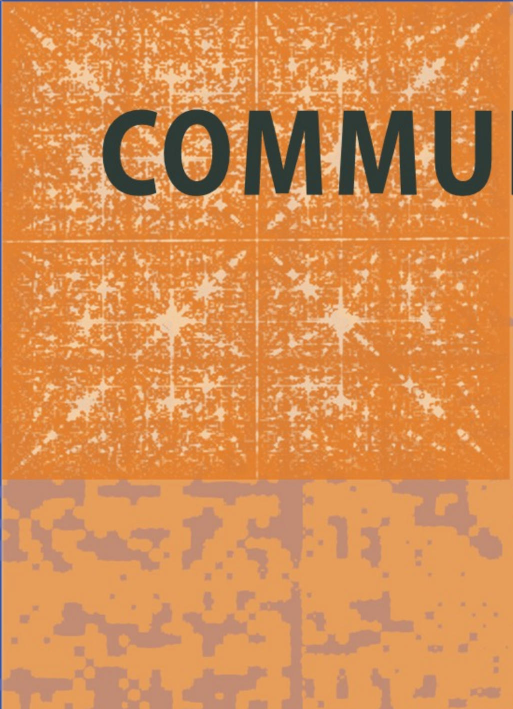
Manfred R. Schroeder

NUMBER THEORY IN SCIENCE AND COMMUNICATION



WITH APPLICATIONS IN CRYPTOGRAPHY, PHYSICS, DIGITAL INFORMATION, COMPUTING, AND SELF-SIMILARITY

Fifth Edition

Springer

Number Theory in Science and Communication

Manfred Schroeder

Number Theory in Science and Communication

With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity

Fifth Edition

Springer

Prof. Dr. Manfred Schroeder
Universität Göttingen
Inst. Physik III
Friedrich-Hund-Platz 1
37077 Göttingen
Germany
mrs17@aol.com

Previously published in Springer Series in Information Sciences as Volume 7

ISBN: 978-3-540-85297-1 e-ISBN: 978-3-540-85298-8

DOI 10.1007/978-3-540-85298-8

Library of Congress Control Number: 2008936508

Cover design: WMXDesign GmbH Heidelberg

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Dedicated to the Memory of
Hermann Minkowski

who added a fourth dimension to our World
and many more to Number Theory

Foreword

"Beauty is the first test: there is no permanent place in the world for ugly mathematics."
– G. H. Hardy

Number theory has been considered since time immemorial to be the very paradigm of pure (some would say useless) mathematics. In fact, the Chinese characters for mathematics are *Number Science*. "Mathematics is the queen of sciences – and number theory is the queen of mathematics", according to Carl Friedrich Gauss, the lifelong Wunderkind, who himself enjoyed the epithet "Princeps Mathematicorum". What could be more beautiful than a deep, satisfying relation between whole numbers? (One is almost tempted to call them wholesome numbers.) In fact, it is hard to come up with a more appropriate designation than their learned name: the integers – meaning the "untouched ones". How high they rank, in the realms of pure thought and aesthetics, above their lesser brethren: the real and complex numbers – whose first names virtually exude unsavory involvement with the complex realities of everyday life!

Yet, as we shall see in this book, the theory of integers can provide totally unexpected answers to real-world problems. In fact, discrete mathematics is taking on an ever more important role. If nothing else, the advent of the digital computer and digital communication has seen to that. But even earlier, in physics the emergence of quantum mechanics and discrete elementary particles put a premium on the methods and, indeed, the spirit of discrete mathematics.

And even in mathematics proper, Hermann Minkowski, in the preface to his introductory book on number theory, *Diophantische Approximationen*, published in 1907 (the year he gave special relativity its proper four-dimensional clothing in preparation for its journey into general covariance and cosmology) expressed his conviction that the "deepest interrelationships in analysis are of an arithmetical nature".

Yet much of our schooling concentrates on analysis and other branches of continuum mathematics to the virtual exclusion of number theory, group theory, combinatorics and graph theory. As an illustration, at a recent symposium on information theory, the author met several young researchers formally trained in mathematics and working in the field of *primality testing*, who – in all their studies up to the Ph.D. – had not heard a single lecture on number theory!

Or, to give an earlier example, when Werner Heisenberg discovered “matrix” mechanics in 1925, he didn’t know what a matrix was (Max Born had to tell him), and neither Heisenberg nor Born knew what to make of the appearance of matrices in the context of the atom. (David Hilbert is reported to have told them to go look for a differential equation with the same eigenvalues, if that would make them happier. They did not follow Hilbert’s well-meant advice and thereby may have missed discovering the Schrödinger wave equation.)

The present book seeks to fill this gap in basic literacy in number theory – not in any formal way, for numerous excellent texts are available – but in a manner that stresses intuition and interrelationships, as well as applications in physics, biology, computer science, digital communication, cryptography and more playful things, such as puzzles, teasers and artistic designs.

Among the numerous applications of number theory on which we will focus in the subsequent chapters are the following:

1) The division of the circle into equal parts (a classical Greek preoccupation) and the implications of this ancient art for modern fast computation and random number generation.
2) The Chinese remainder theorem (another classic, albeit *far* Eastern) and how it allows us to do coin tossing over the telephone (and many things besides).
3) The design of concert hall ceilings to scatter sound into broad lateral patterns for improved acoustic quality (and wide-scatter diffraction gratings in general).
4) The precision measurement of delays of radar echoes from Venus and Mercury to confirm the general relativistic slowing of electromagnetic waves in gravitational fields (the “fourth” – and last to be confirmed – effect predicted by Einstein’s theory of general relativity).
5) Error-correcting codes (giving us distortion-free pictures of Jupiter and Saturn and their satellites).
6) “Public-key” encryption and deciphering of secret messages. These methods also have important implications for computer security.
7) The creation of artistic graphic designs based on prime numbers.
8) How to win at certain parlor games by putting the Fibonacci number systems to work.
9) The relations between Fibonacci numbers and the regular pentagon, the Golden ratio, continued fractions, efficient approximations, electrical networks, the “squared” square, and so on – almost ad infinitum.

I dedicated this book to Hermann Minkowski because he epitomizes, in my mind, the belief in the usefulness outside mathematics of groups and number theory. He died young and never saw these concepts come to full fruition in general relativity, quantum mechanics and some of the topics touched upon here. I am therefore glad that the town of Göttingen is moving to honor its former resident on the occasion of the 100th anniversary of his doctorate on 30 July, 1885 (under F. Lindemann, in now transcendental Königsberg). The late Lilly Rüdenberg, née Minkowski (born in Zürich, while her father was teaching a still unknown Albert there), communicated valuable information in preparation for this late recognition.

Preface to the Fifth Edition

With the advent of quantum cryptography, the endlessly debated Einstein-Podolski-Rosen (EPR) paradox of quantum mechanics has left the sphere of pure speculation and entered the domain of tangible applications. Accordingly, new material on some of the number-theoretic aspects of quantum cryptography (Sects. 9.7–9.9) has been added in this edition.

Also, a number of minor topics, such as a "stamp coverage" problem, related to the work on linear forms by F.G. Frobenius (Sect. 7.3), have been included. There are also two new chapters – on Permutations (Chap. 9) and Covering Sets (Chap. 23).

Göttingen and Berkeley Heights
July 2008

Manfred R. Schroeder

Preface to the Fourth Edition

On February 23, 2005, the city of Göttingen, where Gauss had spent most of his life, commemorated the 150th anniversary of his death.

The day began with a brief ceremony at his grave, not far from his observatory, and continued with the opening of a comprehensive exhibition including the original of his famous notebook and, besides much of his astronomical and geodesic gear, one of the author's quadratic-residue diffusors as described in this book (Chap. 16).

The *Göttinger Tageblatt* published a series of articles entitled *My Gauss*, to which the author was invited to contribute a summary of his "encounters" with the great man. It read, in translation, as follows:

Already as a young boy I had an abiding interest in mathematics. While my classmates were playing softball in the street, I was ensconsed in my upstairs cubicle "pulling little roots", as my mother called it. Among other endeavors, I tried to sum the infinite sequence of reciprocal squares: $1 + \frac{1}{4} + \frac{1}{9} + \cdots$. But I didn't succeed. Gauss made Euler's solution, $\pi^2/6$, later widely known. The rate of convergence of this sequence played a crucial role in my computer graphic "Prime Spectrum" (Fig. 4.9) conceived several decades later.

In high school too I was always inclined towards mathematics. Occasionally I had to help the teacher along when, for example, his explanation of the Doppler-effect threatened do dissolve in thin air.

But my grandmother always thought I should become a physicist and in 1947, after returning from the war, I registered for physics in Göttingen, where Max von Laue and Werner Heisenberg were the shining stars. But in my soul I always remained a mathematician. My parents lived then in Hamm (Westphalia) and during one trip home (4 hours by "express" train) I wanted to calculate the average distance between adjacent primes. After 4 hours I had the answer: the average distance grows with the natural logarithm. For 20, this gives $\log 20 = 3.00$ and, indeed, the spacings between the three primes near 20 (17, 19 and 23) are 2 and 4, or 3 on average (Sect. 4.1).

Gauss had discovered the logarithm in the distribution of primes at a very young age but he has never published a proof. – A formal proof did not appear until a hundred years later.

Althoug originally "married" to theoretical physics in Göttingen, I eventually alighted on the acoustics of concert halls, the topic of my dissertation. Soon I noticed that the spatial transmision function in larger enclosures is a complex Gaussian process, which then (in 1954) was a sensation. I even became a little famous and soon I found myself at the prestigious Bell Laboratories, famous for the invention of the transistor, information theory and error-correcting codes. At Bell, I concentrated on human speech and its efficient representation for transatlantic communication, later applied successfully to mobile telephony (cell phones) and the internet. The criteria I used were based on Gauss' method of least squares which he had pioneered in the rediscovery of the asteroid *Ceres*.

But deep down I never lost my attachment to the prime numbers. I "discovered" that the probability that two arbitrarily selected integers are relative prime approaches $6/\pi^2$ or about 0.608, i.e. the reciprocal of the aforementioned infinite sum $1+\frac{1}{4}+\frac{1}{9}+\cdots$. – It is always surprising to see how different problems in mathematics are interwoven!

But my most consequential encounter with Gauss still lay in the future. In September 1962, in the presence of Jacqueline Kennedy, Philharmonic Hall, the first installment of the Lincoln Center for the Performing Arts in New York City, was ceremoniously inaugurated. But the acoustics of the hall (now called Avery Fisher Hall) left something to be desired: there were disturbing echoes, the celli were difficult to hear, and, worst, the new hall lacked in "intimacy," the feeling of being enveloped by the music. Detailed studies with my students in Göttingen confirmed that for good acoustics, a hall had to have strong *lateral* reflections arriving at a listener's head. But how to realize this goal in a modern wide enclosure with a low ceiling? It occurred to me that the lateral sound in such halls could be enhanced by a structured ceiling that scatters the sound from the orchestra into a broad lateral pattern. But what should such a ceiling look like?

During a talk by the French mathematician André Weil, brother of Simone Weil, in 1977, celebrating the 200th anniversary of Gauss' birth, I fainally saw the light: the ceiling structure should be based on Gauss sums and quadratic residues. Such structures, known as "Schroeder diffusors," are now in worldwide use for improving acoustic quality in concert halls, recording studios and even private homes.

In this fourth edition new material has been added on the Fibonacci numbers (Sects. 5.6, 5.7, 5.8).

Newly rediscovered divisibility tests for 7, 13, 17, and 19 can be found in Sect. 6.1.

An intriguing application of the greatest common divisor is in Sect. 2.4.

A new viewpoint on the decimal system is given in Sect. 2.9.

The section on the two-squares theorem is expanded by a prescription of how to actually decompose a given prime (congruent 1, modulo 4) into two squares. Generalizations of the form $p=x^2+cy^2$, with $c=\pm2,\pm3,\ldots$ are also considered (Sect. 7.9).

Sections 20.6 (Deterministic Primality Testing) and 20.8 (Factoring with Elliptic Curves) have been updated to incorporate recent progress.

Finally, the Introduction is amplified by a consideration of the “Family of Numbers” (p. 4).

Göttingen and Zürich
July 2005

Manfred R. Schroeder

Preface to the Third Edition

Number theory has not rested on its laurels since the appearance of the second edition in 1985. Fermat's last theorem has finally been proven – some 350 years after its bold pronouncement. And Andrew Wiles will before long receive the coveted Wolfskehl Prize of the Göttingen Academy of Sciences. Elliptic curves, which played such a large role in the proof, have given new impetus to fast factoring (Sect. 20.8). Even faster factoring is in the offing if ever quantum computers can be persuaded to cohere long enough for the results to be read out (Sect. 20.8).

Additive number theory has found new applications in exact models in statistical mechanics. This new edition has therefore been amplified by several topics from additive number theory: prime clusters and prime spacings (Sects. 4.8–12), the Goldbach conjecture (Sect. 4.13), and the sum of three primes (Sect. 4.14). "Golomb rulers", which started out as a curiosity, have found important applications in radio astronomy and signal processing (Sect. 30.6). A new application of the two-squares theorem allows the creation of circularly polarized acoustic waves (Sect. 7.9).

Much has also happened in the application of number theory to dynamical systems (Sects. 5.10.1 and 5.10.2).

Galois ("maximum-length") sequences have been turned into multidimensional arrays with surprising applications in X-ray astronomy (Sect. 14.10). New gratings and antenna arrays based on the number-theoretic logarithm show directional patterns with suppressed broadside radiation (Sect. 16.10).

Finally, certain aperiodic sequences of integers have given rise to intriguing rhythms and "baroque" melodies that make us recall the Italian composer Domenico Scarlatti (Sect. 31.4).

Commensurate with its practical importance, the chapters on cryptography and data security have been amplified (Chaps. 10 and 13).

Since the appearance of the second edition several conferences have been devoted to the practical applications of number theory, such as the Les Houches Winter School "Number Theory in Physics" [J. M. Luck et al. *Number Theory in Physics*, Springer 1990] and the symposium of the American Mathematical Society on the "Unreasonable Effectiveness of Number Theory". [S. A. Burr (Ed.): *The Unreasonable Effectiveness of Number Theory*, Am. Math. Soc. 1991].

Göttingen and Berkeley Heights
March 1997

Manfred R. Schroeder

Preface to the Second Edition

The first edition quickly shrank on Springer's shelves, giving me a welcome opportunity to augment this volume by some recent forays of number theory into new territory. The most exciting among these is perhaps the discovery, in 1984, of a *new state of matter*, sharing important properties with both perfect crystals and amorphous substances, without being either one of these. The atomic structure of this new state is intimately related to the Golden ratio and a certain *self-similar* (rab)bit sequence that can be derived from it. In fact, certain generalizations of the Golden ratio, the "Silver ratios" – numbers that can be expressed as periodic continued fractions with period length one – lead one to postulate quasicrystals with "forbidden" 8- and 12-fold symmetries and additional quasicrystals with 5-fold symmetry whose lattice parameters are generated by the Lucas numbers (Sect. 32.1).

The enormously fruitful concept of self-similarity, which pervades nature from the distribution of atoms in matter to that of the galaxies in the universe, also occurs in number theory. And because self-similarity is such a pretty subject, in which *Cantor* and *Julia sets* join *Weierstrass functions* to create a new form of art (distinguished by *fractal Hausdorff dimensions*), a brief celebration of this strangely attractive union seems all but self set.

It is perhaps symptomatic that with *set theory* still another abstract branch of mathematics has entered the real world. Who would have thought that such utterly mathematical constructions as *Cantor sets*, invented solely to reassure the sceptics that sets could have both measure zero and still be uncountable, would make a difference in any practical arena, let alone become a pivotal concept? Yet this is precisely what happened for many natural phenomena from gelation, polymerization and coagulation in colloidal physics to nonlinear systems in many branches of science. Percolation, dendritic growth, electrical discharges (lightnings and *Lichtenberg* figures) and the composition of glasses are best described by set-theoretic *fractal dimensions*.

Or take the weird functions Weierstrass invented a hundred years ago purely to prove that a function could be both everywhere continuous and yet nowhere differentiable. The fact that such an analytic pathology describes something in the real world – nay, is *elemental* to understanding the basins of attraction of strange attractors, for one – gives one pause. These exciting new themes are sounded in Chap. 32.

Physicists working in *deterministic chaos* have been touting the Golden ratio g as the "most irrational" of irrational numbers; and now they gladden us with yet another kind of new number: the *noble numbers*, of which g (how aptly named!) is considered the noblest. *Nonlinear dynamical systems* governed by these "new" numbers (whose continued-fraction expansions end in infinitely many 1's) show the greatest resistance to the onslaught of chaotic motion (such as turbulence). The rings of Saturn and, indeed, the very stability of the solar system are affected by these numbers. This noble feat, too, merits honorable mention (in Sects. 5.3 and 32.1).

Another topic, newly treated and of considerable contemporary import, is *error-free computation*, based on Farey fractions and p-adic "Hensel codes" (Sect. 5.12).

Other recent advances recorded here are applications of the *Zech logarithm* to the design of optimum *ambiguity functions* for radar, new *phase-arrays* with unique radiation patterns, and *spread-spectrum communication* systems (Sects. 27.7–9).

A forthcoming Italian translation occasioned the inclusion of a banking puzzle (Sect. 5.11) and other new material on Fibonacci numbers (in Sect. 1.1 and Chap. 32).

Murray Hill and Göttingen *Manfred R. Schroeder*
April 1985

Acknowledgements

I thank Martin Kneser, Don Zagier, Henry Pollak, Jeff Lagarias and Hans Werner Strube for their careful reviews of the manuscript.

Wilhelm Magnus, Armin Kohlrausch, Paul C. Spandikow, Herbert Taylor and Martin Gardner spotted several surviving slips.

Joseph L. Hall and AT of T Bell Laboratories computed the spectra of the self-similar sequences shown in Chap. 32. Peter Meyer, University of Göttingen, prepared the auditory paradox, described in Sect. 32.2, based on a "fractal" signal waveform.

Norma Culviner's red pencil and yellow paper slips have purged the syntax of original sin. Angelika Weinrich, Irena Schönke and Elvire Hung, with linguistic roots far from English soil, and Dorothy Caivano have typed pages and fed computers to near system overflow. Gisela Kirschmann-Schröder and Liane Liebe prepared the illustrations with professional care and love. Anny Schroeder piloted the ms. through the sea of corrections inundating her desk from two continents.

Dr. Helmut Lotsch of Springer-Verlag deserves recognition for his leadership in scientific publication. Working with Reinhold Michels and Claus-Dieter Bachem from the production department in Heidelberg proved a pure delight.

William O. Baker and Arno Penzias fostered and maintained the research climate at Bell Laboratories in which an endeavor such as the writing of this book could flourish. Max Mathews provided support when it was needed. John R. Pierce inspired me to write by his writings.

Contents

Part I A Few Fundamentals

1 Introduction . . . 3

The Family of Numbers . . . 6

1.1 Fibonacci, Continued Fractions and the Golden Ratio . . . 8

1.2 Fermat, Primes and Cyclotomy . . . 11

1.3 Euler, Totients and Cryptography . . . 13

1.4 Gauss, Congruences and Diffraction . . . 14

1.5 Galois, Fields and Codes . . . 16

2 The Natural Numbers . . . 21

2.1 The Fundamental Theorem . . . 21

2.2 The Least Common Multiple . . . 22

2.3 Planetary "Gears" . . . 23

2.4 The Greatest Common Divisor . . . 23

2.5 Human Pitch Perception . . . 25

2.6 Octaves, Temperament, Kilos and Decibels . . . 26

2.7 Coprimes . . . 28

2.8 Euclid's Algorithm . . . 28

2.9 The Decimal System Decimated . . . 29

3 Primes . . . 31

3.1 How Many Primes are There? . . . 31

3.2 The Sieve of Eratosthenes . . . 32

3.3 A Chinese Theorem in Error . . . 33

3.4 A Formula for Primes . . . 34

3.5 Mersenne Primes . . . 35

3.6 Repunits . . . 39

3.7 Perfect Numbers . . . 39

3.8 Fermat Primes . . . 41

3.9 Gauss and the Impossible Heptagon . . . 42

4 The Prime Distribution 45
4.1 A Probabilistic Argument 45
4.2 The Prime-Counting Function $\pi(x)$ 47
4.3 David Hilbert and Large Nuclei 51
4.4 Coprime Probabilities 52
4.5 Primes in Progressions 56
4.6 Primeless Expanses 58
4.7 Squarefree and Coprime Integers 59
4.8 Twin Primes 59
4.9 Prime Triplets 61
4.10 Prime Quadruplets and Quintuplets 62
4.11 Primes at Any Distance 63
4.12 Spacing Distribution Between Adjacent Primes 66
4.13 Goldbach's Conjecture 67
4.14 Sum of Three Primes 68

Part II Some Simple Applications

5 Fractions: Continued, Egyptian and Farey 73
5.1 A Neglected Subject 73
5.2 Relations with Measure Theory 77
5.3 Periodic Continued Fractions 78
5.4 Electrical Networks and Squared Squares 81
5.5 Fibonacci Numbers and the Golden Ratio 83
5.6 Fibonacci, Rabbits and Computers 87
5.7 Fibonacci and Divisibility 90
5.8 Generalized Fibonacci and Lucas Numbers 90
5.9 Egyptian Fractions, Inheritance and Some Unsolved Problems 94
5.10 Farey Fractions 95
5.10.1 Farey Trees 97
5.10.2 Locked Pallas 100
5.11 Fibonacci and the Problem of Bank Deposits 102
5.12 Error-Free Computing 103

Part III Congruences and the Like

6 Linear Congruences 111
6.1 Residues 111
6.2 Some Simple Fields 114
6.3 Powers and Congruences 115

7 Diophantine Equations 119
7.1 Relation with Congruences 119
7.2 A Gaussian Trick 120
7.3 A Stamp Problem 122
7.4 Nonlinear Diophantine Equations 124
7.5 Triangular Numbers 125
7.6 Pythagorean Numbers 127
7.7 Exponential Diophantine Equations 128
7.8 Fermat's Last "Theorem" 129
7.9 The Demise of a Conjecture by Euler 130
7.10 A Nonlinear Diophantine Equation in Physics and the Geometry of Numbers 131
7.11 Normal-Mode Degeneracy in Room Acoustics (A Number-Theoretic Application) 135
7.12 Waring's Problem 136

8 The Theorems of Fermat, Wilson and Euler 139
8.1 Fermat's Theorem 139
8.2 Wilson's Theorem 140
8.3 Euler's Theorem 141
8.4 The Impossible Star of David 143
8.5 Dirichlet and Linear Progression 144

9 Permutations, Cycles and Derangements 147
9.1 Permutations 147
9.2 Binomial Coefficients 148
9.3 The Binomial and Related Distributions 149
9.4 Permutation Cycles 150
9.5 Derangements 152
9.6 Ascents and Descents 152
9.7 Quantum Decrypting 153
9.8 Decrypting without Factoring 155
9.9 Quantum Cryptography 156
9.10 One-Time Pads 156
9.11 The Bennet-Brossard Key Distribution Scheme (BB84) 157

Part IV Cryptography and Divisors

10 Euler Trap Doors and Public-Key Encryption 161
10.1 A Numerical Trap Door 163
10.2 Digital Encryption 164
10.3 Public-Key Encryption 165
10.4 A Simple Example 167
10.5 Repeated Encryption 168
10.6 Summary and Encryption Requirements 169

11 The Divisor Functions 171
11.1 The Number of Divisors 171
11.2 The Average of the Divisor Function 174
11.3 The Geometric Mean of the Divisors 174
11.4 The Summatory Function of the Divisor Function 175
11.5 The Generalized Divisor Functions 175
11.6 The Average Value of Euler's Function 176

12 The Prime Divisor Functions 179
12.1 The Number of Different Prime Divisors 179
12.2 The Distribution of $\omega(n)$ 182
12.3 The Number of Prime Divisors 185
12.4 The Harmonic Mean of $\Omega(n)$ 188
12.5 Medians and Percentiles of $\Omega(n)$ 190
12.6 Implications for Public-Key Encryption 191

13 Certified Signatures 193
13.1 A Story of Creative Financing 193
13.2 Certified Signature for Public-Key Encryption 193

14 Primitive Roots 195
14.1 Orders 195
14.2 Periods of Decimal and Binary Fractions 197
14.3 A Primitive Proof of Wilson's Theorem 201
14.4 The Index – A Number-Theoretic Logarithm 201
14.5 Solution of Exponential Congruences 202
14.6 What is the Order T_m of an Integer m Modulo a Prime p? 204
14.7 Index "Encryption" 205
14.8 A Fourier Property of Primitive Roots and Concert Hall Acoustics 206
14.9 More Spacious-Sounding Sound 207
14.10 Galois Arrays for X-Ray Astronomy 209
14.11 A Negative Property of the Fermat Primes 211

15 Knapsack Encryption 213
15.1 An Easy Knapsack 213
15.2 A Hard Knapsack 214

Part V Residues and Diffraction

16 Quadratic Residues 219
16.1 Quadratic Congruences 219
16.2 Euler's Criterion 220
16.3 The Legendre Symbol 221
16.4 A Fourier Property of Legendre Sequences 223
16.5 Gauss Sums 224
16.6 Pretty Diffraction 225

16.7 Quadratic Reciprocity 226
16.8 A Fourier Property of Quadratic-Residue Sequences 227
16.9 Spread Spectrum Communication 229
16.10 Generalized Legendre Sequences Obtained Through Complexification of the Euler Criterion 230

Part VI Chinese and Other Fast Algorithms

17 The Chinese Remainder Theorem and Simultaneous Congruences . . . 235
17.1 Simultaneous Congruences 235
17.2 The Sino-Representation: A Chinese Number System 236
17.3 Applications of the Sino-Representation 237
17.4 Discrete Fourier Transformation in Sino 239
17.5 A Sino-Optical Fourier Transformer 240
17.6 Generalized Sino-Representation 241
17.7 Fast Prime-Length Fourier Transform 242

18 Fast Transformation and Kronecker Products 245
18.1 A Fast Hadamard Transform 245
18.2 The Basic Principle of the Fast Fourier Transforms 248

19 Quadratic Congruences 251
19.1 Application of the Chinese Remainder Theorem (CRT) 251

Part VII Pseudoprimes, Möbius Transform, and Partitions

20 Pseudoprimes, Poker and Remote Coin Tossing 255
20.1 Pulling Roots to Ferret Out Composites 255
20.2 Factors from a Square Root 257
20.3 Coin Tossing by Telephone 258
20.4 Absolute and Strong Pseudoprimes 260
20.5 Fermat and Strong Pseudoprimes 262
20.6 Deterministic Primality Testing 263
20.7 A Very Simple Factoring Algorithm 264
20.8 Factoring with Elliptic Curves 265
20.9 Quantum Factoring 265

