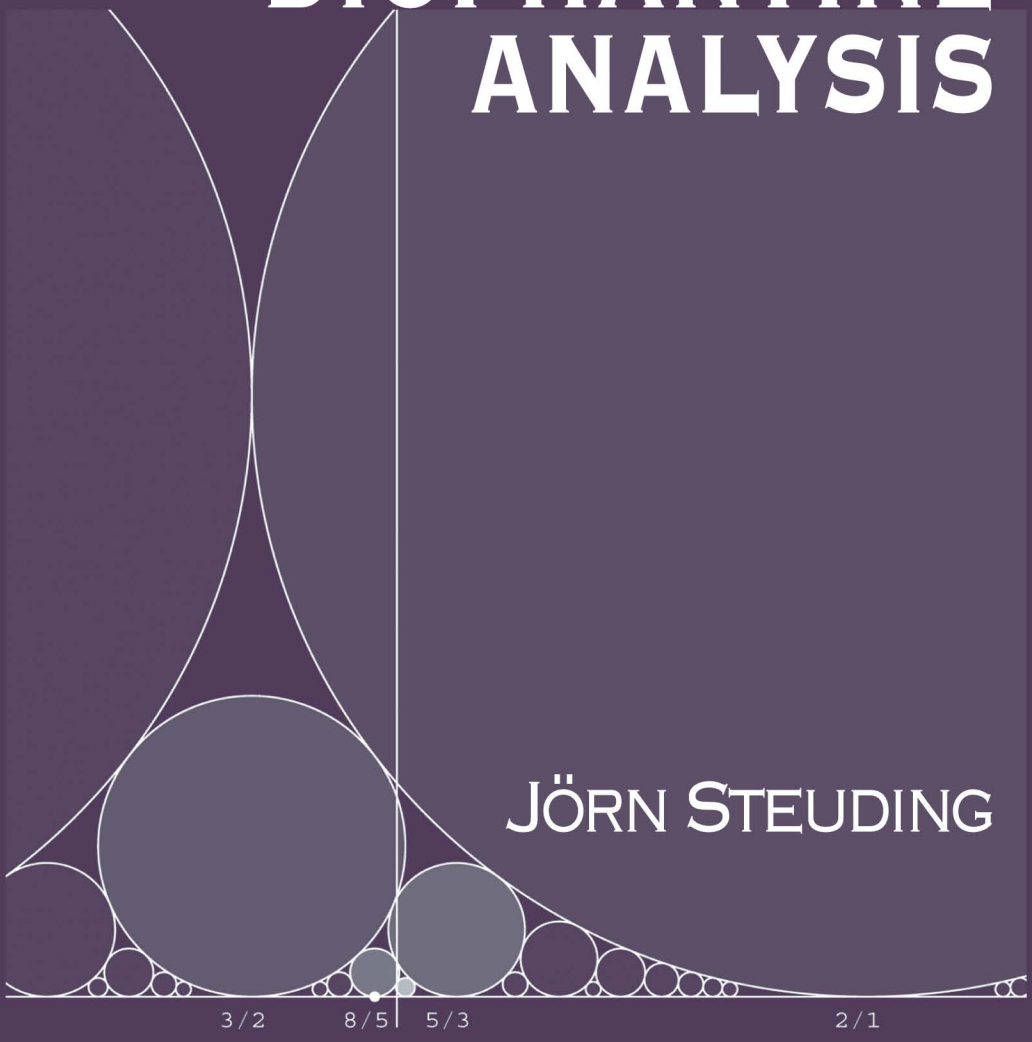
DISCRETE MATHEMATICS AND ITS APPLICATIONS
Series Editor KENNETH H. ROSEN

DIOPHANTINE ANALYSIS



JÖRN STEUDING

Chapman & Hall/CRC
Taylor & Francis Group

DISCRETE MATHEMATICS AND ITS APPLICATIONS
Series Editor KENNETH H. ROSEN

DIOPHANTINE ANALYSIS

DISCRETE MATHEMATICS AND ITS APPLICATIONS

Series Editor
Kenneth H. Rosen, Ph.D.

Juergen Bierbrauer, Introduction to Coding Theory

Kun-Mao Chao and Bang Ye Wu, Spanning Trees and Optimization Problems

Charalambos A. Charalambides, Enumerative Combinatorics

Charles J. Colbourn and Jeffrey H. Dinitz, The CRC Handbook of Combinatorial Designs

Steven Furino, Ying Miao, and Jianxing Yin, Frames and Resolvable Designs: Uses, Constructions, and Existence

Randy Goldberg and Lance Riek, A Practical Handbook of Speech Coders

Jacob E. Goodman and Joseph O'Rourke, Handbook of Discrete and Computational Geometry, Second Edition

Jonathan Gross and Jay Yellen, Graph Theory and Its Applications

Jonathan Gross and Jay Yellen, Handbook of Graph Theory

Darrel R. Hankerson, Greg A. Harris, and Peter D. Johnson, Introduction to Information Theory and Data Compression, Second Edition

Daryl D. Harms, Miroslav Kraetzl, Charles J. Colbourn, and John S. Devitt, Network Reliability: Experiments with a Symbolic Algebra Environment

Derek F. Holt with Bettina Eick and Eamonn A. O'Brien, Handbook of Computational Group Theory

David M. Jackson and Terry I. Visentin, An Atlas of Smaller Maps in Orientable and Nonorientable Surfaces

Richard E. Klima, Ernest Stitzinger, and Neil P. Sigmon, Abstract Algebra Applications with Maple

Patrick Knupp and Kambiz Salari, Verification of Computer Codes in Computational Science and Engineering

William Kocay and Donald L. Kreher, Graphs, Algorithms, and Optimization

Donald L. Kreher and Douglas R. Stinson, Combinatorial Algorithms: Generation Enumeration and Search

Charles C. Lindner and Christopher A. Rodgers, Design Theory

Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone, Handbook of Applied Cryptography

Continued Titles

Richard A. Mollin, Algebraic Number Theory

Richard A. Mollin, Codes: The Guide to Secrecy from Ancient to Modern Times

Richard A. Mollin, Fundamental Number Theory with Applications

Richard A. Mollin, An Introduction to Cryptography

Richard A. Mollin, Quadratics *Richard A. Mollin,* RSA and Public-Key Cryptography

