S(1)$ is true. For any $k \geq 1$, we assume that $S(k)$ is true and then deduce that

$$\begin{aligned} 1+2+3+\cdots+(k+1) &= \underbrace{(1+2+3+\cdots+k)}\quad +\quad (k+1) \\ &= \qquad \frac{k(k+1)}{2} \qquad + \qquad (k+1) \quad \text{as } S(k) \text{ is true} \\ &= \qquad\qquad \frac{(k+1)(k+2)}{2}\ ; \end{aligned}$$

that is, $S(k+1)$ is true. Hence, by the principle of induction, we deduce that $S(n)$ is true for all integers $n \geq 1$.

To highlight the technique of induction with more examples, we develop the theory of sums of powers of integers (for example, we prove a statement which gives a formula for $1^2 + 2^2 + \cdots + n^2$ for each integer $n \geq 1$) in section 0.1 and give formulas for the values of the terms of recurrence sequences (like the Fibonacci numbers) in section 0.2.

[1] There are other, easier, proofs of this assertion, but induction will be the only viable technique to prove some of the more difficult theorems in the course, which is why we highlight the *technique* here.

Induction and the least counterexample: Induction can be slightly disguised. For example, sometimes one proves that a statement $T(n)$ is true for all $n \geq 1$, by supposing that it is false for some n and looking for a contradiction. If $T(n)$ is false for some n, then there must be a least integer m for which $T(m)$ is false. The trick is to use the assumption that $T(m)$ is false to prove that there exists some smaller integer k, $1 \leq k < m$, for which $T(k)$ is also false. This contradicts the minimality of m, and therefore $T(n)$ must be true for all $n \geq 1$. Such proofs are easily reformulated into an induction proof:

Let $S(n)$ be the statement that $T(1), T(2), \ldots, T(n)$ all hold. The induction proof then works for if $S(m-1)$ is true, but $S(m)$ is false, then $T(m)$ is false and so, by the previous paragraph, $T(k)$ is false for some integer k, $1 \leq k \leq m-1$, which contradicts the assumption that $S(m-1)$ is true.

A beautiful example is given by the statement, "*Every integer* > 1 *has a prime divisor.*" (A *prime* number is an integer > 1, such that the only positive integers that divide it are 1 and itself.) Let $T(n)$ be the statement that n has a prime divisor, and let $S(n)$ be the statement that $T(2), T(3), \ldots, T(n)$ all hold. Evidently $S(2) = T(2)$ is true since 2 is prime. We suppose that $S(k)$ is true (so that $T(2), T(3), \ldots, T(k)$ all hold). Now:

Either $k+1$ is itself a prime number, in which case $T(k+1)$ holds and therefore $S(k+1)$ holds.

Or $k+1$ is not prime, in which case it has a divisor d which is not equal to either 1 or $k+1$, and so $2 \leq d \leq k$. But then $S(d)$ holds by the induction hypothesis, and so there is some prime p, which divides d, and therefore divides $k+1$. Hence $T(k+1)$ holds and therefore $S(k+1)$ holds.

(The astute reader might ask whether certain "facts" that we have used here deserve a proof. For example, if a prime p divides d, and d divides $k+1$, then p divides $k+1$. We have also assumed the reader understands that when we write "d divides $k+1$" we mean that when we divide $k+1$ by d, the remainder is zero. One of our goals at the beginning of the course is to make sure that everyone interprets these simple facts in the same way, by giving as clear definitions as possible and outlining useful, simple deductions from these definitions.)

Preliminary Chapter on Induction

Induction is an important proof technique in number theory. This preliminary chapter gives the reader the opportunity to practice its use, while learning about some intriguing number-theoretic concepts.

0.1. Fibonacci numbers and other recurrence sequences

The *Fibonacci numbers*, perhaps the most famous sequence of integers, begin with

$$F_0 = 0,\ F_1 = 1,\ F_2 = 1,\ F_3 = 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ 233, \ldots.$$

The Fibonacci numbers appear in many places in mathematics and its applications.[1] They obey a rule giving each term of the Fibonacci sequence in terms of the recent history of the sequence:

$$F_n = F_{n-1} + F_{n-2} \quad \text{for all integers } n \geq 2.$$

We call this a *recurrence relation*. It is not difficult to find a formula for F_n:

$$F_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right) \quad \text{for all integers } n \geq 0, \tag{0.1.1}$$

where $\frac{1+\sqrt{5}}{2}$ and $\frac{1-\sqrt{5}}{2}$ each satisfy the equation $x+1 = x^2$. Having such an explicit formula for the Fibonacci numbers makes them easy to work with, but there is a problem. It is not obvious from this formula that every Fibonacci number is an integer; however that does follow easily from the original recurrence relation.[2]

[1] Typically when considering a biological process whose current state depends on its past, such as evolution, and brain development.

[2] It requires quite sophisticated ideas to decide whether a given complicated formula like (0.1.1) is an integer or not. Learn more about this in appendix 0F on symmetric polynomials.

Exercise 0.1.1. (a) Use the recurrence relation for the Fibonacci numbers, and induction to prove that every Fibonacci number is an integer.
(b) Prove that (0.1.1) is correct by verifying that it holds for $n = 0, 1$ and then, for all larger integers n, by induction.

Exercise 0.1.2. Use induction on $n \geq 1$ to prove that
(a) $F_1 + F_3 + \cdots + F_{2n-1} = F_{2n}$ and
(b) $1 + F_2 + F_4 + \cdots + F_{2n} = F_{2n+1}$.

The number $\phi = \frac{1+\sqrt{5}}{2}$ is called the *golden ratio*; one can show that F_n is the nearest integer to $\phi^n/\sqrt{5}$.

Exercise 0.1.3. (a) Prove that ϕ satisfies $\phi^2 = \phi + 1$.
(b) Prove that $\phi^n = F_n\phi + F_{n-1}$ for all integers $n \geq 1$, by induction.

Any sequence $x_0,\ x_1,\ x_2, \ldots$, for which the terms x_n, with $n \geq 2$, are defined by the equation

$$\boxed{x_n = ax_{n-1} + bx_{n-2} \ \text{ for all } n \geq 2,} \tag{0.1.2}$$

where a, b, x_0, x_1 are given, is called a *second-order linear recurrence sequence*. Although this is a vast generalization of the Fibonacci numbers one can still prove a formula for the general term, x_n, analogous to (0.1.1): We begin by factoring the polynomial

$$x^2 - ax - b = (x - \alpha)(x - \beta)$$

for the appropriate $\alpha, \beta \in \mathbb{C}$ (we had $x^2 - x - 1 = (x - \frac{1+\sqrt{5}}{2})(x - \frac{1-\sqrt{5}}{2})$ for the Fibonacci numbers). If $\alpha \neq \beta$, then there exist coefficients c_α, c_β for which

$$x_n = c_\alpha\alpha^n + c_\beta\beta^n \ \text{ for all } n \geq 0. \tag{0.1.3}$$

(In the case of the Fibonacci numbers, we have $c_\alpha = 1/\sqrt{5}$ and $c_\beta = -1/\sqrt{5}$.) Moreover one can determine the values of c_α and c_β by solving the simultaneous equations obtained by evaluating the formula (0.1.3) at $n = 0$ and $n = 1$, that is,

$$c_\alpha + c_\beta = x_0 \quad \text{and} \quad c_\alpha\alpha + c_\beta\beta = x_1.$$

Exercise 0.1.4. (a) Prove (0.1.3) is correct by verifying that it holds for $n = 0, 1$ (with x_0 and x_1 as in the last displayed equation) and then by induction for $n \geq 2$.
(b) Show that c_α and c_β are uniquely determined by x_0 and x_1, provided $\alpha \neq \beta$.
(c) Show that if $\alpha \neq \beta$ with $x_0 = 0$ and $x_1 = 1$, then $x_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}$ for all integers $n \geq 0$.
(d) Show that if $\alpha \neq \beta$ with $y_0 = 2, y_1 = a$ with $y_n = ay_{n-1} + by_{n-2}$ for all $n \geq 2$, then $y_n = \alpha^n + \beta^n$ for all integers $n \geq 0$.

The $\{x_n\}_{n\geq 0}$ in (c) is a *Lucas sequence*, and the $\{y_n\}_{n\geq 0}$ in (d) its *companion sequence*

Exercise 0.1.5.[3] (a) Prove that $\alpha = \beta$ if and only if $a^2 + 4b = 0$.
(b)† Show that if $a^2 + 4b = 0$, then $\alpha = a/2$ and $x_n = (cn + d)\alpha^n$ for all integers $n \geq 0$, for some constants c and d.
(c) Deduce that if $\alpha = \beta$ with $x_0 = 0$ and $x_1 = 1$, then $x_n = n\alpha^{n-1}$ for all $n \geq 0$.

Exercise 0.1.6. Prove that if $x_0 = 0$ and $x_1 = 1$, if (0.1.2) holds, and if α is a root of $x^2 - ax - b$, then $\alpha^n = \alpha x_n + bx_{n-1}$ for all $n \geq 1$.

[3] In this question, and from here on, induction should be used at the reader's discretion.

0.2. Formulas for sums of powers of integers

When Gauss was ten years old, his mathematics teacher aimed to keep his class quiet by asking them to add together the integers from 1 to 100. Gauss did this in a few moments, by noting if one adds that list of numbers to itself, but with the second list in reverse order, then one has

$$1 + 100 = 2 + 99 = 3 + 98 = \cdots = 99 + 2 = 100 + 1 = 101.$$

That is, twice the asked-for sum equals 100 times 101, and so

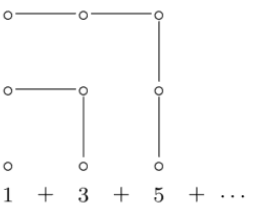
$$1 + 2 + \cdots + 100 \;=\; \frac{1}{2} \times 100 \times 101.$$

This argument generalizes to adding up the natural numbers less than any given N, yielding the formula[4]

$$\sum_{n=1}^{N-1} n \;=\; \frac{(N-1)N}{2}. \tag{0.2.1}$$

The sum on the left-hand side of this equation varies in length with N, whereas the right-hand side does not. The right-hand side is a formula whose value varies but has a relatively simple structure, so we call it a *closed form* expression. (In the prerequisite section, we gave a less interesting proof of this formula, by induction.)

Exercise 0.2.1. (a) Prove that $1 + 3 + 5 + \cdots + (2N-1) = N^2$ for all $N \geq 1$ by induction.
(b) Prove the formula in part (a) by the young Gauss's method.
(c) Start with a single dot, thought of as a 1-by-1 array of dots, and extend it to a 2-by-2 array of dots by adding an appropriate row and column. You have added 3 dots to the original dot and so $1 + 3 = 2^2$.



In general, draw an N-by-N array of dots, and add an additional row and column of dots to obtain an $(N+1)$-by-$(N+1)$ array of dots. By determining how many dots were added to the number of dots that were already in the array, deduce the formula in (a).

Let $S = \sum_{n=1}^{N-1} n^2$. Using exercise 0.2.1 we can write each square, n^2, as the sum of the odd positive integers $\leq 2n$. Therefore $2m-1$ appears $N-m$ times in the sum for S, and so

$$S = \sum_{m=1}^{N-1} (2m-1)(N-m) = -N \sum_{m=1}^{N-1} 1 + (2N+1) \sum_{m=1}^{N-1} m - 2S.$$

[4]This same idea appears in the work of Archimedes, from the third century B.C. in ancient Greece.

Using our closed formula for $\sum_m m$, we deduce, after some rearrangement, that

$$\sum_{n=1}^{N-1} n^2 = \frac{(N-1)N(2N-1)}{6},$$

a closed formula for the sum of the squares up to a given point. There is also a closed formula for the sum of the cubes:

$$\sum_{n=1}^{N-1} n^3 = \left(\frac{(N-1)N}{2}\right)^2. \tag{0.2.2}$$

This is the square of the closed formula (0.2.1) that we obtained for $\sum_{n=0}^{N-1} n$. Is this a coincidence or the first hint of some surprising connection?

Exercise 0.2.2. Prove these last two formulas by induction.

These three examples suggest that there are closed formulas for the sums of the kth powers of the integers, for every $k \geq 1$, but it is difficult to guess exactly what those formulas might look like. Moreover, to hope to prove a formula by induction, we need to have the formula at hand.

We will next find a closed formula in a simpler but related question and use this to find a closed formula for the sums of the kth powers of the integers in appendix 0A. We will go on to investigate, in section 7.34 of appendix 7I, whether there are other amazing identities for sums of different powers, like

$$\sum_{n=1}^{N-1} n^3 = \left(\sum_{n=1}^{N-1} n\right)^2.$$

0.3. The binomial theorem, Pascal's triangle, and the binomial coefficients

The *binomial coefficient* $\binom{n}{m}$ is defined to be the number of different ways of choosing m objects from n. (Therefore $\binom{n}{m} = 0$ whenever $m < 0$ or $m > n$.) From this definition we see that the binomial coefficients are all integers. To determine $\binom{5}{2}$ we note that there are 5 choices for the first object and 4 for the second, but then we have counted each pair of objects twice (since we can select them in either order), and so $\binom{5}{2} = \frac{5\times 4}{2}$. It is arguably nicer to write 5×4 as $\frac{5\times4\times3\times2\times1}{3\times2\times1} = \frac{5!}{3!}$ so that $\binom{5}{2} = \frac{5!}{3!2!}$. One can develop this proof to show that, for any integers $0 \leq m \leq n$, one has the very neat formula[5]

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}, \text{ where } r! = r \cdot (r-1) \cdots 2 \cdot 1. \tag{0.3.1}$$

From this formula alone it is not obvious that the binomial coefficients are integers.

Exercise 0.3.1. (a) Prove that $\binom{n+1}{m} = \binom{n}{m} + \binom{n}{m-1}$ for all integers m, and all integers $n \geq 0$.
(b) Deduce from (a) that each $\binom{n}{m}$ is an integer.

[5] We prefer to work with the closed formula 27!/(15!12!) rather than to evaluate it as 17383860, since the three factorials are easier to appreciate and to manipulate in subsequent calculations, particularly when looking for patterns.

Pascal's triangle is a triangular array in which the $(n+1)$st row contains the binomial coefficients $\binom{n}{m}$, with m increasing from 0 to n, as one goes from left to right:

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ 1 & 5 & & 10 & & 10 & & 5 & 1 \\ 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \end{array}$$

$$\ldots \text{ etc.}$$

The addition formula in exercise 0.3.1(a) yields a rule for obtaining a row from the previous one, by adding any two neighboring entries to give the entry immediately below. For example the third entry in the bottom row is immediately below 5 and 10 (to either side) and so equals $5 + 10 = 15$. The next entry is $10 + 10 = 20$, etc.

The *binomial theorem* states that if n is an integer ≥ 1, then

$$\boxed{(x+y)^n = \sum_{m=0}^{n} \binom{n}{m} x^{n-m} y^m.}$$

Exercise 0.3.2.† Using exercise 0.3.1(a) and induction on $n \geq 1$, prove the binomial theorem.