21 The Möbius Function and the Möbius Transform 267
21.1 The Möbius Transform and Its Inverse 268
21.2 Proof of the Inversion Formula 269
21.3 Second Inversion Formula 270
21.4 Third Inversion Formula 271
21.5 Fourth Inversion Formula 271
21.6 Riemann's Hypothesis and the Disproof of the Mertens Conjecture 271
21.7 Dirichlet Series and the Möbius Function 272

22 Generating Functions and Partitions 275
22.1 Generating Functions 275
22.2 Partitions of Integers 277
22.3 Generating Functions of Partitions 278
22.4 Restricted Partitions 279

23 From Error Correcting Codes to Covering Sets 283
23.1 Covering Sets in Coding Theory 283
23.2 Discrete Covering Sets 284

Part VIII Cyclotomy and Polynomials

24 Cyclotomic Polynomials 289
24.1 How to Divide a Circle into Equal Parts 289
24.2 Gauss's Great Insight 292
24.3 Factoring in Different Fields 296
24.4 Cyclotomy in the Complex Plane 297
24.5 How to Divide a Circle with Compass and Straightedge 298
24.5.1 Rational Factors of $z^N - 1$ 299
24.6 An Alternative Rational Factorization 301
24.7 Relation Between Rational Factors and Complex Roots 301
24.8 How to Calculate with Cyclotomic Polynomials 303

25 Linear Systems and Polynomials 305
25.1 Impulse Responses 305
25.2 Time-Discrete Systems and the z Transform 306
25.3 Discrete Convolution 306
25.4 Cyclotomic Polynomials and z Transform 307

26 Polynomial Theory 309
26.1 Some Basic Facts of Polynomial Life 309
26.2 Polynomial Residues 310
26.3 Chinese Remainders for Polynomials 312
26.4 Euclid's Algorithm for Polynomials 313

Part IX Galois Fields and More Applications

27 Galois Fields 317
27.1 Prime Order 317
27.2 Prime Power Order 317
27.3 Generation of GF(2^4) 320
27.4 How Many Primitive Elements? 321
27.5 Recursive Relations 321
27.6 How to Calculate in GF(p^m) 324
27.7 Zech Logarithm, Doppler Radar and Optimum Ambiguity Functions 324

27.8 A Unique Phase-Array Based on the Zech Logarithm 327
27.9 Spread-Spectrum Communication and Zech Logarithms 328

28 Spectral Properties of Galois Sequences 331
28.1 Circular Correlation 331
28.2 Application to Error-Correcting Codes and Speech Recognition . . 333
28.3 Application to Precision Measurements 335
28.4 Concert Hall Measurements 336
28.5 The Fourth Effect of General Relativity 336
28.6 Toward Better Concert Hall Acoustics 338
28.7 Higher-Dimensional Diffusors 343
28.8 Active Array Applications 344

29 Random Number Generators 347
29.1 Pseudorandom Galois Sequences 348
29.2 Randomness from Congruences 349
29.3 “Continuous” Distributions 350
29.4 Four Ways to Generate a Gaussian Variable 351
29.5 Pseudorandom Sequences in Cryptography 352

30 Waveforms and Radiation Patterns 355
30.1 Special Phases 356
30.2 The Rudin-Shapiro Polynomials 358
30.3 Gauss Sums and Peak Factors 359
30.4 Galois Sequences and the Smallest Peak Factors 361
30.5 Minimum Redundancy Antennas 363
30.6 Golomb Rulers 365

31 Number Theory, Randomness and “Art” 367
31.1 Number Theory and Graphic Design 367
31.2 The Primes of Gauss and Eisenstein 369
31.3 Galois Fields and Impossible Necklaces 370
31.4 “Baroque” Integers 374

Part X Self-Similarity, Fractals and Art

32 Self-Similarity, Fractals, Deterministic Chaos and a New State of Matter 379
32.1 Fibonacci, Noble Numbers and a New State of Matter 382
32.2 Cantor Sets, Fractals and a Musical Paradox 388
32.3 The Twin Dragon: A Fractal from a Complex Number System 394
32.4 Statistical Fractals 395
32.5 Some Crazy Mappings 397
32.6 The Logistic Parabola and Strange Attractors 399
32.7 Conclusion 403

Glossary of Symbols 405

References 409

Name Index 419

Subject Index 423

Part I
A Few Fundamentals

Chapter 1
Introduction

> *"Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk."*
> *– Leopold Kronecker*

Hermann Minkowski, being more modest than Kronecker, once said "The primary source (Urquell) of all of mathematics are the integers." Today, integer arithmetic is important in a wide spectrum of human activities and natural phenomena amenable to mathematic analysis.

Integers have repeatedly played a crucial role in the evolution of the natural sciences. Thus, in the 18th century, Lavoisier discovered that chemical compounds are composed of fixed proportions of their constituents which, when expressed in proper weights, correspond to the ratios of *small integers*. This was one of the strongest hints to the existence of atoms, but chemists for a long time ignored the evidence and continued to treat atoms merely as a conceptual convenience devoid of any physical meaning. (Ironically, it was from the statistical laws of *large* numbers in Einstein's analysis of Brownian motion at the beginning of our own century, that the irrefutable reality of atoms and molecules finally emerged.)

In the analysis of optical spectra, certain integer relationships between the wavelengths of spectral lines emitted by excited atoms gave early clues to the *structure of atoms*, culminating in the creation of matrix mechanics in 1925, an important year in the growth of *integer physics*.

Later, the near-integer ratios of atomic weights suggested to physicists that the *atomic nucleus* must be made up of an *integer* number of similar nucleons. The *deviations* from integer ratios led to the discovery of elemental *isotopes*.

And finally, small divergencies in the atomic weight of pure isotopes from exact integers constituted an early confirmation of Einstein's famous equation $E = mc^2$, long before the "mass defects" implied by these integer discrepancies blew up into widely visible mushroom clouds.

On a more harmonious theme, the role of integer ratios in musical scales has been widely appreciated ever since Pythagoras first pointed out their importance. The occurence of integers in biology – from plant morphology to the genetic code – is pervasive. It has even been hypothesized that the North American 17-year cicada selected its life cycle because 17 is a prime number, prime cycles offering better

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_1, © Springer-Verlag Berlin Heidelberg 2009

protection from predators than nonprime cycles. (The suggestion that the 17-year cicada "knows" that 17 is a *Fermat* prime has yet to be touted though.)

Another reason for the resurrection of the integers is the penetration of our lives achieved by that 20th-century descendant of the abacus, the *digital computer*. (Where did all the slide rules go? Ruled out of most significant places by the ubiquitous pocket calculator, they are sliding fast into restful oblivion. – Collectors of antiques take note!)

An equally important reason for the recent revival of the integer is the congruence of *congruential arithmetic* with numerous modern developments in the natural sciences and digital communications – especially "secure" communication by cryptographic systems. Last but not least, the proper protection and security of computer systems and data files rests largely on congruence relationships.

In congruential arithmetic, what counts is not a numerical value per se, but rather its remainder or *residue* after division by a *modulus*. For example, in wave interference (be it of ripples on a lake or of electromagnetic fields on a hologram plate) it is not path differences that determine the resulting interference pattern, but rather the residues after dividing by the wavelength. For perfectly periodic events, there is no difference between a path difference of half a wavelength or one and a half wavelengths: in either case the interference will be destructive.

One of the most dramatic consequences of congruential arithmetic is the existence of the chemical elements as we know them. In 1913, Niels Bohr postulated that certain integrals associated with electrons in "orbit" around the atomic nucleus should have integer values – a requirement that 10 years later became comprehensible as a wave interference phenomenon of the newly discovered *matter* waves: essentially, integer-valued integrals meant that path differences are divisible by the electron's wavelength without leaving a remainder.

But in the man-made world, too, applications of congruential arithmetic abound. In binary representation and error-correcting codes, the important Hamming distance is calculated from the sum (or difference) of corresponding places *modulo* 2. In other words, $1+1$ equals 0 – just as two half-wavelength path differences add up to no path difference at all.

Another example of a man-made application of congruential arithmetic is a public-key encryption method in which messages represented by integers are raised to a given power and only the residue, modulo a preselected encryption modulus, is transmitted from sender to receiver. Thus, in a simplified example, if the message was 7, the encrypting exponent 3 and the modulus 10, then the transmitted message would not be $7^3 = 343$, but only the last digit: 3. The two most significant digits are discarded. Of course, in a serious application, the prime factors of the modulus are so large (greater than 10^{200}, say) that knowledge of their product gives no *usable* clue to the factors, which are needed for decryption.

Thus, we are naturally led to the most important concept of number theory: the distinction between *primes* and nonprimes, or *composites*, and to the properties of divisibility, the greatest common divisor and the least common multiple – precursor to the art of factoring.

These topics form the theme of the first four chapters of the book. The reader will find here an emphasis on *applications* (musical scales and human pitch perception, for example). Also, the useful *probabilistic viewpoint* in number theory is introduced here (Chap. 4) and applied to "derivations" of the prime density and coprime probability.

The book continues with a discussion of *continued fractions, Fibonacci series* and some of their *endless* applications (Chap.5).

In Chap. 6 *congruences* are introduced and put to work on *Diophantine equations* – both linear and nonlinear (Chap. 7). Again, there are interesting applications both within mathematics (*Minkowski's* geometry of numbers) and without.

In Chap. 8, in preparation for one of our main themes – *cryptography* (Chaps. 10, 13, 15) – we become acquainted with the pivotal *theorems of Fermat and Euler*.

Some of the basic concepts of *factoring* – the various *divisor functions* – are taken up in Chaps. 11 and 12, again stressing the probabilistic viewpoint.

Chapter 14 is devoted to the important concepts of *order, primitive roots* and the number theoretic *index* function. Among the applications encountered are period lengths of periodic fractions and, surprisingly, a prescription for better *concert hall acoustics*.

In Chap. 16 we are introduced to *quadratic residues, Gauss sums* and their *Fourier* spectrum properties.

The *Chinese remainder theorem* (Chap. 17) allows us to tackle *simultaneous congruences* and *quadratic congruences* (Chap. 19). The so-called *Sino-representation* of integers leads to important *fast algorithms*, a theme that is continued in Chap. 18 where we examine the basic principle (*Kronecker* factorization, direct products) behind such indispensable algorithms as the *fast Hadamard* (FHT) and *Fourier transforms* (FFT).

Since in encryption there will always be a need for *very large primes* beyond the "primality horizon", that need may have to be satisfied by *pseudoprimes* which we define in Chap. 20. The techniques of testing the so-called *strong* pseudoprimes are related to such "games" as coin tossing over the telephone and *certified digital mail*.

Chapter 22 introduces *generating functions* and *partitions*, an important topic in *additive* number theory, which is a difficult field and poses many unsolved problems.

In preparation for our excursion into the art and science of dividing the circle (Chap. 24), we acquaint ourselves with the *Möbius* function (Chap. 21) and then put *polynomials* (Chaps. 25 and 26) to use to construct *finite number fields* (Chap. 27). Some of their numerous applications based on the spectrum and correlation properties of *"Galois" sequences* are discussed in Chap. 28 (*error-correcting codes, precision measurements and antenna theory*), Chap. 29 (*random number generators*) and Chap. 30 (special *waveforms* for *sonar, radar* and *computer speech synthesis*). Minimum redundancy arrays, important in radio astronomy, underwater sound detection and real-time diagnostic tomography, are also mentioned here.

The book concludes with a brief excursion into the world of *fractals* and *artistic design* based on the aesthetically desirable mixture of regularity and randomness inherent in the distribution of primes and primitive polynomials.

In the sections which follow some of these topics are introduced informally, together with the leading dramatis personae. The reader interested in widening her or his arithmetic horizon is referred to [1.1–28].

The Family of Numbers

Whole(some) Numbers

Before approaching prime and other numbers, we should take a brief look at the family of numbers in which they grew up. There are, first of all, the natural numbers or positive 'whole numbers' – 1, 2, 3, 4 and so forth – which, according to the German mathematician Leopold Kronecker, were created by God. Good to know! All the rest, Kronecker decreed, is *Menschenwerk* (the work of mere mortals).

Of course, the integers werde not revealed in one fell swoop. First came the *positive* integers (although not a mathematical text, recall the *Book of Numbers*).

But soon people, especially traders, felt a need for *negative* numbers (think of overdrawn bank accounts). But the zero so familiar to us remained unrecognized for a long long time. What a difficult time our distant ancestors had with 'nothing' is perhaps best illustrated by the following tale. The ancient Egyptians, perennial pyramid builders, had a neat formula for the volume V of an *unfinished* pyramid,

$$V = \frac{h}{3}(a^2 + ab + b^2), \tag{1}$$

where h is the height, and a und b are the sides of the base and the top square, respectively. Of course, for $a = b$, they got the volume of the cube: ha^2. Check!

The above formula ist quite an accomplishment. Even with modern calculus it takes a clever trick or two to derive it. And of course, the Egyptians didn't have calculus, modern or otherwise.

Admittedly, the formula for unfinished pyramids was important because it allowed them to calculate, on the basis of the added volume, necessary material and the weekly wages (if any). But, beyond belief, they never seem to have hit upon the simple formula for the finished pyramid,

$$V = \frac{h}{3}a^2, \tag{2}$$

obtained by setting b equal to zero in their equation. Where they afraid of zero? Did they 'equate' it with nothingness, the fearsome void? After all, as the pyramid nears completion, b gets smaller and smaller and, upon completion, b vanishes. It sounds so easy to our modern minds. But what do you do, if you don't have zero? This little story just goes to show again how misleading it can be to argue an ancient case from hindsight.

How dangerous ‘nothing’ or ‘nobody’ can be is nicely illustrated by the account of Odysseus and the Cyclops whom Odysseus told his name was *Outis*, Greek for nobody. Later when the itinerant hero attacked the one-eyed monster, this cried ‘Help, help. Nobody is killing me’. And nobody came.

Rational Numbers

The positive and negative whole numbers, supplemented by zero, constitute the *integers* $\ldots -3,-2,-1,0,1,2,\ldots$, literally (from Latin) the ‘untouched ones’. But they didn’t remain ‘untouched’ for long. Already the old Greeks used them to describe *ratios*, like the side ratio 3/4 of a 3-by-4 rectangle. Or the ratio of the rectangle’s diagonal (5 according to Pythagoras) to the longer side: 5/4, a *rational number*.

Maybe the Greeks, so fond of ratios, should have stopped then and there. But no, they had to venture forth and ask for the ratio of the diagonal of a 1-by-1 *square* to its side – and almost lost their minds. Obviously, the desired ratio is greater than 1 but less than 2. So what is it? Let’s try m/n, where m/n is the ratio of two integers *with no common factors*, say 7/5. The square of this ratio, 49/25, once more following Pythagoras, must equal $1^2+1^2=2$. Thus, we can see that 7/5 was not a bad guess – a bit to small, but close. Really, we are looking for two numbers m and n such that $m^2=2n^2$, implying that m must be even, say $m=2k$. With k we have $m^2=4k^2=2n^2$ or $2k^2=n^2$. Now it is the turn of n to be even. Hence the solution of the Greek’s problem is two even integers that have no common factors. Unfortunately, two even integers do have a common divisor, namely 2. Hence there are no integers m and n such that $(m/n)^2=2$. End of proof.

The Greeks were devasted. Here they had found a manifest length, the diagonal of a square, that to their way of reckoning had no length. There was no common yardstick for measuring both the sides and the diagonal of a square. And they had had such an enduring faith in ‘yardsticks’! Their enduring dictum ‘everything is number’ endured no more.

It took a long time for mathematicians, not only in Greater Greece, to simply call such numbers as the square root of 2, written $\sqrt{2}$, *irrational numbers*. And it wasn’t until the 19th century that the German mathematician Richard Dedekind (1831–1916) showed that irrational numbers were nothing to be afraid of and that one could calculate with them without running the risk of contradiction. Thus, for example, the product of the diagonal of the 1-by-1 square, $\sqrt{2}$, and the diagonal of the 2-by-2 square, which equals $\sqrt{2^2+2^2}=\sqrt{8}$, is $\sqrt{2}\times\sqrt{8}=\sqrt{16}=4$. In other words, ‘no yardstick times no yardstick equals a yardstick.’ Try it on your pocket calculator (but don’t throw it away if the answer is 3.999999999).

Transcendental Numbers

Irrational numbers, such as $\sqrt{2}$, are given by algebraic equations like $x^2=2$ or $y^2+y=1$. the latter equation has the positive solution $y=\frac{1}{2}(\sqrt{5}-1)$, an irrational number known as the golden mean. (In fact, loose-lipped physicists sometimes call the golden mean 'the most irrational number' – for excellent reasons, as we shall see.)

What algebraic equation gives π as a solution? Again we are stymied. the number π transcends the algebraic irrationals and is therefore called *transcendental*, a fact that was not proved until 1882 by Ferdinand Lindemann[1] (This proof should have put a full stop to the syndicate of circle squares, but some people are still trying.)

Another well-known transcendental number is Napier's number e, the base of the natural logarithmus. And there are many more: in fact, the number of transcendentals is not only infinite but *uncountably* infinite. In other words, you can't make a list of them, not even an infinitely long list. By contrast, the algebraic numbers, like $\sqrt{2}$ and the golden mean, form a countably infinite set.

Some algebraic numbers are difficult to approximate by rationals. By contrast, there are transcendental numbers that come dangerously close to being mistaken for a rational number. Take the 'Liouville number'

$$L = 0.110001000000000000000000010\ldots$$

whose only nonzero digits are 1s with the nth 1 appearing in the $(n!)$th place. Liouville was able to show that L cannot be the solution of any algebraic equation. L was thus the first solid evidence for the existence of transcendental numbers, even before e and π were shown to be transcendental. The 'degree of irrationality' plays a significant role in the physics of nonlinear dynamical systems.

Another numerical distinction important in modern physics is the difference between *normal numbers* and nonnormals. For normal numbers in a given base all digits and digit combinations are equally likely. For example π is believed to be a normal number in the decimal system and the first billion digits bear this out, but there is no proof. Similarly, $\sqrt{2}$ appears to be normal in decimal notation but, according to recent extensive tests, less so than π. And e seems to be less normal in decimal than either π or $\sqrt{2}$. Very Curious!

1.1 Fibonacci, Continued Fractions and the Golden Ratio

In 1202, the Italian mathematician Fibonacci (also known as Leonardo da Pisa) asked a simple question. Suppose we have a pair of newly born rabbits who, after maturing, beget another pair of rabbits. These children, after they mature, beget another pair. So we have first one pair of rabbits, then two pairs, and then three pairs. How will this continue, if it does, supposing that each new pair of rabbits, after one season of maturing, will beget another pair each and every breeding season thereafter? To make things simple, Fibonacci also assumed that rabbits never die.

[1] Lindemann's greatest scientific contribution may have been to physics rather then mathematics by scaring away the young Heisenberg who originally wanted to study mathematics under Lindemann.

Obviously, the number of rabbit pairs in the nth breeding season will be equal to the number of pairs one season earlier (because none have died) *plus* the number of rabbits two seasons earlier because all of those rabbits are now mature and each pair produces a new pair.

Calling the number of rabbit pairs in the nth season F_n (F as in Fibonacci), we then have the recursion formula

$$F_n = F_{n-1} + F_{n-2}.$$

Starting with one immature pair of rabbits $(F_1 = 1, F_2 = 1)$, it is easy to calculate the number of rabbit pairs each successive season:

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ \ldots,$$

where each number is the sum of its two predecessors. Nothing could be much simpler than this – and few things probably seem as useless. At least this is what one might be led to believe by the unreality and triviality of this rabbit story. And yet this apparently harmless example of numerology touches more domains of mathematics and more diverse applications than many other simply defined collections of numbers.

For one, the ratio of successive numbers approaches the so-called Golden ratio $g = 1.618\ldots$, often considered the ideal proportion in art. (Strangely, the very same ratio also plays a role in modern theories of deterministic chaos, where g is considered the “most irrational” of all ratios.)

The Golden ratio is defined in geometry by sectioning a straight line segment in such a way that the ratio of the total length to the longer segment equals the ratio of the longer to the shorter segment. In other words, calling the total length of the line l, and the longer segment a, g is determined by the equation

$$g := \frac{l}{a} = \frac{a}{l-a}.$$

Substituting g for l/a on the right side of the equation yields

$$g = \frac{1}{g-1},$$

which leads to the quadratic equation for g:

$$1 + g - g^2 = 0.$$[2]

The only positive solution (length ratios are positive) is

$$g = \frac{1+\sqrt{5}}{2},$$

[2] The connoisseur may recognize the similarity between this equation, which defines the Golden ratio, and the generating function of the Fibonacci sequence: $1/(1-x-x^2)$.

which also happens to be the ratio of a diagonal to a side in the regular pentagon. Since no higher roots than square roots ($\sqrt{5}$) appear in this ratio, it follows that the regular pentagon can be constructed by straightedge and compass alone. So Fibonacci's rabbits have now brought us into contact with one of the classical problems of geometry – the problem of circle division or cyclotomy, a problem that will occupy us again and again in this book because of its importance for fast computational algorithms and numerous other applications.

Incidentally, there is also a fast algorithm making use of g to compute F_n without repeated additions:

$$F_n \approx \frac{g^n}{\sqrt{5}},$$

where the wavy equal sign here means "take the nearest integer".

The remarkable numbers F_n can also be used to represent any positive integer N in a unique way by the simple rule that N is written as the greatest F_n not exceeding N and a (positive) remainder. Then this remainder is represented by the greatest Fibonacci number not exceeding *it* and another remainder, and so forth. For example, with the aid of the above sequence,

$$\begin{aligned} 83 &= 55+28 \\ 28 &= 21+7 \\ 7 &= 5+2. \end{aligned}$$

Thus, in the Fibonacci system,

$$83 = 55+21+5+2.$$

Note that no two of the Fibonacci numbers occurring in the expansion of 83 are adjacent F_n. This is a general rule and distinguishes the Fibonacci system from the binary system where adjacent powers of 2 are often needed. (For example, $83 = 64+16+2+1$, where 2 and 1 are adjacent powers of 2.)

This property of the Fibonacci system has two interesting (and somewhat unexpected) consequences. The first is that in fast search algorithms it is often more efficient to base sequential searches (for the zero of a given function, for example) on Fibonacci numbers rather than binary segmentation. As a second consequence, the Fibonacci system leads to a winning strategy for a game of Fibonacci Nim. [Rules of the game: 1) Take at least one chip, but not more than twice the number of chips just removed by your opponent. 2) On the first move, don't take all the chips. 3) He who takes the last chips wins.] In our numerical example, if you have before you a pile of 83 chips, take away 2 chips (the least term in the decomposition of 83) to start your winning strategy.

Interestingly, the ratios of successive Fibonacci numbers are given by a very simple *continued fraction* involving only the integer 1. To wit: 1 is the ratio of the first two Fibonacci numbers. Then $1+1/1 = 2$ is the ratio of the next two Fibonacci numbers. Next $1+1/(1+1/1) = 3/2$ gives the next two numbers. Repeating this process yields successively 5/3, 8/5, 13/8, etc. and converges on the Golden ratio. Numbers, like the Golden ratio, whose continued fraction expansion ends in infinitely many 1's have recently been christened "noble" numbers, because such numbers characterize nonlinear dynamic systems that are most resistant to chaos.

Continued fractions, treated in detail in Chap. 5, are a powerful mathematical tool playing a role of increasing importance in modern physics and other quantitative sciences. Indeed they are much more "natural" and inherently mathematical than, say, decimal fractions.

1.2 Fermat, Primes and Cyclotomy

No one doubts that 9 plus 16 equals 25, or, somewhat more elegantly,

$$3^2 + 4^2 = 5^2.$$

In fact, even the Pythagoreans knew that

$$x^2 + y^2 = z^2, \tag{1.1}$$

has solutions for many values of x, y, z other than the (trivial) $x = y = z = 0$ and the above solution $x = 3$, $y = 4$, $z = 5$; another example is $x = 5$, $y = 12$, $z = 13$. Such triplets of numbers, of which there are infinitely many, are called *Pythagorean numbers*. The problem of finding *all* such numbers is treated in Diophantus's *Arithmetica*, a Latin translation of which was published in 1621.

About fifteen years later, *Pierre de Fermat* (1601–1665), son of French leather merchants, accomplished Toulouse jurist, eminent part-time mathematician and inventor of the *principle of least action* (Fermat's principle) scribbled in his personal copy of Diophantus's book the following now famous note: "I have discovered a truly remarkable proof of this theorem which this margin is too small to contain." By "this theorem" Fermat meant that

$$x^n + y^n = z^n \tag{1.2}$$

has no integral solution with $xyz \neq 0$ for $n > 2$. "This theorem" is now known as Fermat's last theorem or simply FLT. Unfortunately, Fermat's note notwithstanding FLT was *not* a theorem, but only a conjecture, because it was only recently proved – in spite of massive mathematical efforts spanning several centuries.

As a result of these efforts, which have spawned much new mathematics, it is now known that up to very large values of the exponent n ($n \approx 10^8$ if xyz and n have no common divisor) there are no solutions of (1.2). But 350 years after Fermat's casual remark, a general proof was still not in sight. In fact, the question of FLT's validity might have remained unresolved forever, because if it is false, it is very unlikely that a counterexample will be found – even with the aid of today's fastest computers. (Of course, a single counterexample would suffice to inter FLT for good.)

However, some progress on FLT *had* been made. In 1922, Lewis Joel Mordell conjectured that (1.2) for $n > 3$ has at most a finite number of solutions. In spite of many attempts by the world's leading mathematicians, it was not until 1983 that a young German mathematician, *Gerd Faltings*, proved the recalcitrant conjecture true.

Fermat was luckier with an earlier edict, now called *Fermat's theorem* (without the "last"). It states that a prime number p divides $a^{p-1} - 1$ without remainder if a and p have no common divisor. Symbolically Fermat's theorem looks as follows

$$p|a^{p-1} - 1, \quad \text{if } (p,a) = 1. \tag{1.3}$$

For example, $p = 7$ divides $2^6 - 1 = 63$.

Although the ancient Chinese thought otherwise, a converse of Fermat's theorem is not true: if $p > 2$ divides $2^{p-1} - 1$, p is not *necessarily* a prime. Since large prime numbers are in great demand these days (for cryptographic purposes, for example) such a converse would have been very useful as a sieve for primes. This is regrettable, but primes are simply not that easy to catch. Nevertheless, Fermat's theorem forms the basis of sophisticated primality tests that are capable of stilling our hunger for primes a hundred or more digits long.

Another great invention of Fermat's are the so-called *Fermat numbers*, defined by

$$F_m := 2^{2^m} + 1, \tag{1.4}$$

which Fermat mistakenly thought to be prime for all $m \geq 0$. Indeed $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$, $F_4 = 65537$ are all prime numbers (called *Fermat primes*) but, as Leonhard Euler showed, $F_5 = 4\,294\,967\,297$ is divisible by 641. This is easily verified with a good pocket calculator – which neither Fermat nor Euler possessed. But Euler knew that $5 \cdot 2^7 + 1 = 641$ was a *potential* factor of F_5 and he could do the necessary calculations in his head (which stood him in good stead when he went blind).