Kenneth H. Rosen, Handbook of Discrete and Combinatorial Mathematics

Douglas R. Shier and K.T. Wallenius, Applied Mathematical Modeling: A Multidisciplinary Approach

Jörn Steuding, Diophantine Analysis

Douglas R. Stinson, Cryptography: Theory and Practice, Second Edition

Roberto Togneri and Christopher J. deSilva, Fundamentals of Information Theory and Coding Design

Lawrence C. Washington, Elliptic Curves: Number Theory and Cryptography

DISCRETE MATHEMATICS AND ITS APPLICATIONS
Series Editor KENNETH H. ROSEN

DIOPHANTINE ANALYSIS

JÖRN STEUDING

Chapman & Hall/CRC
Taylor & Francis Group
Boca Raton London New York Singapore

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20110713

International Standard Book Number-13: 978-1-4200-5720-1 (eBook - PDF)

Contents

Preface

The book is devoted to the theory of diophantine approximations and the theory of diophantine equations with emphasis on interactions between these subjects. Many diophantine problems have simple formulations but are extremely hard to attack. For instance, consider the rather simple looking equation

$$X^n + Y^n = Z^n,$$

where n is a positive integer and which has to be solved in integers x, y, z. Integer solutions x, y, z which satisfy $xyz = 0$ are — for obvious reasons — called *trivial*. In the case $n = 2$ there are many other solutions, e.g.,

$$3^2 + 4^2 = 5^2, \quad 5^2 + 12^2 = 13^2,$$

already known by the ancient Greeks. However, the situation becomes totally different when $n \geq 3$. For such exponents, Pierre de Fermat, a French lawyer and mathematician in the seventeenth century, claimed to be able to prove that there are no solutions other than trivial ones; however, he never published his proof. In Fermat's time publishing proofs was not very common; usually, mathematicians wrote letters to other mathematicians announcing what they could prove and asked whether the other could do the same. Fermat is famous for this method of doing math. But for more than three centuries no one succeeded in proving Fermat's statement, known as Fermat's last theorem (since it was the last of his statements to be proved).

In the twentieth century, diophantine analysis emerged from its very beginnings to an extraordinary modern and powerful theory (with three Fields medalists: Roth, Baker, Faltings); one example for this success story is Wiles' solution of Fermat's last theorem, the final proof that there are no solutions other than the trivial ones. However, there are still a lot of open problems and conjectures which turn diophantine analysis into an active and attractive field of interest for researchers and students.

The book's motivation is contained in its introduction. Here we present some basic principles of diophantine analysis. The second chapter deals with classical approximation theorems and the third chapter is devoted to the theory of continued fractions. Then we give a more detailed account on certain topics of the classic theory (Chapters 5 and 6) and present some of its applications (Chapters 4 and 7). Chapter 8 gives a short introduction to the geometry of numbers and its applications. The following two chapters deal with transcendental numbers and Roth's theorem with applications to Thue equations. In Chapter 11 we present the recent *abc*-conjecture and discuss its importance. We conclude with a short introduction to p-adic numbers and their applications to diophantine equations (Chapters 12–14).

Certain topics like elliptic curves or the metric theory of continued fractions are touched but not entirely included because an appropriate presentation would be beyond the scope of this book. However, there exist many

excellent books describing these areas in detail and we refer to some of them for further reading.

Modern topics (which were only considered in a few earlier textbooks) are Apéry's celebrated proof of the irrationality of $\zeta(3)$ (Chapter 4), the polynomial Pell equation (parts of Chapters 6 and 11), and the *abc*-conjecture with its plenty of applications (Chapter 11). Furthermore, we touch several topics which are of interest in related fields of discrete mathematics: factoring methods for large integers with continued fractions (Chapter 7), the LLL-lattice reduction algorithm (Chapter 8), error-free computing with p-adic numbers (Chapter 12), and factoring polynomials with Hensel lifting (Chapter 13).

This book primarily aims at advanced undergraduate students or graduates who may want to learn the fundamentals of this subject. In some sense, diophantine analysis may be regarded as algebraic geometry over the spectrum of a Dedekind domain. This modern point of view is much beyond the scope of this introduction but we shall keep in mind the idea of using the geometry of the objects under observation to learn something about their arithmetic nature. Our approach is elementary; with exception of the proof of Roth's theorem, only a small background in algebra and number theory is needed. Many of the results are presented with respect to their historical development. I believe that the best way to learn math is to look at how it was developed. However, only practice makes perfect. For this purpose, at the end of each chapter, several exercises and problems of different degrees of difficulty are posed, the advanced ones indicated by an asterisk $*$. Furthermore, I recommend using a computer algebra package like Mathematica or Maple; excellent introductions to experimental mathematics are Bressoud & Wagon [**29**] and Vivaldi [**165**], respectively.

The concept of the book is based upon a course which I gave at Frankfurt University in the winter term 2002/2003. I am very grateful to the audience, in particular, Iqbal Lahseb, Thomas Müller, and Matthias Völz. I want to express my appreciation to my colleagues at the departments of the universities in Frankfurt, Vilnius, Šiauliai, and Madrid; and the many people from whom I learned this subject, in particular, my academic teachers Georg Johann Rieger, Wolfgang Schwarz, and Jürgen Wolfart. I am very much indebted to Jürgen Sander and Hessel Posthuma for inspiring conversations and careful reading of the manuscript. My thanks also go to Christian Beck, Artūras Dubickas, Carsten Elsner, Ernesto Girondo, and Harald König for their interest, help, and valuable remarks. Further, I want to thank Helena Redshaw, Bob Stern, and Kenneth Rosen from CRC Press for their encouragement, and the editorial staff at CRC Press for their support during the preparation of this book.

Finally, and most importantly, I want to thank Rasa.