Notice that one can read off the coefficients of $(x+y)^n$ from the $(n+1)$st row of Pascal's triangle; for example, reading off the bottom row above (which is the 7th row down of Pascal's triangle), we obtain

$$(x+y)^6 = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6.$$

In the previous section we raised the question of finding a closed formula for the sum of n^k, over all positive integers $n < N$. We can make headway in a related question in which we replace n^k with a different polynomial in n of degree k, namely the binomial coefficient

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

This is a polynomial of degree k in n. For example, we have $\binom{n}{3} = \frac{n^3}{6} - \frac{n^2}{2} + \frac{n}{3}$, a polynomial in n of degree 3. We can identify a closed formula for the sum of these binomial coefficients, over all positive integers $n < N$, namely:

$$\sum_{n=0}^{N-1} \binom{n}{k} = \binom{N}{k+1} \tag{0.3.2}$$

for all N and $k \geq 0$. For $k = 2, N = 6$, this can be seen in the following diagram:

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & (1) & & 2 & & 1 & & & \\
 & & 1 & & (3) & & 3 & & 1 & & \\
 & 1 & & 4 & & (6) & & 4 & & 1 & \\
1 & & 5 & & 10 & & (10) & & 5 & & 1 \\
 & 1 & & 6 & & 15 & & \mathbf{(20)} & & 15 & & 6 & & 1
\end{array}$$

so that $1 + 3 + 6 + 10$ equals 20.

Exercise 0.3.3. Prove (0.3.2) for each fixed $k \geq 1$, for each $N \geq k + 1$, using induction and exercise 0.3.1. You might also try to prove it by induction using the idea behind the illustration in the last diagram.

If we instead display Pascal's triangle by lining up the initial 1's and then summing the diagonals,

$$\begin{array}{ccccccc}
1 & & & & & & \\
1 & & 1 & & & & \\
1 & & 2 & & 1 & & \\
1 & & 3 & & 3 & & (1) \\
1 & & 4 & & (6) & & 4 & & 1 \\
1 & & (5) & & 10 & & 10 & & \dots \\
(1) & & 6 & & 15 & & \dots
\end{array}$$

the sums are $1, 1, 1 + 1, 1 + 2, 1 + 3 + 1, 1 + 4 + 3, 1 + 5 + 6 + 1, \dots$ which equal $1, 1, 2, 3, 5, 8, 13, \dots$, the Fibonacci numbers. It therefore seems likely that

$$F_n = \sum_{k=0}^{n-1} \binom{n-1-k}{k} \text{ for all } n \geq 1. \tag{0.3.3}$$

Exercise 0.3.4. Prove (0.3.3) for each integer $n \geq 1$, by induction using exercise 0.3.1(a).

Articles with further thoughts on factorials and binomial coefficients

[1] Manjul Bhargava, *The factorial function and generalizations*, Amer. Math. Monthly **107** (2000), 783–799 (preprint).

[2] John J. Watkins, chapter 5 of *Number theory. A historical approach*, Princeton University Press, 2014.

Additional exercises

Exercise 0.4.1. (a) Prove that for all $n \geq 1$ we have

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

(b) Deduce that $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$ for all $n \geq 1$.
(c) Deduce that $F_{n+1}^2 - F_{n+1}F_n - F_n^2 = (-1)^n$ for all $n \geq 0$.

Exercise 0.4.2.† Deduce from (0.1.1) that the Fibonacci number F_n is the nearest integer to $\phi^n/\sqrt{5}$, for each integer $n \geq 0$, where the constant $\phi := \frac{1+\sqrt{5}}{2}$. This *golden ratio* appears in art and architecture when attempting to describe "perfect proportions".

Exercise 0.4.3. Prove that $F_n^2 + F_{n+3}^2 = 2(F_{n+1}^2 + F_{n+2}^2)$ for all $n \geq 0$.

Exercise 0.4.4. Prove that for all $n \geq 1$ we have

$$F_{2n-1} = F_{n-1}^2 + F_n^2 \quad \text{and} \quad F_{2n} = F_{n+1}^2 - F_{n-1}^2.$$

Exercise 0.4.5. Use (0.1.1) to prove the following:
(a) For every r we have $F_n^2 - F_{n+r}F_{n-r} = (-1)^{n-r}F_r^2$ for all $n \geq r$.
(b) For all $m \geq n \geq 0$ we have $F_mF_{n+1} - F_{m+1}F_n = (-1)^nF_{m-n}$.

Exercise 0.4.6. Let $u_0 = b$ and $u_{n+1} = au_n$ for all $n \geq 0$. Give a formula for all u_n with $n \geq 0$.

The expression 011010 is a *string of* 0*'s and* 1*'s*. There are 2^n strings of 0's and 1's of length n as there are two possibilities for each entry. Let A_n be the set of strings of 0's and 1's of length n which contain no two consecutive 1's. Our example 011010 does not belong to A_6 as the second and third characters are consecutive 1's, whereas 01001010 is in A_8. Calculations reveal that $|A_1| = 2$, $|A_2| = 3$, and $|A_3| = 5$, data which suggests that perhaps $|A_n| = F_{n+2}$, the Fibonacci number.

Exercise 0.4.7.† (a) If $0w$ is a string of 0's and 1's of length n, prove that $0w \in A_n$ if and only if $w \in A_{n-1}$.
(b) If $10w$ is a string of 0's and 1's of length n, prove that $10w \in A_n$ if and only if $w \in A_{n-2}$.
(c) Prove that $|A_n| = F_{n+2}$ for all $n \geq 1$, by induction on n.

Exercise 0.4.8.† Prove that every positive integer other than the powers of 2 can be written as the sum of two or more consecutive integers.

Exercise 0.4.9. Prove that $\binom{n}{m}\binom{n-m}{a-m} = \binom{a}{m}\binom{n}{a}$ for any integers $n \geq a \geq m \geq 0$.

Exercise 0.4.10.† Suppose that a and b are integers and $\{x_n : n \geq 0\}$ is the second-order linear recurrence sequence given by (0.1.2) with $x_0 = 0$ and $x_1 = 1$.
(a) Prove that for all non-negative integers m we have

$$x_{m+k} = x_{m+1}x_k + bx_mx_{k-1} \text{ for all integers } k \geq 1.$$

(b) Deduce that

$$x_{2n+1} = x_{n+1}^2 + bx_n^2 \quad \text{and} \quad x_{2n} = x_{n+1}x_n + bx_nx_{n-1} \quad \text{for all natural numbers } n.$$

Exercise 0.4.11. Suppose that the sequences $\{x_n : n \geq 0\}$ and $\{y_n : n \geq 0\}$ both satisfy (0.1.2) and that $x_0 = 0$ and $x_1 = 1$, whereas y_0 and y_1 might be anything. Prove that

$$y_n = y_1x_n + by_0x_{n-1} \text{ for all } n \geq 1.$$

Exercise 0.4.12. Suppose that $x_0 = 0$, $x_1 = 1$, and $x_{n+2} = ax_{n+1} + bx_n$. Prove that for all $n \geq 1$ we have
(a) $(a + b - 1)\sum_{j=1}^{n} x_j = x_{n+1} + bx_n - 1$;
(b) $a(b^nx_0^2 + b^{n-1}x_1^2 + \cdots + bx_{n-1}^2 + x_n^2) = x_nx_{n+1}$;
(c) $x_n^2 - x_{n-1}x_{n+1} = (-b)^{n-1}$.

Exercise 0.4.13. Suppose that $x_{n+2} = ax_{n+1} + bx_n$ for all $n \geq 0$.
(a) Show that

$$\begin{pmatrix} x_{n+2} & x_{n+1} \\ x_{n+1} & x_n \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} x_2 & x_1 \\ x_1 & x_0 \end{pmatrix} \text{ for all } n \geq 0.$$

(b) Deduce that $x_{n+2}x_n - x_{n+1}^2 = c(-b)^n$ for all $n \geq 0$ where $c := x_2x_0 - x_1^2$.
(c) Deduce that $x_{n+1}^2 - ax_{n+1}x_n - bx_n^2 = -c(-b)^n$.

Other number-theoretic sequences can be obtained from linear recurrences or other types of recurrences. Besides the Fibonacci numbers, there is another sequence of integers that is traditionally denoted by $(F_n)_{n\geq 0}$: These are the *Fermat numbers*, $F_n = 2^{2^n} + 1$ for all $n \geq 0$ (see sections 3.11 of appendix 3A, 5.1, 5.25 of appendix 5H, etc.).

Exercise 0.4.14. Show that if $F_0 = 3$ and $F_{n+1} = F_n^2 - 2F_n + 2$, then $F_n = 2^{2^n} + 1$ for all $n \geq 0$.

Exercise 0.4.15. (a) Show that if $M_0 = 0$, $M_1 = 1$, and $M_{n+2} = 3M_{n+1} - 2M_n$ for all integers $n \geq 0$, then $M_n = 2^n - 1$ for all integers $n \geq 0$. The integer M_n is the nth *Mersenne number* (see exercise 2.5.16 and sections 4.2, 5.1, etc.).

(b) Show that if $M_0 = 0$ with $M_{n+1} = 2M_n + 1$ for all $n \geq 0$, then $M_n = 2^n - 1$.

Exercise 0.4.16.‡ We can reinterpret exercise 0.4.3 as giving a recurrence relation for the sequence $\{F_n^2\}_{n\geq 0}$, where F_n is the nth Fibonacci number; that is,

$$F_{n+3}^2 = 2F_{n+2}^2 + 2F_{n+1}^2 - F_n^2 \text{ for all } n \geq 0.$$

Here F_{n+3}^2 is described in terms of the last three terms of the sequence; this is called a *linear recurrence of order* 3. Prove that for any integer $k \geq 1$, the sequence $\{F_n^k\}_{n\geq 0}$ satisfies a linear recurrence of order $k + 1$.

How to proceed through this book. It can be challenging to decide what proof technique to try on a given question. There is no simple guide—practice is what best helps decide how to proceed. Some students find Zeitz's book [**Zei17**] helpful as it exhibits all of the important techniques in context. I like Conway and Guy's [**CG96**] since it has lots of great questions, beautifully discussed with great illustrations, and introduces quite a few of the topics from this book.

A paper that questions one's assumptions is

[1] Richard K. Guy *The strong law of small numbers*, Amer. Math. Monthly, **95** (1988), 697–712.

Appendix 0A. A closed formula for sums of powers

In chapter 0, we discussed closed form expressions for sums of powers. We will prove here that there is such a formula for the sum of the kth power of the integers up to a given point, developing themes from earlier in this chapter.

0.5. Formulas for sums of powers of integers, II

Our goal in this section is to use our formula (0.3.2) for summing binomial coefficients, to find a formula for summing powers of integers. For example, since

$$n^3 = 6\binom{n}{3} + 6\binom{n}{2} + \binom{n}{1},$$

we can use (0.3.2) with $k = 3$, 2, and 1, respectively, to obtain

$$\begin{aligned}\sum_{n=0}^{N-1} n^3 &= 6\sum_{n=0}^{N-1}\binom{n}{3} + 6\sum_{n=0}^{N-1}\binom{n}{2} + \sum_{n=0}^{N-1}\binom{n}{1} \\ &= 6\binom{N}{4} + 6\binom{N}{3} + \binom{N}{2}.\end{aligned}$$

Summing these three multiples of binomial coefficients gives the formula for the sum of the cubes of the integers up to $N-1$, which we encountered in section 0.2. To make this same technique work to sum n^k, for arbitrary integer $k \geq 1$, we need to show that x^k can be expressed as a sum of fixed multiples of the binomial coefficients $\binom{x}{k}, \ldots, \binom{x}{1}$, where by $\binom{x}{k}$ we mean the polynomial

$$\binom{x}{k} = \frac{x(x-1)\cdots(x-(k-1))}{k!}.$$

Notice that if we substitute $x = n$ into this expression, we obtain the binomial coefficient $\binom{n}{k}$.

Proposition 0.5.1. *Any polynomial $f(x) \in \mathbb{Z}[x]$ of degree $k \geq 0$ can be written as a sum of integer multiples of the binomial coefficients $\binom{x}{k}, \ldots, \binom{x}{1}, \binom{x}{0}$.*

Proof. By induction on k. The result is immediate for $k = 0$. Otherwise, suppose that $f(x)$ has leading coefficient ax^k; then subtract $a \cdot k! \cdot \binom{x}{k}$, which also has leading coefficient ax^k. The resulting polynomial, $g(x) = f(x) - a \cdot k! \cdot \binom{x}{k}$, has degree $k-1$ so can be written as $c_0\binom{x}{0} + \cdots + c_{k-1}\binom{x}{k-1}$ by the induction hypothesis. But then $f(x) = c_0\binom{x}{0} + \cdots + c_k\binom{x}{k}$, with $c_k = a \cdot k!$, as desired. □

In particular, there are integers $c_0,\ c_1, \ldots, c_k$ for which

$$x^k \ = \ c_k\binom{x}{k} + \cdots + c_1\binom{x}{1} + c_0\binom{x}{0}. \tag{0.5.1}$$

One can then immediately deduce, from (0.3.2), that

$$\begin{aligned}\sum_{n=0}^{N-1} n^k &= c_k \sum_{n=0}^{N-1}\binom{n}{k} + \cdots + c_1 \sum_{n=0}^{N-1}\binom{n}{1} + c_0 \sum_{n=0}^{N-1}\binom{n}{0} \\ &= \ c_k\binom{N}{k+1} \ + \cdots + \quad c_1\binom{N}{2} \ + \ c_0\binom{N}{1}.\end{aligned}$$

Expanding out the binomial coefficients, this gives the desired closed form expression for $\sum_{n=0}^{N-1} n^k$, a polynomial in N of degree $k+1$.

There is a difficulty. We proved that the c_j exist but did not show how to determine them. We can do this by successively substituting in $x = 0$, then $x = 1$, then $\ldots, x = k-1$ into (0.5.1), since one obtains

$$0^k = c_k \cdot 0 + \cdots + c_1 \cdot 0 + c_0,$$

and so $c_0 = 0$; then

$$1^k \ = \ c_k \cdot 0 + \cdots + c_2 \cdot 0 + c_1 + c_0,$$

and so $c_1 = 1$; and then $c_2 = 2^k - 2$, $c_3 = 3^k - 3 \cdot 2^k + 3$, etc. We end this appendix with a particularly challenging exercise.

Exercise 0.5.1.‡ (a) Establish that (0.5.1) holds with

$$c_m \ = \ m^k - \binom{m}{1}(m-1)^k + \binom{m}{2}(m-2)^k - \cdots + (-1)^{m-2}\binom{m}{m-2}2^k + (-1)^{m-1}m,$$

for all $m \geq 1$ and for all $k \geq 1$. The integers $c_m/m!$ are the *Stirling numbers of the second kind*, usually denoted by $S_2(k, m)$. They arise in several interesting combinatorial settings; for example, $S_2(k, m)$ is the number of ways to partition a set of k objects into m non-empty subsets.

(b) Deduce that, for any given integer $k \geq 0$, there exist rational numbers $a_0, a_1, \ldots, a_{k+1}$ for which $\sum_{n=0}^{N-1} n^k = a_0 + a_1 N + \cdots + a_{k+1}N^{k+1}$ for all integers $N \geq 1$.

Exercise 0.5.2. Prove that $c_j/j!$ is an integer for all $j \geq 0$ in (0.5.1).

Exercise 0.5.3.† Let $f(x) \in \mathbb{C}[x]$. Prove that $f(n)$ is an integer for all integers n if and only if $f(x) = \sum_m a_m\binom{x}{m}$ where the a_m are all integers.

We will return to this topic, finding an elegant description of the rational numbers a_j by introducing the Bernoulli numbers in the next appendix, appendix 0B.

Appendix 0B. Generating functions

The *generating function* (or *generating series*) of a given sequence of numbers $a_0, a_1, \ldots$ is the *power series*

$$a_0 + a_1 x + a_2 x^2 + \cdots$$

involving a variable x, where the nth term is $a_n x^n$. We now see how generating functions allow us to provide alternative, elegant proofs of the results of this chapter. We begin with an alternative proof of (0.3.2) that exhibits the power of constructing generating functions.