To this day, no Fermat prime larger than F_4 has been found. At present the smallest Fermat number whose primality status remains unknown is F_{20}, a 315653-digit number. On the other hand, a numerical monster such as F_{3310} is known to be composite. In fact, modern factoring algorithms have shown it to be divisible by $5 \cdot 2^{3313} + 1$. This is no small achievement, because F_{3310} is unimaginably large, having more than 10^{990} digits! (This is not to be confused with the extremely large number 10^{990} which, however, is miniscule by comparison with F_{3310}.)

On March 30, 1796, the Fermat primes, until then largely a numerical curiosity (the mathematical sleeper of the century, so to speak), were raised from dormancy and took on a new beauty embracing number theory and geometry. On that day the young Carl Friedrich Gauss showed that a circle could be divided into n equal parts by straightedge and compass alone, if n was a Fermat prime. Progress on this problem had eluded mathematicians since classical Greece, in spite of many efforts over the intervening 2000 years.

Since the time of Gauss the art of *cyclotomy*, the learned name for dividing the circle, has been studied and perfected in numerous other number fields with spectacular successes. Among the many ingenious applications are methods for designing highly effective error-correcting codes and fast computational algorithms, for making precision measurements of exceedingly weak general relativistic effects, and even designing unique phase gratings for scattering sound in concert halls to achieve better acoustic quality. Sometimes the range of applications is even more dazzling than the mathematical invention itself!

1.3 Euler, Totients and Cryptography

Fibonacci, the lone mathematical luminary of the European Dark Ages, and Fermat, man of the Renaissance, were followed in the 18th century by an almost baroque flowering of mathematical genius, of whom Leonhard Euler was the most colourful and prolific centrepiece. Euler's collected works, currently numbering 70 published volumes, are still incomplete. Besides being foremost in geometry and analysis, and a pioneer in the most varied applications of mathematics to physics and astronomy, he added much to the theory of numbers.

Among his myriad contributions while working at the Academies of St. Petersburg and Berlin, few epitomize his genius as concisely as the formula

$$\mathrm{e}^{\mathrm{i}\pi} = -1, \tag{1.5}$$

which he not only discovered as a mathematical identity but which also comprises three of Euler's enduring notational inventions: e for the base of the natural logarithms, i for the square root of -1, and π for the ratio of a circle's circumference to its diameter.

One of Euler's fundamental contributions to number theory was a generalization of Fermat's theorem (1.3) which he extended to composite (nonprime) exponents m:

$$m | a^{\phi(m)} - 1, \tag{1.6}$$

if m and a have no common divisor. Here $\phi(m)$ is *Euler's totient function*, which counts how many integers from 1 to $m-1$ have no common divisor with m. If m is prime, then $\phi(m)$ equals $m-1$, and we are back to Fermat's theorem (1.3). But for composite m we obtain totally new and often unexpected results. For example, for $m = 10$, $\phi(m) = 4$ and we have

$$10 | a^4 - 1.$$

In other words, the fourth power of *any* number not containing the factors 2 or 5 has 1 as the last digit, to wit: $3^4 = 81$, $7^4 = 2401$, $9^4 = 6561$, $13^4 = 28\,561$, and so on.

Euler's totient function pervades much of number theory, including such statistical concepts as the probability that two numbers selected at random will have no common divisor. This probability tends to the average value of $\phi(m)/m$, which in turn approaches $6/\pi^2 \approx 0.61$ even for relatively small ranges of m ($m = 1$ to 10, say). The same asymptotic average of $\phi(m)/m$, incidentally, also describes the probability that a random integer will not contain a square as a factor. In fact, even among the 10 smallest natural numbers (1 to 10) the proportion of "squarefree" numbers (1, 2, 3, 5, 6, 7, i. e., 6 out of 10) is already close to the asymptotic value $6/\pi^2$.

Recently, Euler's function has taken on a new and most practical significance: it is the basis of a promising *public-key encryption* system. In public-key encryption the encrypting keys of all participants are, as the name implies, publicly known. Thus, anyone wishing to send a secret message to Mr. Sorge can encrypt it with Sorge's published key. And Sorge can, of course, decrypt the message intended for him, because he has the proper "inverse" key. But no one else can do this, since the published key contains no useful cue to the inverse key. If the key consists of an "exponent" k and a "modulus" m (typically the product of two 100-digit numbers), the inverse exponent k' is in principle computable from the expression

$$k^{\phi(\phi(m))-1}. \tag{1.7}$$

But in practice this is impossible because, to calculate $\phi(m)$, the factors of m must be known. And those factors are *not* published – nor are they easily derived from m. Even with today's fastest factoring algorithms it could take aeons to find the factors of properly chosen 200-digit numbers. And random guessing is no help either, because $\phi(m)$ fluctuates between $m-1$ (for prime m) and values small compared to m. In other words, the chances of a random guess are about 1 in 10^{100}.

On the other hand, for one who *knows* the factors of m, Euler's function quickly gives the answer with the aid of (1.7). Did Euler ever dream of such an application when he invented the totient function?

1.4 Gauss, Congruences and Diffraction

Carl Friedrich Gauss, the son of a poor labourer in the German ducal city of Brunswick, was one of the greatest mathematicians in history. Yet he almost missed becoming a mathematician because he excelled in languages, being fluent at an early age in classical Greek and Latin, in addition to French and his native German. But his desire to study philology came to an abrupt end the day he discovered how to divide a circle into 17 equal parts. Gauss was not yet 19 years old when he made this epochal discovery. From his extensive reading, Gauss knew that no progess in "cyclotomy" had been made since the time of the ancient Greeks, though no effort had been spared in the course of two millennia. Gauss therefore reasoned that he must be a pretty good mathematician, and resolved to devote his life to mathematics – without, however, forsaking his love of literature and languages. In fact, when he was in his sixties he added another foreign language, Russian, to his repertoire enabling him to read Pushkin in the original as well as some innovative mathematical papers that had never been translated into a Western language.

One of Gauss's many mathematical inventions was the congruence notation. Instead of writing $m|a$ for "m divides a", Gauss introduced a kind of equation, a *congruence:*

$$a \equiv 0 \pmod{m}, \tag{1.8}$$

which reads "a is congruent 0 modulo m". This notation can be immediately generalized to

$$a \equiv b \pmod{m}, \tag{1.9}$$

which means that $a-b$ is divisible by m or, equivalently, a divided by m leaves the remainder b.

Just as calculus was much advanced by the seductive notation of Leibniz (the differential quotient and the integral sign), so the congruence notation of Gauss had an import beyond expectation. One reason was that the congruence notation brings out *remainders*, and remainders are often more important than the rest of a problem. This is true not only in wave interference, as has already been pointed out, but also in many other situations in science, technology and daily life. For example, the trick of many cryptographic systems is to work with remainders: the secret information is locked up in remainders, so to speak, to be released only with a secret decrypting key.

One of the mathematical entities that occupied Gauss a long time are certain complex sums like

$$S(n) = \sum_{k=0}^{n-1} \mathrm{e}^{\mathrm{i}2\pi k^2/n} \tag{1.10}$$

now called Gauss sums. It is relatively easy to show that $|S(n)|^2 = n$ for odd n. But it took even the great Gauss, to his own dismay, years to prove his guess of the value of $S(n)$ itself.

The sequence defined by the individual terms in the sum (1.10), which repeats with a period of n if n is prime, has a noteworthy correlation property: the periodic correlation is strictly zero for nonzero shifts. As a consequence, the Fourier spectrum of such a periodic sequence has components whose magnitude are all alike. Such spectra are called flat or *white*. And of course the terms of the sequence itself also have equal magnitudes, namely 1. These facts have led to numerous interesting applications in physics and communications.

If the above sequence represents wave amplitudes at uniform intervals along a linear spatial coordinate then, at some distance, the wave will break up into many wavelets with (nearly) equal energies! Thus, if the terms of a Gauss sum are put to proper use in what may be called quadratic-residue phase gratings, then coherent light, radar beams or sound waves can be very effectively scattered. For light this leads to the ultimate in frosted glass, with interesting applications in coherent optical processing. For radar and sonar detection this means less visibility. In noise control, such phase gratings permit one to disperse offending sound where it cannot be absorbed (along highways, for example). And in concert hall acoustics, the possibility of diffusing sound without weakening it means better acoustic quality for more musical enjoyment.

1.5 Galois, Fields and Codes

We all hear in school about fractions or *rational numbers*, easy objects of adding, subtracting, multiplying and dividing (except by zero). Like whole numbers, the rational numbers form a countable, albeit infinite, set. Such a corpus of numbers is called a *field*.

Interestingly, there are also *finite fields*. Consider the seven numbers 0, 1, 2, 3, 4, 5, 6. If we disregard differences that are multiples of 7 (i. e., if we identify 3, 10, -4, and so forth), we can add, subtract and multiply to our heart's content and stay within these seven numbers. For example $5+4=2$, $3-6=4$, $2\cdot 4=1$. These operations obey the customary commutative, associative and distributive laws. To illustrate, $4\cdot(5+4)=4\cdot 5+4\cdot 4=6+2=1$; and we obtain the same result if we first add then multiply: $4\cdot(2)=1$.

But what happens with division; for example, what is 1 divided by 2, $1/2$, in this finite number field? It must be one of the numbers 0, 1, 2, 3, 4, 5, 6, but which? In other words, what number times 2 equals 1? By trial and error (or consulting one of the above examples) we find $1/2=4$ and, of course, $1/4=2$, and consequently $3/2=3\cdot 4=5$.

How can we find reciprocals in a systematic way; what is the general law? The answer is

$$1/n=n^5$$

for $n\neq 0$. (As usual, 0 has no inverse.) Thus, $1/3=3^5=5$, and therefore $1/5=3$. Likewise $1/6=6^5=6$, i. e., 6 is its own inverse! But this is not unusual because the rational numbers, too, have a self-inverse other than 1, namely -1. (And incidentally, in our finite number system -1 is equivalent to 6.)

While it may seem odd that the sum of 2 odd numbers is sometimes odd ($3+5=1$, for example) and sometimes even $(1+5=6)$ or that $1/4$ of $1/4$ is $1/2$, all of these results are nevertheless perfectly consistent, and it is *consistency* that counts in mathematics.

But what good is consistency? Are finite fields just consistent nonsense? The answer is an emphatic *no*. For one thing, finite fields have held a secure place in algebra ever since the young Frenchman *Evariste Galois* used them to show the conditions under which algebraic equations have solutions in radicals (roots).

Such “radical” solutions had been known for quadratic equations since antiquity and for cubic and quartic equations since the Renaissance. But curiously enough, no one had been able to find a formula in radicals for the general equation of the fifth degree or higher. Then, in 1830, the 19-year-old Galois, generalizing a result obtained by the very able Norwegian mathematician Niels Abel, gave the general and definitive answer to a problem that had baffled many generations of mathematicians. Unfortunately, Galois was killed in a duel (probably a political provocation in amorous disguise) before he turned 21, but his name will be enshrined forever in the annals of algebra and immortalized by the addition to the mathematical vocabulary of such important terms as *Galois group* and *Galois field*, the latter a synonym for finite field.

Of particular importance in modern digital applications are the finite fields having p^m members, where p is some prime number, $p = 2$ being the case most relevant to computers, communication and encryption. The simplest such field with $m > 1$ has $2^2 = 4$ elements (members). But these elements cannot be the *scalar* numbers (0, 1, 2, 3, say) if we want to interpret them as usual. Because then we would have, for example, $2 \cdot 2 = 4$; and 4 is congruent 0 modulo 4. As a consequence the equation $2x = 0$ has *two* solutions in such a “nonfield”: $x = 0$ and $x = 2$. On the other hand, the equation $2x = 1$ has *no* solution; in other words, 2 has no inverse. This is no way to construct a field!

How then are we to construct a field with four elements? The problem is reminiscent of the situation in quantum mechanics when physicists tried to convert the Schrödinger wave equation into a relativistically invariant form. In order to preserve both quantum essence and Lorentz invariance, P. A. M. Dirac discovered that he had to replace the Schrödinger scalar ψ-function by a vector function with four components, representing two possible spins and two electric charges (spin “up” or “down” and electron or positron).

Similarly, with our four-element field, the solution is to choose *vectors* or “*n*-tuples” as field elements. For a four-element field we may try *binary* vectors with two components, each component being 0 or 1:

$$\begin{array}{cc} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{array}$$

Addition is componentwise, *without* carries. The 00-element is the zero-element of the field, and the 10 is the one-element. The two remaining elements must be multiplicative inverses of each other:

$$0\,1 \cdot 1\,1 = 1\,0.$$

Any nonzero element raised to the third power must equal the one-element 10. This fixes the multiplication table of the field:

$$\begin{array}{c|ccc} \cdot & 1\,0 & 0\,1 & 1\,1 \\ \hline 1\,0 & 1\,0 & 0\,1 & 1\,1 \\ 0\,1 & 0\,1 & 1\,1 & 1\,0 \\ 1\,1 & 1\,1 & 1\,0 & 0\,1\,. \end{array} \tag{1.11}$$

Here the entry $01 \cdot 01 = 11$ is the only nonredundant one; all others follow or are fixed *a priori*.

Instead of proving the consistency of the resulting algebra, we give a simple example of the distributive law:

$$0\,1 \cdot (1\,1 + 1\,0) = 1\,0 + 0\,1 = 1\,1,$$

which yields the same result as adding *before* multiplying:

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

The choice of symbols for our finite field is of course arbitrary. To save ink, we might write simply 0 instead of 00 and 1 instead of 10. However, for 01 we need a new symbol, say ω. We must then identify 11 with ω^2 because of our multiplication table (1.11). With this notation, one of the entries of (1.11) tells us that $\omega \cdot \omega^2 = \omega^3 = 1$. Furthermore, $1 + \omega + \omega^2 \,\widehat{=}\, 10 + 01 + 11 = 00 \,\widehat{=}\, 0$.

This is all we need to know about our new symbol ω to calculate in our four-element field GF(4). There is a convenient interpretation of ω in terms of *complex* numbers. Because $\omega^3 = 1$, ω must be a third root of 1, and, because $\omega \neq 1$, it must be either $\exp(2\pi\mathrm{i}/3)$ or $\exp(-2\pi\mathrm{i}/3)$. This is of course no accident: the three nonzero elements of GF(4) form a cyclic group of order three, of which the three roots of 1 are another representation. (However, this analogy does not carry over to addition. For example, our $-\omega$ equals ω. To wit: $-\omega = 0 - \omega \overset{\wedge}{=} 00 - 01 = 01 \overset{\wedge}{=} \omega$.)

The ω-representation of GF(4) is used in Sect. 31.3 to construct a pretty necklace with four different kinds of pearls.

Of course, a four-element Galois field is not very exciting, but the eight-element field GF(2^3) has interesting applications. But first we must give a rule for its construction. We start with the zero- and one-elements.

$$\begin{gathered} 0\,0\,0 \\ 1\,0\,0 \end{gathered}$$

and generate the following two elements by rightward shifts, calling them g^1 and g^2, respectively:

$$\begin{aligned} g^1 &= 0\,1\,0 \\ g^2 &= 0\,0\,1. \end{aligned}$$

Then we continue shifting, and for each 1 "disappearing" on the right, we add (modulo 2) a 1 to each of the two left places. Thus, our field elements are

$$\begin{array}{lcl} g^{-\infty} & = & 0\,0\,0 \\ g^{0} & = & 1\,0\,0 \\ g^{1} & = & 0\,1\,0 \\ g^{2} & = & 0\,0\,1 \\ g^{3} & = & 1\,1\,0 \\ g^{4} & = & 0\,1\,1 \\ g^{5} & = & 1\,1\,1 \\ g^{6} & = & 1\,0\,1 \\ & & \overline{} \\ g^{7} & = & 1\,0\,0 \;=\; g^{0} \\ g^{8} & = & 0\,1\,0 \;=\; g^{1}, \quad \text{etc.} \end{array} \tag{1.12}$$

Another (equivalent) method of generating $\mathrm{GF}^*(2^3)$, i. e., the nonzero elements of $\mathrm{GF}(2^3)$, is to focus on the *columns* of (1.12). The columns are generated by the recursion

$$a_{n+3} = a_{n+1} + a_n a \tag{1.13}$$

with the *initial conditions* 100, 010, 001, respectively. All columns have period $2^3 - 1 = 7$ and are identical within a cyclic shift. Each of the seven nonzero triplets occurs exactly once per period.

An important application of $\mathrm{GF}(2^m)$ occurs in binary error-correcting codes. One of the simplest codes of block lengths m is obtained by using m information bits as the initial condition computing the remaining $2^m - m - 1$ “check bits” with an appropriate recursion. For $m = 3$, for example, the three information bits 111 (say) are supplemented by the four check bits, which according to (1.13) are 0010. Thus, the total seven-bit code word for the information bits 111 is

$$1\;1\;1\;0\;0\;10. \tag{1.14}$$

Except for the 0000000-word, all code words are cyclic shifts of (1.14). Using our finite field algebra, it is easy to show that the “Hamming distance” (defined as the modulo 2 component-wise sum) between any two code words so generated is 4. Thus, we have constructed a code capable of *correcting* a single error and *detecting* a second error. The occurrence and location of a single error is found by the occurrence and position of a single 1 in the sum of the erroneous code word with all possible code words.

Many other powerful codes can be derived from Galois-field algebra. For example, interchanging the roles of information and check bits in the above example results in the famous Hamming error-correcting code of block length 7.

Other interesting applications arise when we choose $+1$ and -1 as elements of $\mathrm{GF}(2^m)$. The above code words then become periodic pulse trains whose power spectrum is “white”; i. e., all finite-frequency components of the Fourier transform have equal magnitudes. Thus, Galois-field theory has given us a method of generating “signals” with *constant power in both Fourier domains* (time and frequency or space and direction). Such signals are of great value in making precision measurements at very low energies – such as in the observation of general relativistic effects in radar echoes from planets and interplanetary spacecraft.

Interestingly, by interchanging the time and frequency domains, the same Galois-field algebra also yields continuous periodic waveforms with near-minimal amplitude ranges and flat power spectra. Such waveforms have found useful application in the synthesis by computer of more human-sounding speech.

And by applying these "Galois" sequences of ±1's is to spatial coordinates, diffraction gratings of high scattering power can be realized. Such gratings are ideal diffusors for all kinds of waves: laser light, radar radiations, noise emissions and musical sound waves. Here is another example of how a relatively abstract mathematical theory – initiated in the early 19th century by a precocious teenager – can yield useful results in a variety of applications for which the theory was *not* constructed. As is so often the case in human affairs, it is the *unexpected* that matters most.

Chapter 2
The Natural Numbers

‘O θεὸζ ἀριθμητίζει
– Carl Friedrich Gauss

Here we encounter such basic concepts as *integers, composites*, and *primes*, and we learn the very fundamental fact that the composites can be represented in a *unique* way as a product of primes.

The *least common multiple* and the *greatest common divisor* of two or more integers may be familiar from high school, but they are ideas that pervade all of number theory. Here we demonstrate some of their basic properties and point to some natural phenomena in the real world of gears, planetary motion, and musical pitch.

If integers can be prime, pairs of integers can be “mutually prime” or *coprime* if they have no common factors, in other words, if their greatest common divisor is 1. Coprimality is another important property of two (or more) integers.

One of the very early tools of number theory is *Euclid’s algorithm*; it allows us to find, in a systematic manner, the greatest common divisor of two integers without solving the often difficult problem of factoring the two integers. As we shall later see, Euclid’s algorithm generalizes to polynomials and allows us to solve important integer equations, the so-called Diophantine equations.

2.1 The Fundamental Theorem

We will speak here of the “whole numbers” or *integers* $\ldots -3, -2, -1, 0, 1, 2, 3, \ldots$, denoted by the letter $\mathbb{Z}$, and more often of the so-called “natural” numbers or positive integers: 1, 2, 3, 4, 5 and so forth. Some of these are divisible by others without leaving a remainder. Thus, $6 = 2 \cdot 3$, i. e., 6 is divisible by 2 and by 3 without a remainder. Such numbers are called *composites*.

Other numbers have no divisors other than 1 and themselves, such as 2, 3, 5, 7, 11, 13, 17, etc. These numbers are called *prime* numbers or simply *primes*. All primes are odd, except 2 – the “oddest” prime (a designation alluding to the special role which 2 plays among the primes). The number 1 is considered neither prime nor

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_2, © Springer-Verlag Berlin Heidelberg 2009

composite. Otherwise some theorems would require very awkward formulations – such as the following.

The *fundamental theorem of arithmetic* states that each natural number n can be uniquely factored into primes:

$$n = p_1^{e_1} p_2^{e_2} \dots p_k^{e_k} \dots p_r^{e_r} = \prod_i p_i^{e_i}. \tag{2.1}$$

Here the order of the factors is considered irrelevant.

Equation (2.1) can be read in two ways:

1) p_i is the ith prime – in which case the exponent e_i has to be zero if p_i is not a factor of n.
2) Only those primes that are factors of n appear in (2.1). We will use either reading of (2.1) and state which if it makes a difference.

There is no corresponding theorem for the *additive* decomposition of natural numbers into primes. This is one of the reasons why additive number theory, for example partitions (Chap. 22), is such a difficult subject. In this book we will be mostly concerned with *multiplicative* number theory, which has many more applications.

2.2 The Least Common Multiple

Two integers n and m have a least common multiple (LCM) $[n,m]$. The LCM is needed to combine two fractions with denominators n and m into a single fraction. In fact, that is where the everyday expression "to find the least common denominator" (of divergent views, for example) comes from. For example, for $n = 6$ and $m = 9$, $[6,9] = 18$.

Example: $\frac{1}{6} + \frac{2}{9} = \frac{3}{18} + \frac{4}{18} = \frac{7}{18}$.

It is easily seen that with n as in (2.1) and

$$m = \prod_i p_i^{f_i}, \tag{2.2}$$

$$[n,m] = \prod_i p_i^{\max(e_i,f_i)}, \tag{2.3}$$

because in the LCM each prime factor p_i must occur at *least* as often as it does in either n or m. Thus, for $n = 6 = 2^1 \cdot 3^1$ and $m = 9 = 3^2$, $[6,9] = 2^1 \cdot 3^2 = 18$.

There are numerous applications of the LCM. Consider two gears with n and m teeth meshing, and suppose we mark with a white dot one of the n teeth on the first gear and one of the m spaces between teeth of the second gear. When the gears turn, how often will the two white dots meet? Perhaps never! But if they meet once, they will meet again for the first time after $[n,m]$ teeth have passed the point of contact, i. e., after the first gear has undergone $[n,m]/n$ (an integer!) revolutions and the second gear an integer $[n,m]/m$ revolutions.

2.3 Planetary "Gears"

Our "gears", of course, could be any of a plethora of other objects that can be modelled as meshing gears even if no teeth are visible. Thus, the revolutions of the planet Mercury around itself and the Sun are locked by gravitational forces as if geared: during two revolutions around the Sun, Mercury revolves three times around itself. (As a consequence, one day on Mercury lasts two Mercury years. Strange gears – and even stranger seasons!) Similarly, the Earth's moon revolves exactly once around itself while completing one orbit around the Earth; that is why it always shows us the same side. On the moon, Earth day (or night) lasts forever.

The "teeth" that keep the moon locked to the Earth are, as in the case of Mercury and the sun, gravitational forces. But these "gravitational teeth" are relatively weak and would not "engage" if unfavourable initial conditions were not damped out by friction such as that provided by the ocean tides on Earth. (Eventually, the Earth day may lock in with the Earth year, which will play havoc with night and day as we know it.)

And not long ago, it was discovered that even the distant planets Pluto and Neptune are coupled to each other strongly enough to be locked into an integer "resonance" (in the astronomer's lingo).

Another question answered by the LCM, although no teeth are in evidence, has to do with the coincidence of dates and weekdays. Because the number of days per year (365) is not divisible by the numer of days per week (7), coincidences of dates and weekdays do not recur from one year to the next. Furthermore, because every fourth year is a leap year, coincidences are not equally spaced in years. However, even without knowing when leap years occur, we can always guarantee that a coincidence will recur after 28 years, 28 being the LCM of 4 and 7. (In the year 2100 the leap day will be dropped, temporarily violating the 28-year cycle.)

Equation (2.3) easily generalizes to more than two integers: the max function in (2.3) then contains as many entries as there are integers whose LCM we want to determine.

As we indicated above when introducing the meshing gear picture, the two white markers may never meet. More learnedly, we would say that a certain linear Diophantine equation (see Chap. 7) has no solution. This can happen only if n and m have a greatest common divisor greater than 1. This brings us to our next topic.

2.4 The Greatest Common Divisor

Another important relation between integers is their greatest common divisor (GCD). For two integers n and m given by (2.1) and (2.2), the GCD is

$$(n,m) = \prod_i p_i^{\min(e_i,f_i)}, \tag{2.4}$$

because for the GCD to divide both n and m it cannot have the factor p_i more often than it is contained in either n or m, whichever is *less*.

Example: $n = 10 = 2^1 \cdot 5^1$ and $m = 25 = 5^2$. Thus $(10,25) = 5$.

Two numbers n and m that have no common factors are called relatively prime, mutually prime or *coprime*. In this case the GCD equals 1.

Example: $(6,35) = (2{\cdot}3, 5{\cdot}7) = 1$.

For any two numbers n and m, the product of the GCD and the LCM equals the product of n and m:

$$(n,m)[n,m] = nm,$$

because whenever the formula (2.4) for the GCD picks the exponent e_i for p_i, the formula (2.3) for the LCM picks the exponent f_i, and vice versa.

Thus,

$$(n,m)[n,m] = \prod_i p_i^{e_i+f_i} = nm. \tag{2.5}$$

Example:

$$(4,10) = (2^2, \underline{2}{\cdot}5) = 2; \quad [4,10] = [\underline{2}^2, 2{\cdot}\underline{5}] = 20; \quad 2\cdot 20 = 4\cdot 10. \ \textit{Check}!$$

The generalization of (2.5) to three integers is

$$(n,m,k)[nm,mk,kn] = nmk, \tag{2.6}$$

which is easily verified. Assume that a given prime p occurring in the prime factorization of the product nmk occurs e_n times in n, e_m times in m and e_k times in k and that, without loss of generality,

$$e_n \le e_m \le e_k.$$

Then the exponent of p in (n,m,k) is e_n, and in $[nm,mk,kn]$ it is $e_m + e_k$. Thus the left side of (2.6) has the prime p with the exponent $e_n + e_m + e_k$, as does the right side of (2.6). The same is true for all primes occurring in nmk. The correctness of (2.6) then follows from the fundamental theorem of arithmetic.

The *dual* of (2.6) is

$$[n,m,k](nm,mk,kn) = nmk, \tag{2.7}$$

which is proved by the same reasoning. Generalizations of (2.6) and (2.7) to more than three factors should be obvious.