Jörn Steuding, Madrid, December 2004

CHAPTER 1

Introduction: basic principles

Diophantine analysis is devoted to diophantine approximations and diophantine equations. Here the word *diophantine* means that we are concerned about integral or rational solutions. The exact meaning will become clear in the sequel. In this introductory chapter we shall learn some of the basic principles of diophantine analysis. These are special methods (as Fermat's method of infinite descent) as well as building bridges to other fields (e.g., the use of real analysis or geometry for understanding underlying arithmetic structures). In order to present these principles we prove some fundamental historical results.

1.1. Who was Diophantus?

Diophantus of Alexandria was a mysterious mathematician who lived around 250 A.D.; it is not known whether Diophantus was Greek or Egyptian and there are even some rumors that his name stands for a collective of authors (like Bourbaki). Virtually nothing more about his life is known than the following conundrum:

> *God granted him to be a boy for the sixth part of his life, and adding a twelfth part to this, He clothed his cheeks with down. He lit him the light of wedlock after a seventh part, and five years after his marriage He granted him a son. Alas! late-born wretched child; after attaining the measure of half his father's life, chill Fate took him. After consoling his grief by this science of numbers for four years he ended his life.* (cf. Singh [**153**])

How old was Diophantus when he died? This riddle is an example for the kind of problems in which Diophantus was interested. Diophantus wrote the influential monography *Arithmetica*, but, unfortunately, only 10 of the 13 books of it survived. He was the first writer who made a systematic study of the solutions of polynomial equations in integers or rationals. The *Arithmetica* is a collection of isolated problems for each of which Diophantus gave a special solution, but in general, not all solutions. To some extent he was aware of some general methods; however, there was no algebraic formalism (as denoting unknowns by symbols) in Diophantus' time. With the rise of algebra around the tenth century in Arabia Diophantus' monography was translated into algebraic language and mathematicians asked for generalizations of diophantine problems. Thus Diophantus' epoch-making

Arithmetica might be viewed as the birth to number theory. The first translation in Europe, by Regiomantus, appeared only in the fifteenth century. Maybe the most important one is that of Bachet from 1621 which was edited with Pierre de Fermat's remarks by his son Samuel de Fermat in 1679 as part of Fermat's collected works. For more details we refer to Schappacher [**140**] and Weil [**172**].*

1.2. Pythagorean triples

We shall take a closer look at one of the problems discussed in Diophantus' book. Consider the so-called **Pythagorean equation**

$$X^2 + Y^2 = Z^2. \tag{1.1}$$

We are interested in solving this equation in integers and, of course, it suffices to consider non-negative integers. Such solutions (x, y, z) are called **Pythagorean triples** in honor of the contributions of Pythagoras (572–492 B.C.). With an integer solution (x, y, z) of (1.1) the triple (ax, ay, az) is also an integer solution provided $a \in \mathbb{Z}$. If we want to get an overview over all Pythagorean triples, it thus makes sense to consider only integer solutions (x, y, z) which are coprime; such solutions are called **primitive**. Given a solution (x, y, z) of (1.1), it is easily seen that the greatest common divisor of x and y, denoted by $\gcd(x, y)$, must divide z. Clearly, the same holds if we interchange x or y and z. Hence, pairwise coprimality of a Pythagorean triple is necessary and sufficient for having a primitive one.

Any solution of the Pythagorean equation in positive real numbers corresponds to a right angular triangle; this is Pythagoras' famous theorem in geometry. Integer solutions of (1.1) were known for quite a long time before Pythagoras. The ancient Babylonians, four millennia ago, were aware of the solutions

$$3^2 + 4^2 = 5^2, \quad 5^2 + 12^2 = 13^2, \quad 8^2 + 15^2 = 17^2$$

and many more. It is assumed that the Babylonians used Pythagorean triples for constructing right angles.

Pythagoras not only gave a mathematical proof for what the Babylonians knew in practice, but also constructed an infinitude of primitive Pythagorean triples by the identity

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

In the third century B.C. Euclid solved the problem of finding all solutions.

Theorem 1.1. *If a and b are positive coprime integers of opposite parity (i.e., a is even and b is odd, or vice versa) such that $a > b$, then the triple (x,y,z), given by*

$$x = a^2 - b^2, \quad y = 2ab, \quad z = a^2 + b^2, \tag{1.2}$$

*Of course, MacTutor's Web page http://www-groups.dcs.st-and.ac.uk/∼history/ also gives plenty of information on the history of mathematics.

is a primitive solution of (1.1). This establishes a bijection between the set of pairs (a, b) satisfying the above conditions and the set of primitive integer solutions of the Pythagorean equation (1.1).

Proof. It is easy to verify that any triple of the form (1.2) solves the Pythagorean equation (1.1):

$$x^2 + y^2 = \left(a^2 - b^2\right)^2 + (2ab)^2 = a^4 + 2a^2b^2 + b^4 = \left(a^2 + b^2\right)^2 = z^2.$$

Clearly, x, y, z are positive integers. If $d = \gcd(x, y, z)$, then d divides $x + z = 2a^2$ and $z - x = 2b^2$. Since a and b are coprime, it follows that either $d = 1$ or $d = 2$. Since a and b have opposite parity, x is odd and thus the case $d = 2$ cannot occur. This shows that (x, y, z) is a primitive Pythagorean triple.

For the converse implication assume that (x, y, z) is a Pythagorean triple. Since x and y are coprime, and y is even, it follows that x and z are odd and coprime. Hence, $\frac{1}{2}(z + x)$ and $\frac{1}{2}(z - x)$ are coprime integers and, by (1.1),

$$\left(\frac{1}{2}y\right)^2 = \left(\frac{1}{2}(z + x)\right) \cdot \left(\frac{1}{2}(z - x)\right).$$

Since the factors on the right have no common divisor, both have to be squares; i.e., there are coprime positive integers a and b such that

$$a^2 = \frac{1}{2}(z + x) \quad \text{and} \quad b^2 = \frac{1}{2}(z - x)$$

(here we used the fundamental theorem of arithmetic). Further,

$$a + b \equiv a^2 + b^2 = z \equiv 1 \bmod 2,$$

so a and b have opposite parity. Now it is easy to deduce the parametrization (1.2).