Exercise 0.6.1. (a) Prove that for every integer $k \geq 0$ one has

$$\frac{1}{(1-t)^{k+1}} = \binom{k}{k} + \binom{k+1}{k} t + \binom{k+2}{k} t^2 + \cdots + \binom{k+m}{k} t^m + \cdots .$$

(b) Prove that (0.3.1) follows by equating the coefficient of t^{N-k-1} on either side of

$$\frac{1}{(1-t)^{k+1}} \cdot \frac{1}{(1-t)} = \frac{1}{(1-t)^{k+2}}.$$

(c) Multiply this identity through by $1-t$ and reprove the formula in exercise 0.3.1.(a) by equating the coefficients on each side.

0.6. Formulas for sums of powers of integers, III

The *Bernoulli numbers*, B_n, are the coefficients in the power series:

$$\frac{X}{e^X - 1} = \sum_{n \geq 0} B_n \frac{X^n}{n!}.$$

They are a sequence of numbers that occur in all sorts of interesting contexts in number theory. The first few Bernoulli numbers are $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$, $B_5 = 0$, $B_6 = \frac{1}{42}$, $B_7 = 0$, $B_8 = -\frac{1}{30}$, $B_9 = 0$, $B_{10} = \frac{5}{66}, \ldots$.

From this data we can make a few guesses as to what they look like in general:

- If n is odd and > 1, then $B_n = 0$. This is easily proved since

$$\sum_{\substack{n\geq 0\\ n \text{ odd}}} 2B_n \frac{X^n}{n!} = \sum_{\substack{n\geq 0\\ n \text{ odd}}} B_n \frac{X^n}{n!} - \sum_{\substack{n\geq 0\\ n \text{ odd}}} B_n \frac{(-X)^n}{n!}$$
$$= \frac{X}{e^X - 1} - \frac{(-X)}{e^{-X} - 1} = \frac{X}{e^X - 1} - \frac{(Xe^X)}{e^X - 1} = -X.$$

  Comparing the coefficients of X^n on either side of this equation, we conclude that $B_1 = -\frac{1}{2}$ and $B_n = 0$ for all odd $n \geq 3$.

- The B_n are rational. We expand the power series

$$1 = \frac{e^X - 1}{X} \cdot \frac{X}{e^X - 1} = \sum_{r\geq 0} \frac{X^r}{(r+1)!} \cdot \sum_{s\geq 0} B_s \frac{X^s}{s!}$$
$$= \sum_{n\geq 0} \left(\sum_{\substack{r,s\geq 0\\ r+s=n}} \frac{(n+1)!}{(r+1)!s!} B_s \right) \frac{X^n}{(n+1)!}$$

  and compare the coefficients of X^n on either side to obtain that $B_0 = 1$ and $\sum_{s=0}^{n} \binom{n+1}{s} B_s = 0$ for each $n \geq 1$. This can be rewritten as

$$B_n = -\frac{1}{n+1} \sum_{s=0}^{n-1} \binom{n+1}{s} B_s \text{ for each } n \geq 1.$$

  We can then deduce by induction on $n \geq 1$ that the B_n are rational, since we have given B_n as a finite sum of rational numbers times Bernoullli numbers B_s with $s < n$.

Next we define the *Bernoulli polynomials*, $B_n(t)$, as the coefficients in the power series:

$$\frac{Xe^{tX}}{e^X - 1} = \sum_{n\geq 0} B_n(t) \frac{X^n}{n!},$$

and therefore $B_n(0) = B_n$. To verify that these are really polynomials, note that

$$\sum_{k\geq 0} B_k(t) \frac{X^k}{k!} = e^{tX} \cdot \frac{X}{e^X - 1} = \sum_{m\geq 0} \frac{(tX)^m}{m!} \cdot \sum_{n\geq 0} B_n \frac{X^n}{n!} = \sum_{m\geq 0} \sum_{n\geq 0} B_n t^m \frac{X^{m+n}}{m!n!}.$$

Here we change variable, writing $k = m + n$, and then the coefficient of $X^k/k!$ is

$$B_k(t) = \sum_{\substack{m,n\geq 0\\ m+n=k}} \frac{k!}{m!n!} B_n t^m = \sum_{n=0}^{k} \binom{k}{n} B_n t^{k-n}. \tag{0.6.1}$$

We have done all this preliminary work so as to prove the following extraordinary formula for the sum of the mth powers of the positive integers $< N$.

Theorem 0.1. *For any integers $k \geq 1$ and $N \geq 1$ we have*

$$\sum_{n=0}^{N-1} n^{k-1} = \frac{1}{k}(B_k(N) - B_k).$$

Proof. If N is an integer ≥ 1, then

$$\sum_{k\geq 0}(B_k(N)-B_k)\frac{X^k}{k!} = \frac{X(e^{NX}-1)}{e^X-1} = X\sum_{n=0}^{N-1}e^{nX}$$
$$= X\sum_{n=0}^{N-1}\sum_{j\geq 0}\frac{(nX)^j}{j!} = \sum_{j\geq 0}\left(\sum_{n=0}^{N-1}n^j\right)\frac{X^{j+1}}{j!}.$$

Therefore for any integer $N\geq 1$ we obtain

$$\sum_{k\geq 0}(B_k(N)-B_k)\frac{X^k}{k!} = \sum_{k\geq 1}\left(k\sum_{n=0}^{N-1}n^{k-1}\right)\frac{X^k}{k!}$$

by letting $k=j+1$. The result follows by comparing the coefficients on both sides. □

Negative powers. A key quantity in number theory is the infinite sum

$$\zeta(k) = 1+\frac{1}{2^k}+\frac{1}{3^k}+\frac{1}{4^k}+\cdots$$

which we define for each integer $k\geq 2$. This is called the *Riemann zeta-function* even though it was first explored by Euler more than a hundred years earlier. Each of these sums is convergent, as each $1/m^k\leq 1/m^2$ for all $m\geq 1, k\geq 2$, and $1/m^2<\int_{m-1}^m dt/t^2$, so that

$$1+\frac{1}{2^k}+\frac{1}{3^k}+\cdots\leq 1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots<1+\int_1^\infty\frac{dt}{t^2}=2.$$

We will make a few observations about the values of these sums:

Exercise 0.6.2. (a) Prove that $\sum_{m\geq 2}\frac{1}{m(m-1)}=1$.
(b) Prove that $\sum_{k\geq 2}\frac{1}{m^k}=\frac{1}{m(m-1)}$.
(c) Deduce that $\sum_{k\geq 2}(\zeta(k)-1)=1$.

Exercise 0.6.3. Let $\mathcal{P}$ be the set of perfect powers >1. Let $\mathcal{N}$ be the set of integers >1 that are not perfect powers (so that $\mathcal{P}\cup\mathcal{N}$ is a partition of the integers >1).
(a) Prove that $\mathcal{P}=\{n^k : n\in\mathcal{N} \text{ and } k\geq 2\}$ and $\{n^k : n\in\mathcal{N} \text{ and } k\geq 1\}=\{m\geq 2\}$.
(b) Prove that $\sum_{P\in\mathcal{P}}\frac{1}{P-1}=\sum_{k\geq 2}\sum_{n\in\mathcal{N}}\sum_{j\geq 1}\frac{1}{n^{jk}}$.
(c) Deduce that $\sum_{P\in\mathcal{P}}\frac{1}{P-1}=1$.
This result was communicated by Goldbach to Euler in 1744.

0.7. The power series view on the Fibonacci numbers

An alternate view on Fibonacci numbers, and indeed all second-order linear recurrence sequences, is via their generating functions. For Fibonacci numbers we study the generating function

$$\sum_{n\geq 0}F_nx^n,$$

which is a power series in x. Remembering that $F_n - F_{n-1} - F_{n-2} = 0$ for all $n \geq 2$ we then have

$$\begin{aligned}(1 - x - x^2) \sum_{n\geq 0} F_n x^n &= F_0 + (F_1 - F_0)x + (F_2 - F_1 - F_0)x^2 \\ &\qquad + \cdots + (F_n - F_{n-1} - F_{n-2})x^n + \cdots \\ &= 0 + (1-0)x + 0 \cdot x^2 + \cdots + 0 \cdot x^n + \cdots \;=\; x.\end{aligned}$$

Hence if $\alpha = \frac{1+\sqrt{5}}{2}$ and $\beta = \frac{1-\sqrt{5}}{2}$, then

$$\begin{aligned}\sum_{n\geq 0} F_n x^n \;&=\; \frac{x}{1-x-x^2} \;=\; \frac{1}{\alpha-\beta}\left(\frac{\alpha x}{1-\alpha x} - \frac{\beta x}{1-\beta x}\right) \\ &= \frac{1}{\alpha-\beta}\left(\sum_{m\geq 1} \alpha^m x^m - \sum_{m\geq 1} \beta^m x^m\right) = \sum_{m\geq 1} \frac{\alpha^m - \beta^m}{\alpha - \beta} x^m,\end{aligned}$$

and the result (0.1.1) follows, again. Note that if $x_n = ax_{n-1} + bx_{n-2}$, then

$$(1 - at - bt^2) \sum_{n\geq 0} x_n t^n = x_0 + (x_1 - ax_0)t.$$

The sequence $\{x_n\}_{n\geq 0}$ is again determined by the values of a, b, x_0, and x_1.

Exercise 0.7.1.† Use this to deduce (0.1.3) when $a^2+4b \neq 0$, and exercise 0.1.5(c) when $a^2+4b = 0$.

Both of these methods generalize to arbitrary linear recurrences of degree n, as follows.

Theorem 0.2. *Suppose that $a_1, a_2, \ldots, a_d$ and $x_0, x_1, \ldots, x_{d-1}$ are given and that*

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_d x_{n-d} \quad \textit{for all } n \geq d.$$

Factor the following polynomial into linear factors as

$$X^d - a_1 X^{d-1} - a_2 X^{d-2} + \cdots - a_{d-1} X - a_d = \prod_{j=1}^{k} (X - \alpha_j)^{e_j}.$$

Then there exist polynomials $P_1, \ldots, P_k$, each P_j of degree $\leq e_j - 1$, such that

$$x_n \;=\; \sum_{j=1}^{k} P_j(n)\alpha_j^n \quad \textit{for all } n \geq 0. \tag{0.7.1}$$

The coefficients of the P_j (and the polynomials P_j themselves) can be determined by solving the linear equations obtained by taking this for $n = 0, 1, 2, \ldots, d-1$.

Exercise 0.7.2.‡ Prove that (0.7.1) holds.

Exercise 0.7.3.‡ Let $(x_n)_{n\geq 0}$ be the sequence which begins $x_0 = 0, x_1 = 1$ and then $x_n = ax_{n-1} + bx_{n-2}$ for all $n \geq 2$. Its *companion sequence*, $(y_n)_{n\geq 0}$, begins $y_0 = 2, y_1 = x_2$ and then $y_n = ay_{n-1} + by_{n-2}$ for all $n \geq 2$. For example, $x_n = 2^n - 1$ has companion sequence $y_n = 2^n + 1$.

(a) Prove that $y_n = \alpha^n + \beta^n$ for all $n \geq 0$ and also that $y_n = x_{2n}/x_n$.
(b) Let $z_0 = -1$ and $z_n = -bz_{n-1}$ for all $n \geq 1$. Give an explicit formula for z_n.
(c) Prove that $x_{m+2n} = y_n x_{m+n} + z_n x_m$ for all $m, n \geq 0$.
(d) Deduce that $F_{n+6} = 4F_{n+3} + F_n$ for all $n \geq 0$.

Appendix 0C. Finding roots of polynomials

In the remaining appendices of this preliminary chapter (chapter 0) we introduce several important themes in number theory that do not often appear in a first course but will be of interest to some readers. We also take some time to introduce some basic notions of algebra that appear (sometimes in disguise) throughout this and subsequent number theory courses. To begin with we discuss the famous question of techniques for factoring polynomials into their linear factors.

The reader knows that the roots of a quadratic polynomial $ax^2 + bx + c = 0$ are

$$\frac{-b \pm \sqrt{\Delta}}{2a}, \quad \text{where} \quad \Delta := b^2 - 4ac,$$

is called the *discriminant* of our polynomial, $ax^2 + bx + c$. The easy way to prove this is to put the equation into a form that is easy to solve: Divide through by a, to get $x^2 + (b/a)x + c/a = 0$, so that the leading coefficient is 1. Next make the change of variable, $y = x + b/2a$, to obtain

$$y^2 - \Delta/4a^2 = 0.$$

Having removed the y^1 term, we can simply take square roots to obtain the possibilities, $\pm\sqrt{\Delta}/2a$, for y, and hence we obtain the possible values for x (since $x = y - b/2a$). *Can one similarly find the roots of a cubic?*

0.8. Solving the general cubic

We can certainly begin solving cubics in the same way as we approached quadratics.

Exercise 0.8.1. Show that the roots of any given cubic polynomial, $Ax^3 + Bx^2 + Cx + D$ with $A \neq 0$, can be obtained from the roots of some cubic polynomial of the form $x^3 + ax + b$, by adding $B/3A$ to each root. Moreover write a and b explicitly as functions of A, B, and C.

We wish to find the roots of $x^3 + ax + b = 0$ for arbitrary a and b (which then allows us to determine the roots of an arbitrary cubic polynomial, by exercise 0.8.1). This does not look so easy since we cannot simply take cube roots unless $a = 0$. Cardano's trick (1545) is a little surprising: Write $x = u + v$ so that

$$x^3 + ax + b = (u + v)^3 + a(u + v) + b = (u^3 + v^3 + b) + (u + v)(3uv + a).$$

This equals 0 when

$$u^3 + v^3 = -b \text{ and } 3uv = -a.$$

These conditions imply the simultaneous equations

$$u^3 + v^3 = -b \text{ and } u^3v^3 = -a^3/27,$$

so that, as a polynomial in X we have

$$(X - u^3)(X - v^3) = X^2 + bX - a^3/27.$$

Using the formula for the roots of a quadratic polynomial yields

$$u^3, v^3 = \frac{-b \pm \sqrt{b^2 + 4a^3/27}}{2} = \frac{-b \pm \sqrt{\Delta/(-27)}}{2}, \tag{0.8.1}$$

where $\Delta := -4a^3 - 27b^2$ is the discriminant of our polynomial, $x^3 + ax + b$. (The definition and some uses of discriminants are discussed in detail in section 2.11 of appendix 2B.)

All real numbers have a unique real cube root, call it t, and then the other cube roots are ωt and $\omega^2 t$, where ω is a cube root of 1; for instance we may take $\omega = e^{2i\pi/3} = \frac{-1+\sqrt{-3}}{2}$. Therefore if U and V are the real cube roots in (0.8.1), so that $-3UV$ is real and therefore equal to a, then the possible solutions to $u^3 + v^3 = -b$ together with $3uv = -a$ are

$$(u, v) = (U, V), (\omega U, \omega^2 V), \text{ and } (\omega^2 U, \omega V).$$

This implies that the roots of $x^3 + ax + b$ are given by

$$U + V, \ \omega U + \omega^2 V, \text{ and } \omega^2 U + \omega V.$$

The roots of a quadratic polynomial were obtained in terms of integers and square roots of integers. We have just seen that the roots of a cubic polynomial can be obtained in terms of integers, square roots, and finally cube roots. How about the roots of a quartic polynomial? Can these be found in terms of integers, square roots, cube roots, and fourth roots? And are there analogous expressions for the roots of quintics and higher degree polynomials?