For more than two integers, some particularly interesting relations between GCD and LCM exist. For example, two such relations are the "distributive law"

$$(k[m,n]) = [(k,m),(k,n)], \tag{2.8}$$

and its *dual*

$$[k,(m,n)] = ([k,m],[k,n]), \tag{2.9}$$

both of which are a direct consequence of the properties of the min and max functions in (2.4) and (2.3).

There is even a very pretty *self-dual* relation:

$$([k,m],[k,n],[m,n]) = [(k,m),(k,n),(m,n)], \tag{2.10}$$

i. e., in the expression appearing on either side of (2.10), the operations LCM and GCD can be completely interchanged without affecting their validity!

However, from a practical point of view there *is* a difference: The right-hand side of (2.10), i. e., doing GCDs before the LCMs, is usually easier to figure out. Thus, (2.10) can be exploited to computational advantage.

It is interesting to note that relations such as (2.6–2.10) occur in many other mathematical fields, such as mathematical logic or set theory, where our LCM corresponds to the set-theoretic *union* $\cup$ and the GCD corresponds to *intersection* $\cap$.

But what, in number theory, corresponds to the set-theoretic relation

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B},$$

where the bar stands for complement?

For additional relations see [2.1].

The GCD appears in the solution to many, seemingly unrelated, problems. For example, take n jugs with capacities of $L_1, L_2, \ldots, L_n$ liters. What amounts k of water (or wine) can be dispensed by these $n/1$ jugs?

Answer: k must be a multiple of the GCD $[L_1, L_2, \ldots, L_n]$. (After T. J. Pfaff and M. M. Tran. *The Pi Mu Epsilon Journal* 12:1 (2004), 37–38.)

2.5 Human Pitch Perception

An interesting and most surprising application of the GCD occurs in human perception of pitch: the brain, upon being presented with a set of harmonically related frequencies, will perceive the GCD of these frequencies as the pitch. Thus, the subjective pitch of the two-tone chord (320 Hz and 560 Hz) is $(320, 560) = 80$ Hz, and *not* the difference frequency (240 Hz).

Upon a frequency shift of $+5$ Hz applied to both frequencies, the GCD drops to 5 Hz; and for an irrational frequency shift, the GCD even drops to 0 Hz. But that is not what the ear perceives as the pitch. Rather it tries to find a *close match* in the range of pitches above 50 Hz. For the frequencies 325 Hz and 565 Hz such a match is given by 81 Hz, which is the GCD of 324 Hz and 567 Hz – close to the two given frequencies.

Note that the concept that pitch is given by the difference frequency or "beat" frequency has been beaten: if both frequencies are shifted by the same amount, their difference remains unchanged. Yet psychoacoustic experiments clearly show that the perceived pitch is increased, from 80 Hz to about 81 Hz in our example, just as the amplified GCD model predicts [2.2].

What this tells us is that the human brain switches on something like a GCD-spectral matching computer program when listening to tone complexes. Fascinating? Indeed. Unbelievable? Well, the brain has been caught doing much trickier things than that.

2.6 Octaves, Temperament, Kilos and Decibels

The Pythagoreans discovered that subdividing the string of a musical instrument into the ratio of small integers resulted in pleasing musical intervals. Thus, dividing the string into 2 equal parts gives a frequency ratio (compared with the full-length string) of 2 : 1 – the musical octave. Shortening the string by one third gives rise to the frequency ratio 3 : 2 – the musical fifth. And dividing the string into 4 equal parts results in the frequency ratio 4 : 3 – the musical fourth.

The Pythagorean musical scale was constructed from these simple ratios. How do they fit together? How many fifths make an integral number of octaves? Or, what is x in

$$\left[\frac{3}{2}\right]^x = 2^y,$$

or, equivalently,

$$3^x = 2^z,$$

where $z = y + x$? The fundamental theorem (Sect. 2.1) tells us that there are *no* integer solutions. But there are *approximate* solutions, even in small integers. Thus,

$$3^5 = 243 \approx 256 = 2^8. \tag{2.11}$$

Consequently, 5 musical fifths equal *about* 3 octaves. To make the octave come out correctly, we would have to tamper with the ratio $3 : 2 = 1.5$, increasing it by about 1% to 1.515..., to achieve a well-tempered temperament.

The fact that $(3/2)^5$, with a little tampering, equals 2^3 also has its effect on the musical fourth: from

$$\left[\frac{3}{2}\right]^5 \approx 2^3$$

follows directly

$$\left[\frac{4}{3}\right]^5 \approx 2^2;$$

in other words, 5 fourths make about 2 octaves. The tampering required on the fourth to make it fit 2 octaves exactly is, as in the case of the fifth, only one part in a hundred.

We shall leave the musical details to J. S. Bach and his well-tempered clavier and ask ourselves the more general question of how we can find approximate integer solutions to equations like $a^x = b^y$ in a more systematic way. The answer: by expanding logarithms into continued fractions, as will be explained in Sect. 5.1. There we learn that for $a = 3$ and $b = 2$, for example, the next best approximation (after $3^5 \approx 2^8$) is $3^{12} \approx 2^{19}$, requiring an adjustment of the musical fifth by only one part in a thousand so that 12 "tampered" fifths will make 7 octaves, thereby avoiding the *Pythagorean comma*. This is of great interest to musicians because it allows the construction of a complete key from ascending fifths (the famous Circle of Fifths).

A much closer numerical coincidence, with important consequences in music, computer memory, photography and power measurements, is the approximation

$$5^3 = 125 \approx 128 = 2^7. \tag{2.12}$$

Musically, this means that 3 major thirds (frequency ratio $= 5:4$) equal about *one* octave:

$$\left[\frac{5}{4}\right]^3 \approx 2,$$

which requires an adjustment of less than 8 parts in a thousand in the major third so that 3 of them match the octave exactly.

Another consequence of (2.12) is that

$$2^{10} = 1024 \approx 10^3.$$

According to international standards, the factor 10^3 is denoted by the prefix *kilo*, as in kilometre. But computer memories are not measured in kilometres or weighed in kilograms; rather they are *addressed*, and the proper form of address is *binary*. As a consequence, memory sizes are usually powers of 2, and in computerese a 256-kilobit memory chip can actually store 262144 bits of information because, to hard- and software types, kilo means 1024 – not 1000.

The near coincidence of 5^3 and 2^7 also shows up among camera exposure times, where $1/125$ of a second is 7 lens-aperture "stops" away from 1 second. But 7 stops correspond to a light energy factor of $2^7 = 128$.

Still another application in which $5^3 \approx 2^7$ is exploited is the field of power or intensity measurement. The preferred logarithmic measure of intensity is the *decibel*,[1] 10 decibels being equal to an intensity ratio of 10 : 1. Thus, twice as much power (of a loudspeaker output, for example) means an extra 3 decibels – almost exactly. (A better figure would be 3.01 decibels, but who can hear a hundredth of a decibel?)

J. R. Pierce, lately of Stanford University, has recently proposed a new musical scale based on dividing the frequency ratio 3 : 1 (instead of the 2 : 1 octave) into 13 (instead of 12) equal parts. This scale matches such simple integer ratios as 5 : 3 *and* 7 : 5 (and 9 : 7) with an uncanny accuracy, resulting from the number-theoretic fluke that certain 13th powers of *both* 5 *and* 7 are very close to integer powers of 3. To wit: $5^{13} = 3.0077^{19}$, and $7^{13} = 3.0037^{23}$. Since the integers appearing in the exponents (13, 19, 23) are also coprime (in fact, all three are prime), it is easy to construct complete musical scales exclusively from the small-integer ratios 5 : 3 and 7 : 5. The basic chords of the new scale, 3 : 5 : 7 and 5 : 7 : 9, are superbly approximated by the equal tempered scale $3^{k/13}$ and were found by M. V. Mathews, A. Reeves, and L. Roberts to provide a strong harmonic foundation for music written in the new scale.

2.7 Coprimes

Two integers are said to be *coprime* if their GCD equals 1. Thus, 5 and 9 are coprime: $(5,9) = 1$, while 6 and 9 are *not* coprime: $(6,9) = 3 \neq 1$.

The *probability* that two "randomly selected" integers will be coprime is $6/\pi^2$ (see Sect. 4.4). This is also the probability that a randomly selected integer is "squarefree" (not divisible by a square).

Of three or more integers it is often said that they are *pairwise coprime* if all possible pairs are coprime. Thus, 2, 5 and 9 are pairwise coprime: $(2,5) = (2,9) = (5,9) = 1$. However, 2, 5 and 8 are *not* pairwise coprime because $(2,8) = 2$, although the three numbers seen as a *triplet* have no common factor. The probability that three randomly selected integers will be pairwise coprime is $0.28\ldots$ (see Sect. 4.4).

2.8 Euclid's Algorithm

If the GCD is so important, how does one go about finding it? Answer: by Euclid's algorithm, which is best illustrated by an example. To find the GCD of 35 and 21, first divide the larger number by the smaller:

$$\frac{35}{21} = 1 + \frac{14}{21},$$

[1] Curiously, one never hears about the full unit, the bel, perhaps because a difference of 10 bel is the difference between the sound of a babbling brook and an earsplitting screech.

and repeat the process on the remainder:

$$\frac{21}{14} = 1 + \frac{7}{14},$$

until the remainder is 0:

$$\frac{14}{7} = 2 + 0,$$

which is guaranteed to happen sooner or later. The GCD is the last divisor, 7 in our case. Thus, $(35,21) = 7$, which is the correct answer.

The philosophy behind Euclid's algorithm is the following. It is easy to show that $(a,b) = (a - kb, b)$, where k is an integer. If $a > b > 0$ and if one picks k as large as possible without making $a - kb$ negative, then $a - kb < b$. Thus, we have reduced the problem of computing the GCD of a and b to that of two smaller numbers, namely $a - kb$ and b. Now b is the larger number of the pair, and it can be reduced by subtracting a proper multiple of $a - kb$. Continuing this simple process generates smaller and smaller number pairs all having the same GCD. Finally we must arrive at two numbers that are multiples of each other and the smaller of the two numbers is the GCD. (If a and b are coprime that "smaller number" is of course 1.) This is how and why the Euclidean algorithm works: it chops large numbers down to manageable size.

2.9 The Decimal System Decimated

One of the greatest arithmetical inventions was that of the 0 as a *place holder*. Thus, in the decimal system, 903 means 9 times 100 plus 3 times 1. Their sensitivity to place is the foundation of the decimal, the binary and similar number systems based on place.

But choosing a base, 10 say, there are other ways to represent the integers which does not depend on the customary place principle while still using only the 10 digits 0 through 9. For example, 100 integers from 0 to 99 can be thought of as the *direct sum* of two sets of integers each of size 10: $\{0,1,2,\ldots,9\}$ and $\{0,10,20,\ldots,90\}$. (Direct sum of two sets means every member of one set is added to every member of the other set.) Representing the integers from 0 through 99 by the direct sum of the above two sets is of course analoguous to the standard decimal representation. And, if the base is a prime, this representation is unique. But if the base, like 10 can be factored, there are other *nonstandard* representations by the direct sum of two sets, each of size 10. For example, the sets $\{0,20,22,40,42,60,62,80,82\}$ and $\{0,1,4,5,8,9,12,13,16,17\}$ are two such sets.

Are there other such sets? And how many? What is the "sister" set of $\{0,1,2,3,4,$ $25,26,27,28,29\}$?

Chapter 3
Primes

As we go to larger and larger integers, primes become rarer and rarer. Is there a largest prime after which all whole numbers are composite? This sounds counterintuitive and, in fact, it isn't true, as Euclid demonstrated a long time ago. Actually, he did it without demonstrating any primes – he just showed that assuming a finite number of primes leads to a neat contradiction.

Primes are found by sieves, not by formulas, the classical sieve having been designed by Eratosthenes in classic Greece. (Formulas that pretend to give only primes are really shams.) Primality testing has advanced to a stage where the primality or compositeness of 100-digit numbers can now be ascertained by computer in less than a minute, without actually giving any of the factors [3.1]. Factoring, on which the security of certain kinds of cryptographic systems depends (Chaps. 10–15), is still very difficult at this writing.

The largest primes known are of a special form called *Mersenne* primes because they don't hide their compositeness too well and, indeed, some were discovered by high-school students. The largest Mersenne prime known (in mid-1983) has 25,962 digits! Mersenne primes lead to even *perfect numbers* and to prime "repunits", meaning repeated units, i. e., numbers consisting exclusively of 1's in any given base system. (The Mersenne primes are repunits in the binary number system.)

Of special interest are the *Fermat primes* of which, in spite of Fermat's expectations, only 5 are known, the largest one being 65537. Each Fermat prime allows the construction of a regular polygon by using only straightedge and compass – Gauss's great discovery made just before he turned nineteen.

3.1 How Many Primes are There?

Again we turn to Euclid, who proved that there are infinitely many primes by giving one of the most succinct indirect proofs of all of mathematics:

Suppose that the number of primes is finite. Then there is a largest prime p_r. Multiply all primes and add 1:

$$N = p_1 p_2 \ldots p_r + 1.$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_3, © Springer-Verlag Berlin Heidelberg 2009

Now N is larger than p_r and thus cannot be a prime because p_r was assumed to be the largest prime. Thus N must have a prime divisor. But it cannot be any of the *known* primes because by construction of N, all known primes divide $N-1$ and therefore leave the remainder 1 when dividing N. In other words, none of the known primes divides N. Thus, there is a prime larger than p_r – a contradiction! We must therefore conclude that there is no largest prime, i. e., that there are infinitely many.

In actual fact, the above construction *often* (but not always) does give a prime. For example:

$$\begin{aligned}
&2+1=3\\
&2\cdot 3+1=7\\
&2\cdot 3\cdot 5+1=31\\
&2\cdot 3\cdot 5\cdot 7+1=211\\
&2\cdot 3\cdot 5\cdot 7\cdot 11+1=2311,
\end{aligned}$$

all of which are prime. *But*

$$2\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13+1=30031=59\cdot 509.$$

Suppose we set $P_1=2$ and call P_{n+1} the largest prime factor of $P_1P_2\ldots P_n+1$. Is the sequence P_n monotonically increasing? No! Both P_9 and P_{10} have 16 decimal digits but P_{10} equals only about $0.3\,P_9$.

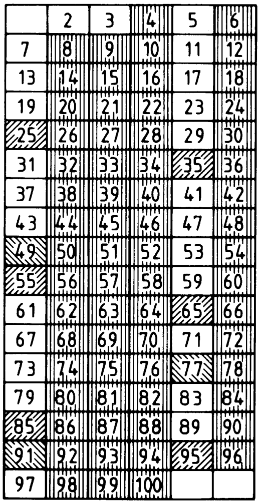
3.2 The Sieve of Eratosthenes

Like gold nuggets, primes are mostly found by sieves – the first one having been designed in ancient Greece by Eratosthenes of Kyrene around 200 B.C. Eratosthenes's sieve idea is charmingly simple.

To find the primes below 100, say, write down the integers from 1 to 100 in order. Then, after 2, cross out every second one $(4,6,8\ldots)$, in other words all the even numbers, because they are divisible by 2 and therefore not prime (except 2 itself). Then, after 3, cross out every third number that is still standing $(9,15,21\ldots)$ because these numbers are divisible by 3 and therefore also not prime. Repeat the crossing out process for every fifth number after 5 and every seventh number after 7. The remaining numbers (except 1, which is not considered a prime) are the 25 primes below 100:

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47,
53, 59, 61, 67, 71, 73, 79, 83, 89, 97.

Roughly speaking, to find the primes below a given integer N, we only have to use sieving primes smaller than $\sqrt{N}$. (This rule would tell us that by sieving with 2, 3, 5 and 7 we will find all primes below $11^2=121$, while actually we have found 4 more primes.)



	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36
37	38	39	40	41	42
43	44	45	46	47	48
49	50	51	52	53	54
55	56	57	58	59	60
61	62	63	64	65	66
67	68	69	70	71	72
73	74	75	76	77	78
79	80	81	82	83	84
85	86	87	88	89	90
91	92	93	94	95	96
97	98	99	100		

Fig. 3.1 The sieve of Eratosthenes (modulo 6)

In applying Eratosthenes's sieve method there is an additional trick that simplifies matters considerably: we write the integers in six columns starting with 1 (Fig. 3.1). Then only the first and the fifth columns (no pun) contain primes because all numbers in the second, fourth, and sixth columns are divisible by 2, and those in the third column are divisible by 3.

To eliminate the numbers divisible by 5 and 7 as well, a few 45° diagonals have to be drawn, as shown in Fig. 3.1.

Things become a little more complicated if we include the next two primes, 11 and 13, in our sieve because the numbers divisible by 11 and 13 follow "knight's-move" patterns as known from chess. But then we have already eliminated all composite numbers below $17 \cdot 19 = 323$. In other words, we have caught the 66 primes up to 317 in our 6-prime sieve.

Sieving may connote a child playing in a sandbox or a gold digger looking for a prime metal, but sieving in number theory is a very respectable occupation and sometimes the only method of finding an elusive prime or unmasking a composite as such. Of course, the sieving algorithms employed today are becoming increasingly sophisticated. We will hear more about the search for primes, especially the urgently needed very large ones, in subsequent chapters.

3.3 A Chinese Theorem in Error

The ancient Chinese had a test for primality. The test said that n is prime iff[1] n divides $2^n - 2$:

$$n|(2^n - 2). \tag{3.1}$$

[1] Here and in the rest of this book, "iff" means *if and only if*. Further, x "divides" y means that x divides y *without leaving a remainder*. As a formula this is written with a vertical bar: $x|y$.

As we shall prove later, (3.1) is indeed true if n is an odd prime (by Fermat's theorem). Of course, for $n = 2$, (3.1) is trivially true.

Example: for $n = 5$, $2^n - 2 = 30$, which is indeed divisible by 5.

Conversely, for odd $n < 341$, if n is *not* prime n does *not* divide $2^n - 2$.

Example: for $n = 15$, $2^n - 2 = 32{,}766$, which is *not* divisible by 15.

Fortunately (for their self-esteem!), the ancient Chinese never tried $n = 341$, which is composite: $341 = 11 \cdot 31$ and yet 341 divides $2^{341} - 2$ without remainder. This might be a bit hard to check by abacus, but the test is within reach of many a programmable pocket calculator. Of course, the calculation does not give the quotient $(2^{341} - 2)/341$, a number 101 digits long, but rather the remainder, which is 0, thereby falsely asserting that $341 = 11 \cdot 31$ is prime.

The rules for calculating high powers efficiently will be given later, together with a "fast" calculator program.

3.4 A Formula for Primes

In 1947, *Mills* [3.2] showed that there is a constant A, such that $\lfloor A^{3^n} \rfloor$ is[2] prime for every n. Here we have a formula that, although it does not generate each and every prime, could be used to generate arbitrarily large primes – for which the sieve methods are less suited.

For anyone who has an appreciation of what a precious thing a prime is, this seems impossible. And indeed, there is trickery at play here, albeit cleverly hidden trickery: determination of the constant A presupposes prior knowledge of the primes! This trickery is explained in the excellent little book by *Nagell* [3.3], but it *is* a bit tricky, and we will illustrate the point by another, not quite so surreptitious, trick. Consider the real constant

$$B = 0.2030005000000070000000000000110\ldots. \tag{3.2}$$

Upon multiplying by 10 and taking the integer part, one obtains 2 – the first prime. Dropping the integer part, multiplying by 100, and taking the integer part then gives 3 – the second prime. In general, after the nth prime has been extracted from B, multiplying by 10^{2^n} and taking the integer part yields the $(n+1)$th prime. Thus, we have specified a (recursive) algorithm[3] for specifying not only primes but *all* the primes, and in proper order at that!

[2] The so-called "Gauss bracket" or "floor function" $\lfloor x \rfloor$ is defined as the largest integer not exceeding x. Thus, $\lfloor 4.9 \rfloor = 4$; but $\lfloor 5.0 \rfloor = 5$. The Gauss bracket (for $x \geq 0$) corresponds to the instruction "take integer part", often designated by INT in computer programs.

[3] Can the reader specify a nonrecursive algorithm, i. e., one that gives the nth prime directly, without calculating all prior ones?

Of course, here the trick is patently transparent: we have simply "seeded" the primes, one after another, into the constant B, interspersing enough 0's so that they do not "run into" each other. In other words, the constant B does not yield any primes that are not already known. (How many 0's between seeded primes are required to guarantee that adjacent primes do not overlap in B? If 0's are considered expensive because they make B very long, that question is not easy to answer and, in fact, requires a little "higher" number theory.)

Apart from (3.2) and Mill's formula, there have been many other prescriptions for generating primes or even "all" primes. Most of these recipes are just complicated sieves in various disguises, one of the few really elegant ones being Conway's Prime Producing Machine (cf. R. K. Guy, Math. Mag. *56*, 26–33 (1983)). Other attempts, making use of Wilson's theorem (Sect. 8.2), are hilarious at best and distinguished by total impracticality. All this (non)sense is reviewed by U. Dudley in a delightful article ("Formulas for Primes", Math. Mag. *56*, 17–22 (1983)).

One of the most astounding algorithms for producing primes, in fact seemingly all of them and in perfect order, is the *Perrin sequence*. This sequence is defined by the recursion $A(n+1) = A(n-1) + A(n-2)$, with the initial condition $A(0) = 3$, $A(1) = 0$, and $A(2) = 2$ (cf. I. Stewart, Sci. Am. *247*, 6, 102–103 (June 1996)). Lucas has proved that whenever n is prime n divides $A(n)$. But the converse also seemed to be true: if n divides $A(n)$, n is prime. But alas, a first counterexample was found for $n = 271441 = 521^2$.

3.5 Mersenne Primes

A *Mersenne number* is a number $M_p = 2^p - 1$, where p is *prime*. If M_p itself is prime, then it is called a *Mersenne prime*. Note that numbers of the form $2^n - 1$, where n is composite, can never be prime because, for $n = pq$,

$$2^n - 1 = (2^p - 1)(2^{p(q-1)} + 2^{p(q-2)} + \ldots + 1), \tag{3.3}$$

However, not all primes p yield Mersenne primes, the first exception being $p = 11$, because $2^{11} - 1 = 2047 = 23 \cdot 89$. Still, there is a fairly simple primality test for numbers of the form $2^p - 1$, the so-called Lucas Test: $2^p - 1$ is prime *iff* (note the double f, meaning *if and only if*) M_p divides S_p $(p > 2)$, where S_n is defined by the recursion

$$S_n = S_{n-1}^2 - 2, \tag{3.4}$$

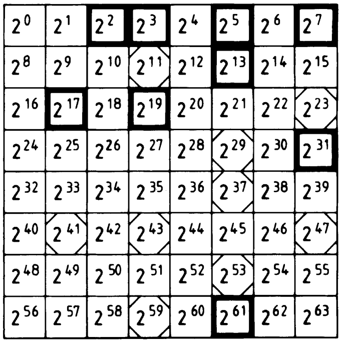
starting with $S_2 = 4$.

Thus, for example, S_{11} is given by the 10th number in the sequence

$$4,\ 14,\ 194,\ 37634, \ldots,$$

which is not divisible by $M_{11} = 2047$. Thus, M_{11} is composite.

Fig. 3.2 Mersenne primes on a checkerboard



2^0	2^1	2^2	2^3	2^4	2^5	2^6	2^7
2^8	2^9	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}
2^{16}	2^{17}	2^{18}	2^{19}	2^{20}	2^{21}	2^{22}	2^{23}
2^{24}	2^{25}	2^{26}	2^{27}	2^{28}	2^{29}	2^{30}	2^{31}
2^{32}	2^{33}	2^{34}	2^{35}	2^{36}	2^{37}	2^{38}	2^{39}
2^{40}	2^{41}	2^{42}	2^{43}	2^{44}	2^{45}	2^{46}	2^{47}
2^{48}	2^{49}	2^{50}	2^{51}	2^{52}	2^{53}	2^{54}	2^{55}
2^{56}	2^{57}	2^{58}	2^{59}	2^{60}	2^{61}	2^{62}	2^{63}

While this test does not reveal any factors, there is another test that *can* give a factor for M_p with $p = 4k+3$: for $q = 2p+1$ prime, $q | M_p$ *iff* $p \equiv 3 \bmod 4$.

Example: $p = 11 = 4 \cdot 2 + 3$; $M_{11} = 2047$ is not divisible by $(p-1)/2 = 5$ and is therefore divisible by $2p+1 = 23$. Check: $2047 = 23 \cdot 89$. Check! Similarly, 47 is discovered as a factor of $M_{23} = 8388607$, etc.

Figure 3.2 shows the first 9 Mersenne primes arranged on a checkerboard. On January 17, 1968, the largest known prime was the Mersenne prime $2^{11213} - 1$, an event that was celebrated with a postmark (Fig. 3.3) from Urbana, Illinois (at no profit to the U.S. Post Office, considering the zero value of the stamp).

In the meantime, much larger Mersenne primes have been found. The record on November 18, 1978, stood at $2^{21701} - 1$, a prime with 6533 decimal digits found by two California high-school students, Laura Nickel and Curt Noll, using 440 hours on a large computer. The next Mersenne prime is $2^{23209} - 1$. By early 1982 the largest known prime was $2^{4497} - 1$, having 13395 digits [3.4].

More recently, another Mersenne prime was discovered by D. Slowinski, the 28th known specimen: $2^{86243} - 1$. Assuming that there are no other Mersenne primes between it and $M(27) = 2^{44497} - 1$, then $2^{86243} - 1$ is, in fact, $M(28)$.

Are there more Mersenne primes beyond $2^{86243} - 1$? The answer is almost certainly *yes*. Curiously, we can even say roughly how large the next Mersenne prime is: 10^{38000} – give or take a dozen thousand orders of magnitude. How can we make such a seemingly outrageous statement?

Fermat and Euler proved that all factors of M_p must be of the form $2kp+1$ and simultaneously of the form $8m \pm 1$. Thus, *potential* factors of M_p are spaced on average $4p$ apart. Assuming that, subject to this constraint, the number of factors of a Mersenne number is governed by a Poisson process, Gillies [3.5] conjectured that of all the primes in the "octave" interval $(x, 2x)$, on average approximately 2

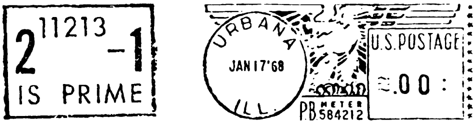


Fig. 3.3 The largest known Mersenne prime on Januar 17, 1968

give Mersenne primes. More precisely, the density of primes near p giving rise to Mersenne primes $M_p = 2^p - 1$ would be asymptotic to

$$\frac{2}{p \ln 2}.$$

In a recent paper S. S. Wagstaff, Jr. ("Divisors of Mersenne primes", Math. Comp. *40*, 385–397 (1983)), following an argument by H. W. Lenstra, Jr., suggested that the expected number of primes p in an octave interval is $\mathrm{e}^{\gamma} = 1.78\ldots$. Thus, the correct asymptotic density would be

$$\frac{\mathrm{e}^{\gamma}}{p \ln 2}.$$

Comparing this with the general prime density for primes near p, $1/\ln p$, we see that of

$$\frac{p}{\mathrm{e}^{\gamma} \log_2 p}$$

primes, one prime on average leads to a Mersenne prime. For $p \approx 100000$, this means that roughly every 3000th prime gives a Mersenne prime. (The appearance of the factor e^{γ} is a consequence of Merten's theorem, see Sect. 12.1, and its relevance to prime sieving.)