It remains to show the one-to-one correspondence between pairs (a, b) and triples (x, y, z). If x and z are given, a^2 and b^2, and consequently a and b, are uniquely determined. Thus, different triples (x, y, z) correspond to different pairs (a, b). The theorem is proved. •

One important invention of mathematics is *considering numbers modulo m, so putting the infinitude of integers into a finite set of residues.* We used this idea of modular arithmetic only a tiny bit in the proof just given (when we investigated the parity) but later on we shall meet it several times.

1.3. Fermat's last theorem

Pierre de Fermat (1607(?)–1665) was a lawyer and government official in Toulouse, and, last but not least, a hobby mathematician. Often his year of birth is dated to be 1601; however, recent investigations make the above given date more reasonable; see Barner [**14**]. When Fermat died, he was one of the most famous mathematicians in Europe although he never had published any mathematical work; his reputation simply grew out of his extensive correspondence with other scientists. Fermat made important contributions to the very beginnings of analytic geometry, probability

theory, and number theory. The reading of Diophantus' *Arithmetica*, in particular, the part on the Pythagorean equation, inspired Fermat to write in his copy of Diophantus' monograph:

> *It is impossible for a cube to be written as a sum of two cubes or a fourth power to be written as the sum of two fourth powers or, in general, for any number which is a power greater than the second to be written as a sum of two like powers. I have a truly marvelous demonstration of this proposition which this margin is too narrow to contain.* (cf. Singh [**153**])

In the modern language of algebra, he claimed to have a proof of

Fermat's last theorem. *All solutions of the equation*

$$X^n + Y^n = Z^n \tag{1.3}$$

in integers x, y, z *are* **trivial**, *i.e.,* $xyz = 0$, *whenever* $n \geq 3$.

Fermat never published a proof and, by the unsuccessful quest for a solution of Fermat's last theorem, mathematicians started to believe that Fermat actually had no proof. However, no counterexample was found. In fact, the above statement is the only result stated by Fermat which could not be proved for quite a long time, and so it became Fermat's *last* theorem. Only recently Wiles [**174**], supported by Taylor and the earlier works of many others, found a proof for Fermat's last theorem. The proof relies on a link to the theory of modular forms. We refer to Edwards [**52**] for the prehistory of attempts to solve Fermat's last theorem, Singh [**153**] for the amazing story of this problem and its final solution, and Washington [**169**] for a brief mathematical discussion of Wiles' breakthrough.

One may ask why the ancient Greeks considered only the quadratic case of the Fermat equation but not the general one. Greek mathematics was inspired by at most three-dimensional geometry and only in the late works of Greek mathematics higher powers occur. They also had an advanced knowledge on divisibility and prime numbers but it seems that they had no idea about the unique prime factorization of the integers.

The exponent in Fermat's equation is crucial. By Theorem 1.1, there are infinitely many solutions when $n = 2$, but by Wiles' proof there are only trivial solutions when $n \geq 3$. This observation due to Fermat is essential for the importance of Fermat's last theorem for diophantine analysis. It is the exponent n which defines the geometric character of the Fermat curve (1.3) and indeed, the corresponding geometric quantity called genus rules the solvability.

1.4. The method of infinite descent

The classification of the Pythagorean triples, Theorem 1.1, can be used to prove that the biquadratic case of Fermat's last theorem has only trivial solutions. However, we start with a slightly more general equation.

Theorem 1.2. *There are no positive integer solutions of the equation*

$$X^4 + Y^4 = Z^2.$$

We give Fermat's original and marvelous

Proof. Suppose that z is the least positive integer for which the equation

$$X^4 + Y^4 = z^2$$

has a solution in positive integers x, y. It follows that x and y are coprime since otherwise we can divide through $\gcd(x, y)^4$, contradicting the minimality of z. Thus at least one of x and y is odd. Since the squares modulo 4 are 0 and 1, it follows that

$$z^2 = x^4 + y^4 \equiv 1 \text{ or } 2 \mod 4.$$

A square cannot be congruent 2 mod 4, so z is odd, and only one of x and y is odd; the other one is even. Without loss of generality we may assume that y is even. Then, by (1.2) of Theorem 1.1,

$$x^2 = a^2 - b^2, \quad y^2 = 2ab, \quad z = a^2 + b^2,$$

where a and b are coprime positive integers of opposite parity. If a is even and b is odd, then $x^2 \equiv -1 \mod 4$, which is impossible. Thus, a is odd and b is even, say $b = 2c$ for some integer c. We observe that

$$\left(\frac{1}{2}y\right)^2 = ac,$$

where a and c are coprime. It follows that $a = u^2$ and $c = v^2$ with some positive coprime integers u, v, where u is odd (since a is odd). This leads to

$$\left(2v^2\right)^2 + x^2 = \left(u^2\right)^2,$$

where no two of the numbers $2v^2, x, u^2$ have a common factor. Applying once more Theorem 1.1 we obtain

$$2v^2 = 2AB \qquad \text{and} \qquad u^2 = A^2 + B^2,$$

where A and B are coprime positive integers. Dividing the v-equation by 2, we get (by the coprimality of A and B) the existence of some positive coprime integers $\mathcal{X}$ and $\mathcal{Y}$ such that $A = \mathcal{X}^2$ and $B = \mathcal{Y}^2$. Substituting this in the u-equation gives

$$\mathcal{X}^4 + \mathcal{Y}^4 = u^2,$$

which is another non-trivial solution of the diophantine equation under consideration. However,

$$u \le u^2 = a \le a^2 < a^2 + b^2 = z,$$

which contradicts the assumption that z was the least solution. This proves the assertion of the theorem. •

The method of proof is called **method of infinite descent**. The simple but ingenious idea of *constructing a smaller solution out of a given one can often be used for proving that certain diophantine equations have no integer solutions.*

It is obvious how Theorem 1.2 solves Fermat's last theorem in the case $n = 4$. It might be possible that Fermat had this argument in mind when he made his statement of having a proof for the general case. However, the case $n = 4$ in Fermat's last theorem is the only *easy* one. Odd exponents cannot be treated as above.