0.9. Solving the general quartic

This is bound to be technically complicated, so much so that it is arguably more interesting to know that it can be done rather than actually doing it, so we just sketch the proofs:

We begin, as above, by rewriting the equation in the form $x^4 + ax^2 + bx + c = 0$. Following Ferrari (1550s) we add an extra variable y to obtain the equation

$$(x^2 + a + y)^2 = (a + 2y)x^2 - bx + ((a + y)^2 - c) \tag{0.9.1}$$

and then select y so as to make the right-hand side the square of a linear polynomial $rx+s \in \mathbb{C}[x]$, in which case $(x^2+a+y)^2 = (rx+s)^2$, so that x is a root of one of the quadratic polynomials

$$(x^2+a+y) \pm (rx+s).$$

A quadratic polynomial is the square of a linear polynomial in x if and only if its discriminant equals 0. The right side of (0.9.1) has discriminant

$$b^2 - 4(a+2y)((a+y)^2 - c),$$

a cubic polynomial in y. We can find the roots, y, of this cubic polynomial by the method explained in the previous section. Given these roots, we can determine the possible values of r and s, and then we can solve for x to find the roots of the original equation.

Example. The roots of $X^4 + 4X^3 - 37X^2 - 100X + 300$. Letting $x = X + 1$ yields $x^4 - 43x^2 - 18x + 360$. Proceeding as above leads to the cubic equation $2y^3 - 215y^2 + 6676y - 64108 = 0$. Dividing through by 2 and then changing variable $y = t + 215/6$ gives the cubic $t^3 - (6169/12)t - (482147/108) = 0$. This has discriminant $-4(6169/12)^3 + 27(482147/108)^2 = -(2310)^2$. Hence $u^3, v^3 = (482147 \pm 27720\sqrt{-3})/216$. Unusually this has an exact cube root in terms of $\sqrt{-3}$; that is, $u, v = \omega^*(-37 \pm 40\sqrt{-3})/6$, where ω^* denotes ω to some power. Now $-3(-37+40\sqrt{-3})/6 \cdot (-37-40\sqrt{-3})/6 = -6169/12 = a$. Therefore we can take $u, v = (-37 \pm 40\sqrt{-3})/6$, and the roots of our cubic are $t = u + v = -37/3$, $\omega u + \omega^2 v = 157/6$, $\omega^2 u + \omega v = -83/6$ so that $y = 47/2, 62, 22$. From these, Ferrari's equation becomes $(x^2 - 39/2) = \pm(2x + 9/2)$ for $y = 47/2$ and so the possible roots $-5, 3; -4, 6$; or $(x^2 + 19) = \pm(9x + 1)$ for $y = 62$ and so the possible roots $-5, -4; 3, 6$; or $(x^2 - 21) = \pm(x + 9)$ for $y = 22$ and so the possible roots $-5, 6; 3, -4$. For each such y we get the same roots $x = 3, -4, -5, 6$, yielding the roots $X = 2, -5, -6, 5$ of the original quartic.

Example. A fun example is to find the fifth roots of unity, other than 1. That is those x satisfying $\frac{x^5-1}{x-1} = x^4 + x^3 + x^2 + x + 1 = 0$. Proceeding as above we find the four roots

$$\frac{\sqrt{5} - 1 \pm \sqrt{-2\sqrt{5} - 10}}{4}, \quad \frac{-\sqrt{5} - 1 \pm \sqrt{2\sqrt{5} - 10}}{4}.$$

0.10. Surds

A *surd* is a square root or a cube root or a higher root, that is, an nth root for some number $n \geq 2$. We have shown above that the roots of degree 2, 3, and 4 polynomials can be determined by taking a combination of surds. We would like to show something similar for polynomials of degree 5 and higher, which is the focus of a course on Galois theory.

Gauss's favorite example of surds was the expression for $\cos \frac{2\pi}{2^k}$, which we denote by $c(k)$. A double angle formula states that $\cos 2\theta = 2\cos^2 \theta - 1$, and so taking $\theta = 2\pi/2^k$ we have

$$c(k-1) = 2c(k)^2 - 1,$$

which may be rewritten as $c(k) = \frac{1}{2}\sqrt{2 + 2c(k-1)}$. Note that $c(k) \geq 0$ for $k \geq 2$ and $c(2) = 0$. Hence

$$c(3) = \frac{1}{2}\sqrt{2}, \quad c(4) = \frac{1}{2}\sqrt{2+\sqrt{2}}, \quad c(5) = \frac{1}{2}\sqrt{2+\sqrt{2+\sqrt{2}}}$$

and so we deduce by induction that

$$\cos\left(\frac{2\pi}{2^k}\right) = \frac{1}{2}\underbrace{\sqrt{2+\sqrt{2+\sqrt{2+\sqrt{2+\cdots+\sqrt{2}}}}}}_{k-2 \text{ times}} \qquad \text{for each } \; k \geq 3.$$

Why does expressing the roots of polynomials in terms of surds seem like a good idea? Are the roots, given explicitly as in the second example above, any more enlightening than simply saying that one has a root of the original equation? We can give arbitrarily good approximations to the value of any given irrational (and rather rapidly using the right software), so what is really the advantage of expressing the roots of polynomials in terms of surds? The answer is more to do with our comfort with certain concepts, and aesthetics, than any intrinsic notion. In the rather sophisticated *Galois theory* there are identifiable differences between these different types of expressions, but such concepts are best left to a more advanced course.

One can learn much more about these beautiful classical themes by studying the first six chapters of [**Tig16**].

References discussing solvability of polynomials

[1] Raymond G. Ayoub, *On the nonsolvability of the general polynomial*, Amer. Math. Monthly 89 (1982), 397–401.

[2] Harold M. Edwards, *The construction of solvable polynomials*, Bull. Amer. Math. Soc. **46** (2009), no. 3, 397–411.

[3] Blair K. Spearman and Kenneth S. Williams, *Characterization of solvable quintics* $x^5 + ax + b$, Amer. Math. Monthly 101 (1994), 986–992.

Appendix 0D. What is a group?

Mathematical objects are often structured into *groups*. It is easiest to prove results for arbitrary groups, so that these results apply for all examples of groups that arise.[6] Many of the main theorems about groups were first proved in a number theory context and then found to apply elsewhere.

0.11. Examples and definitions

The main examples of groups that you have encountered so far are *additive groups* such as the integers, the rationals, the complex numbers, polynomials of a given degree, and matrices of given dimensions; and *multiplicative groups* such as the non-zero rationals, the non-zero complex numbers, and invertible square matrices of given dimension. (The integers mod p, a notion we will introduce in chapter 2, also give rise to both an additive and a multiplicative group.)

We will now give the definition of a group—keep in mind the objects named in the last paragraph and the usual operations of addition and multiplication:

A *group* is defined to be a set of objects G and an operation, call it $*$, such that:

(i) If $a, b \in G$, then $a * b \in G$. We say that G is *closed* under $*$.

(ii) If $a, b, c \in G$, then $(a * b) * c = a * (b * c)$; that is, when multiplying three elements of G together it does not matter which pair we multiply first. We say that G is *associative*.

(iii) There exists an element $e \in G$ such that for every $a \in G$ we have $a * e = e * a = a$. We call e the *identity element* of G for $*$. (For a group in which "$*$" is

[6] One can waste a lot of energy giving the same proof, with minor variations, in each situation where a group arises. Gauss wrote *Disquisitiones* before the abstract notion of a group was formulated and therefore *does* give very similar proofs in different places when dealing with different examples of groups.

much like addition we typically denote the identity by 0; for a group in which "$*$" is much like multiplication we typically denote the identity by 1.)

(iv) For every $a \in G$ there exists $b \in G$ such that $a * b = b * a = e$. We call b the *inverse* of a. (For a group in which "$*$" is much like addition we write $-a$ for the inverse of a; for a group in which "$*$" is much like multiplication we write a^{-1} for the inverse of a.)

One can check that the examples of groups given above satisfy these criteria. We have given examples of both finite and infinite groups. Notice that neither the integers nor the polynomials form multiplicative groups, because there is no inverse to 2 in the integers and no inverse to x amongst the polynomials.

There is one familiar property of numbers and polynomials that is not used in the definition of a group, and that is that $a * b = b * a$, that a and b *commute*. Although this often holds, there are some simple counterexamples, for instance most pairs of 2-by-2 matrices *do not commute*: For example,

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1 & 2 \end{pmatrix} \quad \text{whereas} \quad \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

We develop the full theory for 2-by-2 matrices in the next section. If all pairs of elements of a group commute, then we call the group *commutative* or *abelian*. Typically we use multiplicative notation for groups that are non-commutative. It will be useful to develop a theory that works for non-commutative, as well as commutative, groups

A given group G can contain other, usually smaller, groups H, which are called *subgroups*. Every group G contains the subgroup given by the identity element, $\{0\}$ (the *trivial subgroup*), and also the subgroup G itself. It can also contain others; any subgroup other than G itself is a *proper subgroup*. For example the additive group of integers mod 6 with elements $\{0, 1, 2, 3, 4, 5\}$ contains the four subgroups

$$\{0\},\ \{0, 3\},\ \{0, 2, 4\},\ \{0, 1, 2, 3, 4, 5\}.$$

The middle two are non-trivial, proper subgroups. Note that every group, and so subgroup, contains the identity element. Infinite groups can also contain subgroups; indeed

$$\mathbb{C} \supset \mathbb{R} \supset \mathbb{Q} \supset \mathbb{Z}.$$

Exercise 0.11.1. Prove that if G a subgroup of $\mathbb{Z}$ under addition, then either $G = \{0\}$ or $G = m\mathbb{Z} := \{mn : n \in \mathbb{Z}\}$ for some integer $m \geq 1$.

0.12. Matrices usually don't commute

Let $\mathcal{M}_2(\mathbb{C})$ be the set of 2-by-2 matrices with entries in $\mathbb{C}$, and then define

$$\mathrm{Comm}(M) := \{A \in \mathcal{M}_2(\mathbb{C}) :\ AM = MA\}$$

for each $M \in \mathcal{M}_2(\mathbb{C})$. It is evident that if M is a multiple of the identity matrix I, then M commutes with all of $\mathcal{M}_2(\mathbb{C})$. Otherwise $\mathrm{Comm}(M)$ forms a 2-dimensional subspace of (the 4-dimensional) $\mathcal{M}_2(\mathbb{C})$, as we now prove:

Proposition 0.12.1. *If M is not a multiple of the identity matrix, then*

$$\mathrm{Comm}(M) = \{rI + sM :\ r, s \in \mathbb{C}\}.$$

Exercise 0.12.1. Let M be an n-by-n matrix.

(a) Prove that if A and B commute with M, then so does $rA + sB$ for any complex numbers r and s. (We call $rA + sB$ a *linear combination* of A and B.)
(b) Prove that M^k commutes with M, for all k.
(c) Deduce that all linear combinations of $I, M, \ldots, M^{n-1}$ belong to $\mathrm{Comm}(M)$.

Exercise 0.12.2. Let $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(a) Prove that M is not a multiple of I if and only if at least one of $a \neq d$, $b \neq 0$, $c \neq 0$ holds.
(b) Prove that if $a \neq d$, then for any matrix A there exists $r, s \in \mathbb{C}$ such that $A - rI - sM$ has zeros down the diagonal.
(c) Prove that if $b \neq 0$, then for any matrix A there exists $r, s \in \mathbb{C}$ such that $A - rI - sM$ has zeros throughout the top row.
(d) Prove that if $c \neq 0$, then for any matrix A there exists $r, s \in \mathbb{C}$ such that $A - rI - sM$ has zeros throughout the first column

Proof of Proposition 0.12.1. It is evident that I and M commute with M, and hence M commutes with any linear combination of I and M by exercise 0.12.1. We now show that these are the only matrices that commute with M.

Let $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If $A \in \mathrm{Comm}(M)$, then $B = A - rI - sM \in \mathrm{Comm}(M)$ for any r and $s \in \mathbb{C}$, by exercise 0.12.1.

If $a \neq d$, then we select r and s as in exercise 0.12.2(a) so that $B = \begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$ for some $x, y \in \mathbb{C}$. As $B \in \mathrm{Comm}(M)$, we have

$$\begin{pmatrix} cx & dx \\ ay & by \end{pmatrix} = \begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = BM = MB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix} = \begin{pmatrix} by & ax \\ dy & cx \end{pmatrix}.$$

Comparing the off-diagonal terms on the left- and right-hand ends of the equation forces $x = y = 0$ (as $a \neq d$), so that $B = 0$, and therefore $A = rI + sM$.

If $a = d$, $b \neq 0$, and $M \neq aI$, then B may be written in the form $B = \begin{pmatrix} 0 & 0 \\ x & y \end{pmatrix}$ for some $x, y \in \mathbb{C}$, by exercise 0.12.2(b), so that

$$\begin{aligned} \begin{pmatrix} 0 & 0 \\ ax + cy & bx + dy \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ x & y \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = BM = MB \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 0 \\ x & y \end{pmatrix} = \begin{pmatrix} bx & by \\ dx & dy \end{pmatrix}. \end{aligned}$$

Comparing the terms in the top row on the left- and right-hand ends of the equation forces $x = y = 0$ (as $b \neq 0$), so that $B = 0$, and therefore $A = rI + sM$.

If $a = d$, $b = 0$, $c \neq 0$, and $M \neq aI$, then we may proceed analogously to the previous paragraph. Alternatively we may note that the result for M^T (the transpose of M) is given by the previous paragraph and then follows for M since $BM = MB$ if and only if $B^T M^T = M^T B^T$.

Finally if $a = d$ and $b = c = 0$, then M is a multiple of I. □

Appendix 0E. Rings and fields

In section 0.11 of appendix 0D we introduced the notion of a group and gave various examples. The real numbers are a set of objects that remarkably support two different groups: There is both the additive group and the multiplicative group acting on the non-zero real numbers, and this partly explains why they play such a fundamental role in mathematics. In this appendix, we formalize these notions and the key differences between the structure of the real numbers and of the integers. This will allow us to better identify the properties of many important types of numbers that arise in number theory.

0.13. Mixing addition and multiplication together: Rings and fields

A set of numbers A which, like the reals, has an additive group on A with identity element 0 and a commutative multiplicative group on $A \setminus \{0\}$ and for which the two groups interact according the *distributive* properties,

$$a \times (b + c) = (a \times b) + (a \times c) \quad \text{and} \quad (a + b) \times c = (a \times c) + (b \times c),$$

is called a *field*. The reals $\mathbb{R}$ are an example, as are $\mathbb{C}$ and $\mathbb{Q}$. A field provides the most convenient situation in which to do arithmetic.

Exercise 0.13.1. Prove that $a \times 0 = 0$ for all $a \in A$, when A is a field.

However the integers, $\mathbb{Z}$, which are also vital to arithmetic, do not form a field: Although they form a group under addition, they do not form a group under multiplication, since not every integer has a multiplicative inverse within the integers (for example, the multiplicative inverse of 2 is $1/2$ which is not an integer). But you can multiply integers together, and the integers possess a multiplicative identity, 1, so they have some of the properties of a multiplicative group, but not all. The integers are an example of a *ring*, which is a set of objects that form an additive

group, are closed under multiplication, and have a multiplicative identity, 1, as well as satisfying the above distributive properties.[7] Thus $\mathbb{Z}$ is a commutative ring.

The set of even integers, $2\mathbb{Z}$, narrowly fails being a ring; it simply lacks the multiplicative identity. The polynomials with integer coefficients, $\mathbb{Z}[x]$, form a ring. Indeed if A is a commutative ring, then $A[x]$ is also a commutative ring.

For a given ring or field A and object α that is not in A, we are often interested in what type of mathematical object is created by adjoining α to A. This may be done in more than one way:

- $A[\alpha]$, which denotes polynomials in α with coefficients in A, that is, expressions of the form $a_0 + a_1\alpha + \cdots + a_d\alpha^d$ for any $d \geq 0$ where each $a_i \in A$.
- $A(\alpha)$, which denotes rational functions in α with coefficients in A, or, more simply, quotients u/v with $u, v \in A[\alpha]$ and $v \neq 0$.