The distribution of primes p that generate Mersenne primes is expressed even more simply if we consider the density of $\log_2 p$: it is constant and should equal e^{γ}. Since $\log_2 p$ very nearly equals $\log_2(\log_2 M_p)$, these statements are equivalent to the following: if $\log_2(\log_2 M(n))$ is plotted as a function of n, we can approximate the empirical "data" by a straight line with a slope of about $1/\mathrm{e}^{\gamma} = 0.56$. In fact, for the 27 smallest Mersenne primes $(2 \le p \le 44497)$ the average slope is 0.57, remarkably close to $1/\mathrm{e}^{\gamma}$. The correlation coefficient between $\log_2(\log_2 M(n))$ and n in this range exceeds 0.95.

Figure 3.4 shows $\log_2(\log_2 M(n))$ versus n for 28 known Mersenne primes (assuming $n = 28$ for $p = 86243$). The great regularity is nothing short of astounding.

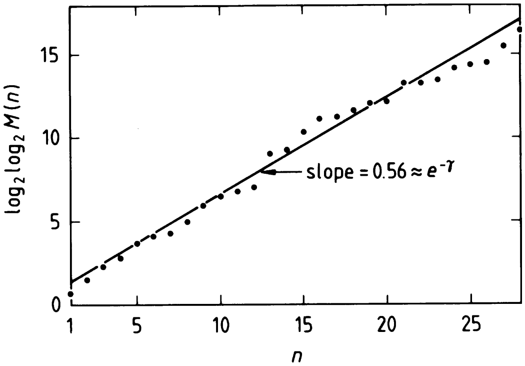


Fig. 3.4 $\log_2(\log_2 M(n))$ versus n

As a Poisson process, the cumulative distribution P of the *intervals* between successive values of $\log_2(\log_2 M(n))$ should go according to the exponential law:

$$P = 1 - \mathrm{e}^{-d/\overline{d}}$$

with $\overline{d} = 1/\mathrm{e}^{\gamma}$. This function is plotted in Fig. 3.5 together with the empirical evidence (in interval ranges of 0.2). Here again, the correspondence with the theoretical result expected from a Poisson process is very good. Specifically, the mean interval (0.57) is close to the standard deviation (0.51) and (beyond the information contained in Fig. 3.5) successive intervals are nearly uncorrelated (correlation coefficient $= -0.17$).

Using 0.56 as the average increment of $\log_2(\log_2 M(n))$ with n, we expect the next Mersenne prime above $2^{86243} - 1$ in the "neighbourhood" of $2^{130000} \approx 10^{38000}$. Of course, to find a prime in this vast "haystack" that gives a Mersenne prime is no small order.

More accurately, we can say that the *probability* of finding the next Mersenne prime either below or above 10^{34000} is about 0.5, and the probability that it exceeds 10^{65000} is less than 10 %. But where is it, *exactly* – not with an uncertainty of thousands of orders of magnitude? Even the fastest number crunchers available today, using the most efficient search algorithms, will have a hey (hay?) day.

Unfortunately, the Mersenne primes are very thinly seeded. Thus, if one is looking for a 50-digit prime among the Mersenne primes, one is out of luck: $2^{127} - 1$ has 39 digits and the *next* Mersenne prime, $2^{521} - 1$, has 157 digits – an awesome gap!

Does a Mersenne prime M_p always yield another Mersenne prime by the formula

$$2^{M_p} - 1\ ?$$

This had been widely conjectured, but a counterexample is now known: the prime $p = 13$ gives a Mersenne prime $M_{13} = 8191$, but $2^{8191} - 1$ is composite. Too bad! The nearest Mersenne primes, $2^{4423} - 1$ and $2^{9689} - 1$, have 1332 and 2917 decimal digits, respectively – leaving another great void.

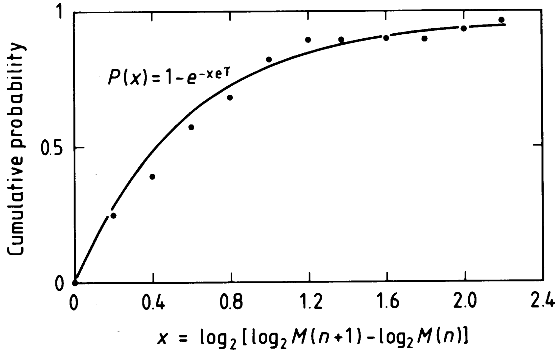


Fig. 3.5 Interval statistics

3.6 Repunits

Expressed as binary numbers, all numbers of the form $2^n - 1$ consist exclusively of 1's, namely exactly n 1's. For example $M_4 = 2^4 - 1 = 15$, or 1111 in binary. Similarly, for any a, $(a^n - 1)/(a-1)$ expressed in base a consists of precisely n 1's and no other digits. Such numbers are called "repunits", and apart from those to the base 2, those to base 10 have been the most widely studied. For a repunit to be prime, n must be prime, but as with the Mersenne numbers, that is not sufficient.

Examples: $(10^5 - 1)/9 = 11111 = 41 \cdot 271$ and $(10^7 - 1)/9 = 1111111 = 239 \cdot 4649$.

More examples for which $(10^9 - 1)/9$ is composite can be created by finding primes $q > 3$ such that $\operatorname{ord}_q 10 = p$ (see Chap. 14 for the definition of ord). In fact, q then divides $(10^p - 1)/9$.

Which prime exponents p give repunit primes and whether there are infinitely many are two of the many unsolved problems of number theory. With some luck $(10^{317} - 1)/9$ was proved prime, but not until most 50 years after $(10^{23} - 1)/9$ was found to be prime. Primality testing of large numbers is not easy and *factoring* is even more difficult! In fact, the factoring of $(10^{71} - 1)/9$ (into two primes with 30 and 41 digits, respectively) had to wait for 1984 machines and algorithms.

3.7 Perfect Numbers

Each Mersenne prime has a companion *perfect number* $P = M_p 2^{p-1}$. A perfect number is a number for which the sum of all divisors (not including P itself) equals P. Thus, for example, $M_2 = 2^2 - 1 = 3$ leads to the perfect number $P = 6$; and indeed, the sum of the divisors of 6: $1 + 2 + 3$ equals 6 itself.

The next Mersenne prime, M_3, equals $2^3 - 1 = 7$ and the corresponding perfect number is 28. Check: $1 + 2 + 4 + 7 + 14 = 28$. Check!

It is easy to see why this is so. Since M_p is by definition prime, the only divisors of the perfect number $P = M_p 2^{p-1}$ are

$$1, 2, \ldots, 2^{p-1}, M_p, 2M_p, \ldots, 2^{p-1} M_p,$$

and their sum equals

$$\sum = 1 + 2 + \ldots + 2^{p-1} + M_p(1 + 2 + \ldots + 2^{p-1}), \quad \text{or} \tag{3.5}$$

$$\sum = (1 + M_p)(2^p - 1) = 2^p M_p = 2P. \tag{3.6}$$

(The factor 2 appears here because we included, in the sum, P itself as a divisor of P.) The remarkable fact that *all even* perfect numbers are of the form $M_p 2^{p-1}$, where M_p is a Mersenne prime, was first *proved* by Leonhard Euler (1707–1783), the great Swiss mathematician (and not just that!) from Basel who worked for most of his life in St. Petersburg, the then new capital of all the Russias.

Because 37 Mersenne primes are known, at the time of this writing, there are exactly 37 known perfect numbers, all of them even, and the largest one having 909526 decimal places. No odd perfect numbers are known and, tantalizingly, it is not known whether there are any such. As of 1971, no odd perfect number had been found among all the numbers up to 10^{36}. P. Hagis, Jr., showed recently that an odd perfect number not divisible by 3 has at least eleven prime factors (Math. Comp. *40*, 399–404 (1983)).

Apart from perfect numbers, there are pseudoperfect numbers (Sect. 5.9) and *amicable* numbers. Amicable numbers come in pairs. The sum of divisors of one amicable number equals its mate and vice versa. The smallest amicable pair is 220 and 284. Another pair is 17296 and 18416. In a sense, perfect numbers are "self-amicable".

In a further generalization, certain number sequences are called *sociable*. In these, each number equals the sum of the divisors of the preceding number, and the first number equals the sum of divisors of the last number. One such five-member sociable group is 14288, 15472, 14536, 14264, 12496. There is a sociable chain of length 28 whose smallest member is 14316.

A frequently used concept in number theory is the sum of some function f taken over all divisors of a number n, *including* n itself. This is usually shown by the following notation:

$$\sum_{d|n} f(d).$$

Using this notation, our statement about perfect numbers P reads

$$\sum_{d|p} d = 2P, \tag{3.7}$$

the factor 2 appearing here because by definition P itself is a divisor of itself and therefore is included in the sum.

There was a time when the present author was much impressed by the fact that the sum of *reciprocal* divisors of P is *always* 2:

$$\sum_{d|p} \frac{1}{d} = 2! \tag{3.8}$$

(Here, for once, the exclamation mark does not do any harm because 2! – read "two factorial" – still equals 2.) However, (3.8) is a "trivial" consequence of (3.7) because, in a sum over all divisors d of a given number n, the divisor d may be replaced by n/d. This reverses the *order* of the terms in such a sum, but does not affect its value:

$$\sum_{d|n} f(d) = \sum_{d|n} f(n/d). \tag{3.9}$$

Indeed, for $n = 6$ and $f(d) = d$,

$$1+2+3+6 = 6+3+2+1. \tag{3.10}$$

Applying (3.8) to (3.6), we have

$$2P = \sum_{d|P} d = \sum_{d|P} \frac{P}{d}, \tag{3.11}$$

confirming (3.8). Check: $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 2$. Check!

It is remarkable that the sum of the reciprocal divisors of a perfect number always equals 2, no matter how large it is. This implies that perfect numbers cannot have too many small divisors, as we already know.

3.8 Fermat Primes

Besides the Mersenne primes $2^p - 1$, which lead to perfect numbers, and of which only 37 are presently known, there is another kind of prime family with even fewer known members: the Fermat primes. Only 5 such primes are currently known.

$$F_n = 2^{2^n} + 1 \quad \text{for} \quad n = 0,\ 1,\ 2,\ 3,\ 4. \tag{3.12}$$

They are $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$.

Incidentally, for $2^m + 1$ to be prime, m must be a power of 2. In fact, for any $a^m + 1$ to be prime, a must be even and $m = 2^n$.

All numbers of the form $2^{2^n} + 1$, whether prime or composite, are called Fermat numbers. They obey the simple (and obvious) recursion

$$F_{n+1} = (F_n - 1)^2 + 1, \quad \text{or} \tag{3.13}$$

$$F_{n+1} - 2 = F_n(F_n - 2), \tag{3.14}$$

which leads to the interesting product

$$F_{n+1} - 2 = F_0 F_1 \ldots F_{n-1}. \tag{3.15}$$

In other words, $F_n - 2$ is divisible by all lower Fermat numbers:

$$F_{n-k} \mid (F_n - 2), \qquad 1 < k \leq n. \tag{3.16}$$

With (3.16) it is easy to prove that all Fermat numbers are coprime to each other, and the reader may wish to show this.

Fermat thought that all Fermat numbers F_n were prime, but Euler showed that $F_5 = 4294967297 = 641 \cdot 6700417$, which can easily be confirmed with a good pocket calculator.

The fact that F_6 and F_7 are also composite is a little harder to show because F_6 has 20 decimal digits and F_7 has 39. Nevertheless, *complete* factorizations of F_6, F_7 and, since 1981, F_8 are now known. Further, it is now known that F_{11}, F_{12} and F_{13}

are composite, and *some* of their factors are known. Another Fermat number known to be composite is F_{73}, which has more than 10^{21} digits! For special primality tests for Fermat numbers, see *Hardy* and *Wright* [3.6].

At present the smallest, and so far most enduring mystery is presented by F_{20}: its primality status remains unknown. However, the latest progress in primality testing, reported by Walter Sullivan in The New York Times in February 1982 and in [3.1], might yet reveal other Fermat primes, although the next candidate, F_{20}, has 315653 digits. One helpful clue which has been utilized in the past is that, if F_n is composite, then it is divisible by $k \cdot 2^{n+2} + 1$ for some k. In fact, Euler knew this, and that is how he discovered the factor $641 = 5 \cdot 2^7 + 1$ in F_5.

In this manner the compositeness of some *very* large Fermat numbers has been established. For example, $5 \cdot 2^{3313} + 1$ is a factor of F_{3310}. By the way, F_{3310} has more than 10^{990} *digits* – not to be confused with the comparatively miniscule number 10^{990}.

3.9 Gauss and the Impossible Heptagon

In March 1796, the Fermat primes suddenly took on a new and overwhelming significance. A precocious teenager from the German ducal town of Brunswick had just discovered that the circle could be divided into 17 equal parts by purely "geometric means", i. e., by straightedge and compass – something that had eluded professional mathematicians and amateurs alike for over two millennia. In fact, nobody had even suspected that such a feat could be possible. After the cases of 2, 3, 4, 5 and 6 had been solved by the ancient Greeks, "everybody" had been working on the "next" case: the regular heptagon (7-gon). But the Brunswick youth proved that that was impossible and that the only regular n-gons that could be constructed were those derivable from the Fermat primes.

The young person, of course, was none other than Carl Friedrich Gauss [3.7], who was himself so impressed by his feat of unlocking a door that had been closed for 2000 years that he decided to become a mathematician rather than a philologist, to which fate his excellence in the classical languages seemed to have "condemned" him.[4]

[4] His love of books and languages never left Gauss for the rest of his life. At the age of 62 he learned yet another foreign language – Russian – and began to read Pushkin in the original. Gauss selected the University of Göttingen rather than his "state" university, Helmstedt, for his studies, mainly because of Göttingen's open library policy. Even in his first semester at Göttingen, Gauss spent much time in the university library, which was well stocked and where he had access to the writings of Newton and Euler and many others of his predecessors. Much of what Gauss read there he had already derived himself, but he still felt that reading was essential – in stark contrast to other scientific geniuses, notably Einstein, who was convinced he could create most of the correct physics from within himself and who is supposed to have said, in jest, that if nature was not the way he felt it ought to be, he pitied the Creator for not seeing the point ("Da könnt' mir halt der liebe Gott leid tun, die Theorie stimmt doch." [3.8]).

We shall return to the important subject of dividing the circle, or *cyclotomy* by its learned name, in several other contexts later in this book. Let it only be said here that for the circle to be divisible into n parts, n must be the product of *different* Fermat primes or 1 and a nonnegative power of 2. Thus, regular polygons of $n = 2,3,4,5,6,8,10,12,15,16,17,20,24,30,\ldots$ sides can be geometrically constructed, while $n = 7,9,11,13,14,18,19,21,22,23,25,\ldots$ are *impossible* to construct in this manner. Here the "impossible" part of Gauss's assertion is as significant as his positive statement.

Chapter 4
The Prime Distribution

"All this is amusing,
though rather elementary[1] *..., Watson."*
– Sir Arthur Conan Doyle
(Sherlock Holmes)

How are the primes distributed among the integers? Here "distribution" is a misleading term because a given positive integer either is a prime or is not a prime – there is nothing chancy about primality. Yet superficially, the occurrence of primes appears to be rather haphazard, and, indeed, many properties can be derived by playing "dumb" and assuming nothing more than that "every other integer is divisible by 2, every third is divisible by 3", etc., and letting complete randomness reign beyond the most obvious. The result of this loose thinking suggests that the average interval between two successive primes near n is about $\ln n$. This is not easy to prove rigorously, especially if one forgoes such foreign tools as complex analysis. Yet fairly simple probabilistic arguments come very close to the truth. In fact, probabilistic thinking as introduced here can reveal a lot about primality and divisibility [4.1], and we shall make ample use of the probabilistic approach throughout this book to gain an intuitive understanding of numerous number-theoretic relationships. For a formal treatment of probability in number theory see [4.2].

4.1 A Probabilistic Argument

Two facts about the distribution of the primes among the integers can be noticed right away:

1) They become rarer and rarer the larger they get.
2) Apart from this regularity in their mean density, their distribution seems rather irregular.

In fact, their occurrence seems so unpredictable that perhaps probability theory can tell us something about them – at least that is what the author thought in his second (or third) semester at the Georg-August University in Göttingen. He had just

[1] In number theory *elementary* methods are often the most difficult, see [4.5].

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_4, © Springer-Verlag Berlin Heidelberg 2009

taken a course in *Wahrscheinlichkeitsrechnung* at “Courant’s” famous Mathematics Institute, and one afternoon in 1948, in the excruciatingly slow “express” train from Göttingen to his parents’ home in the Ruhr, he started putting some random ideas to paper. His train of thought ran roughly as follows.

The probability that a given “arbitrarily” selected integer is divisible by p_i is $1/p_i$. In fact, starting with 1, precisely every p_ith number is divisible by p_i (every third is divisible by 3, every fifth by 5 and so forth). Thus, the “probability” that a given selected number is *not* divisible by p_i is $1 - 1/p_i$.

Assuming that divisibility by different primes is an *independent*[2] property, the probability that x is not divisible by any prime below it is given by the product

$$W(x) \approx \left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\ldots \approx \prod_{p_1<x}\left(1-\frac{1}{p_i}\right). \tag{4.1}$$

If x is not divisible by any prime below it, it is, of course, not divisible by *any* smaller number, i. e., x is prime.

More strictly, we could limit the product to primes $p_1 < \sqrt{x}$ (see Sect. 3.2 on the sieve of Eratosthenes). In fact, in that 1948 train the author did limit the product to primes smaller than the square root of x. But since the end result is not much affected, we will not bother about this “refinement”.

If one feels uncomfortable with a product, it can be quickly converted into a sum by taking (naturally) logarithms:

$$\ln W(x) \approx \sum_{p_i<x} \ln\left(1-\frac{1}{p_i}\right). \tag{4.2}$$

If one does not like the natural logarithm on the right-hand side, expanding it and breaking off after the first term does not make much difference, especially for the larger primes:

$$\ln W(x) \approx -\sum_{p_i<x}\frac{1}{p_i}. \tag{4.3}$$

There is something about the sum that is still bothersome: it is not over consecutive integers, but only over the primes. How can one convert it into a sum over *all* integers below x? Again, one can use a probability argument: a given term $1/n$ in the sum occurs with probability $W(n)$. Thus, let us write (and this is the main trick here):

$$\ln W(x) \approx -\sum_{n=2}^{x}\frac{W(n)}{n}. \tag{4.4}$$

[2] Simultaneous independence for all primes is never exactly true, but there is near independence that suffices for our argument.

By now sums may have become boring, and one wishes the sum were an integral. Thus, we write with our now customary nonchalance:

$$\ln W(x) \approx -\int_2^x \frac{W(n)}{n}\,\mathrm{d}n. \tag{4.5}$$

The next thing that may strike one as offensive is the minus sign on the right-hand side. Introducing the *average distance* $A(x) = 1/W(x)$ between primes, we get a positive expression:

$$\ln A(x) \approx \int_2^x \frac{\mathrm{d}n}{nA(n)}. \tag{4.6}$$

Now, suddenly, the integral has served its purpose and can go; most people would rather solve differential equations than integral ones. Differentiating will of course be the appropriate integral vanishing trick:

$$\frac{A'(x)}{A(x)} \approx \frac{1}{xA(x)}, \quad \text{or} \tag{4.7}$$

$$A'(x) \approx \frac{1}{x}. \tag{4.8}$$

And the unexpected has happened: we have an answer (fortuitously correct)! The average distance between primes ought to be

$$A(x) \approx \ln x, \tag{4.9}$$

and the mean density becomes

$$W(x) \approx \frac{1}{\ln x}. \tag{4.10}$$

Example: $x = 20$, $\ln 20 \approx 3.00$, and, indeed, the average spacing of the 3 primes closest to 20, namely 17, 19 and 23, is exactly 3.

Around $x = 150$, the average spacing should be about 5, and in the neighbourhood of $x = 10^{50}$, every 115th number, on average, is a prime.

4.2 The Prime-Counting Function $\pi(x)$

If we accept the estimate (4.10) of the average prime density, the number of primes smaller than or equal to x, usually designated by $\pi(x)$, is approximated by the "integral logarithm":

$$\pi(x) \approx \int_2^x \frac{\mathrm{d}x'}{\ln x'} =: Li(x), \tag{4.11}$$

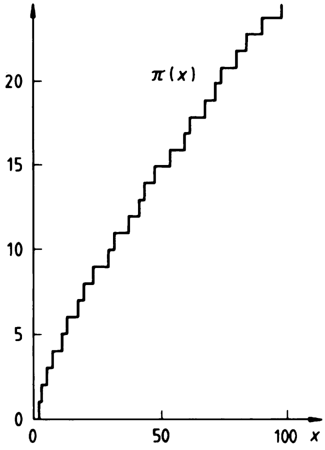


Fig. 4.1 The prime-counting function $\pi(x)$ for $2 \le x \le 100$

where the sign $=:$ indicates that the notation $Li(x)$ is *defined* by the integral on the left.

The prime-counting function $\pi(x)$ is plotted in Fig. 4.1 for $x \le 100$. Every time x equals a prime, $\pi(x)$ jumps up by 1. But apart from the "jumpiness" of $\pi(x)$, a smoother, slightly concave trend is also observable. This smoothness becomes more obvious when we plot $\pi(x)$ for x up to 55,000 as in Fig. 4.2. On this scale, the jumpiness has disappeared completely.

The inadequacy of Gauss's original estimate $\pi(x) \approx x/\ln x$ is illustrated by Fig. 4.3. By contrast, the integral logarithm, which we "derived" above (and which was also conjectured by Gauss) gives seemingly perfect agreement with $\pi(x)$ in the entire range plotted in Fig. 4.4.

However, even $Li(x)$, labelled "Gauss" in Fig. 4.5, shows noticeable deviations when we expand the ordinate by a factor 10^4 as was done in that figure (see [4.3]). In fact, for $x = 10^7$, the excess of $Li(x)$ over $\pi(x)$ is about 300 and remains positive for all $x < 10^9$. Nevertheless, $\pi(x) - Li(x)$ has infinitely many zeros, at least one of which occurs below $x = 10^{10^{10^{34}}}$; in fact, it may be near $x = 10^{370}$. (A number such as $10^{10^{10^{34}}}$, introduced by S. Skewes in 1933, was once considered a large number. But *much much much larger* numbers have now become important in connection with Gödel's famous "incompleteness" theorem [4.4].)

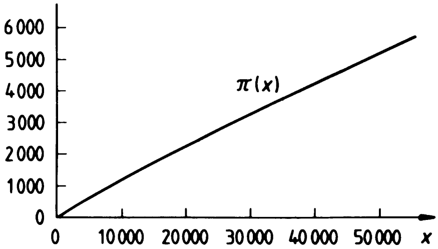


Fig. 4.2 Same as Fig. 4.1 but $x \le 55{,}000$

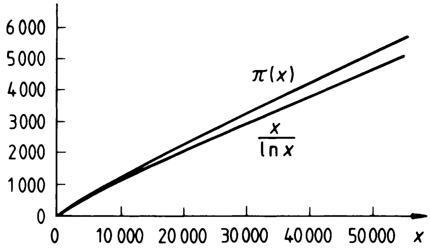


Fig. 4.3 $\pi(x)$ and $x/\ln x$

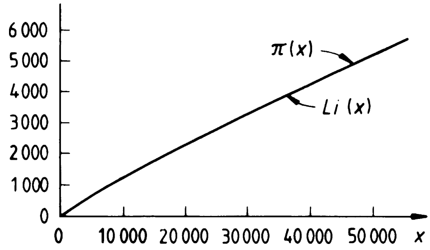


Fig. 4.4 $\pi(x)$ and the integral logarithm $Li(x)$

Legendre, independently of Gauss, gave the following formula in 1778:

$$\pi(x) \approx \frac{x}{\ln x - 1.08366}, \tag{4.12}$$

a closer approximation than (4.11) up to about $x = 4 \cdot 10^6$, as can be seen in Fig. 4.5. However, above $x = 5 \cdot 10^6$ Legendre's formula begins to go to pieces. (Expanding $Li(x)$ gives 1 as the constant in (4.12), but Legendre missed that.)

Either formula (4.11) or Legendre's (4.12) says that there are about $7.9 \cdot 10^{47}$ 50-digit primes – plenty to go around for the "trap-door" encryption schemes to be discussed later in Chap. 10.

In our "derivation" of $\pi(x)$, we considered primes up to x and pointed out that consideration of primes up to $\sqrt{x}$ would have sufficed. This idea was further pursued by Bernhard Riemann, who showed that

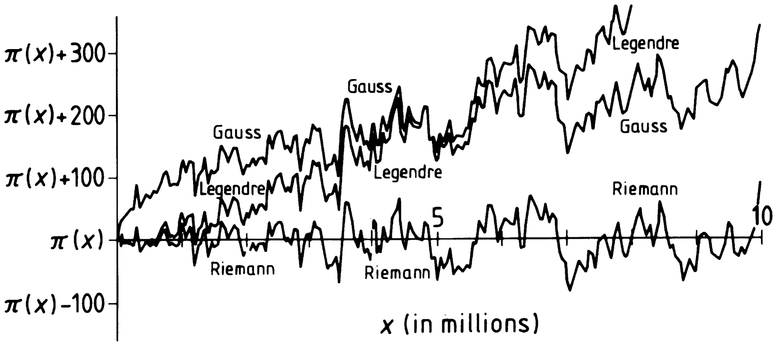


Fig. 4.5 Comparison of formulas by Gauss: $Li(x)$, Legendre: $x/(\ln x - 1.08366)$, and Riemann: $Li(x) - 1/2Li(x^{1/2}) - 1/3Li(x^{1/3}) - \ldots$ (Courtesy of D. Zagier)

$$\pi(x) \approx R(x) := Li(x) - \tfrac{1}{2}Li\left(\sqrt{x}\right) - \tfrac{1}{3}Li\left(\sqrt[3]{x}\right) - \ldots. \tag{4.13}$$

Figure 4.5 demonstrates how good an approximation $R(x)$ is; the curve labelled "Riemann" does not seem to have any deviant trend up to $x = 10^7$.

The closeness of $R(x)$ to $\pi(x)$ is further emphasized by Table 4.1, which shows that, even for $x = 10^9$, the error of $R(x)$ is only 79 (out of $5 \cdot 10^7$).