1.5. Cantor's paradise

Usually, the set of positive integers $\mathbb{N}$ is introduced by the Peano axioms. Adding the neutral element zero and the inverse elements with respect to addition, we obtain the set of integers $\mathbb{Z}$. Further, incorporating the inverse elements with respect to multiplication we get the field of rational numbers $\mathbb{Q}$. Hence, a number is said to be **rational** if it can be represented as a quotient of two integers, the denominator being non-zero; all other numbers (in the set $\mathbb{R}$ of real and the set $\mathbb{C}$ of complex numbers, respectively) are called **irrational**.

It is believed that we are living in a finite universe: there are about 10^{80} atoms in our universe. So our world can be described using only rational numbers. However, we cannot *understand* our world without a larger set of numbers. The Pythagorean equation (1.1) led to one of the great breakthroughs in ancient Greek mathematics. Hippasus, a pupil of Pythagoras, discovered that the set of rational numbers is too small for the simple geometry of triangles and squares. In fact, he proved that the length of the diagonal of a unit square is irrational:

$$\sqrt{2} = \sqrt{1^2 + 1^2} \notin \mathbb{Q}.$$

Nowadays this is taught at school and so we may omit a proof (a rather new and simple proof of this fact was given by Estermann; see Exercise 1.7). But for the Pythagoras school it was the death of its philosophy that all natural phenomena could be explained in integers. It is said that Pythagoras sentenced Hippasus to death by drowning (cf. Singh [**153**]). This unkind act could not stop the mathematical progress. In order to solve polynomial equations and to determine the merits in analysis, mathematicians invented various types of *new* numbers:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

In fact, this is only the top of an iceberg; we shall learn about another type of numbers in Chapter 12. We refer the interested reader to the collection [**51**] of excellent surveys on numbers of all kinds. However, to begin with we shall only consider the set $\mathbb{R}$ of real numbers.

An infinite set is called **countable** if there exists a bijection onto $\mathbb{N}$; otherwise the set is said to be **uncountable**. It is easy to see that any union of countably many countable sets is again countable. $\mathbb{Q}$ is countable as shown by the following one-to-one mapping from $\mathbb{N}$ to the set of positive rationals:

$$\mathbb{Q}^+ \ni \frac{m}{n} \quad \longleftrightarrow \quad n + \frac{1}{2}(m+n-1)(m+n-2) \quad \text{for } m, n \in \mathbb{N}.$$

The real numbers represent a quite different type of infinite set than $\mathbb{Q}$.

Theorem 1.3. *The set $\mathbb{R}$ of real numbers is uncountable.*

This is a famous result of Cantor and it simply shows that almost all real numbers are irrational. Its proof relies on his marvelous **diagonalization argument**.

Proof. It suffices to prove the assertion for the subset of real numbers lying in the interval $(0, 1]$.

Suppose the contrary, that is, $\mathbb{R}$ is countable. Let $\{r_1, r_2, \ldots\}$ be a listing of all real numbers in $(0, 1]$. Using the decimal fraction expansion, every r_n can be written as

$$r_n = 0.a_{n1}a_{n2}a_{n3}\ldots,$$

where the digits a_{nk} are integers satisfying $0 \leq a_{nk} \leq 9$. If we assume additionally that we do not allow any infinite sequence of zeros at the end, this decimal expansion is unique (e.g., $0.1 = 0.09999\ldots$). Now define some $r = 0.b_1b_2b_3\ldots$ by choosing $b_n \in \{1, \ldots, 8\}$ different from a_{nn} (the diagonal entry in our list) for each n. Then, r is a real number in $(0, 1)$ which does not appear in the list of the r_n (since r differs in the nth entry: $b_n \neq a_{nn}$). This is the desired contradiction. •

Theorem 1.3 marks the beginning of modern set theory. Hilbert once said in tribute to Cantor that *no one will drive us from the paradise that Cantor has created.*

1.6. Irrationality of e

One of the most fundamental functions in analysis (and natural sciences) is the exponential function given by the infinite series

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!};$$

as usual, we will sometimes also write e^x for $\exp(x)$. A special role plays **Euler's number**

$$\mathrm{e} := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = 2.71828\,18284\ldots.$$

The exponential series converges very fast. For instance, taking into account the first 15 terms gives the approximation above.

Theorem 1.4. e *is irrational.*

This result dates back to Euler in 1737, resp. Lambert in 1760; however, their approach via continued fractions is rather difficult (we will return to this subject later). The following simple proof would easily have been possible in Euler's time.

Proof. Suppose the contrary; then there exist positive integers a, b such that $\mathrm{e} = \frac{a}{b}$. Let m be an integer $\geq b$. Then b divides $m!$ and the number

$$\alpha := m! \left(\mathrm{e} - \sum_{n=0}^{m} \frac{1}{n!} \right) = a\frac{m!}{b} - \sum_{n=0}^{m} \frac{m!}{n!}$$

is an integer (term by term). We have

$$\alpha = \sum_{n=m+1}^{\infty} \frac{m!}{n!} < \frac{1}{m+1} \sum_{k=0}^{\infty} \left(\frac{1}{m+1} \right)^k .$$

By the formula for the infinite geometric series, we can easily bound the right-hand side and find

$$0 < \alpha < \frac{1}{m+1} \cdot \frac{1}{1 - \frac{1}{m+1}} = \frac{1}{m} \leq 1.$$

Since the interval $(0, 1)$ is free of integers, this contradicts the fact that α is integral. The theorem is proved. •

This proof reveals an important principle in the diophantine toolbox: *the series converges so fast that the limit cannot be of a restricted arithmetic nature!*

The question whether a given real number is irrational might seem to be simple at first glance. Actually, this is a rather difficult problem. For instance, it is unknown whether the **Euler–Mascheroni constant** is irrational:

$$\gamma := \lim_{N \to \infty} \left(\sum_{n=1}^{N} \frac{1}{n} - \log N \right) = 0.57721 \ldots \notin \mathbb{Q} \ ?$$

1.7. Irrationality of π

Another important constant is the ratio π of the circumference to the diameter of a circle. We define

$$\pi \ = \ 3.14159\,26535\,89793 \ldots$$

to be the least positive root of the sine function. Our next aim is

Theorem 1.5. *π and π^2 are irrational.*

The first proof of the irrationality of π was given by Lambert in 1761, also by using continued fractions. The proof which we shall give now, as all other known proofs, is slightly more difficult than the one just given for e. This holds true for other questions concerning these two fundamental numbers. It seems that e has somehow more *structure* than π. Our short but tricky proof is due to Niven [**124**].