For example we will prove in section 3.4 that $\sqrt{2}$ is irrational (that is, $\sqrt{2} \notin \mathbb{Q}$, and so $\sqrt{2} \notin \mathbb{Z}$), so we would like to understand the sets $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Z}[\sqrt{2}]$.

Exercise 0.13.2. (a) Prove that $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} : \ a, b \in \mathbb{Z}\}$ and that $\mathbb{Z}[\sqrt{2}]$ is a ring.
(b) Prove that $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}$ and that $\mathbb{Q}(\sqrt{2})$ is a field.

0.14. Algebraic numbers, integers, and units, I

If $f(x) = \sum_{j=0}^{d} f_j x^j$ where $f_d \neq 0$, then $f(x)$ has *degree* d and *leading coefficient* f_d. We say that $f(x)$ is *monic* if $f_d = 1$. If all of the coefficients, f_j, of $f(x)$ are integers, then we write $f(x) \in \mathbb{Z}[x]$. The *content* of $f(x) \in \mathbb{Z}[x]$ is the largest integer that divides all of its coefficients. If m is the content of $f(x)$, then $f(x) = mg(x)$ for some $g(x) \in \mathbb{Z}[x]$ of content 1. Obviously m divides every value $f(n)$ with $n \in \mathbb{Z}$ but there could be other integers that also have this property. For example, $f(x) = x^2 + x + 2$ has content 1, but 2 divides $f(n)$ for every integer n.

We call $\alpha \in \mathbb{C}$ an *algebraic number* if it is a root of a polynomial $f(x) \in \mathbb{Z}[x]$, with integer coefficients. If f is monic, then α is an *algebraic integer*. We call $f(x)$ the *minimal polynomial* for α if f is the polynomial with integer coefficients, of smallest degree, with positive leading coefficient and of content 1, for which $f(\alpha) = 0$. Minimal polynomials are irreducible in $\mathbb{Z}[x]$, for if $f(x) = g(x)h(x)$ with $g(x), h(x) \in \mathbb{Z}[x]$, then $g(\alpha)h(\alpha) = f(\alpha) = 0$ so that either $g(\alpha) = 0$ or $h(\alpha) = 0$; and therefore α is a root of a polynomial of lower degree than f, contradicting minimality.

Exercise 0.14.1. Let $f(x)$ be the minimal polynomial of an algebraic number α.
(a) Prove that if $g(x)$ is a polynomial with integer coefficients for which $g(\alpha) = 0$, then $f(x)$ divides $g(x)$. (You may use Proposition 2.10.1 of appendix 2B in your proof.)
(b) Prove that if $f(x)$ divides $g(x) \in \mathbb{Z}[x]$ and g is monic, then f is monic. Deduce that if $g(\alpha) = 0$, then α is an algebraic integer.
(c) Prove that if $g(\alpha) = 0$ and g is irreducible, then $g = \kappa f$ for some constant $\kappa \neq 0$.
(d) Prove that $f(x)$ is the only minimal polynomial of α.
(e) Prove that $(x - \alpha)^2$ does not divide $f(x)$.

[7] For the sake of comparison, a ring does not necessarily have two of the properties of a field: The numbers in a ring do not necessarily have a multiplicative inverse, and they do not necessarily commute when multiplying them together.

Exercise 0.14.2. Prove that if α is an algebraic number and a root of $f(x) \in \mathbb{Z}[x]$ where f has leading coefficient a, then $a\alpha$ is an algebraic integer.

Exercise 0.14.3. What are the algebraic integers in $\mathbb{Q}$?

Exercise 0.14.4. (a) Prove that $\mathbb{Z}[\sqrt{d}]$ is a subset of the algebraic integers.
(b) Prove that $\mathbb{Z}[\sqrt{2}]$ is the set of algebraic integers in $\mathbb{Q}(\sqrt{2})$.
(c) Prove that $\frac{1+\sqrt{5}}{2}$ is an algebraic integer.

If α is an algebraic integer, then so is $m\alpha+n$ for any integers m, n, for if $f(x)$ is the minimal polynomial of α and has degree d, then $F(x) := m^d f(\frac{x-n}{m})$ is a monic polynomial in $\mathbb{Z}[x]$ with root $m\alpha + n$.

If α is a non-zero algebraic number, with minimal polynomial $f(x)$ of degree d, then $1/\alpha$ is a root of $x^d f(1/x)$.

Exercise 0.14.5. (a) Prove that $1/\alpha$ has minimal polynomial $x^d f(1/x)$.
(b) Prove that α and $1/\alpha$ are both algebraic integers if and only if f is monic and $f(0) = 1$ or -1. In this case α and $1/\alpha$ are called *units*.

Another way to view this is that α is a unit if and only if α divides 1, for if $\beta = 1/\alpha$, then $\alpha\beta = 1$ and α and β are both algebraic integers.

Exercise 0.14.6. Suppose that α and β are algebraic integers such that α divides β, and β divides α. Prove that there exists a unit u for which $\beta = u\alpha$

In the section 0.17 of appendix 0F we will prove that if α and β are algebraic numbers, then so are $\alpha + \beta$ and $\alpha\beta$. Moreover if α and β are algebraic integers, then so are $\alpha + \beta$ and $\alpha\beta$.

Exercise 0.14.7. (a) Prove that if α is an algebraic number, then $\mathbb{Q}(\alpha)$ is a field.
(b) Prove that if α is an algebraic integer, then $\mathbb{Z}[\alpha]$ is a ring.

Do there exist numbers α that are *irrational* (that is, that are not rational), so that, for instance, $\mathbb{Q}(\alpha)$ is not the same thing as $\mathbb{Q}$? To determine what numbers are irrational we should first classify, in a useful way, the rational numbers. The minimal polynomial of a rational number p/q with $(p, q) = 1$ is $qx - p \in \mathbb{Z}[x]$. Therefore one way to show that an algebraic number is irrational is to show that its minimal polynomial has degree > 1. Therefore given a polynomial, say x^2-2, we have to decide whether it is the minimal polynomial for some number, or perhaps prove that it is irreducible (so that it cannot have a rational root). We will develop number-theoretic tools to do this. Another way to find irrational numbers is to perhaps show that there are numbers that are not the roots of *any* polynomial in $\mathbb{Z}[x]$. Such numbers are not only irrational but are not even algebraic numbers and so are called *transcendental*. It is not too difficult to prove that transcendental numbers exist (by the "diagonalization argument" given in section 11.16 of appendix 11D), but it is rather more subtle to determine an actual transcendental number (though we will do so, using number-theoretic ideas in chapter 11).

For much much more on university level algebra, much of which stems from number theory, the reader might care to look at the excellent textbook [**DF04**] by Dummit and Foote or the more advanced but number theoretic [**IR90**].

Appendix 0F. Symmetric polynomials

It is difficult to work with algebraic numbers since one cannot necessarily evaluate them precisely. For example the golden ratio, $\frac{1+\sqrt{5}}{2}$, can easily be well-approximated, but how can you determine its precise value (since it is irrational)?

We can often avoid working with the actual algebraic numbers themselves, but rather work with the set of all of the roots of the minimal polynomial. For example, the formula (0.1.1) for the nth Fibonacci number involves both the golden ratio and the other root of its minimum polynomial $x^2 - x - 1$. It was Sir Isaac Newton who recognized that a function that is symmetric in all of the roots of a given polynomial is a rational number.

0.15. The theory of symmetric polynomials

We say that $P(x_1, x_2, \dots, x_n)$ is a *symmetric polynomial* if

$$P(x_k, x_2, \dots, x_{k-1}, x_1, x_{k+1}, \dots, x_n) = P(x_1, x_2, \dots, x_n) \text{ for each } k.$$

Here we swapped x_1 and x_k and kept everything else the same.

Exercise 0.15.1. Show that for any permutation σ of $1, 2, \dots, n$ and any symmetric polynomial P we have $P(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = P(x_1, x_2, \dots, x_n)$.

Theorem 0.3 (The fundamental theorem of symmetric polynomials)**.** *For a given monic polynomial $f(x) = \sum_{i=0}^{d} a_i x^i$ with integer coefficients, each symmetric polynomial in the roots of f, with integer coefficients, can be expressed as a polynomial in the a_i with rational coefficients.*

Proof. Let $f(x) = \prod_{i=1}^{d}(x - \alpha_i)$. We begin by proving the claim for the

$$s_k := \sum_{i=1}^{d} \alpha_i^k \quad \text{for each} \quad k \geq 0.$$

Multiplying out $f(x) = \prod_{i=1}^{d}(x - \alpha_i)$ we have

$$\sum_i \alpha_i = -a_1,\ \sum_{i<j} \alpha_i\alpha_j = a_2,\ \sum_{i<j<k} \alpha_i\alpha_j\alpha_k = -a_3,\ \ldots,\ \alpha_1\alpha_2\ldots\alpha_n = \pm a_n.$$

Then, since $\frac{f'(x)}{f(x)} = \sum_{i=1}^{d} \frac{1}{x-\alpha_i}$, we have

$$\frac{\sum_{j=0}^{d} j a_j x^{d-j}}{\sum_{i=0}^{d} a_i x^{d-i}} = \frac{x^{d-1}}{x^{d-1}} \cdot \frac{f'(1/x)}{xf(1/x)} = \sum_{i=1}^{d} \frac{1}{1-\alpha_i x} = \sum_{i=1}^{d}\sum_{k\geq 0} (\alpha_i x)^k = \sum_{k\geq 0} s_k x^k.$$

This implies that

$$\sum_{j=0}^{d} (d-j) a_{d-j} x^j = \sum_{i=0}^{d} a_{d-i} x^i \cdot \sum_{k\geq 0} s_k x^k = \sum_{N\geq 0} \left(\sum_{\substack{i+k=N \\ 0\leq i\leq d}} a_{d-i} s_k \right) x^N.$$

Comparing the coefficients of x^N, we obtain (as $a_d = 1$)

$$s_N = -\sum_{i=1}^{\min\{d,N\}} a_{d-i} s_{N-i} + \begin{cases} (d-N)a_{d-N} & \text{if } N < d, \\ 0 & \text{if } N \geq d. \end{cases}$$

Hence, by induction on N, we see that the s_N are polynomials in the a_j.

We now sketch a proof of Newton's result for arbitrary symmetric polynomials, by showing that every symmetric polynomial in the roots of f, with integer coefficients, can be written as a polynomial in the s_k with integer coefficients. We proceed by induction on the number r of variables in the monomials of the symmetric polynomial; that is, we select the monomials $c\alpha_{i_1}^{k_1}\alpha_{i_2}^{k_2}\ldots\alpha_{i_r}^{k_r}$ in f, with each $k_i \geq 1$, for which r is maximal. In the $r = 1$ case, our polynomial is simply a linear combination of the s_k. Suppose that $r > 1$. If the k_i are distinct in such a monomial, we subtract $cs_{k_1}s_{k_2}\ldots s_{k_r}$ and we are left with various cross terms but, in all of which, two or more of the variables α_j are equal. If the k_i are not all distinct, then we subtract $s_{k_1}s_{k_2}\ldots s_{k_r}/\prod_i m_i!$ where $m_i := \#\{j : k_j = i\}$, to obtain various cross terms, with the same property. Hence, in the remaining expression, each monomial contains fewer variables and the result follows by induction. □

Exercise 0.15.2. If f is not monic, develop analogous results by working with $g(x)$ defined by $g(a_d x) = a_d^{d-1} f(x)$.

Example. Look at $\sum_{i,j,k} \alpha_i\alpha_j^2\alpha_k^3$. Subtract $s_1s_2s_3$ and we have to account for the cases where $i = j$ or $i = k$ or $j = k$. Hence what remains is

$$-\sum_{i,k} \alpha_i^3\alpha_k^3 - \sum_{i,j} \alpha_i^4\alpha_j^2 - \sum_{i,j} \alpha_i\alpha_j^5 + 2s_6,$$

where in the first sum we have $i = j$, in the second $i = k$, in the third $j = k$, and in the last $i = j = k$ (the coefficients being chosen by inclusion-exclusion). Proceeding

the same way again we have

$$\sum_{i,j} \alpha_i^4 \alpha_j^2 = s_4 s_2 - s_6, \quad \sum_{i,j} \alpha_i \alpha_j^5 = s_1 s_5 - s_6, \quad \text{and} \quad \sum_{i,k} \alpha_i^3 \alpha_k^3 = (s_3^2 - s_6)/2,$$

the last since in s_3^2 the cross term $\alpha_i^3 \alpha_k^3$ appears also as $\alpha_k^3 \alpha_i^3$. Collecting this all together yields

$$\sum_{i,j,k} \alpha_i \alpha_j^2 \alpha_k^3 = s_1 s_2 s_3 - s_1 s_5 - s_2 s_4 - s_3^2/2 + 9s_6/2.$$

Throughout these calculations, the sum of the indices in each term is 6, the degree of the original polynomial.

0.16. Some special symmetric polynomials

If α and β are the two roots of a monic quadratic polynomial with integer coefficients, then $x_n = (\alpha^n - \beta^n)/(\alpha - \beta)$ is a symmetric function in α and β and hence must be an integer by the fundamental theorem of symmetric polynomials. (We saw in exercise 0.1.4(c) that this is the nth term of the general second-order linear recurrence sequence that starts $0, 1$.)

If α is a root of an irreducible polynomial $f(x) = a\prod_{i=1}^{d}(x - \alpha_i)$, then there are two symmetric polynomials of particular interest:

The *trace* of α is $\alpha_1 + \alpha_2 + \cdots + \alpha_d$, the sum of the roots of f.

The *norm* of α is $\alpha_1 \alpha_2 \ldots \alpha_d$, the product of the roots of f.

By the fundamental theorem of symmetric polynomials, the trace and the norm of an irreducible polynomial are rational numbers.

Exercise 0.16.1. Show that if $f(t) = \prod_{i=1}^{k}(t - \alpha_i) \in \mathbb{Z}[t]$, then $\prod_{j=1}^{d} f'(\alpha_j)$ is an integer, by using the theory of symmetric polynomials.

Using the product rule we see that

$$f'(t) \;=\; a\sum_{j=1}^{k} \prod_{\substack{1\le i\le k\\ i\ne j}} (t - \alpha_i) \text{ and so } f'(\alpha_j) \;=\; a \prod_{\substack{1\le i\le k\\ i\ne j}} (\alpha_j - \alpha_i).$$

We deduce that

$$\prod_{j=1}^{d} f'(\alpha_j) \;=\; a^d \prod_{1\le i<j\le d} (-(\alpha_i - \alpha_j)^2).$$

This is a symmetric polynomial in the roots α_i, and so by Newton's fundamental theorem of symmetric polynomials it must be a rational number.

Let's evaluate this product for the quadratic polynomial $ax^2 + bx + c$. If this has roots α and β, then $ax^2 + bx + c = a(x - \alpha)(x - \beta)$ and so

$$a(\alpha + \beta) = -b \quad \text{and} \quad a\alpha\beta = c.$$

Therefore

$$a^2(\alpha - \beta)^2 \;=\; (a(\alpha + \beta))^2 - 4a(a\alpha\beta) \;=\; b^2 - 4ac,$$

the discriminant of the polynomial, $ax^2 + bx + c$.

For the cubic polynomial $x^3 + ax + b = (x - \alpha)(x - \beta)(x - \gamma)$ we have

$$\alpha + \beta + \gamma = 0, \quad \alpha\beta + \alpha\gamma + \beta\gamma = a, \quad \text{and} \quad \alpha\beta\gamma = -b.$$

But then $\gamma = -(\alpha + \beta)$ so that $\alpha^2 + \alpha\beta + \beta^2 = -a$ and $\alpha\beta(\alpha + \beta) = b$. Therefore

$$(-1)^3((\alpha - \beta)(\gamma - \beta)(\alpha - \gamma))^2 \;=\; -((\alpha - \beta)(\alpha + 2\beta)(2\alpha + \beta))^2 \;=\; 4a^3 + 27b^2,$$

which is the discriminant of the polynomial $x^3 + ax + b$.