It is interesting to note that it was not until 1896, almost a hundred years after Gauss's and Legendre's conjectures, that Hadamard and de la Vallée Poussin proved the "Prime Number Theorem" in the form

$$\lim_{x\to\infty} \frac{\pi(x)\ln(x)}{x} = 1 \tag{4.14}$$

using "analytic" methods, i. e., mathematical tools from outside the domain of integers. The first "elementary" proof not using such tools did not come until 1948 and is due to *Erdös* [4.5] and Selberg. This illustrates the vast gap between obtaining an easy estimate, as we have done in the preceding pages, and a hard proof.

Perhaps one of the most surprising facts about $\pi(x)$ is that there "exists" an *exact* formula, given by a limiting process of analytic functions $R_k(x)$:

$$\pi(x) = \lim_{k\to\infty} R_k(x), \quad \text{where} \tag{4.15}$$

$$R_k(x) := R(x) - \sum_{l=-k}^{k} R(x^{\rho_l}). \tag{4.16}$$

Here ρ_l is the lth zero of the Riemann zetafunction [4.6]:

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{4.17}$$

Figure 4.6 shows $\pi(x)$ and the two approximations $R_{10}(x)$ and $R_{29}(x)$, the latter already showing a noticeable attempt to follow the jumps of $\pi(x)$.

Table 4.1 Comparison of prime-counting function $\pi(x)$ and Riemann's approximation $R(x)$

x	$\pi(x)$	$R(x)$
100000000	5761455	5761552
200000000	11078937	11079090
300000000	16252325	16252355
400000000	21336326	21336185
500000000	26355867	26355517
600000000	31324703	31324622
700000000	36252931	36252719
800000000	41146179	41146248
900000000	46009215	46009949
1000000000	50847534	50847455

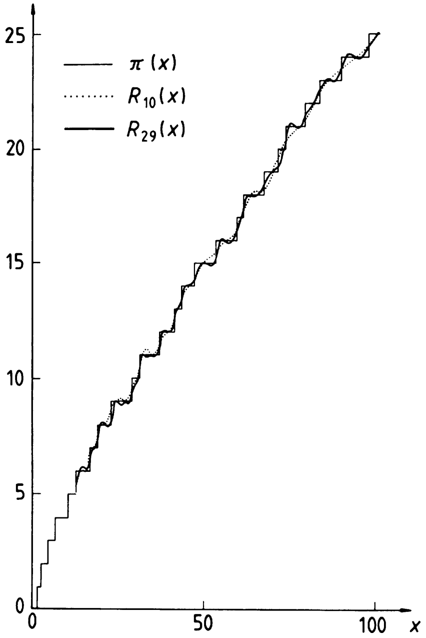


Fig. 4.6 Riemann's approximation to $\pi(x)$. [After H. Riesel, G. Göhl: Math. Comp. *24*, 969–983 (1970)]

Table 4.2 The first five zeros of the zetafunction with real part equal to $1/2$

$\rho_1 = \frac{1}{2} + 14.134725\,i$
$\rho_2 = \frac{1}{2} + 21.022040\,i$
$\rho_3 = \frac{1}{2} + 25.010856\,i$
$\rho_4 = \frac{1}{2} + 30.424878\,i$
$\rho_5 = \frac{1}{2} + 32.935057\,i$

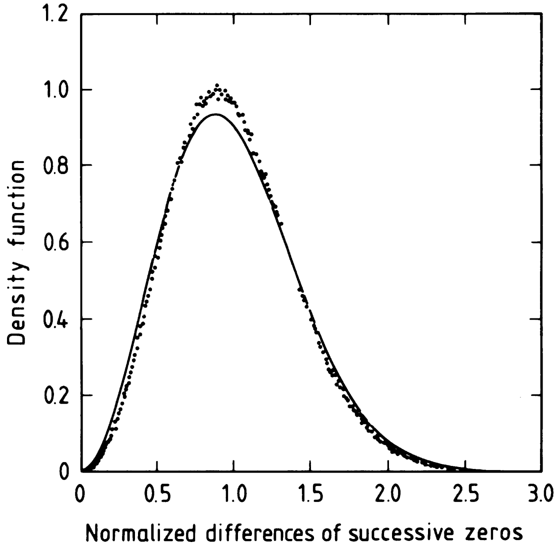
The zeros for $l = 1, 2, \ldots, 5$ of $\zeta(s)$ are shown in Table 4.2. The real parts are all equal to $1/2$. In fact, more than 100 years ago Riemann enunciated his famous hypothesis that *all* complex zeros of $\zeta(s)$ have real part $1/2$. Riemann thought at first that he had a proof, but the Riemann Hypothesis (and the so-called Extended Riemann Hypothesis, abbreviated ERH) has remained unproved to this day, although hundreds of millions of zeros have been calculated, all with real part $1/2$. In fact, the ERH is *so* widely believed today that a sizable edifice is based on it, and will collapse when the first $\mathrm{Re}(\rho_l) \neq 1/2$ makes its appearance.

(In late 1984, a *possible* proof was presented by Matsumoto in Paris. Mind boggling! If it can only be confirmed…)

4.3 David Hilbert and Large Nuclei

In conclusion, we mention that David Hilbert once conjectured that the zeros of the Riemann zetafunction were distributed like the eigenvalues of a certain kind of random Hermitian matrix. This same kind of matrix, incidentally, later gained

Fig. 4.7 Interval distribution between successive zeros ($x = 1/2$) of zetafunction. (——) Conjecture by Hilbert (Courtesy of A. Odlyzko, Bell Laboratories)



prominence in the physics of large atomic nuclei, where its eigenvalues correspond to the energy levels of the nucleons (protons and neutrons) [4.7]. In physics the resulting distribution of energy level differences is called the *Wigner distribution* after Eugene Wigner, who derived it. It is shown in Fig. 4.7 as a solid line. The dots are the results of computer calculations by Andrew Odlyzko of Bell Laboratories (private communications) of the zeros of the Riemann zetafunction around $x = 10^8$. Since the density of zeros increases logarithmically with their distance from the real line, the spacing of zeros normalized by their average spacing is shown. The close agreement between the solid line (Hilbert) and the dots (Odlyzko) shows how close Hilbert's conjecture, made almost a century ago, is.

Even so, there are noticeable differences between Hilbert's conjecture and the numerical data. Was Hilbert off? Of course not. More recent calculations by Odlyzko of hundreds of millions of zeros around the 10^{20}th zero show no discernible differences with the conjecture. In other words, convergence to the asymptotic result is very very slow. But this is not unusual for number theory where not a few results go with the twice or thrice iterated logarithm, and $\ln\ln\ln 10^{20}$ is just a little more than 1 (1.34 to be more exact). While for some problems in physics 3 is already a large number ("almost infinity" in the physicist's book, reminiscent of the sayings "three's a crowd" or "period three means chaos"), even such a monster as 10^{20} is not all that large in some corners of number theory.

4.4 Coprime Probabilities

What is the probability that two arbitrarily and independently selected numbers from a large range do not have a common divisor, i. e., that they are coprime? The probability that one of them is divisible by the prime p_i is, as we have seen, $1/p_i$, and the probability that both of them are divisible by the same prime, assuming independence, is $1/p_i^2$. Thus, the probability that they are *not* both divisible by p_i equals $1-1/p_i^2$. If we assume divisibility by different primes to be independent, then the probability of coprimality becomes

$$W_2 \approx \prod_{p_i}\left(1-\frac{1}{p_i^2}\right), \quad \text{or} \tag{4.18}$$

$$\frac{1}{W_2} \approx \prod_{p_i}\frac{1}{1-1/p_i^2} = \prod_{p_i}\left(1+\frac{1}{p_i^2}+\frac{1}{p_i^4}+\dots\right), \tag{4.19}$$

where we have expanded the denominator into an infinite geometric series.

Now if, for simplicity, we extend the product over *all* primes, then – as Euler first noted – the result is quite simple: one obtains exactly every reciprocal square integer once (this follows from the *unique* decomposition of the integers into prime factors). Thus,

$$\frac{1}{W_2} \approx \sum_{n=1}^{\infty}\frac{1}{n^2} = \zeta(2) = \frac{\pi^2}{6}, \tag{4.20}$$

and the probability of coprimality W_2 should tend towards $6/\pi^2 \approx 0.608$ for large numbers.

The probability that a randomly selected integer n is "squarefree" (not divisible by a square) also tends to $6/\pi^2$. The reasoning leading to this result is similar to that applied above to the coprimality of two integers: for an integer to be squarefree it must not be divisible by the same prime p_i more than once. Either it is not divisible by p_i or, if it is, it is not divisible again. Thus,

$$\text{Prob}\left\{p_i^2 \nmid n\right\} = \left(1-\frac{1}{p_i}\right)+\frac{1}{p_i}\left(1-\frac{1}{p_i}\right) = 1-\frac{1}{p_i^2}.$$

Taking the product over all p_i (assuming again independence of the divisibility by different primes) gives the above expression for $W_2 \approx 6/\pi^2$.

How fast is this asymptotic value reached? The sum over the reciprocal squares in (4.20) converges quite rapidly and the value of $6/\pi^2$ might already hold for the coprimality and squarefreeness of small numbers. In fact, 61 of the first 100 integers above 1 are squarefree and of the 100 number pairs made up from the integers 2 to 11, exactly 60 are coprime. This is the closest possible result because the answer has to be an even number and 62 is further away from $600/\pi^2$ than 60.

Figure 4.8 shows a computer-generated plot of coprimality in the range from 2 to 256: a white dot is plotted if its two coordinates are coprime. As expected, the density of white dots is quite uniform. All kinds of interesting micropatterns can be observed, and a number of long-range structures at angles whose tangents are simple ratios: 0, 1/2, 1, 2, etc., are also visible. Does such a plot pose new questions or suggest new relationships for number theory?

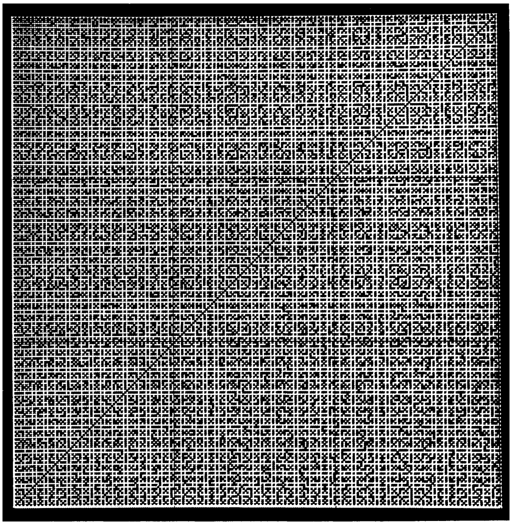


Fig. 4.8 The coprimality function, a simple number-theoretic function, in the range $2 \le x \le 256$ and $2 \le y \le 256$. A white dot is shown if $(x,y) = 1$. Whenever $(x,y) > 1$, there is no dot (*black*)

When the author first had this plot prepared (by Suzanne Hanauer at Bell Laboratories), he thought that a two-dimensional Fourier transform should make an interesting picture because the Fourier transformation brings out periodicities. And, of course, divisibility *is* a periodic property.

Figure 4.9 shows the result, which with its prominent starlike pattern would make a nice design for a Christmas card (and has, in fact, been so used). What is plotted here (as increasing brightness) is the *magnitude* of the two-dimensional discrete Fourier transform of the number-theoretic function $f(n,m)$, for $n,m = 1,2,\ldots,256$ with $f = 1$ if the GCD $(n,m) = 1$ and $f = -1$ otherwise.

Since the original function is symmetric around the 45° diagonal, so is the Fourier transform. Since only magnitude is plotted, there is another symmetry axis: the

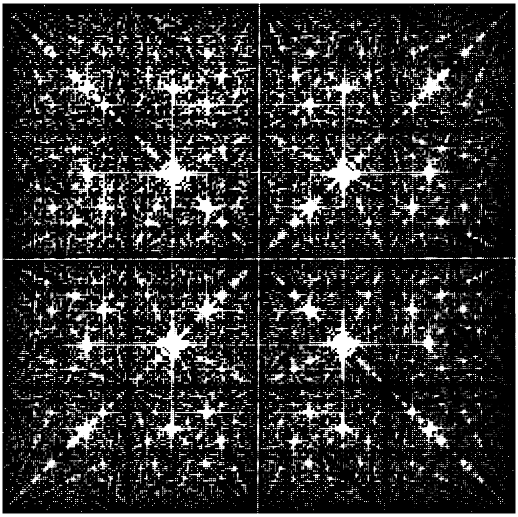


Fig. 4.9 The magnitude of the Fourier transform (simulated by increasing brightness) of the number-theoretic function shown in Fig. 4.8. The presence of a white dot, $(x,y) = 1$, is interpreted as $+1$, and the absence of a white dot as -1

-45° diagonal. In addition, there are *near* symmetries about the horizontal and vertical axes which are not so easy to explain. We leave it as an exercise to the reader to explain both this near symmetry and each of the stars in Fig. 4.9. (See also [4.8].)

The coprimality probability for more than two randomly selected integers is obtained in the manner that led to (4.20). The general result that k integers are coprimes is

$$W_k := \text{Prob}\{(n_1, n_2 \dots, n_k) = 1\} \approx [\zeta(k)]^{-1}, \tag{4.21}$$

where $\zeta(k)$ is Riemann's zetafunction as defined in (4.17). For $k = 3$ one obtains $W_3 \approx 0.832$, and for $k = 4$, $W_4 \approx 90/\pi^4 = 0.9239\dots$. (The actual proportions in the range from 2 to 101 are 0.85 and 0.93, respectively.)

The probabilities that a randomly selected integer is not divisible by a cube, a fourth power, and in general by a kth power, also tend towards (4.21). Thus, roughly 84% of all integers are "cubefree".

A somewhat more difficult problem is posed by the probability of *pairwise* coprimality of three (or more) randomly selected integers. The probability that none of k integers has the prime factor p_i is

$$\left(1 - \frac{1}{p_i}\right)^k$$

and that exactly one has p_i as a factor is

$$\frac{k}{p_i}\left(1 - \frac{1}{p_i}\right)^{k-1}.$$

The sum of these two probabilities is the probability that at *most* one of the integers has p_i as a factor. The product over all primes p_i then approximates the probability that the k integers are pairwise coprime, i. e., that

$$(n_j, n_m) = 1 \quad \text{for all} \quad j \neq m.$$

The reader may want to show that for $k = 3$, this probability can be written

$$\frac{36}{\pi^4} \prod_{p_i} \left(1 - \frac{1}{(p_i + 1)^2}\right) = 0.28\dots.$$

Thus, only about 28% of three randomly selected integers are *pairwise* coprime. (Compare this with the above result $W_3 \approx 0.832$.)

Jobst von Behr of Hamburg, who read the first edition of this book, generalized this problem by considering the probability $P_k(d)$ that the greatest common divisor (GCD) of k integers equal $d > 1$. By Monte Carlo computation on his home computer he obtained numerical results that looked suspiciously like

$$P_k(d) = d^{-k} \zeta^{-1}(k).$$

Can the reader of this edition prove this seductively simple scaling law? Summing over all d gives of course 1, as it should for a proper probability.

For the probability that the GCD of k random integers is *even* the above formula gives 2^{-k}; is this in conformity with elementary probability?

Here is another charming problem amenable to the probabilistic viewpoint: In the prime factor decomposition of a randomly selected integer $n > 1$, what is the probability $P(p)$ that the smallest prime divisor equals p? The probability (density) that n is divisible by p is of course $1/p$. For p to be the *smallest* prime factor, all prime smaller than p must *not* divide n. Hence

$$P(p) = \frac{1}{p} \prod_{q<p} \left(1 - \frac{1}{q}\right)$$

where the product is extended over all primes q smaller than p. For example, for $p = 7$, $P(7) = 4/105$, i.e roughly 4% of all integers have 7 as their smallest prime factor.

Now let us *sum* the above expression over *all* primes, 2, 3, 5, 7, 11, ..., so we get

$$P = \sum_p \frac{1}{p} \prod_{q<p} \left(1 - \frac{1}{q}\right)$$

the probability that an integer $n > 1$ has some prime number as its smallest prime divisor which, given the sets $P(p)$ are pairwise disjoint, equals of course 1! Every integer > 1 has prime divisors one of which is necessarily the smallest.

To what extent does the above result for P depend on the exact values of the primes? For the (very rough) approximation $m \log m$ for the mth prime, the sum converges to 1.5 (instead of 1).

4.5 Primes in Progressions

A famous theorem by Dirichlet (1837), Gauss's successor in Göttingen, states: there are infinitely many primes in every *linear progression*

$$a \cdot n + b, \qquad n = 1, 2, 3, \ldots, \tag{4.22}$$

provided the constants a and b are coprime: $(a, b) = 1$. Thus, for example, with $a = 10$ and $b = 1$, 3, 7 or 9, we see that there are infinitely many primes whose last digit is 1, 3, 7 or 9. (In fact, as we shall see later, these four kinds of primes occur in equal proportion.)

The longest sequence known in early 1982 for which $a \cdot n + b$ gives primes for *consecutive* n is the progression

$$223092870 \cdot n + 2236133941,$$

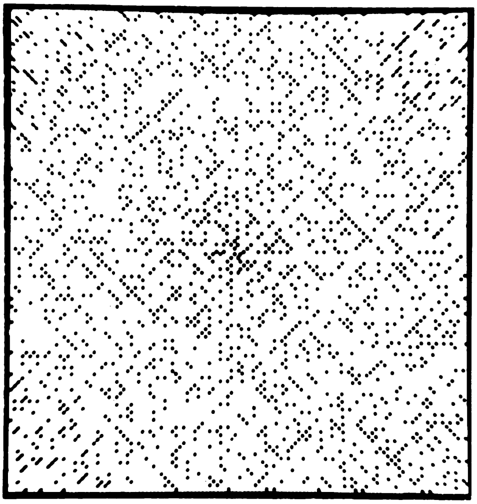


Fig. 4.10 Primes (*dots*) plotted on a spiral. Many primes fall on straight lines

which is prime for sixteen consecutive values: $n = 0, 1, 2, \ldots, 15$. Since then a 19-member progression has been discovered.

The record for a *quadratic* progression stands at 80 consecutive primes, namely

$$n^2 + n + 41 \quad \text{for } n = -40, -39, \ldots, 0, \ldots, 39. \tag{4.23}$$

This is remarkable because it would ordinarily take a polynomial in n of degree 80 to get 80 primes for consecutive values of n.

Many primes are of the form $4n^2 + an + b$, which makes them lie on straight lines if n is plotted along a square spiral. This fact is illustrated by so plotting the primes (see Figs. 4.10, 4.11).

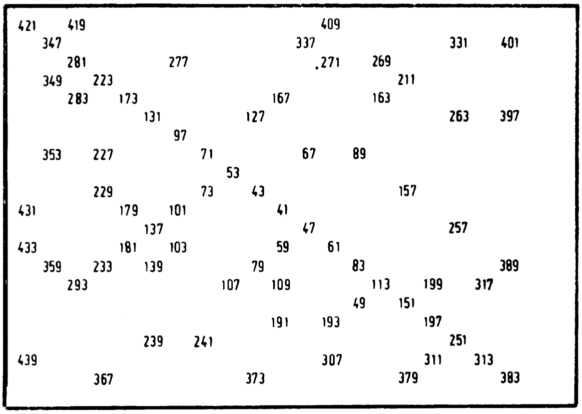


Fig. 4.11 The primes between 41 and 439 plotted on a square spiral beginning with 41 in the centre. Note the "solid-prime" diagonal

However, there is no polynomial, no matter how high its degree, which yields primes for all values of n. If there was a polynomial in n of finite degree r generating primes for all n, then $r+1$ primes would determine the $r+1$ coefficients of the polynomial and infinitely many other primes could then be calculated from these $r+1$ primes. The location of primes among the integers is simply too unpredictable to be "caught" by something as regular and finite as a polynomial.

4.6 Primeless Expanses

On the other hand, there is always a prime between n^3 and $(n+1)^3-1$ for large enough n. This fact was exploited by W. H. Mills (1947) to construct a constant A such that $\lfloor A^{3^n} \rfloor$ is prime for all n [4.9, p. 160]. But of course, the Mills expression is not a polynomial, and as we remarked before (Sect. 3.4), the primes thus generated have been "smuggled" into A first.

Somewhat paradoxically, there are also *arbitrarily* large intervals without a single prime! For example, the one *million consecutive* integers

$$(10^6+1)!+n, \qquad n=2,3,4,\ldots,1{,}000{,}001 \tag{4.24}$$

are all composite! In fact, there is even a set of one million somewhat smaller consecutive integers that are all composite, namely those in which the additive term n in the above expression is replaced by $-n$.

In *relative terms*, the primeless expanse of one million integers is, of course, rather small. A (weak) upper bound on the relative size of primeless ranges is an "octave" of integers; i. e., there is always a prime p in the range n and $2n$ (inclusive):

$$n < p \leq 2n, \tag{4.25}$$

or, equivalently, each prime is less than twice its predecessor [4.10]:

$$p_{k+1} < 2p_k. \tag{4.26}$$

Check: $3 < 2 \cdot 2$, $5 < 2 \cdot 3$, $7 < 2 \cdot 5$, $11 < 2 \cdot 7$, etc.

In fact, the number of primes in the interval from n to $2n$ is of the same order as those below n. This follows directly from the asymptotic expression for $\pi(x)$:

$$\pi(x) \approx \frac{x}{\ln x}, \tag{4.27}$$

so that

$$\begin{aligned} \pi(2x)-\pi(x) &\approx \frac{2x}{\ln x+\ln 2} - \frac{x}{\ln x} \\ &\approx \frac{x}{\ln x} - \frac{2x\ln 2}{\ln^2 x}. \end{aligned} \tag{4.28}$$

4.7 Squarefree and Coprime Integers

The probability that a given integer is squarefree approaches $6/\pi^2$ (see Sect. 4.4) and $6/\pi^2$ is also the asymptotic probability that two randomly chosen integers are coprime. Are these two properties independent? No! Among 500 random integers von Behr found 205 that were both squarefree *and* coprime (to another random integer), instead of only $500 \cdot 36/\pi^4 \approx 185$ if these two properties were independent. Thus, there seems to be a *positive* correlation between squarefreeness and coprimality. The reader may wish to show that the joint probability equals $36/\pi^4$ times a peculiar product, which is larger than 1:

$$\prod_i \left(1 + \frac{1}{p_i^3 + p_i^2 - p_i - 1} \approx 1.16\right).$$

4.8 Twin Primes

Primes not infrequently come in pairs called twin primes, like 11 and 13 or 29 and 31. How often does it happen? An *estimate* [4.2] shows their density to be proportional to $1/(\ln x)^2$, i.e. the square of the density of single primes, suggesting that they may occur independently. But one must be careful here because prime triplets of the form $(x, x+2, x+4)$ can never happen (other than the triplet 3, 5, 7), since one member of such a triplet is always divisible by 3. (Reader: try to show this – it is easy.)

On the other hand, triplets of the form $(x, x+2, x+6)$ or $(x, x+4, x+6)$ are not forbidden and do happen, for example, 11, 13, 17 or 13, 17, 19. Is their asymptotic density proportional to $1/(\ln x)^3$? In number theory, what is not explicitly forbidden often occurs, and often occurs randomly – resembling total chaos rather than neat order.

Let us try to estimate the density of twin primes. The probability that a natural number x is not divisible by a prime $p < x$ equals about $1 - 1/p$. The probability $W(x)$ that x is prime is therefore approximately

$$W(x) = \prod_p^{x'} \left(1 - \frac{1}{p}\right). \tag{4.29}$$

Here p ranges over all primes below some "cut-off" value x', where x' is about $x^{0.5}$. On the other hand the prime number theorem [4.10, Theorem 6] tells us that, asymptotically,

$$W(x) = \frac{1}{\ln x}. \tag{4.30}$$

How are (4.29) and (4.30) related? According to Mertens' Theorem [4.10, Theorem 429]

$$\prod_{p}^{x}\left(1-\frac{1}{p}\right) \to \frac{\mathrm{e}^{-\gamma}}{\ln x}, \tag{4.31}$$

where $\gamma = 0.5772$ is Euler's constant. To have (4.29) agree with (4.30) we set $x' = x^{\mathrm{e}^{-\gamma}} \approx x^{0.56}$.

While such heuristic estimates may appear rather rough, the numerical evidence is quite reassuring. Thus, for $x = 10^3$, (4.29) yields $W(x) = 0.139$, whereas (4.30) and an actual count of primes around $x = 10^3$ give 0.145 and 0.144, respectively. For $x = 10^6$, the corresponding results are 0.0723, 0.0724 and 0.0726. For $x = 10^9$, we get 0.0482, 0.0483 and 0.0484.

The probability $W_2(x)$ that both x *and* $x+2$ are primes is obtained from the following two "inequalities" or rather *incongruences*:

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -2 \quad \mathrm{mod}\, p. \tag{4.32}$$

(For the definition and rules of congruences see Chap. 6.) For $p = 2$, both parts of (4.32) amount to just *one* condition (x must be odd) yielding the probability factor $\frac{1}{2}$. For $p > 2$, the two parts of (4.32) are two independent conditions, yielding the probability factor $(1 - 2/p)$. Thus,

$$W_2(x) = \frac{1}{2}\prod_{p>2}^{x'}\left(1-\frac{2}{p}\right). \tag{4.33}$$

To connect (4.33) with (4.29) we rewrite the product in (4.33) as follows

$$\prod_{p>2}^{x'}\left(1-\frac{2}{p}\right) = \frac{\prod_{p>2}^{x'}\left(1-\frac{2}{p}\right)}{\prod_{p>2}^{x'}\left(1-\frac{1}{p}\right)^2}\prod_{p>2}^{x'}\left(1-\frac{1}{p}\right)^2. \tag{4.34}$$

Here the ratio of the two products converges for large x' to $0.66016\ldots$, called the twin-prime constant. The remaining product equals, according to (4.29) and (4.30), $4/\ln^2 x$. Thus,

$$W_2(x) = \frac{1.32032}{\ln^2 x}, \tag{4.35}$$

showing the expected trend with $1/\ln^2 x$. Note, that the sum over all x diverges, implying an (unproven) infinity of twin primes.

Although there is no mathematical *proof* for an infinity of twin primes, (4.35) is supported by excellent numerical evidence. (Curiously, Kummer, the great mathematician of early Fermat fame, obtained, for some reason, an erroneous answer for $W_2(x)$ that, in the pre-computer age, remained long undetected.)

Interestingly, while the sum of $1/p$ over all primes diverges (albeit very slowly – namely like $\ln(\ln x)$), the sum of $1/p_2$, where p_2 is the smaller member of a twin-prime pair, *converges*. Thus, although there probably are infinitely many twin primes (according to our heuristic estimate (4.35)), their density is not sufficient to make $\sum 1/p_2$ diverge. To what value does the $\sum 1/p_2$ converge? Numerical evidence obtained by summing $1/p_2$ up to $p_2 = 1299451$ gives 0.9652. Adding to this the integral of $1/p$ over the density (4.35) beyond 1299451 yields approximately 1.06 for $\sum 1/p_2$.