Proof. We start with some preliminaries. For $n \in \mathbb{N}$ define the function

$$f_n(x) = \frac{1}{n!} x^n (1-x)^n. \tag{1.4}$$

It is obvious that

$$0 < f_n(x) < \frac{1}{n!} \qquad \text{for} \quad 0 < x < 1. \tag{1.5}$$

By the binomial theorem,

$$(1-x)^n = \sum_{j=0}^{n} \binom{n}{j} (-x)^j.$$

Since the binomial coefficients are integers (this follows immediately from their combinatorial meaning), we have

$$f_n(x) = \frac{1}{n!} \sum_{j=n}^{2n} c_j x^j,$$

where the c_j are integers (actually, they are equal to $\pm\binom{n}{j}$ but we do not need this information). The functions $f_n(x)$ share a symmetry of type $f(x) = f(1-x)$. Differentiation of this functional equation leads to

$$f_n^{(k)}(x) = (-1)^k f_n^{(k)}(1-x),$$

where $f^{(k)}$ denotes the kth derivative of f. Taking into account the Taylor series expansion (or dumb computation) we deduce

$$(-1)^k f_n^{(k)}(1) = f_n^{(k)}(0) = \begin{cases} 0 & \text{if} \quad 0 \le k < n, \\ \frac{k!}{n!} c_k & \text{if} \quad n \le k \le 2n. \end{cases} \tag{1.6}$$

Note that the values in (1.6) are all integers.

Now we are in the position to prove the theorem. Obviously, it suffices to show that π^2 is irrational. Assume that $\pi^2 = \frac{a}{b}$ with positive integers a and b. We consider the polynomial

$$F_n(x) := b^n \left(\pi^{2n} f_n(x) - \pi^{2n-2} f_n^{(2)}(x) \pm \ldots + (-1)^n f_n^{(2n)}(x) \right).$$

Since $b^n \pi^{2k} = b^{n-k} a^k \in \mathbb{Z}$ for $0 \le k \le n$, it follows from (1.6) that $F_n(0), F_n(1) \in \mathbb{Z}$. A short calculation shows

$$\left(F_n'(x)\sin(\pi x) - \pi F_n(x)\cos(\pi x)\right)' = \pi^2 a^n f_n(x)\sin(\pi x).$$

Taking into account $\sin \pi = \sin 0 = 0$ this yields

$$\mathcal{I}_n := \pi a^n \int_0^1 f_n(x)\sin(\pi x)\,\mathrm{d}x = F_n(0) + F_n(1).$$

In view of our previous observation it follows that $\mathcal{I}_n$ is an integer. On the other side with regard to (1.5) we get

$$0 < \mathcal{I}_n < \pi \frac{a^n}{n!}.$$

Since the exponential series for $\exp(a)$ converges, $n!$ grows faster than a^n as $n \to \infty$, and thus the right-hand side is <1 for sufficiently large n. This contradicts $\mathcal{I}_n \in \mathbb{Z}$ and the theorem is proved. •

Again this proof is very interesting: *a problem concerning the arithmetic nature of a given real number is solved by the construction of an appropriate sequence of polynomials with respect to its analytic behavior!*

This method of proof can also be applied to prove the irrationality of the exponential function at any non-zero rational value (see Exercise 1.11).

1.8. Approximating with rationals

In 1682, the astronomer and mathematician Huygens (1629–1695) built an automatic planetarium. In one year Earth covers $359^o45'40''30'''$ and Saturn covers $12^o13'34''18'''$, which gives the ratio

$$\frac{77\,708\,431}{2\,640\,858} = 29.42544\ldots\ .$$

Huygens had to construct a gear mechanism which materializes this ratio well. It makes sense to search for approximations which allow a *good* approximation with only a *few* teeth for the gears. So Huygens was looking for *small* integers whose ratio is sufficiently close to the preceding one. The first idea for such an approximation might be $\frac{294}{10} = \frac{147}{5}$ coming from the decimal fraction expansion. However, this does not approximate sufficiently good; the error is $0.02544\ldots$, so more than two percent. Can we do better? Huygens could; he found the rational approximation $\frac{206}{7}$. The error of this approximation is

$$\frac{206}{7} - \frac{77\,708\,431}{2\,640\,858} = 0.00312\ldots\ ,$$

which is less than $40'$ in a century! How did Huygens find this excellent approximation?

Almost all real numbers are irrational. If we have to deal with such numbers, the situation is even worse than in Huygens' case. Computers cannot work with irrationals! In fact, a computer even has problems working with rational numbers; however, there are several strategies to overcome this problem (we will meet this theme in Section 12.8). Fortunately, for most problems it is sufficient to have an *approximate* solution. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, it is natural to search for rational approximations. However, $\mathbb{Q}$ has the disadvantage of being very *thin* in $\mathbb{R}$.

We return to the famous constant π. It is interesting to see which rational approximations to π were used in ancient times:

- The Rhind Papyrus ($\approx$ 1650 B.C.): $\pi \approx 4\left(\frac{8}{9}\right)^2 = 3.16\ldots$;
- Old Testament ($\approx$ 1000 B.C.): $\pi \approx 3$;
- Archimedes (287–212 B.C.): $\pi \approx \frac{22}{7} = 3.142\ldots$;
- Tsu Chung Chi ($\approx$ 500 A.D.): $\pi \approx \frac{355}{113} = 3.14159\,29\ldots$.

How could they find these approximations in those Dark Ages without computers? Their approaches used the underlying geometry.