A beautifully symmetric function is given by the Vandermonde matrix. The 3-by-3 version is

$$\begin{pmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{pmatrix},$$

which has determinant $(x-y)(y-z)(z-x)$. This is not quite a symmetric function since swapping any two variables multiplies the determinant by -1. (This is also apparent when swapping any two columns of the matrix.) One intuitive way to see that $(x - y)(y - z)(z - x)$ is the determinant is by showing that each factor separately and together divides the determinant. To begin with, if $x = y$, then the determinant equals 0 as the first two columns are now equal, which implies that $x - y$ must be a factor of the determinant. Similarly $x - z$ and $z - y$ also divide the determinant. If x, y, and z are variables, then these expressions do not have any common factors, and so their product divides the determinant. This product has degree three (adding the degrees of the variables), as does the determinant, so they can differ by at most a constant factor. The constant factor can be determined by checking the coefficient of a particular monomial on both sides. For example $x^0y^1z^2$ only arises in the determinant from multiplying out the terms of the main diagonal and therefore has coefficient 1, and one can equally look for how this monomial arises in the product.[8]

Exercise 0.16.2. Use the same argument to explain that the determinant of *Vandermonde matrix* V, where $V_{i,j} = \alpha_i^{j-1}$, $1 \le i, j \le d$, is $\prod_{1 \le i < j \le d}(\alpha_i - \alpha_j)$.

Exercise 0.16.3. Prove Theorem 0.2 when each $e_j = 1$ (assuming exercise 0.16.2).

Now

$$(V^TV)_{i,k} = \sum_{j=1}^{d} (V^T)_{i,j}V_{j,k} = \sum_{j=1}^{d} V_{j,i}V_{j,k} = \sum_{j=1}^{d} \alpha_j^{i-1}\alpha_j^{k-1} = s_{i+k-2}$$

for $1 \le i, k \le d$. Hence $(\det V)^2$ is the determinant of the matrix with (i, k)th entry s_{i+k-2}.

Exercise 0.16.4.‡ (This question requires some knowledge of linear algebra.) Suppose that M is an n-by-n matrix.

(a) Prove that if M is a diagonal matrix in which all the diagonal entries are distinct, then $\mathrm{Comm}(M)$ equals the set of diagonal matrices.

(b) Use exercise 0.16.2 to show that the set of diagonal matrices is then given by $\{a_0I + a_1M + \cdots + a_{n-1}M^{n-1} : \text{ each } a_j \in \mathbb{C}\}$.

[8] Diehard algebraists might be uncomfortable with this discussion since we ignore ideals that arise from the gcds of the polynomial factors, but these details can all be justified.

(c) Now let M, N, and T be n-by-n matrices with T invertible. Prove that M and N commute if and only if $T^{-1}MT$ and $T^{-1}NT$ commute.
(d) Prove that if M is an n-by-n matrix with n distinct eigenvalues, then $\mathrm{Comm}(M) = \{a_0 I + a_1 M + \cdots + a_{n-1}M^{n-1} : \text{ each } a_j \in \mathbb{C}\}$.

0.17. Algebraic numbers, integers, and units, II

We are now in a position to prove some of the claims of section 0.14 of appendix 0E. Suppose that α and β are algebraic integers with minimal polynomials f and g. Then

$$\prod_{\substack{u:\ f(u)=0 \\ v:\ g(v)=0}} (x-(u+v)) = \prod_{u:\ f(u)=0} g(x-u),$$

by exercise 0.14.1(d). This is a symmetric polynomial in the roots u of f and so, by the fundamental theorem of symmetric polynomials, this is a monic polynomial with integer coefficients having root $\alpha+\beta$, and so $\alpha+\beta$ is an algebraic integer.

Exercise 0.17.1. (a)† Prove that if $\alpha \neq 0$ and β are algebraic integers, then $\alpha\beta$ is also an algebraic integer.
(b) Prove that if $\alpha \neq 0$ and β are algebraic numbers, then $\alpha+\beta$ and $\alpha\beta$ are algebraic numbers.

Exercise 0.17.2. Prove that if $\alpha_1, \ldots, \alpha_k$ are algebraic numbers, then $\mathbb{Q}(\alpha_1, \ldots, \alpha_k)$ is a field. These are the *number fields*.

Let $\overline{\mathbb{Q}}$ denote the set of all algebraic numbers. Evidently if K is any number field, then $K \subset \overline{\mathbb{Q}}$. It is not difficult to prove that $\overline{\mathbb{Q}}$ is itself a field. Similarly if A is the set of all algebraic integers, then A is a ring and the algebraic integers inside a given number field K form a subring, which is precisely $K \cap A$. However identifying the elements of $K \cap A$ explicitly can be rather more challenging, as we saw in exercise 0.14.4.

Rather more interestingly, the roots of any polynomial with coefficients in $\overline{\mathbb{Q}}$ all belong to $\overline{\mathbb{Q}}$.

Proposition 0.17.1. *Suppose that $f(x) \in \overline{\mathbb{Q}}[x]$ and that $f(\rho) = 0$. Then $\rho \in \overline{\mathbb{Q}}$. We say that $\overline{\mathbb{Q}}$ is* algebraically closed.

Proof. Suppose that $f(x) = a_0 + a_1 x + \cdots + a_d x^d$ so that each a_j is an algebraic number. Suppose that a_j has minimal polynomial $g_j(x)$; and let A_j be the set of roots of $g_j(x)$. Then $f(x)$ divides the polynomial

$$F(x) := \prod_{\alpha_0 \in A_0} \prod_{\alpha_1 \in A_1} \cdots \prod_{\alpha_d \in A_d} (\alpha_0 + \alpha_1 x + \cdots + \alpha_d x^d)$$

which is a symmetric polynomial in the elements of each A_j with $0 \leq j \leq d$ and therefore belongs to $\mathbb{Q}[x]$ by the law of symmetric polynomials. Any root of $f(x)$ is a root of $F(x)$ and therefore must be an algebraic number. □

For further development of these ideas see chapter 8 of [**Tig16**].

Appendix 0G. Constructibility

0.18. Constructible using only compass and ruler

The ancient Greeks were interested in what could be constructed using only a straight edge (sometimes called an "unmarked ruler", or just plain "ruler") and a compass. Three questions stumped them:

Quadrature of the circle:

Draw a square that has area equal to that of a given circle.

To draw a square whose area is π (the same area as a circle of radius 1), we need to be able to draw a square with sides of length x, where

$$x \text{ is a root of the equation } x^2 - \pi.$$

Duplication of the cube:

Construct a cube that has twice the volume of a given cube.

If the original cube has side length 1 (and so volume 1), we would need to be able to construct a cube with sides of length x, where

$$x \text{ is a root of the equation } x^3 - 2.$$

Trisection of the angle:

Construct an angle which is one third the size of a given angle.

Constructing an angle θ is as difficult as constructing a right-angled triangle containing that angle, that is, the triangle with side lengths $\sin\theta,\ \cos\theta,\ 1$. Therefore if we start with angle 3θ and wish to determine the angle θ, then we will need to be able to determine $\cos\theta$ from $\cos 3\theta$ and $\sin 3\theta$. But these are linked by the formula $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$; that is, we need to find the root $x = 2\cos\theta$ of $x^3 - 3x - A$ where $A = 2\cos 3\theta$. For example, if $\theta = \pi/9$, we will need to be able to construct a right-angled triangle with a side of length $x/2$ where

$$x \text{ is a root of the equation } x^3 - 3x - 1.$$

We need to understand the algebra of points that are constructed from given points and lengths by "ruler and compass". Our tools are:

• *An unmarked ruler*, which allows us to draw the line between any two given points and to extend that line as far as we like.

• *A compass*, which allows us to draw the circle centered on one given point, of radius a given length, or the distance between two given points.

Proposition 0.18.1. *Given a set of points on lines and a set of lengths, any new points that can be constructed from these, using only ruler and compass, will have coordinates that can be determined as the roots of degree-one or degree-two polynomials, whose coefficients are rational functions of the already known coordinates.*

Proof. The lines are defined by pairs of points: Given the points $A = (a_1, a_2)$ and $B = (b_1, b_2)$ the line between them is $(b_1 - a_1)(y - a_2) = (b_2 - a_2)(x - a_1)$.

Exercise 0.18.1. Show that the coefficients of the equation of this line can be determined by a degree-one equation in already known coordinates.

Exercise 0.18.2. Prove that any two (non-parallel) lines intersect in a point that can be determined by a degree-one equation in the coefficients of the equations of the lines.

Given a length r and a point $C = (c_1, c_2)$, we can draw the circle $(x - c_1)^2 + (y - c_2)^2 = r^2$ centered at C of radius r.

Exercise 0.18.3. Prove that the points of intersection of this circle with a given line can be given by a degree-two equation in already known coordinates.

Exercise 0.18.4. Prove that the points of intersection of two circles can be given by a degree-two equation in already known coordinates.

Combining all these exercises implies Proposition 0.18.1. □

We sketch here how one uses Proposition 0.18.1 to show that the Greeks were stumped by their three questions for good reason—none of the three were possible. Proposition 0.18.1 implies that we can draw a square that has area equal to that of a given circle if and only if π can be obtained in terms of a (finite) succession of roots of linear or quadratic polynomials whose coefficients are already constructed. If this can be done, then π would be the root of some polynomial (perhaps of high degree); in other words π would be an algebraic number. However Lindemann proved, in 1882, that π is transcendental (as we will discuss in more detail in section 11.17 of appendix 11D).

If α is obtained from a (finite) succession of roots of linear or quadratic polynomials whose coefficients are already constructed, then α is not only an algebraic number but one can show that its minimal polynomial has degree which is a power of 2. Both $x^3 - 2$ and $x^3 - 3x - 1$ are irreducible (which can be shown using Theorem 3.4; see exercise 3.4.4), and so these are the minimum polynomials for their roots (by exercise 0.14.1(c)). Therefore one cannot duplicate the cube, nor trisect the angle $\pi/3$, since the roots of these irreducible polynomials of degree three do not have minimum polynomials that have degrees that are a power of 2.

For further development of these ideas see section 13.3 of [**DF04**] or section 9.11 of [**IR90**].

The Euclidean algorithm

1.1. Finding the gcd

Most readers will know the Euclidean algorithm, used to find the greatest common divisor (gcd) of two given integers. For example, to determine the greatest common divisor of 85 and 48, we begin by subtracting the smaller from the larger, 48 from 85, to obtain $85 - 48 = 37$. Now $\gcd(85, 48) = \gcd(48, 37)$, because the common divisors of 48 and 37 are precisely the same as those of 85 and 48, and so we apply the algorithm again to the pair 48 and 37. So we subtract the smaller from the larger to obtain $48 - 37 = 11$, so that $\gcd(48, 37) = \gcd(37, 11)$. Next we should subtract 11 from 37, but then we would only do so again, and a third time, so let's do all that in one go and take $37 - 3 \times 11 = 4$, to obtain $\gcd(37, 11) = \gcd(11, 4)$. Similarly we take $11 - 2 \times 4 = 3$, and then $4 - 3 = 1$, so that the gcd of 85 and 48 is 1. This is the Euclidean algorithm that you might already have seen,[1] but did you ever prove that it really works?

To do so, we will first carefully define terms that we have implicitly used in the above paragraph, perhaps mathematical terms that you have used for years (such as "divides", "quotient", and "remainder") without a formal definition. This may seem pedantic but the goal is to make sure that the rules of basic arithmetic are really established on a sound footing.

Let a and b be given integers. We say that a *is divisible by* b, or that b divides a,[2] if there exists an integer q such that $a = qb$. For convenience we write "$b \mid a$".[3,4] We now set an exercise for the reader to check that the definition allows one to manipulate the notion of division in several familiar ways.

Exercise 1.1.1. In this question, and throughout, we assume that a, b, and c are integers.
(a) Prove that if b divides a, then either $a = 0$ or $|a| \geq |b|$.

[1] There will be a formal discussion of the Euclidean algorithm in appendix 1A.
[2] One can also say a *is a multiple of* b or b *is a divisor of* a or b *is a factor of* a.
[3] And if b does not divide a, we write "$b \nmid a$".
[4] One reason for giving a precise mathematical definition for division is that it allows us to better decide how to interpret questions like, "What is 1 divided by 0?" or "What is 0 divided by 0?"

(b) Deduce that if $a|b$ and $b|a$, then $b = a$ or $b = -a$ (which, in future, we will write as "$b = \pm a$").
(c) Prove that if a divides b and c, then a divides $bx + cy$ for all integers x, y.
(d) Prove that a divides b if and only if a divides $-b$ if and only if $-a$ divides b.
(e) Prove that if a divides b, and b divides c, then a divides c.
(f) Prove that if $a \neq 0$ and ac divides ab, then c divides b.

Next we formalize the notion of "dividing with remainder".

Lemma 1.1.1. *If a and b are integers, with $b \geq 1$, then there exist unique integers q and r, with $0 \leq r \leq b-1$, such that $a = qb + r$. We call q the "*quotient*", and r the "*remainder*".*

Proof by induction. We begin by proving the existence of q and r. For each $b \geq 1$, we proceed by induction on $a \geq 0$. If $0 \leq a \leq b-1$, then the result follows with $q = 0$ and $r = a$. Otherwise assume that the result holds for $0, 1, 2, \ldots, a-1$, where $a \geq b$. Then $a - 1 \geq a - b \geq 0$ so, by the induction hypothesis, there exist integers Q and r, with $0 \leq r \leq b-1$, for which $a - b = Qb + r$. Therefore $a = qb + r$ with $q = Q + 1$.

If $a < 0$, then $-a > 0$ so we have $-a = Qb + R$, for some integers Q and R, with $0 \leq R \leq b-1$, by the previous paragraph. If $R = 0$, then $a = qb$ where $q = -Q$ (and $r = 0$). Otherwise $1 \leq R \leq b-1$ and so $a = qb + r$ with $q = -Q - 1$ and $1 \leq r = b - R \leq b - 1$, as required.

Now we show that q and r are unique. If $a = qb + r = Qb + R$, then b divides $(q - Q)b = R - r$. However $0 \leq r, R \leq b-1$ so that $|R - r| \leq b-1$, and $b \mid R - r$. Therefore $R - r = 0$ by exercise 1.1.1(a), and so $Q - q = 0$. In other words $q = Q$ and $r = R$; that is, the pair q, r is unique. □

An easier, but less intuitive, proof. We can add a multiple of b to a to get a positive integer. That is, there exists an integer n such that $a + nb \geq 0$; any integer $n \geq -a/b$ will do. We now subtract multiples of b from this number, as long as it remains positive, until subtracting b once more would make it negative. In other words we now have an integer $a - qb \geq 0$, which we denote by r, such that $r - b < 0$; in other words $0 \leq r \leq b - 1$. □

Exercise 1.1.2. Suppose that $a \geq 1$ and $b \geq 2$ are integers. Show that we can write a in base b; that is, show that there exist integers $a_0, a_1, \ldots \in [0, b-1]$ for which $a = a_d b^d + a_{d-1} b^{d-1} + a_1 b + a_0$.

We say that d is a *common divisor* of integers a and b if d divides both a and b. We are mostly interested in the *greatest common divisor* of a and b, which we denote by $\gcd(a, b)$, or more simply as (a, b).[5,6]

We say that a is *coprime* with b, or that a and b are *coprime integers*, or that a and b are *relatively prime*, if $(a, b) = 1$.