What can we deduce about twin primes with a spacing of 4 instead of 2? Instead of the incongruences (4.32) we now have

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -4 \quad \bmod p. \tag{4.36}$$

Again, for $p > 2$, two remainders, 0 and 4, are forbidden. Thus, the density $W_4(x)$ of $(x, x+4)$ twins equals the density of $W_2(x)$ of $(x, x+2)$ twins:

$$W_4(x) = W_2(x). \tag{4.37}$$

If both x and $x+4$ are primes, $x+2$ cannot be prime. In fact, $x+2$ must be divisible by 3. Thus, x and $x+4$ are true twins, i. e. primes with no intervening primes.

4.9 Prime Triplets

There are two kinds of prime triplets, those with a 2;4 spacing pattern and others with a 4;2 pattern. We first consider the 2;4 pattern, that is those cases for which x, $x+2$ and $x+6$ are primes. Thus the following three incongruences must be obeyed

$$x \not\equiv 0, \quad x \not\equiv -2, \quad x \not\equiv -6 \qquad \bmod p. \tag{4.38}$$

For $p = 2$, this amounts again to just one condition: x must be odd, which yields the probability factor $\frac{1}{2}$. For $p = 3$, the incongruences (4.38) impose two conditions, namely neither x nor $x+2$ must be divisible by 3, yielding the probability factor $\frac{1}{3}$. For $p > 3$, all three incongruences of (4.38) are "active", eliminating three out of p cases and yielding the probability factor $1 - 3/p$. The probability $W_{2;4}(x)$ for a prime triplet with differences between successive primes of 2 and 4, respectively, is therefore

$$W_{2;4}(x) = \frac{1}{6} \prod_{p>3}^{x'} \left(1 - \frac{3}{p}\right), \tag{4.39}$$

or

$$W_{2;4}(x) = \frac{1}{6} \frac{\prod_{p>3}^{x'} \left(1 - \frac{3}{p}\right)}{\prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^3} \prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^3. \tag{4.40}$$

Here the ratio of the two products converges for large x' to 0.63516... (which might be called the triple-prime constant). With (4.29) and (4.30), the remaining product is seen to approach $27 \ln^3 x$. Thus, approximately,

$$W_{2;4}(x) = \frac{2.858}{\ln^3 x}. \tag{4.41}$$

For prime triplets with x, $x+4$ and $x+6$ prime, the incongruences are

$$x \not\equiv 0, \quad x \not\equiv -4, \quad x \not\equiv -6 \qquad \bmod p, \tag{4.42}$$

leading to equinumerous sets of restrictions as (4.38). Thus, the corresponding probability $W_{4;2}(X)$ is given by

$$W_{4;2}(x) = W_{2;4}(x). \tag{4.43}$$

4.10 Prime Quadruplets and Quintuplets

As with prime triplets, there are two different spacing patterns for quadruplets of close primes: $2;4;2$ and $4;2;4$. Curiously, numerical evidence suggests that the latter pattern $(4;2;4)$ seems to be twice as numerous as the other pattern $(2;4;2)$. We would like to understand why.

For the $2;4;2$ pattern the four incongruences are

$$x \not\equiv 0, \quad x \not\equiv -2, \quad x \not\equiv -6, \quad x \not\equiv -8 \qquad \bmod p. \tag{4.44}$$

The *effective* number of incongruences for $p = 2$ equals one; for $p = 3$ the number equals two; for $p > 3$ all four incongruences are active. The probability of the $2;4;2$ pattern is therefore

$$W_{2;4;2}(x) = \frac{1}{6} \prod_{p>3}^{x'} \left(1 - \frac{4}{p}\right), \tag{4.45}$$

or

$$W_{2;4;2}(x) = \frac{1}{6} \frac{\prod_{p>3}^{x'} \left(1 - \frac{4}{p}\right)}{\prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^4} \prod_{p>3}^{x'} \left(1 - \frac{1}{p}\right)^4. \tag{4.46}$$

Here the ratio of the two products converges to 0.307496... (the "quadruplet constant") and the remaining factor tends to $81/\ln^4 x$. Thus

$$W_{2;4;2}(x) = \frac{4.15}{\ln^4 x}. \tag{4.47}$$

For the spacing pattern $4;2;4$ the four congruences are

$$x \not\equiv 0, \quad x \not\equiv -4, \quad x \not\equiv -6, \quad x \not\equiv -10 \qquad \bmod p. \tag{4.48}$$

Here, for $p = 5$, the last incongruence is automatically fulfilled by the first ($x \not\equiv 0$). Thus there are only *three* (instead of four) active incongruences for $p = 5$. The probability factor for $p = 5$ therefore doubles from $(1 - 4/5)$ to $(1 - 3/5)$. As a result, we have

$$W_{4;2;4}(x) = 2W_{2;4;2}(x), \tag{4.49}$$

confirming the numerical evidence.

The next closely spaced cluster of primes is a quintuplet with the spacing pattern $2;4;2;4$, such as 5, 7, 11, 13, 17 which repeats at 11, 13, 17, 19, 23 and 101, 103, 107, 109, 113, etc. Another possible quintuplet type has the spacing pattern $4;2;4;2$, such as 7, 11, 13, 17, 19, which repeats at 97, 101, 103, 107, 109 but not again until 1867, 1871, ... The reader should have no difficulty deriving the asymptotic densities for these quintuplets following the recipe of Sects. 4.8–4.10.

However, not all patterns are possible, as we have already seen in Sect. 4.9 with prime triplets: the spacing pattern $2;2$ is impossible after 3, 5, 7 because one member of such a triplet must be divisible by 3. Another way to demonstrate the impossibility of the spacing pattern $2;2$ uses the appropriate incongruences: for x, $x+2$, $x+4$ to be prime. These are

$$x \not\equiv 0, \quad x \not\equiv -2, \quad x \not\equiv -4 \qquad \bmod p.$$

For $p = 3$, the last incongruence may be written as $x \not\equiv -1 \bmod 3$. Thus *all three* possible remainders modulo 3 $(0, -2, -1)$ are forbidden and 3, 5, 7 is the only triplet with the $2;2$ pattern.

In a similar manner, we can show that the sextet with the spacing pattern $2;4;2;4;2$ occurs only once (5, 7, 11, 13, 17, 19) and never again. The six incongruences for the sextet considered are

$$x \not\equiv 0,\ -2,\ -6,\ -8,\ -12,\ -14 \quad \bmod p.$$

For $p = 5$ we may write

$$x \not\equiv 0,\ -2,\ -1,\ -3,\ -2,\ -4 \quad \bmod 5,$$

which excludes *all five* possible classes of remainders modulo 5.

What if any spacing pattern of length 7, 8 etc. is unique in the sense that, like 2, 3 and 3, 5, 7 and 5, 7, 11, 13, 17, 19, it occurs only once and never again?

4.11 Primes at Any Distance

In Sect. 4.8 we derived the formula $W_2(x) = 1.32032/\ln^2 x$ for the asymptotic density of twin primes with a distance of 2 (such as 3 and 5 etc.). We also noted that twin primes with a distance of 4 (like 7 and 11 etc.) are equally probable: $W_4(x) = W_2(x)$.

What is the density of pairs of odd primes, not necessarily adjacent, with a distance of $\Delta = 6$ (like 5 and 11) or a distance of $\Delta = 8$ (like 3 and 11) or any other distance Δ irrespective of any intervening primes?

For $\Delta = 6$, the incongruences (4.32) have to be replaced by

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -6 \qquad \bmod p. \tag{4.50}$$

For $p = 2$ and $p = 3$, only one of these two incongruences is active, yielding a probability factor $(1 - 1/2)\ (1 - 1/3) = 1/3$. For $p > 3$ both incongruences are active yielding probability factors $(1 - 2/p)$. Hence, with (4.33),

$$W_6(x) = \frac{1}{3}\prod_{p>3}^{x'}\left(1 - \frac{2}{p}\right) = \prod_{p>2}^{x'}\left(1 - \frac{2}{p}\right) = 2W_2(x), \tag{4.51}$$

i.e. pairs of primes with a distance $\Delta = 6$ are twice as numerous as twin primes.

For $\Delta = 8$, the second of the two incongruences (4.50) becomes $x \not\equiv -8 \bmod p$. Since 8 does not contain any prime factors $p > 2$, we can proceed as in the derivation of the twin-prime density with $\Delta = 2$ and $\Delta = 4$. The result is

$$W_8(x) = W_2(x). \tag{4.52}$$

In fact, for any $\Delta = 2^k$, $k = 1, 2, 3\ldots$, we obtain

$$W_{2^k}(x) = W_2(x). \tag{4.53}$$

More generally, for arbitrary Δ, only the prime factors of Δ are important. Of the two incongruences

$$x \not\equiv 0 \quad \text{and} \quad x \not\equiv -\Delta \qquad \bmod p \tag{4.54}$$

only one is active for all odd prime factors q of Δ. Thus

$$W_\Delta(x) = \frac{1}{2}\prod_q \frac{1 - \frac{1}{q}}{1 - \frac{2}{q}} \prod_{p>2}^{x'}\left(1 - \frac{2}{p}\right), \tag{4.55}$$

where q are the different odd prime factors of Δ and p are *all* primes above 2 and below x'. With (4.33) we get

$$W_\Delta(x) = \Pi(\Delta)W_2(x), \quad \text{where} \quad \Pi(\Delta) = \prod_{\substack{q>2 \\ q|\Delta}} \frac{q-1}{q-2}, \tag{4.56}$$

a charming generalization of our earlier result (4.51) that $W_6(x)$ equals $2W_2(x)$. But W_6 is not the largest value. For $\Delta = 30$ we get $W_{30} = 2.66W_2$. In fact, there is no upper bound for $\Pi(\Delta)$ because the product over q diverges if all possible primes are included as prime factors of Δ. More specifically, the product $\Pi(\Delta)$ diverges as $1.349 \ln q_{\max}$, where $q_{\max}$ is the largest prime factor of Δ. Are there two different coprime values of Δ having the same $\Pi(\Delta)$? Let the reader decide – or rather find out.

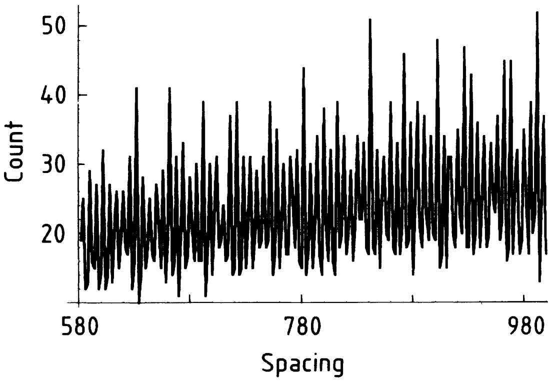


Fig. 4.12 The number of primes with a given spacing

Of the two factors in (4.56) one is slowly varying, representing a monotonic trend, $W_2(x)$, whereas the product $\Pi(\Delta)$ fluctuates appreciably with changing values of Δ. Figure 4.12 shows a plot of $\Pi(\Delta)$ as a function of Δ. For every value of Δ that is divisible by 3, i.e. every third value, the quotient $(q-1)/(q-2)$ contributes a factor of 2 to the product. Similarly, for every value of Δ divisible by 5, i.e. every 5th value, a factor of $4/3$ is contributed – and so on for 7, 11 etc.

These inherent periodicities are brought out nicely by the Fourier transform of $\Pi(\Delta)$: see Fig. 4.13, which shows a pronounced peak at one third the sampling frequencies, corresponding to $p = 3$, and smaller peaks corresponding to $p = 5$ and 7 and their "harmonics". As we shall see, $\Pi(\Delta)$ has a certain universal air about it because it also governs the sum of two primes (cf. Sect. 4.13).

$\Pi(\Delta)$ is also related to Euler's Φ function $\Phi(m)$ (see Sect. 8.3) and in fact resembles $m/\Phi(m)$:

$$\frac{m}{\Phi(m)} = \prod_{p|m} \frac{p}{p-1},$$

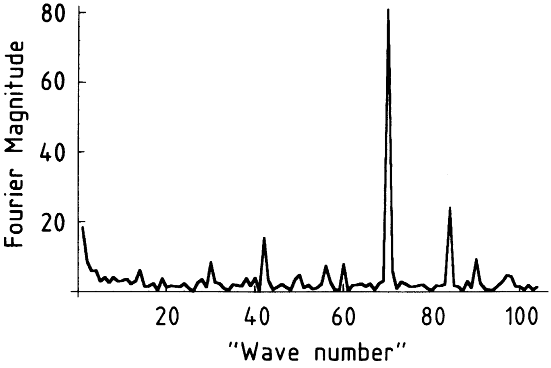


Fig. 4.13 The magnitude of the Fourier transform showing the preponderance of spacings at multiples of 6

whereas

$$\Pi(m) = \prod_{\substack{p>2 \\ p|m}} \frac{p-1}{p-2}. \tag{4.57}$$

4.12 Spacing Distribution Between Adjacent Primes

The prime number theorem prescribes the average spacing $\bar{s}$ between adjacent primes. For $n \gg 1$, this spacing is about $\ln n$. (For example, for $n = 20$, $\bar{s} \approx \ln 20 \approx 3$.)

All spacings between odd primes are of course even. If the distribution of spacings were otherwise unconstrained, then, given an average spacing, the maximum-entropy principle [4.11] would tell us that the distribution of $k = s/2$ is the geometric distribution:

$$d(k) = \frac{1}{\bar{k}} \left(\frac{\bar{k}}{1+\bar{k}} \right)^k, \quad \bar{k} = \frac{1}{2} \ln n, \qquad k = 1, 2, \ldots . \tag{4.58}$$

In reality, the distribution of spacings is anything but unconstrained. We already know (4.37) that the frequencies of the spacings $s = 2$ (i.e. $k = 1$) and $s = 4$ (i.e. $k = 2$) are equal, while (4.58) would predict a ratio of $d(1)/d(2) = 1 + 1/\bar{k}$. Figure 4.14 shows a plot of the logarithm of $d(k)$ in the range p_{20000} to p_{40000} for $k = 1$ to 36. There is the predicted overall linear trend: $\ln d(k) \sim r/k$ with $r = -\ln(1 + 1/\bar{k})$. But the "bumps" at $k = 3$ (i.e. $s = 6$), $k = 15$ (i.e. $s = 30$) and other places (especially multiples of 3) are also visible.

The average slope obtained by regression of the data in Fig. 4.14 is $r = -0.148$, in good agreement with the theoretical value of $r = -0.146$.

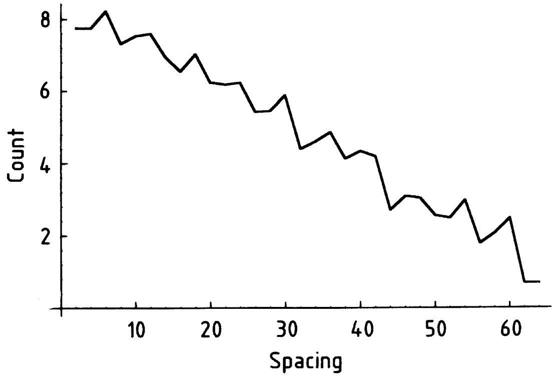


Fig. 4.14 The distribution of the spacings between *adjacent* primes

4.13 Goldbach's Conjecture

One of the most enduring (if not endearing) mathematical conjectures, made by the Russian mathematician Christian Goldbach (1690–1764) in a letter to Leonhard Euler, is the famous *Goldbach conjecture*. It asserts that every even number $n > 4$ is the sum of two odd primes. Some progress has been made on related weaker assertions, and the Goldbach conjecture itself has been *numerically* confirmed up to very large numbers. But alas, even if it had been shown to hold up to $10^{10^{10}}$, there would be no guarantee that it would not fail for $10^{10^{10}} + 2$.

Let us try our heuristic muscle at this recalcitrant conjecture. We want to get a feeling for the number of representations

$$x + y = n,$$

where x and y are odd primes and $n > 4$ is even. We want to count permuted representations, such as $3 + 5 = 5 + 3 = 8$, as only *one case*. Without loss of generality we assume $x \leq n/2$.

The number of odd primes below $n/2$ equals approximately $n/(2 \ln n)$. Each such prime is a candidate for $x + y = n$ and contributes to the count in question if $y = n - x$ is also prime. What is the probability of $n - x$ being prime, given that x is an odd prime?

To answer this question, we have to distinguish two cases. For a given potential prime divisor p of $n - x$: does p divide n or does it not? In the first case $(p|n)$ the two incongruences

$$x \not\equiv 0 \quad \text{and} \quad n - x \not\equiv 0 \qquad \bmod p \tag{4.59}$$

amount to only *one* condition because, for $p|n$, $n - x \equiv x \bmod p$. Thus the probability factor is $(1 - 1/p)$. In the other case $(p \not| n)$ and $p > 2$, the incongruences (4.59) are independent of each other. The probability factor is therefore $(1 - 2/p)$. For $p = 2$ the factor is $1/2$. Multiplying these probability factors yields

$$\prod_{p|n} \left(1 - \frac{1}{p}\right) \prod_{\substack{p>2 \\ p|n}} \left(1 - \frac{2}{p}\right) \cdot \frac{1}{2}.$$

This factor has to be divided by the unrestricted probability factor

$$\prod_{p} \left(1 - \frac{1}{p}\right)^2$$

and multiplied by the *a priori* probability that $n - x$ is prime. Since $n - x$ is already odd (because n is even and x is odd), this probability equals

$$\frac{2}{\ln(n-x)} \approx \frac{2}{\ln n}$$

(because $x \leq n/2$).

Thus, the total estimated count $C_2(n)$ of cases for which both x and $n-x$ are prime is given by

$$C_2(n) = \prod_{p>2} \frac{\left(1-\frac{2}{p}\right)}{\left(1-\frac{1}{p}\right)^2} \cdot \prod_{\substack{p>2\\ p|n}} \frac{p-1}{p-2} \cdot \frac{n}{\ln^2 n}. \tag{4.60}$$

Here the first factor approaches our old friend, the twin-prime constant 0.66016... (4.34). This factor combined with the last factor $(n/\ln^2 n)$ is nothing but the twin-prime density $W_2(n)$, see (4.35), multiplied by 1/2. Thus

$$C_2(n) = \frac{n}{2} W_2(n) \prod_{\substack{p>2\\ p|n}} \frac{p-1}{p-2}. \tag{4.61}$$

To simplify this formula even further, we recall the function $W_\Delta(n)$ [see (4.55)]. With $\Delta = n$

$$C_2(n) = \tfrac{n}{2} W_n(n). \tag{4.62}$$

Thus the number of Goldbach representations $C_2(n)$ equals, within a factor $n/2$, the density of prime pairs $W_n(n)$ with a spacing of n in the neighbourhood of n.

Another interesting aspect of this result is the product in (4.61). Suppose n is a multiple of 3, then this product, for $p=3$, contributes a factor of 2 to the count. If n is divisible by 5, the product contributes a factor 4/3, which is still appreciably larger than 1. Thus, we see that the count $C_2(n)$ has a pronounced periodic component with a period of 3 (and weaker periodicities with periods of 5, 7, 11, etc.), just like $\Pi(n)$ (see Figs. 4.12 and 4.13).

4.14 Sum of Three Primes

Another problem from additive number theory that Goldbach posed in 1742 in correspondence with Euler, then at St. Petersburg, concerned the sum of *three* primes. Goldbach asked whether every sufficiently large odd n can be written as

$$n = p_1 + p_2 + p_3. \tag{4.63}$$

Positive proof that this is indeed possible had to wait almost 200 years – until 1937 when I. M. Vinogradov [4.12] furnished a proof based on Fourier-like trigonometric sums. Trigonometric sums have played an important role in number theory ever since, forging a strong link between additive and multiplicative number theory.

How did Vinogradov get from (4.63) to trigonometric sums? The number of cases including permutations $C_3(n)$ for which (4.63) holds can be written as follows

$$C_3(n) = \sum_{p_1<n} \sum_{p_2<n} \sum_{p_3<n} \int_0^1 \exp\left[2\pi \mathrm{i}(p_1 + p_2 + p_3 - n)x\right] \mathrm{d}x, \tag{4.64}$$

because when (4.63) holds the integral equals 1 and the triple sum is augmented by 1; otherwise the integral is 0 and nothing is added.

As a next step, Vinogradov converted the triple sum into a single sum:

$$C_3(n) = \int_0^1 \sum_{p<n} \left[\exp(2\pi \mathrm{i} xp)\right]^3 \exp(-2\pi \mathrm{i} n)\, \mathrm{d}x, \tag{4.65}$$

which he was able to convert into the product of $(1 - 1/(p^2 - 3p + 3))$ taken over all prime divisors of n.

The final result of Vinogradov's method is that the number of representations $C_3(n)$ of an odd integer as the sum of three primes equals, asymptotically,

$$C_3(n) = C \prod_{p|n} \left(1 - \frac{1}{p^2 - 3p + 3}\right) \frac{n^2}{\ln^3 n}. \tag{4.66}$$

This formula resembles (4.60) for the sum of *two* primes, except that the slowly varying factor $n/\ln^2 n$ (the "trend") has been replaced by $n^2/\ln^3 n$ and the product over $p|n$ is over a quadratic expression in p. However, like $\Pi(n)$ [see (4.57)] this product, too, depends sensitively on the small prime factors of n. Thus, if n is divisible by 3, the term for $p = 3$ contributes a factor 2/3 to the product.

Part II
Some Simple Applications

Chapter 5
Fractions: Continued, Egyptian and Farey

Continued fractions are one of the most delightful and useful subjects of arithmetic, yet they have been continually neglected by our educational factions. Here we discuss their applications as approximating fractions for rational or *irrational numbers* and *functions*, their relations with *measure theory* (and deterministic chaos!), their use in *electrical networks* and in solving the *"squared square"*; and the *Fibonacci* and *Lucas numbers* and some of their endless applications.

We also mention the (almost) useless *Egyptian fractions* (good for designing puzzles, though, including *unsolved* puzzles in number theory) and we resurrect the long-buried *Farey fractions*, which are of considerable contemporary interest, especially for *error-free computing*.

Among the more interesting recent applications of Farey series is the reconstruction of periodic (or nearly periodic) functions from "sparse" sample values. Applied to two-dimensional functions, this means that if a motion picture or a television film has sufficient structure in space and time, it can be reconstructed from a fraction of the customary picture elements ("pixels"). ("Sufficient structure" in spacetime implies that the reconstructions might not work for a blizzard or a similar "snow job".)

5.1 A Neglected Subject

Continued fractions (CFs) play a large role in our journey through number theory [5.1]. A simple continued fraction

$$b_0 + \cfrac{1}{b_1 + \cfrac{1}{b_2 + \cfrac{1}{b_3 + \dots}}} \quad , \tag{5.1}$$

a typographical nightmare if there ever was one, is usually written as follows: $[b_0; b_1, b_2, b_3, \ldots]$. Here the b_m are integers. A finite simple CF then looks like this:

$$[b_0; b_1, \ldots, b_n]. \tag{5.2}$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_5, © Springer-Verlag Berlin Heidelberg 2009

If a finite or infinite CF is broken off after $k < n$, then $[b_0;b_1,\ldots,b_k] = A_k/B_k$ is called the approximating fraction or convergent of order k. Here A_k and B_k are coprime integers; they obey the recursion

$$A_k = b_k A_{k-1} + A_{k-2}, \tag{5.3}$$

with $A_0 = b_0, A_{-1} = 1$ and $A_{-2} = 0$. The B_k are derivable from the same recursion:

$$B_k = b_k B_{k-1} + B_{k-2}, \tag{5.4}$$

with $B_0 = 1$ and $B_{-1} = 0$.

As the order of the approximating fractions increases, so does the degree of approximation to the true value of the fraction, which is approached alternately from above and below.

CFs are unique if we outlaw a 1 as a final entry in the bracket. Thus 1/2 should be written as $[0;2]$ and not $[0;1,1]$. In general, if a 1 occurs in the last place, it can be eliminated by adding it to the preceding entry.

Continued fractions are often much more efficient in approximating rational or irrational numbers than ordinary fractions, including decimals. Thus,

$$r = \frac{964}{437} = [2;\quad 4,\quad 1,\quad 5,\quad 1,\quad 12], \tag{5.5}$$

and its approximating fraction of order 2, $[2;4,1] = 11/5$, approaches the final value within 3 parts in 10^3.

One interesting application of CFs is to answer such problems as “when is the power of the ratio of small integers nearly equal to a power of 2?”, a question of interest in designing cameras, in talking about computer memory and in the tuning of musical instruments (Sect. 2.6). For example, what integer number of musical *major thirds* equals an integral number of *octaves*, i. e., when is

$$\left(\frac{5}{4}\right)^n \approx 2^m, \quad \text{or}$$

$$5^n \approx 2^{m+2n}?$$

By taking logarithms to the base 2, we have

$$\log_2 5 \approx \frac{m}{n} + 2.$$

The fundamental theorem tells us that there is no exact solution; in other words, $\log_2 5$ is irrational. With the CF expansion for $\log_2 5$ we find

$$\begin{aligned} &\log_2 5 = 2.3219\cdots = [2;3,9,\ldots] \quad \text{or} \\ &\log_2 5 \approx 2 + \tfrac{1}{3}, \end{aligned}$$

yielding $m = 1$ and $n = 3$. Check:

$$\left(\frac{5}{4}\right)^3 = 1.953\cdots \approx 2.$$

In other words, the well-tempered third-octave $2^{1/3}$ matches the major third within 0.8 % or 14 musical cents. (The musical cent is defined as 1/1200 of an octave. It corresponds to less than 0.6 Hz at 1 kHz, roughly twice the just noticeable pitch difference.)

The next best CF approximation gives

$$\log_2 5 \approx 2 + \tfrac{9}{28},$$

or $m = 9$, $n = 28$, a rather unwieldy result.

Because $\log_{10} 2$, another frequently occurring irrational number, is simply related to $\log_2 5$:

$$\log_{10} 2 = 1/(1 + \log_2 5),$$

it has a similar CF expansion:

$$\log_{10} 2 = [0; 3, 3, 9, \ldots].$$

It, too, is well approximated by breaking off before the 9. This yields

$$\log_{10} 2 \approx [0; 3, 3] = \tfrac{3}{10},$$

a well-known result (related to the fact that $2^{10} \approx 10^3$).

Some irrational numbers are particularly well approximated. For example, the widely known first-degree approximation to π, namely, $[3; 7] = 22/7$, comes within 4 parts in 10^3. The second-order approximation $[3; 7, 16] = 355/113$, known to the early Chinese, approaches π within 10^{-7}.