We shall briefly sketch how Archimedes came to his approximation. He considered a circle of radius one, in which he inscribed a regular polygon of 96 sides, and circumscribed a regular polygon with the same number of sides, such that the first polygon has its vertices on the circle; the second

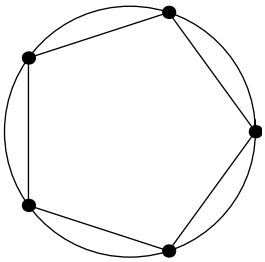


FIGURE 1.1. The pentagon inscribed in the unit circle has area $\frac{5}{2}\sin\frac{2\pi}{5} = 2.37764\ldots$, which gives a *poor* lower bound for π.

one, the midpoints of its edges. Comparing the perimeters led him to the inequality

$$\frac{223}{71} < \pi < \frac{22}{7}. \tag{1.7}$$

Alternatively, one can also consider the areas. This method of exhaustion can be used to find as good rational approximations to π as we please; however, this algorithm is not very efficient. Recently, Kanada & Takahashi computed π up to more than 206 billion digits, based on fast converging series, so calculus replaces geometry. Such a precision is beyond any use in applications (the Planck constant 10^{-33} is the smallest unit in quantum mechanics) but interesting from a mathematical point of view. To remember the first decimals, we recommend the rhyme

> *Now I want a drink, alcoholic of course, after the heavy lectures involving quantum mechanics!*

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, for any real number there exist infinitely many rational approximations and we can approximate with any assigned degree of accuracy. But what are *good* and what are *bad* approximations among them? Thinking back to Huygens' gears, we find that a natural measure for a rational approximation is its denominator.

Let α be any real number; then we say that $\frac{p}{q} \in \mathbb{Q}$ with $q \geq 1$ is a **best approximation** to α if

$$|q\alpha - p| < |Q\alpha - P| \qquad \text{for all} \quad Q < q, \tag{1.8}$$

where $P, Q \in \mathbb{Z}$. Necessarily a best approximation $\frac{p}{q}$ is a reduced fraction (that is, p and q are coprime). Dividing inequality (1.8) by q shows

$$\left|\alpha - \frac{p}{q}\right| < \frac{Q}{q}\left|\alpha - \frac{P}{Q}\right| < \left|\alpha - \frac{P}{Q}\right|$$

(since $Q < q$). Consequently, a best approximation $\frac{p}{q}$ to α is the nearest rational number with denominator $\leq q$. However, the converse does not hold as we shall see in the following section.

The first best approximations to π are

$$\frac{3}{1},\ \frac{22}{7},\ \frac{333}{106},\ \frac{355}{113},\ \frac{1\,03993}{33102},\ \ldots \quad \to \quad \pi. \tag{1.9}$$

It is remarkable that the fractions given by Archimedes and Tsu Chung Chi are best approximations. This means that they could not do better than they did. Given a real number we want to approximate, there is usually no geometric information which we can use to find an appropriate rational approximation (as we did in the case of π). So we are faced with the problem of finding an *efficient* and *universal* algorithm which provides the best approximations to any given real number.

1.9. Linear diophantine equations

In honor of Diophantus we speak about

- **Diophantine approximations** when we search for rational approximations to rational or irrational numbers;
- **Diophantine equations** when we investigate polynomial equations for solutions in integers (or rationals).

As we shall show now these areas are linked in both directions.

We consider linear diophantine equations in two variables. For instance, we may ask for solutions of the equation

$$106X - 333Y = 1 \tag{1.10}$$

in integers; here and in the sequel we denote variables by capitals and corresponding solutions by small letters. One approach to answering this question offers Euclid's algorithm.

We recall some facts from elementary number theory. By **division with remainder**, for any positive integers a, b with $b \le a$ there exist integers q, r such that

$$a = bq + r \qquad \text{with} \quad 0 \le r < b.$$

Now define $r_{-1} := a, r_0 := b$. Then, successive application of division with remainder yields the **Euclidean algorithm**:

$$\begin{array}{lcll} & & \mathsf{For} \quad n = 0, 1, \ldots \quad \mathsf{do} & \\ r_{n-1} & = & q_{n+1} r_n + r_{n+1} & \mathsf{with} \quad 0 \le r_{n+1} < r_n. \end{array} \tag{1.11}$$

Since the sequence of remainders is strictly decreasing, the algorithm terminates, and by simplest divisibility properties the last non-vanishing remainder r_m is equal to the greatest common divisor of a and b,

$$r_m = \mathsf{gcd}(a, b).$$

It should be noted that the Euclidean algorithm is *very fast*; more precisely, the number of steps m is bounded by a polynomial in the input length.

Reading the Euclidean algorithm backwards, we can substitute any r_{n+1} in terms of r_{n-1} and r_n one after the other down to $n = 0$. This yields a

representation of r_m as a linear combination of $r_{-1} = a$ and $r_0 = b$. Thus we get an explicit integral solution of the linear equation

$$bX - aY = \mathsf{gcd}(a, b), \tag{1.12}$$

say x_0, y_0. It is easily seen that then all integral solutions to the latter diophantine equation are given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + k \frac{1}{\mathsf{gcd}(a, b)} \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{for} \quad k \in \mathbb{Z}. \tag{1.13}$$

From here it is only a small step to Bezout's theorem:

Theorem 1.6. *The linear diophantine equation*

$$bX - aY = c \tag{1.14}$$

with integers a, b, c is solvable if and only if $\mathsf{gcd}(a, b)$ divides c; in the case of solvability, the set of solutions is given by (1.13).

Proof. Given any integer solution x, y of (1.12), if $\mathsf{gcd}(a, b)$ divides c, then the numbers

$$\mathcal{X} = \frac{c}{\mathsf{gcd}(a, b)}\, x \qquad \text{and} \qquad \mathcal{Y} = \frac{c}{\mathsf{gcd}(a, b)}\, y$$

solve (1.14). For the converse implication simply note that, for any integers x, y, the number $ax + by$ is divisible by $\mathsf{gcd}(a, b)$. The theorem is proved. •

We return to our example (1.10). Applying the Euclidean algorithm shows

$$22 \cdot 106 - 7 \cdot 333 = 1.$$

With regard to our observations above we find that the set of integer solutions to (1.10) is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 22 \\ 7 \end{pmatrix} + k \begin{pmatrix} 333 \\ 106 \end{pmatrix} \qquad \text{for} \quad k \in \mathbb{Z}.$$