[5] In the UK this is known as the *highest common factor* of a and b and is written $\mathrm{hcf}(a, b)$.

[6] When $a = b = 0$, every integer is a divisor of 0, so there is no greatest divisor, and therefore $\gcd(0, 0)$ is undefined. There are often one or two cases in which a generally useful mathematical definition does not give a unique value. Another example is 0 divided by 0, which we explore in exercise 1.7.1. For aesthetic reasons, some authors choose to assign a value which is consistent with the theory in one situation but perhaps not in another. This can lead to artificial inconsistencies which is why we choose to leave such function-values undefined.

Corollary 1.1.1. *If $a = qb + r$ where a, b, q, and r are integers, then*

$$\gcd(a, b) = \gcd(b, r).$$

Proof. Let $g = \gcd(a, b)$ and $h = \gcd(r, b)$. Now g divides both a and b, so g divides $a - qb = r$ (by exercise 1.1.1(c)). Therefore g is a common divisor of both r and b, and therefore $g \leq h$. Similarly h divides both b and r, so h divides $qb + r = a$ and hence h is a common divisor of both a and b, and therefore $h \leq g$. We have shown that $g \leq h$ and $h \leq g$, which together imply that $g = h$. □

Corollary 1.1.1 justifies the method used to determine the gcd of 85 and 48 in the first paragraph of section 1.1 and indeed in general:

Exercise 1.1.3. Use Corollary 1.1.1 to prove that the Euclidean algorithm indeed yields the greatest common divisor of two given integers. (You might prove this by induction on the smallest of the two integers.)

Exercise 1.1.4. Prove that $(F_n, F_{n+1}) = 1$ by induction on $n \geq 0$.

1.2. Linear combinations

The Euclidean algorithm can also be used to determine a linear combination[7] of a and b, over the integers, which equals $\gcd(a, b)$; that is, one can always use the Euclidean algorithm to find integers u and v such that

$$au + bv = \gcd(a, b). \tag{1.2.1}$$

Let us see how to do this in an example, by finding integers u and v such that $85u + 48v = 1$; remember that we found the gcd of 85 and 48 at the beginning of section 1.1. We retrace the steps of the Euclidean algorithm, but in reverse: The final step was that $1 = 1 \cdot 4 - 1 \cdot 3$, a linear combination of 4 and 3. The second to last step used that $3 = 11 - 2 \cdot 4$, and so substituting $11 - 2 \cdot 4$ for 3 in $1 = 1 \cdot 4 - 1 \cdot 3$, we obtain

$$1 = 1 \cdot 4 - 1 \cdot 3 = 1 \cdot 4 - 1 \cdot (11 - 2 \cdot 4) = 3 \cdot 4 - 1 \cdot 11,$$

a linear combination of 11 and 4. This then implies, since we had $4 = 37 - 3 \cdot 11$, that

$$1 = 3 \cdot (37 - 3 \cdot 11) - 1 \cdot 11 = 3 \cdot 37 - 10 \cdot 11,$$

a linear combination of 37 and 11. Continuing in this way, we successively deduce, using that $11 = 48 - 37$ and then that $37 = 85 - 48$,

$$\begin{aligned} 1 = \ & 3 \cdot 37 - 10 \cdot (48 - 37) \ = 13 \cdot 37 - 10 \cdot 48 \\ = \ & 13 \cdot (85 - 48) - 10 \cdot 48 \ = 13 \cdot 85 - 23 \cdot 48; \end{aligned}$$

that is, we have the desired linear combination of 85 and 48.

To prove that this method always works, we will use Lemma 1.1.1 again: Suppose that $a = qb + r$ so that $\gcd(a, b) = \gcd(b, r)$ by Corollary 1.1.1, and that we have $bu - rv = \gcd(b, r)$ for some integers u and v. Then

$$\gcd(a, b) = \gcd(b, r) = bu - rv = bu - (a - qb)v = b(u + qv) - av, \tag{1.2.2}$$

[7] A *linear combination* of two given integers a and b, over the integers, is a number of the form $ax + by$ where x and y are integers. This can be generalized to yield a linear combination $a_1x_1 + \cdots + a_nx_n$ of any finite set of integers, $a_1, \ldots, a_n$. Linear combinations are a key concept in linear algebra and appear (without necessarily being called that) in many courses.

the desired linear combination of a and b. This argument forms the basis of our proof of (1.2.1), but to give a complete proof we proceed by induction on the smaller of a and b:

Theorem 1.1. *If a and b are positive integers, then there exist integers u and v such that*

$$au + bv = \gcd(a, b).$$

Proof. Interchanging a and b if necessary we may assume that $a \geq b \geq 1$. We shall prove the result by induction on b. If $b = 1$, then b only has the divisor 1, so that

$$\gcd(a, 1) = 1 = 0 \cdot a + 1 \cdot 1.$$

We now prove the result for $b > 1$: If b divides a, then

$$\gcd(b, a) = b = 0 \cdot a + 1 \cdot b.$$

Otherwise b does not divide a and so Lemma 1.1.1 implies that there exist integers q and r such that $a = qb + r$ and $1 \leq r \leq b - 1$. Since $1 \leq r < b$ we know, by the induction hypothesis, that there exist integers u and v for which $bu - rv = \gcd(b, r)$. The result then follows by (1.2.2). □

We now establish various useful properties of the gcd:

Exercise 1.2.1. (a) Prove that if d divides both a and b, then d divides $\gcd(a, b)$.
(b) Deduce that d divides both a and b *if and only if* d divides $\gcd(a, b)$.
(c) Prove that $1 \leq \gcd(a, b) \leq |a|$ and $|b|$.
(d) Prove that $\gcd(a, b) = |a|$ if and only if a divides b.

Exercise 1.2.2. Suppose that a divides m, and b divides n.
(a) Deduce that $\gcd(a, b)$ divides $\gcd(m, n)$.
(b) Deduce that if $\gcd(m, n) = 1$, then $\gcd(a, b) = 1$.

Exercise 1.2.3. Show that Theorem 1.1 holds for any integers a and b that are not both 0. (It is currently stated and proved only for positive integers a and b.)

Corollary 1.2.1. *If a and b are integers for which $\gcd(a, b) = 1$, then there exist integers u and v such that*

$$au + bv = 1.$$

This is one of the most useful results in mathematics and has applications in many areas, including in safeguarding today's global communications. For example, we will see in section 10.3 that to implement RSA, a key cryptographic protocol that helps keep important messages safe in our electronic world, one uses Corollary 1.2.1 in an essential way. More on that later, after developing more basic number theory.

Exercise 1.2.4. (a) Use exercise 1.1.1(c) to show that if $au + bv = 1$, then $(a, b) = (u, v) = 1$.
(b) Prove that $\gcd(u, v) = 1$ in Theorem 1.1.

Corollary 1.2.2. *If $\gcd(a, m) = \gcd(b, m) = 1$, then $\gcd(ab, m) = 1$.*

Proof. By Theorem 1.1 there exist integers r, s, u, v such that

$$au + mv = br + ms = 1.$$

Therefore

$$ab(ur) + m(bvr + aus + msv) = (au + mv)(br + ms) = 1,$$

and the result follows from exercise 1.2.4(a). □

Corollary 1.2.3. *We have* $\gcd(ma, mb) = m \cdot \gcd(a, b)$ *for all integers* $m \geq 1$.

Proof. By Theorem 1.1 there exist integers u, v such that $au+bv = \gcd(a, b)$. Now $\gcd(ma, mb)$ divides ma and mb so it divides $mau + mbv = m \cdot \gcd(a, b)$. Similarly $\gcd(a, b)$ divides a and b, so that $m \cdot \gcd(a, b)$ divides ma and mb, and therefore $\gcd(ma, mb)$ by exercise 1.2.1(a). The result follows from exercise 1.1.1(b), since the gcd is always positive. □

Exercise 1.2.5. (a) Show that if A and B are given integers, not both 0, with $g = \gcd(A, B)$, then $\gcd(A/g, B/g) = 1$.

(b) Prove that any rational number u/v where $u, v \in \mathbb{Z}$ with $v \neq 0$ may be written as r/s where r and s are coprime integers with $s > 0$. This is called a *reduced fraction.*

1.3. The set of linear combinations of two integers

Theorem 1.1 states that the greatest common divisor of two integers is a linear combination of those two integers. This suggests that it might be useful to study *the set of linear combinations*

$$I(a, b) := \{am + bn : \ m, n \in \mathbb{Z}\}$$

of two given integers a and b.[8] We see that $I(a, b)$ contains $0,\ a,\ b,\ a + b,\ a + 2b,\ 2b + a,\ a - b,\ b - a, \ldots$ and any sum of integer multiples of a and b, so that $I(a, b)$ is closed under addition. Let $I(a) := I(a, 0) = \{am : \ m \in \mathbb{Z}\}$, the set of integer multiples of a. We now prove that $I(a, b)$ can be described as the set of integer multiples of $\gcd(a, b)$, a set which is easier to understand:

Corollary 1.3.1. *For any given non-zero integers* a *and* b*, we have*

$$\{am + bn : \ m, n \in \mathbb{Z}\} = \{gk : \ k \in \mathbb{Z}\}$$

where $g := \gcd(a, b)$*; that is,* $I(a, b) = I(g)$*. In other words, there exist integers* m *and* n *with* $am + bn = c$ *if and only if* $\gcd(a, b)$ *divides* c*.*

Proof. By Theorem 1.1 we know that there exist $u, v \in \mathbb{Z}$ for which $au + bv = g$. Therefore $a(uk)+b(vk) = gk$ so that $gk \in I(a, b)$ for all $k \in \mathbb{Z}$; that is, $I(g) \subset I(a, b)$. On the other hand, as g divides both a and b, there exist integers A, B such that $a = gA$, $b = gB$, and so any $am+bn = g(Am+Bn) \in I(g)$. That is, $I(a, b) \subset I(g)$. The result now follows from the two inclusions. □

It is instructive to see how this result follows directly from the Euclidean algorithm: In our example, we are interested in $\gcd(85, 48)$, so we will study $I(85, 48)$, that is, the set of integers of the form

$$85m + 48n.$$

[8] This is usually called the *ideal* generated by a and b in $\mathbb{Z}$ and denoted by $\langle a, b\rangle_{\mathbb{Z}}$. The notion of an ideal is one of the basic tools of modern algebra, as we will discuss in appendix 3D.

The first step in the Euclidean algorithm was to write $85 = 1 \cdot 48 + 37$. Substituting this in above yields

$$85m + 48n = (1 \cdot 48 + 37)m + 48n = 48(m + n) + 37m,$$

and so $I(85, 48) \subset I(48, 37)$. In the other direction, any integer in $I(48, 37)$ can be written as

$$48a + 37b = 48a + (85 - 48)b = 85b + 48(a - b),$$

and so belongs to $I(85, 48)$. Combining these last two statements yields that

$$I(85, 48) = I(48, 37).$$

Each step of the Euclidean algorithm leads to a similar equality, and so we get

$$I(85, 48) = I(48, 37) = I(37, 11) = I(11, 4) = I(4, 3) = I(3, 1) = I(1, 0) = I(1).$$

To truly justify this we need to establish an analogous result to Corollary 1.1.1:

Lemma 1.3.1. *If $a = qb+r$ where a, b, q, and r are integers, then $I(a, b) = I(b, r)$.*

Proof. We begin by noting that

$$am + bn = (qb + r)m + bn = b(qm + n) + rm$$

so that $I(a, b) \subset I(b, r)$. In the other direction

$$bu + rv = bu + (a - qb)v = av + b(u - qv)$$

so that $I(b, r) \subset I(a, b)$. The result follows by combining the two inclusions. □

We have used the Euclidean algorithm to find the gcd of any two given integers a and b, as well as to determine integers u and v for which $au + bv = \gcd(a, b)$. The price for obtaining the actual values of u and v, rather than merely proving the existence of u and v (which is all that was claimed in Theorem 1.1), was our somewhat complicated analysis of the Euclidean algorithm. However, if we *only wish to prove* that such integers *u and v exist*, then we can do so with a somewhat easier proof: [9]

Non-constructive proof of Theorem 1.1. Let h be the smallest positive integer that belongs to $I(a, b)$, say $h = au + bv$. Then $g := \gcd(a, b)$ divides h, as g divides both a and b.

Now $a = a \cdot 1 + b \cdot 0$ so that $a \in I(a, b)$, and $1 \leq h \leq a$ by the definition of h. Therefore Lemma 1.1.1 implies that there exist integers q and r, with $0 \leq r \leq h-1$, for which $a = qh + r$. Therefore

$$r = a - qh = a - q(au + bv) = a(1 - qu) + b(-qv) \in I(a, b),$$

which contradicts the minimality of h, unless $r = 0$; that is, h divides a. An analogous argument reveals that h divides b, and so h divides g by exercise 1.2.1(a).

[9] We will now prove the *existence* of u and v by showing that their non-existence would lead to a contradiction. We will develop other instances, as we proceed, of both constructive and non-constructive proofs of important theorems.

Which type of proof is preferable? This is somewhat a matter of taste. The non-constructive proof is often shorter and more elegant. The constructive proof, on the other hand, is practical—that is, it *gives* solutions. It is also "richer" in that it develops more than is (immediately) needed, though some might say that these extras are irrelevant.

Which type of proof has the greatest clarity? That depends on the *algorithm* devised for the constructive proof. A compact algorithm will often cast light on the subject. But a cumbersome one may obscure it. In this case, the Euclidean algorithm is remarkably simple and efficient ([**Sha85**, p. 11]).

Hence g divides h, and h divides g, and g and h are both positive, so that $g = h$ as desired. □

We say that the integers a, b, and c are *relatively prime* if $\gcd(a, b, c) = 1$. We say that they are *pairwise coprime* if $\gcd(a, b) = \gcd(a, c) = \gcd(b, c) = 1$. For example, $6, 10, 15$ are relatively prime, but they are not pairwise coprime (since each pair of integers has a common factor > 1).

Exercise 1.3.1. Suppose that a, b, and c are non-zero integers for which $a + b = c$.
(a) Show that a, b, c are relatively prime if and only if they are pairwise coprime.
(b) Show that $(a, b) = (a, c) = (b, c)$.
(c) Show that the analogy to (a) is false for integer solutions a, b, c, d to $a + b = c + d$ (perhaps by constructing a counterexample).

1.4. The least common multiple

The *least common multiple*[10] of two given integers a and b is defined to be the smallest positive integer that is a multiple of both a and b. We denote this by $\operatorname{lcm}[a, b]$ (or simply $[a, b]$). We now prove the counterpart to exercise 1.2.1(a):

Lemma 1.4.1. $\operatorname{lcm}[a, b]$ *divides integer* m *if and only if* a *and* b *both divide* m.

Proof. Since a and b divide $\operatorname{lcm}[a, b]$, if $\operatorname{lcm}[a, b]$ divides m, then a and b both divide m, by exercise 1.1.1(e).

On the other hand suppose a and b both divide m, and write $m = q \operatorname{lcm}[a, b] + r$ where $0 \le r < \operatorname{lcm}[a, b]$. Now a and b both divide m and $\operatorname{lcm}[a, b]$ so they both divide $m - q \operatorname{lcm}[a, b] = r$. However $\operatorname{lcm}[a, b]$ is defined to be the smallest positive integer that is divisible by both a and b, which implies that r must be 0. Therefore $\operatorname{lcm}[a, b]$ divides m. □

The analogies to exercise 1.2.1(d) and Corollary 1.2.3 for lcms are given by the following two exercises:

Exercise 1.4.1. Prove that $\operatorname{lcm}[m, n] = n$ if and only if m divides n.

Exercise 1.4.2. Prove that $\operatorname{lcm}[ma, mb] = m \cdot \operatorname{lcm}[a, b]$ for any positive integer m.