It might be a little surprising to see that the solutions x, y of (1.10) yield *good* approximations $\frac{x}{y}$ to $\frac{333}{106}$. We may rewrite each solution as

$$\frac{x}{y} = \frac{333}{106} + \frac{1}{106y}, \tag{1.15}$$

and since the second term is rather small, tending to zero as $|y| \to \infty$, the solutions x, y to (1.10) yield better and better approximations to $\frac{333}{106}$; of course, having Huygens' approximation problem in mind, we are only interested in approximations with a denominator less than 106. Any fraction $\frac{P}{Q}$ with $1 \le Q < 106$ satisfies

$$\left| Q \frac{333}{106} - P \right| = Q \left| \frac{333}{106} - \frac{P}{Q} \right| = Q \frac{|106P - 333Q|}{106Q} \ge \frac{1}{106},$$

equality holding if and only if P, Q is a solution of (1.10). Thus, we cannot approximate $\frac{333}{106}$ better than with a fraction coming from a solution of (1.10). Moreover, the solution x, y with minimal $|y|$ yields a best approximation $\frac{x}{y}$ to $\frac{333}{106}$, here $\frac{22}{7}$ (notice that both rationals are best approximations

to π). Of course, this holds more generally for linear diophantine equations of the form (1.12).

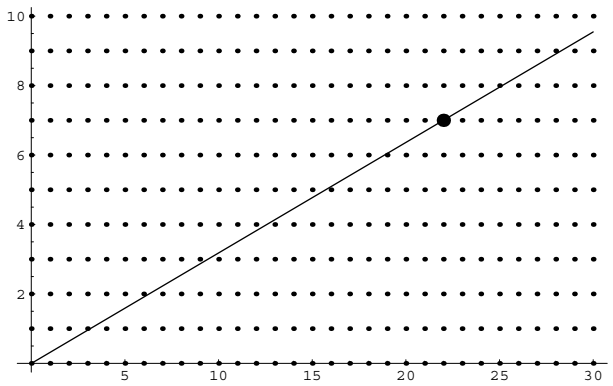


FIGURE 1.2. Integer lattice points (x, y) on the straight line $106X - 333Y = 1$ provide good rational approximations $\frac{x}{y}$ to $\frac{333}{106}$.

In some sense, we have replaced the diophantine equation by an appropriate *diophantine inequality*. So in place of using the Euclidean algorithm backwards, we can also solve the linear diophantine equation (1.12) by searching for best approximations to $\frac{a}{b}$. This was first discovered by Indian mathematicians, namely, Aryabhata around 550 A.D. and Bhaskara around 1150. However, this example marks only the very beginning of the interplay between diophantine approximations and diophantine equations: *certain diophantine equations can be investigated by studying a related problem in the theory of diophantine approximations!*

Exercises

1.1. *How old was Diophantus when he died?*

1.2. *For a Pythagorean triple (x, y, z), show that xyz is divisible by 60.*

1.3.* Consider the recursion defined by $a_1 = 3, c_1 = 5$,

$$a_{n+1} = 3a_n + 2c_n + 1 \qquad \text{and} \qquad c_{n+1} = 4a_n + 3c_n + 2.$$

Prove that $(a_n, a_n + 1, c_n)$ is a Pythagorean triple. Show that this recursion yields all Pythagorean triples of the form $(a, a + 1, c)$.

The first assertion is due to Ryden and the second one was found by Hering; for a more general recursion which constructs *all* primitive Pythagorean triples out of $(3, 4, 5)$, see Gollnick et al. [**67**].

1.4. *Show that it suffices to prove Fermat's last theorem for exponents n being prime and $n = 4$.*

1.5.* *i) Show that the equation*

$$X^4 - 4Y^4 = \pm Z^2$$

has no solutions in positive integers.

Hint: Use Fermat's method of infinite descent.

ii) Deduce from i) that the area of a Pythagorean triangle is not the square of an integer.

1.6.* *Show that $x = y = z = 0$ is the only integral solution of the equation*

$$X^3 + 2Y^3 + 4Z^3 - 34XYZ = 0.$$

1.7. *Let $\mathcal{S}$ be the set of those positive integers n for which $n\sqrt{2}$ is an integer. If $\sqrt{2}$ were rational, $\mathcal{S}$ would be not empty. For the least element of $\mathcal{S}$, say k, consider the number $(\sqrt{2}-1)k$. Deduce from*

$$(\sqrt{2}-1)k\sqrt{2} = 2k - k\sqrt{2}$$

that the set $\mathcal{S}$ is empty and that $\sqrt{2}$ is irrational. Generalize this argument! This idea for proving the irrationality of $\sqrt{2}$ is due to Estermann.

1.8. *Prove that $\mathbb{Q}$ is dense in $\mathbb{R}$.*

1.9.* *Using Archimedes' exhaustion method, improve inequality (1.7). Prove that the area of the regular n-gon with vertices on the unit circle is equal to $\frac{n}{2}\sin\frac{2\pi}{n}$ and show that this tends to π as $n \to \infty$.*

1.10.* For computation of the first digits of π it is convenient to make use of analysis. **Machin's formula** states

$$\frac{\pi}{4} = 4\arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right),$$

where arcus tangent is given by the power series

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} \mp \ldots = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{2n+1}.$$

Deduce Machin's formula from the addition formula of the tangent,

$$\tan(x+y) = \frac{\tan(x)+\tan(y)}{1-\tan(x)\tan(y)}$$

and compute the first ten digits of the decimal fraction of π.

1.11.* With the same notation as in the proof of Theorem 1.5 let q be a positive integer and define

$$G_n(x) = q^{2n}f_n(x) - q^{2n-1}f_n'(x) \pm \ldots + f_n^{(2n)}(x).$$

Prove that $(\exp(qx)G_n(x))' = q^{2n+1}\exp(qx)f_n(x)$, and deduce

$$\int_0^1 q^{2n+1}f_n(x)\exp(qx)\,\mathrm{d}x = \exp(q)G_n(1) - G_n(0).$$

Following the proof of Theorem 1.5, use the above identities to prove that $\exp(q)$ is irrational for any $0 \neq q \in \mathbb{Q}$.