1.5. *Continued fractions*

Another way to write Lemma 1.1.1 is that for any given integers $a \ge b \ge 1$ with $b \nmid a$, there exist integers q and r, with $b > r \ge 1$, for which

$$\frac{a}{b} \;=\; q + \frac{r}{b} \;=\; q + \frac{1}{\frac{b}{r}}.$$

This is admittedly a strange way to write things, but repeating this process with the pair of integers b and r, and then again, will eventually lead us to an interesting representation of the original fraction a/b. Working with our original example, in which we found the gcd of 85 and 48, we can represent $85 = 48 + 37$ as

$$\frac{85}{48} = 1 + \frac{1}{\frac{48}{37}},$$

[10] Sometimes called the *lowest common multiple*.

and the next step, $48 = 37 + 11$, as

$$\frac{48}{37} = 1 + \frac{1}{\frac{37}{11}}, \text{ so that } \frac{85}{48} = 1 + \frac{1}{\frac{48}{37}} = 1 + \frac{1}{1 + \frac{1}{\frac{37}{11}}}.$$

The remaining steps of the Euclidean algorithm may be rewritten as

$$\frac{37}{11} = 3 + \frac{1}{\frac{11}{4}}, \quad \frac{11}{4} = 2 + \frac{1}{\frac{4}{3}}, \quad \text{and} \quad \frac{4}{3} = 1 + \frac{1}{3},$$

so that

$$\frac{85}{48} = 1 + \frac{1}{1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{1 + \frac{1}{3}}}}}.$$

This is the *continued fraction* for $\frac{85}{48}$ and is conveniently written as $[1, 1, 3, 2, 1, 3]$. Notice that this is the sequence of quotients a_i from the various divisions; that is,

$$\frac{a}{b} \;=\; [a_0,\ a_1,\ a_2,\ \ldots,\ a_k] := \; a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots + \frac{1}{a_k}}}}.$$

The a_i's are called the *partial quotients* of the continued fraction.

Exercise 1.5.1. (a) Show that if $a_k > 1$, then $[a_0,\ a_1, \ldots,\ a_k] = [a_0,\ a_1, \ldots,\ a_k - 1, 1]$.
(b) Prove that the set of positive rational numbers are in 1-1 correspondence with the finite length continued fractions that do not end in 1.

We now list the rationals that correspond to the first few entries in our continued fraction $[1, 1, 3, 2, 1, 3]$. We have $[1] = 1, [1, 1] = 2$, and

$$1 + \frac{1}{1 + \frac{1}{3}} = \frac{7}{4}, \qquad 1 + \frac{1}{1 + \frac{1}{3 + \frac{1}{2}}} = \frac{16}{9}, \qquad 1 + \frac{1}{1 + \frac{1}{3 + \frac{1}{2 + \frac{1}{1}}}} = \frac{23}{13}.$$

These yield increasingly good approximations to $85/48 = 1.770833\ldots$, that is, in decimal notation,

$$1,\ 2,\ 1.75,\ 1.777\ldots,\ 1.7692\ldots.$$

We call these p_j/q_j, $j \geq 1$, the *convergents* for the continued fraction, defined by

$$\frac{p_j}{q_j} = [a_0,\ a_1,\ a_2,\ \ldots,\ a_j],$$

since they converge to $a/b = p_k/q_k$ for some k. Do you notice anything surprising about the convergents for 85/48? In particular the previous one, namely 23/13? When we worked through the Euclidean algorithm we found that $13 \cdot 85 - 23 \cdot 48 = 1$ — could it be a coincidence that these same numbers show up again in this new context? In section 1.8 of appendix 1A we show that this is no coincidence; indeed we always have

$$p_j q_{j-1} - p_{j-1} q_j = (-1)^{j-1},$$

so, in general, if $u = (-1)^{k-1} q_{k-1}$ and $v = (-1)^k p_{k-1}$, then

$$au + bv = 1.$$

When one studies this in detail, one finds that the continued fraction is really just a convenient reworking of the Euclidean algorithm (as we explained it above)

for finding u and v. Bachet de Meziriac[11] introduced this method to Renaissance mathematicians in the second edition of his brilliantly named book *Pleasant and delectable problems which are made from numbers* (1624). Such methods had been known from ancient times, certainly to the Indian scholar Āryabhata in 499 A.D., probably to Archimedes in Syracuse (Greece) in 250 B.C., and possibly to the Babylonians as far back as 1700 B.C.[12]

1.6. *Tiling a rectangle with squares*[13]

Given a 48-by-85 rectangle we will tile it, greedily, with squares. The largest square that we can place inside a 48-by-85 rectangle is a 48-by-48 square. This 48-by-48 square goes from top to bottom of the rectangle, and if we place it at the far right, then we are left with a 37-by-48 rectangle to tile, on the left.

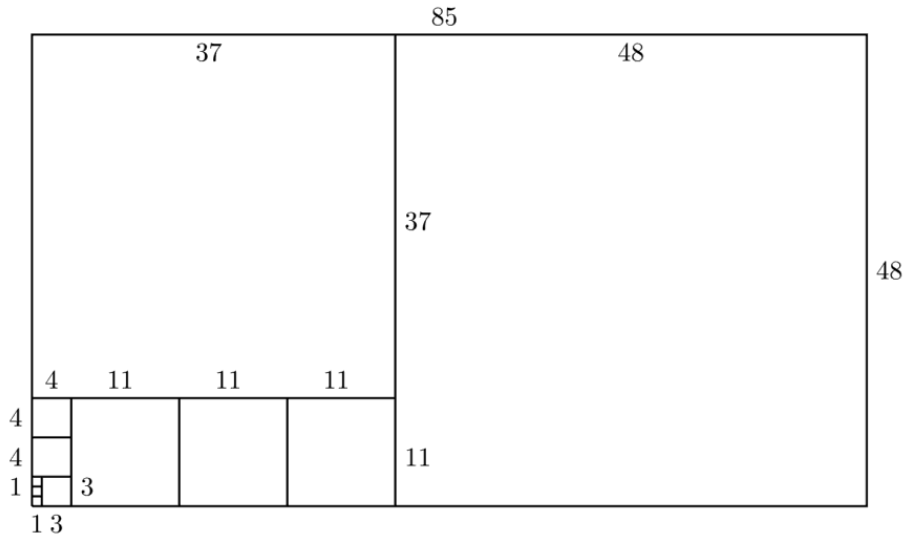


Figure 1.1. Partitioning a rectangle into squares, using the Euclidean algorithm.

If we place a 37-by-37 square at the top of this rectangle, then we are left with an 11-by-37 rectangle in the bottom left-hand corner. We can now place three 11-by-11 squares inside this, leaving a 4-by-11 rectangle. We finish this off with two 4-by-4 squares, one 3-by-3 square, and finally three 1-by-1 squares.

[11] The celebrated editor and commentator on Diophantus, whom we will meet again in chapter 6.

[12] There are Cuneiform clay tablets from this era that contain related calculations. It is known that after conquering Babylon in 331 B.C., Alexander the Great ordered his archivist Callisthenes and his tutor Aristotle to supervise the translation of the Babylonian astronomical records into Greek. It is therefore feasible that Archimedes was introduced to these ideas from this source. Indeed, Pythagoras's Theorem may be misnamed as the Babylonians knew of integer-sided right-angled triangles like $3, 4, 5$ and $5, 12, 13$ more than one thousand years before Pythagoras (570–495 B.C.) was born.

[13] Thanks to Dusa MacDuff and Dylan Thurston for bringing my attention to this beautiful application.

The area of the rectangle can be computed in terms of the areas of each of the squares; that is,

$$85 \cdot 48 = 1 \cdot 48^2 + 1 \cdot 37^2 + 3 \cdot 11^2 + 2 \cdot 4^2 + 1 \cdot 3^2 + 3 \cdot 1^2.$$

What has this to do with the Euclidean algorithm? Hopefully the reader has recognized the same sequence of numbers and quotients that appeared above, when we computed the $\gcd(85, 48)$. This is no coincidence. At a given step we have an a-by-b rectangle, with $a \geq b \geq 1$, and then we can remove q b-by-b squares, where $a = qb + r$ with $0 \leq r \leq a - 1$ leaving an r-by-b rectangle, and so proceed by induction.

Exercise 1.6.1. Given an a-by-b rectangle show how to write $a \cdot b$ as a sum of squares, as above, in terms of the partial quotients and convergents of the continued fraction for a/b.

Exercise 1.6.2. (a) Use this to show that $F_{n+1}F_n = F_n^2 + F_{n-1}^2 + \cdots + F_0^2$, where F_n is the nth Fibonacci number (see section 0.1 for the definition and a discussion of Fibonacci numbers and exercise 0.4.12(b) for a generalization of this exercise).
(b)† Find the correct generalization to more general second-order linear recurrence sequences.

Additional exercises

Exercise 1.7.1. (a) Does 0 divide 0? (Use the definition of "divides".)
(b) Show that there is no unique meaning to $0/0$.
(c) Prove that if b divides a and $b \neq 0$, then there is a unique meaning to a/b.

Exercise 1.7.2. Prove that if a and b are not both 0, then $\gcd(a, b)$ is a positive integer.

Exercise 1.7.3.† Prove that if m and n are coprime positive integers, then $\frac{(m+n-1)!}{m!n!}$ is an integer.

Exercise 1.7.4. Suppose that $a = qb + r$ with $0 \leq r \leq b - 1$.
(a) Let $[t]$ be the *integer part* of t, that is, the largest integer $\leq t$. Prove that $q = [a/b]$.
(b) Let $\{t\}$ to be the *fractional part* of t, that is, $\{t\} = t - [t]$. Prove that $r = b\{r/b\} = b\{a/b\}$.

(Beware of these functions applied to negative numbers: e.g., $[-3.14] = -4$ *not* -3, and $\{-3.14\} = .86$ *not* $.14$.)

Exercise 1.7.5.† (a) Show that if n is an integer, then $\{n + \alpha\} = \{\alpha\}$ and $[n + \alpha] = n + [\alpha]$ for all $\alpha \in \mathbb{R}$.
(b) Prove that $[\alpha + \beta] - [\alpha] - [\beta] = 0$ or 1 for all $\alpha, \beta \in \mathbb{R}$, and explain when each case occurs.
(c) Deduce that $\{\alpha\} + \{\beta\} - \{\alpha + \beta\} = 0$ or 1 for all $\alpha, \beta \in \mathbb{R}$, and explain when each case occurs.
(d) Show that $\{\alpha\} + \{-\alpha\} = 1$ unless α is an integer in which case it equals 0.
(e) Show that if $a \in \mathbb{Z}$ and $r \in \mathbb{R} \setminus \mathbb{Z}$, then $[r] + [a - r] = a - 1$.

Exercise 1.7.6. Suppose that d is a positive integer and that $N, x > 0$.
(a) Show that there are exactly $[x]$ positive integers $\leq x$.
(b) Show that kd is the largest multiple of d that is $\leq N$, where $k = [N/d]$.
(c) Deduce that there are exactly $[N/d]$ positive integers $n \leq N$ which are divisible by d.

Exercise 1.7.7. Prove that $\sum_{k=0}^{n-1} [a + \frac{k}{n}] = [na]$ for any real number a and integer $n \geq 1$.

Exercise 1.7.8. Suppose that $a + b = c$ and let $g = \gcd(a, b)$. Prove that we can write $a = gA$, $b = gB$, and $c = gC$ where $A + B = C$, where A, B, and C are pairwise coprime integers.

Exercise 1.7.9. Prove that if $(a, b) = 1$, then $(a + b, a - b) = 1$ or 2.

Exercise 1.7.10.† Prove that for any given integers $b > a \geq 1$ there exists an integer solution u, w to $au - bw = \gcd(a, b)$ with $0 \leq u \leq b - 1$ and $0 \leq w \leq a - 1$.

Exercise 1.7.11.† Show that if $\gcd(a,b)=1$, then $\gcd(a^k,b^\ell)=1$ for all integers $k,\ell\geq 1$.

Exercise 1.7.12. Let m and n be positive integers. What fractions do the two lists $\frac{1}{m},\dots,\frac{m-1}{m}$ and $\frac{1}{n},\dots,\frac{n-1}{n}$ have in common (when the fractions are reduced)?

Exercise 1.7.13. Suppose m and n are coprime positive integers. When the fractions $\frac{1}{m},\frac{2}{m},\dots,\frac{m-1}{m},\frac{1}{n},\dots,\frac{n-1}{n}$ are put in increasing order, what is the shortest distance between two consecutive fractions?

Given a 7-liter jug and a 5-liter jug one can measure 1 liter of water as follows: Fill the 5-liter jug, and pour the contents into the 7-liter jug. Fill the 5-liter jug again, use this to fill the 7-liter jug, so we are left with 3 liters in the 5-liter jug and the 7-liter jug is full. Empty the 7-liter jug, pour the contents of the 5-liter jug into the 7-liter jug, and refill the 5-liter jug. We now have 3 liters in the 7-liter jug. Fill the 7-liter jug using the 5-liter jug; we have poured 4 liters from the 5-liter jug into the 7-liter jug, so that there is just 1 liter left in the 5-liter jug! Notice that we filled the 5-liter jug 3 times and emptied the 7-liter jug twice, and so we used here that $3\times 5-2\times 7=1$. We have wasted 2×7 liters of water in this process.

Exercise 1.7.14. (a) Since $3\times 7-4\times 5=1$ describe how we can proceed by filling the 7-liter jug each time rather than filling the 5-liter jug.
(b) Can you measure 1 liter of water using a 25-liter jug and a 17-liter jug?
(c)† Prove that if m and n are positive coprime integers then you can measure one liter of water using an m liter jug and an n liter jug?
(d) Prove that one can do this wasting less than mn liters of water.

Exercise 1.7.15. Can you weigh 1 lb of tea using scales with 25-lb and 17-lb weights?

The definition of a set of linear combinations can be extended to an arbitrary set of integers (in place of the set $\{a,b\}$); that is,

$$I(a_1,\dots,a_k):=\{a_1m_1+a_2m_2+\cdots+a_km_k:\ m_1,m_2,\dots,m_k\in\mathbb{Z}\}.$$

Exercise 1.7.16. Show that $I(a_1,\dots,a_k)=I(g)$ for any non-zero integers $a_1,\dots,a_k$, where we have $g=\gcd(a_1,\dots,a_k)$.

Exercise 1.7.17.† Deduce that if we are given integers $a_1,a_2,\dots,a_k$, not all zero, then there exist integers $m_1,m_2,\dots,m_k$ such that

$$m_1a_1+m_2a_2+\cdots+m_ka_k=\gcd(a_1,a_2,\dots,a_k).$$

We say that the integers $a_1,a_2,\dots,a_k$ are *relatively prime* if $\gcd(a_1,a_2,\dots,a_k)=1$. We say that they are *pairwise coprime* if $\gcd(a_i,a_j)=1$ whenever $i\neq j$. Note that $6,10,15$ are relatively prime, but not pairwise coprime (since each pair of integers has a common factor >1).

Exercise 1.7.18. Prove that if $g=\gcd(a_1,a_2,\dots,a_k)$, then $\gcd(a_1/g,a_2/g,\dots,a_k/g)=1$.

Exercise 1.7.19.† (a) Prove that $abc=[a,b,c]\cdot\gcd(ab,bc,ca)$.
(b)‡ Prove that if $r+s=n$, then

$$a_1\cdots a_n=\operatorname{lcm}\left[\prod_{i\in I}a_i:\ I\subset\{1,\dots,n\},\ |I|=r\right]\cdot\gcd\left(\prod_{j\in J}a_j:\ J\subset\{1,\dots,n\},\ |J|=s\right).$$