Euler [5.2] discovered that the CF expansion of $\mathrm{e} = 2.718281828\ldots$, unlike that of π, has a noteworthy regularity:

$$\mathrm{e} = [2;\ 1,\ 2,\ 1,\ 1,\ 4,\ 1,\ 1,\ 6,\ 1,\ 1,\ 8,\ 1,\ \ldots], \tag{5.6}$$

but converges initially very slowly because of the many 1's. In fact, the CF for the *Golden section or Golden ratio* $g = [1; 1, 1, 1, 1 \ldots]$, which contains infinitely many 1's, is the most slowly converging CF. It is therefore sometimes said, somewhat irrationally, that g is the "most irrational" number. In fact, for a given order of rational approximation the approximation to g is worse than for any other number. Because of this property, up-to-date physicists who study what they call "deterministic chaos" in nonlinear systems often pick the Golden ratio g as a parameter (e. g., a frequency ratio) to make the behaviour "as aperiodic as possible". A strange application of number theory indeed!

CFs are also useful for approximating *functions*. Thus, in a generalization of our original bracket notation, permitting noninteger entries,

$$\tan z = \left[\frac{z}{1-}\,\frac{z^2}{3-}\,\frac{z^2}{5-}\,\ldots\right] := \cfrac{z}{1-\cfrac{z^2}{3-\cfrac{z^2}{5}}\ldots} \tag{5.7}$$

yields the second-order approximation [5.3]

$$\tan z = z\,\frac{15-z^2}{15-6z^2}. \tag{5.8}$$

For $z = \pi/4$, this is about 0.9998 (instead of 1). By contrast, the three-term power series for $\tan z$, $\tan z = z + z^3/3 + z^5/5$, makes an error that is 32 times larger. The reason for the superiority of the CF over the power-series expansion is quite obvious. As we can see from (5.8), the CF expansion makes use of polynomials not only in the numerator but also in the *denominator*. (Not making use of this degree of freedom is as if a physicist or engineer tried to approximate the behaviour of a *resonant system* by zeros of analytic functions only, rather than by zeros and *poles*: it is possible, but highly inefficient.)

Equally remarkable is the approximation of the error integral by a CF. The third-order approximation

$$\int_0^z e^{-x^2}\,dx \approx \frac{49140 + 3570z^3 + 739z^5}{49140 + 19950z^2 + 2475z^4} \tag{5.9}$$

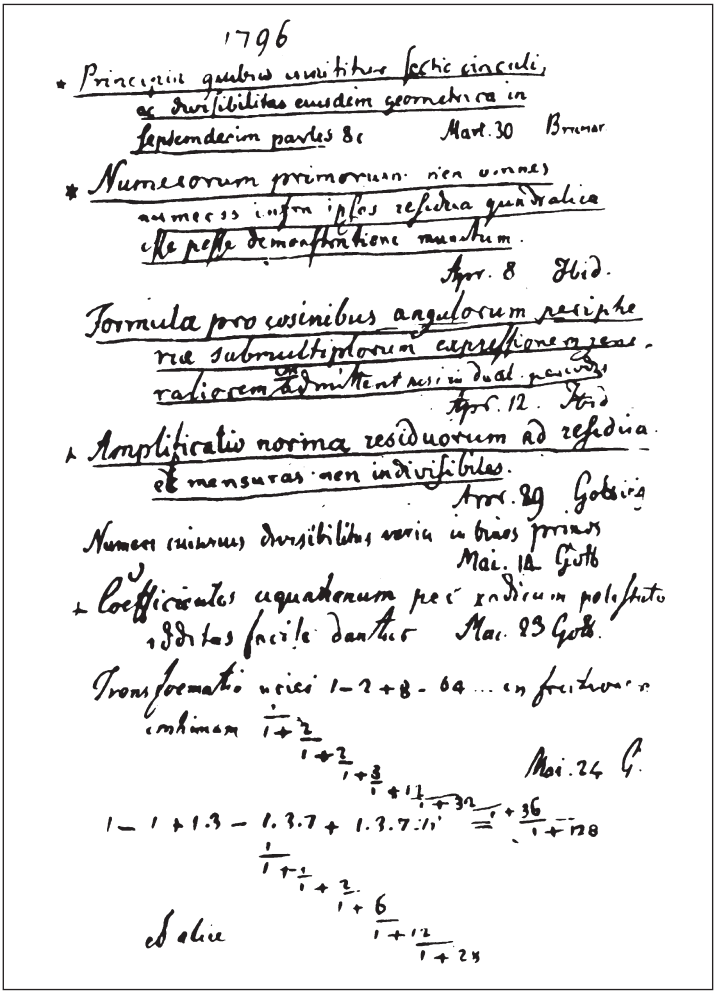
makes an error of only 1.2 % for $z = 2$, as opposed to a power series including terms up to z^9 which overshoots the true value by 110 %.

Incidentally, the fact that e has so regular a CF representation as (5.6), while π does not, does not mean that there is no regular relationship between CFs and some relative of π. In fact, the (generalized) CF expansion of $\arctan z$ for $z = 1$ leads to the following neat CF representation:

$$\frac{\pi}{4} = \left[\frac{1}{1+}\,\frac{1}{2+}\,\frac{9}{2+}\,\frac{25}{2+}\,\frac{49}{2+}\,\ldots\right].$$

Gauss, the prodigious human calculator, used CFs profusely; even on the first page of his new notebook begun on the occasion of the discovery of the regular 17-gon, CFs make their appearance (Fig. 5.1).

Why are CFs treated so negligently in our high (and low) schools? Good question, as we shall see when we study their numerous uses.



1796

- Principia quibus innititur sectio circuli, ac divisibilitas eiusdem geometrica in septemdecim partes &c. Mart. 30 Brunsv.

- Numerorum primorum non omnes numeros infra ipsos residua quadratica esse posse demonstratione munitum. Apr. 8 Ibid.

Formula pro cosinibus angulorum peripheriae submultiplorum expressionem generaliorem non admittent nisi in dual. periodos Apr. 12. Ibid

- Amplificatio normae residuorum ad residua et mensuras non indivisibiles. Apr. 29 Gottingae

Numeri cuiusvis divisibilitas varia in binos primos Mai. 14 Gott.

- Coefficientes aequationum per radicum potestatum summas facile dantur Mai. 23 Gott.

Transformatio seriei 1 − 2 + 8 − 64 ... in fractionem continuam

$$\cfrac{1}{1+\cfrac{2}{1+\cfrac{2}{1+\cfrac{8}{1+\cfrac{12}{1+\cfrac{32}{1+\cfrac{36}{1+128}}}}}}}$$

Mai. 24 G.

$$1-1+1\cdot3-1\cdot3\cdot7+1\cdot3\cdot7\cdot15 = \cfrac{1}{1+\cfrac{1}{1+\cfrac{2}{1+\cfrac{6}{1+\cfrac{12}{1+24}}}}}$$

et alia

Fig. 5.1 First page of Gauss's notebook, begun in his native city of Brunswick when he was only 18. The first entry concerns the epochal "geometrical" construction of the regular 17-gon which convinced him that he should become a mathematician. The last entry on this page, written like the three preceding ones in Göttingen, shows his early interest in continued fractions

5.2 Relations with Measure Theory

Consider the CF

$$\alpha = [0; a_1, a_2, \ldots]. \tag{5.10}$$

In 1828, Gauss established that for almost all α in the open interval $(0,1)$ the probability

$$W\left([0;a_n,a_{n+1},\ldots] < x\right) \tag{5.11}$$

tends to $\log_2(1+x)$ as n goes to infinity. Gauss also showed that the probability

$$W(a_n = k) \to \log_2\left[1+\frac{1}{k(k+2)}\right], \tag{5.12}$$

i. e., the probabilities for $a_n = 1,2,3,\ldots$ decline as 0.42, 0.17, 0.09, …, in contrast to the equal probabilities of the 10 digits for "most" decimal digits. *Khinchin* [5.4] showed in 1935 that for almost all real numbers the geometric mean

$$(a_1a_2\ldots a_n)^{\frac{1}{n}} \to \prod_{k=1}^{\infty}\left[1+\frac{1}{k(k+2)}\right]^{\log_2 k} = 2.68545\ldots, \tag{5.13}$$

and that the denominators of the approximating fractions

$$(B_n)^{\frac{1}{n}} \to \mathrm{e}^{\pi^2/12\ln 2} = 3.27582\ldots. \tag{5.14}$$

These strange constants are reminiscent of the magic numbers that describe period doubling for strange attractors in deterministic chaos. And perhaps there is more than a superficial connection here.

5.3 Periodic Continued Fractions

As with periodic decimals, we shall designate (infinite) periodic CFs like $[1;2,2,2,\ldots]$ by a bar over the period:

$$[1;\overline{2}]. \tag{5.15}$$

Incidentally, $[1;\overline{2}]$ has the value $\sqrt{2}$. In general, periodic CFs have values in which square (but no higher) roots appear.

An integer that is a nonperfect square, whose square root has a periodic, and therefore *infinite*, CF, has an irrational square root. However, there are simpler proofs that $\sqrt{2}$, say, is irrational without involving CFs. Here is a simple indirect proof: suppose $\sqrt{2}$ is rational:

$$\sqrt{2} = \frac{m}{n}, \tag{5.16}$$

where m and n are coprime:

$$(m,n) = 1, \tag{5.17}$$

i. e., the fraction for $\sqrt{2}$ has been "reduced" (meaning the numerator m and the denominator n have no common divisor). Squaring (5.16) yields

$$2n^2 = m^2. \tag{5.18}$$

Thus, m must be even:

$$m = 2k, \tag{5.19}$$

or, with (5.18),

$$2n^2 = 4k^2, \tag{5.20}$$

which implies that n is also even. Consequently,

$$(m,n) > 1, \tag{5.21}$$

contradicting (5.17). Thus, there are *no* integers n,m such that $\sqrt{2} = m/n$; in other words: $\sqrt{2}$ is irrational. Q.E.D.

An even shorter proof of the irrationality of $\sqrt{2}$ goes as follows. Suppose $\sqrt{2}$ is rational. Then there is a *least* positive integer n such that $n\sqrt{2}$ is an integer. Set $k = (\sqrt{2}-1)n$. This is a positive integer *smaller* than n, but

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

is the difference of two different integers and so is a positive integer. Contradiction: n was supposed to be the smallest positive integer such that multiplying it by $\sqrt{2}$ gives an integer! Using this kind of proof, for which s can one show that $\sqrt{s}$ is irrational? What modification(s) does the proof require?

Another exhibition example for CF expansion which we have already encountered is

$$[1;\overline{1}] = \tfrac{1}{2}\left(1+\sqrt{5}\right), \tag{5.22}$$

the famous Golden ratio g: if a distance is divided so that the ratio of its total length to the longer portion equals g, then the ratio of the longer portion to the shorter one also equals g. By comparison with (5.13) we see that the expansion coefficients in the continued fraction of g, being all 1, are 2.68... times smaller than the geometric mean over (almost) all numbers.

Golden *rectangles* have played a prominent role in the pictorial arts, and Fig. 5.2 illustrates the numerous appearances of g in a painting by Seurat. Figure 5.3 shows an infinite sequence of "golden rectangles" in which the sides have ratio g. To construct this design, lop off a square from each golden rectangle to obtain the next smaller golden rectangle.

The Golden ratio, involving as it does the number 5 – a Fermat prime – is also related, not surprisingly, to the regular pentagon, as illustrated in Fig. 5.4. It is easily verified (from Pythagoras) that $\overline{AB}/\overline{AT} = g$. Thus, the Golden ratio emerges as the ratio of the diagonal of the regular pentagon to its side.

Finally, the Golden ratio g emerged as the noblest of "noble numbers", the latter being defined by those (irrational) numbers whose continued fraction expansion

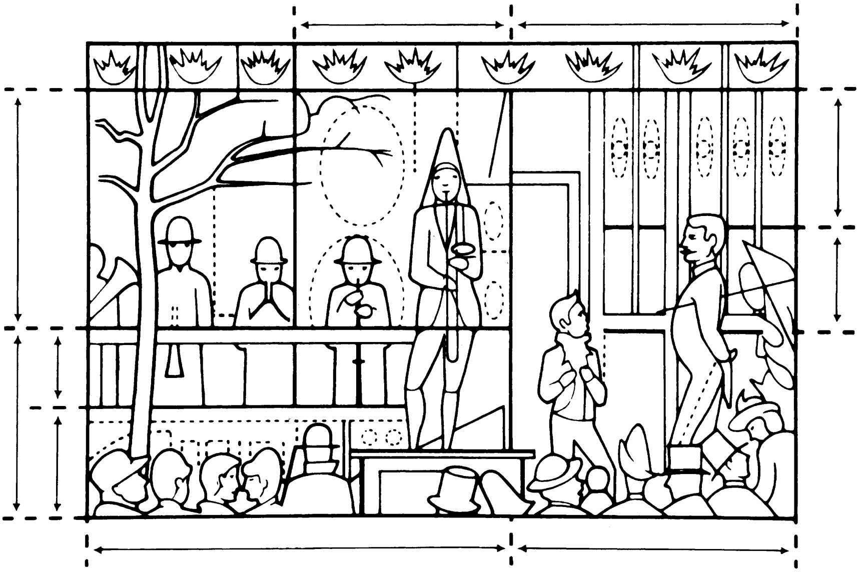


Fig. 5.2 Golden ratios in a painting by Seurat

ends in infinitely many 1's. In fact, the CF of g, see (5.22), has *only* 1's, whence also its nickname "the most irrational number" (because no irrational has a CF approximation that converges more slowly than that for g).

The designation *noble numbers* stems from the fact that in many nonlinear dynamical systems "winding numbers" (the frequency ratios of orbits in phase space) that equal noble numbers are the most resistant against the onset of chaotic motion, which is ubiquitous in nature. (Think of turbulence – or the weather, for that matter.)

Cassini's divisions in the rings of Saturn are a manifestation of what happens when, instead of noble numbers, base numbers reign: rocks and ice particles constituting the rings, whose orbital periods are in simple rational relation with the periods of other satellites of Saturn, are simply swept clean out of their paths by the resonance effects between commensurate orbital periods. In fact, the very stability of the entire solar system depends on the nobility of orbital period ratios.

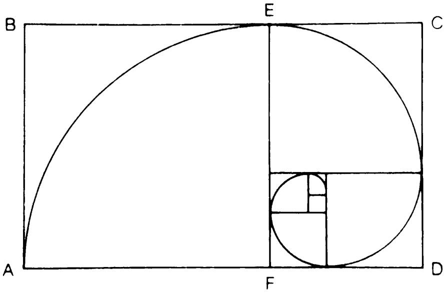


Fig. 5.3 The beginning of an infinite sequence of Golden rectangles. Some sea shells are said to use this construction

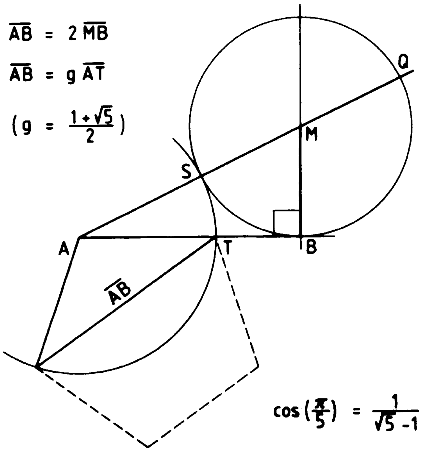

$\overline{AB} = 2\,\overline{MB}$
$\overline{AB} = g\,\overline{AT}$
$(g = \frac{1+\sqrt{5}}{2})$
Q
M
S
A
T
B
$\overline{AB}$
$\cos(\frac{\pi}{5}) = \frac{1}{\sqrt{5}-1}$

Fig. 5.4 The Golden ratio and the construction of the regular pentagon

A double pendulum in a gravitational field is a particularly transparent nonlinear system. As the nonlinearity is increased (by slowly "turning on" the gravitation), the last orbit to go chaotic is the one with a winding number equal to $1/(1+g) = [0;2,1,1,1,\ldots]$, a *very* noble number!

For physical systems a winding number $w < 1$ is often equivalent to the winding number $1-w$. Suppose $w = [0;a_1,a_2,\ldots]$, what is the CF for $1-w$? The reader will find it easy to show that

$$1-[0;a_1,a_2,\ldots] = [0;1,a_1-1,a_2,\ldots]. \tag{5.23}$$

Thus, if w is noble, so is $1-w$. Note that if $a_1-1=0$, we need to invoke the rule

$$[\ldots,a_m,0,a_{m+2},\ldots] = [\ldots,a_m+a_{m+2},\ldots],$$

which assures that the CF for $1-(1-w)$ equals that for w.

Among the most exciting nonlinear systems where CF expansions have led to deep insights are the *fractional quantization* in a two-dimensional electron gas [5.5] and "Frustrated instabilities" in active optical resonators (lasers) [5.6].

5.4 Electrical Networks and Squared Squares

One of the numerous practical fields where CFs have become entrenched – and for excellent reasons – are electrical networks.

What is the input impedance Z of the "ladder network" shown in Fig. 5.5 when the R_k are "series" impedances and the G_k are "shunt" admittances? A moment's thought will provide the answer in the form of a CF:

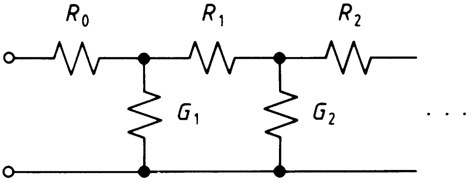


Fig. 5.5 Electrical ladder network. Input impedance is given by continued fraction $[R_0;G_1,R_1,G_2,\ldots]$

$$Z = [R_0;G_1,R_1,G_2,R_2,\ldots].$$

Here, in the most general case, the R_k and G_k are complex-valued rational functions of frequency.

If all R_k and G_k are 1-ohm resistors, the final value of Z for an infinite network, also called the characteristic impedance Z_0, will equal

$$g = \tfrac{1}{2}\left(1+\sqrt{5}\right)\text{ohm}.$$

The application of CFs to electrical networks has, in turn, led to the solution of a centuries-old teaser, the so-called Puzzle of the Squared Square, i. e., the problem of how to divide a square into unequal squares with integral sides. This problem had withstood so many attacks that a solution was widely believed impossible [5.7]. Thus, the first solution, based on network theory, created quite a stir when it appeared (Fig. 5.6).

In the meantime Littlewood has given a solution for a 112 by 112 square, which is the smallest squared square found so far.

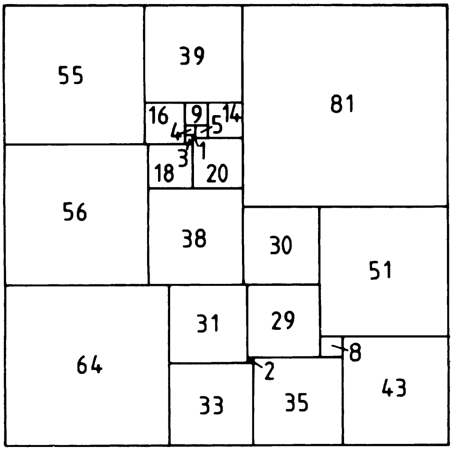


Fig. 5.6 The first squared square, a solution based on the theory of electrical networks and continued fractions (courtesy E. R. Wendorff)

5.5 Fibonacci Numbers and the Golden Ratio

Another close relative of CFs are the Fibonacci numbers [5.8], defined by the recursion

$$F_n = F_{n-1} + F_{n-2}, \quad \text{with} \quad F_0 = 0 \quad \text{and} \quad F_1 = 1, \tag{5.24}$$

which is identical with the CF recursion for the case $b_k = 1$. The first Fibonacci numbers are 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ... where each number is the sum of its two predecessors.

The ratio of two successive F_n approaches the Golden ratio $g = (1+\sqrt{5})/2$, which is easily verified in terms of CFs. From the recursion (5.24) it follows that

$$F_{n+1}/F_n = [1;\, \underbrace{1,\ldots,1}_{n-1\ 1'\text{s}}\,] \qquad (n > 1), \tag{5.25}$$

where the right-hand side of (5.25) is the approximating fraction to the Golden ratio. Equation (5.25) also implies that successive F_n are coprime:

$$(F_n, F_{n+1}) = 1, \qquad n > 0. \tag{5.26}$$

Also, the product of F_n $(n > 1)$ and its predecessor differs by ±1 from the product of their two neighbours:

$$F_{n-1}F_n - F_{n-2}F_{n+1} = (-1)^n. \tag{5.27}$$

Examples: $21 \cdot 34 = 13 \cdot 55 - 1$; $34 \cdot 55 = 21 \cdot 89 + 1$.

The reader may wish to prove his or her prowess by proving these simple statements. Equations such as (5.27) often provide quick answers to a certain class of problems such as the "banking" puzzle described in Sect. 5.11.

A simple alternative recursion for F_n is

$$F_n = 1 + \sum_{k=1}^{n-2} F_k. \tag{5.28}$$

Because of the internal structure of the F_n, which relates each F_n to its *two* predecessors, the odd-index F_n can be obtained from the even-index F_n alone:

$$F_{2n+1} = 1 + \sum_{k=1}^{n} F_{2k}. \tag{5.29}$$

It is sometimes said that there is no *direct* (nonrecursive) formula for the F_n, meaning that all predecessors F_k, with $k < n$, have to be computed first. This statement is true, however, only if we restrict ourselves to the integers. If we extend our number field to include square roots, we get the surprising direct formula, discovered by A. de Moivre in 1718 and proved ten years later by *Nicolas Bernoulli*:

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1}{2}(1+\sqrt{5})\right)^n - \left(\frac{1}{2}(1-\sqrt{5})\right)^n\right], \tag{5.30}$$

and if we admit even complex transcendental expressions, we obtain a very compact formula:

$$F_n = \mathrm{i}^{n-1}\frac{\sin(nz)}{\sin z}, \quad z = \frac{\pi}{2} + \mathrm{i}\ln\left(\frac{1+\sqrt{5}}{2}\right). \tag{5.31}$$

In (5.30), the first term grows geometrically, while the second term alternates in sign and decreases geometrically in magnitude because

$$-1 < \tfrac{1}{2}(1-\sqrt{5}) < 0.$$

In fact, the second term is so small, even for small n, that it can be replaced by rounding the first term to the nearest integer:

$$F_n = \left\lfloor \frac{1}{\sqrt{5}}\left[\frac{1}{2}(1+\sqrt{5})\right]^n + \frac{1}{2}\right\rfloor. \tag{5.32}$$

The result (5.30) is most easily obtained by solving the homogeneous *difference equation* (5.24) by the *Ansatz*

$$F_n = x^n. \tag{5.33}$$

This converts the difference equation into an algebraic equation:

$$x^2 = x + 1. \tag{5.34}$$

(This is akin to solving *differential* equations by an *exponential Ansatz*.)

The two solutions of (5.34) are

$$x_1 = \frac{1}{2}(1+\sqrt{5}) \quad \text{and} \quad x_2 = \frac{1}{2}(1-\sqrt{5}). \tag{5.35}$$

The general solution for F_n is then a linear combination:

$$F_n = ax_1^n + bx_2^n, \tag{5.36}$$

where with the initial conditions $F_0 = 0$ and $F_1 = 1$,

$$a = -b = \frac{1}{\sqrt{5}}. \tag{5.37}$$

Equations (5.35–5.37) taken together yield the desired nonrecursive formula (5.30).

Equation (5.30) can be further compacted by observing that $x_1 = -1/x_2 = g$, so that

$$\sqrt{5}F_n = g^n - (-g)^{-n}. \tag{5.38}$$

The right side of (5.38) can be converted into a trigonometric function by setting

$$t = \mathrm{i} \ln g, \tag{5.39}$$

yielding

$$\sqrt{5} F_n = 2\mathrm{i}^{n-1} \sin\left(\frac{\pi}{2} + t\right) n, \tag{5.40}$$

which is identical with (5.31) because

$$\sin\left(\frac{\pi}{2} + \mathrm{i}\ln g\right) = \frac{1}{2}\sqrt{5}, \tag{5.41}$$

a noteworthy formula in itself.

There are also numerous relations between the binomial coefficients and the Fibonacci numbers. The reader might try to prove the most elegant of these:

$$F_{n+1} = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n-k}{k}. \tag{5.42}$$

In other words, summing diagonally upward in Pascal's triangle yields the Fibonacci numbers. (Horizontal summing, of course, gives the powers of 2.)

There is also a suggestive matrix expression for the Fibonacci numbers:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}, \tag{5.43}$$

which is obviously true for $n = 1$ and is easily proved by induction. Since the determinant on the left equals -1, it follows immediately that

$$F_{n+1} F_{n-1} - F_n^2 = (-1)^n, \tag{5.44}$$

which generalizes to

$$F_{n+k} F_{m-k} - F_n F_m = (-1)^n F_{m-n-k} F_k, y \tag{5.45}$$

where any negative-index F_n are defined by the "backward" recursion

$$F_{n-1} = F_{n+1} - F_n, \tag{5.46}$$

giving

$$F_{-n} = -(-1)^n F_n. \tag{5.47}$$

Summation leads to some interesting relationships, for example

$$\sum_{k=1}^{n} F_k^2 = F_n F_{n+1},$$

which is easily proved by induction.

Products of reciprocals, too, have noteworthy sums:

$$\sum_{k=2}^{n} \frac{1}{F_{k-1}F_{k+1}} = 1 - \frac{1}{F_n F_{n+1}}.$$

Ratios of successive Fibonacci numbers have very simple continued fractions. Instead of (5.25) we may write:

$$\frac{F_n}{F_{n+1}} = [1,1,\ldots,1,2],$$

where the number of 1's equals $n-2$.

By adding Fibonacci numbers, the positive integers can be represented *uniquely*, provided each F_n $(n>1)$ is used at most once and no two adjacent F_n are ever used. Thus, in the so-called *Fibonacci number system*,

$$\begin{aligned}
3 &= 3\\
4 &= 3+1\\
5 &= 5\\
6 &= 5+1\\
7 &= 5+2\\
8 &= 8\\
9 &= 8+1\\
10 &= 8+2\\
11 &= 8+3\\
12 &= 8+3+1\\
1000 &= 987+13 \quad \text{etc.}
\end{aligned}$$

A simple algorithm for generating the Fibonacci representation of m is to find the largest F_n not exceeding m and repeat the process on the difference $m-F_n$ until this difference is zero.

The Fibonacci number system answers such questions as to where to find 0's or 1's or double 1's in the following family of binary sequences:

$$\begin{aligned}
&0\\
&1\\
&1\,0\\
&1\,0\,1\\
&1\,0\,1\,1\,0 \quad \text{etc.,}
\end{aligned}$$

where the next sequence is obtained from the one above by appending the one above *it*. (See Sect. 32.1.)

Another application of the Fibonacci number system is to nim-like games: from a pile of n chips the first player removes any number $m_1 \neq n$ of chips; then the second player takes $0 < m_2 \leq 2m_1$ chips. From then on the players alternate, never taking less than 1 or more than twice the preceding "grab". The last grabber wins.

What is the best first grab? We have to express n in the Fibonacci system:

$$n = F_{k_1} + F_{k_2} + \cdots + F_{k_r}.$$

The best initial move is then to take

$$F_{k_j} + \cdots + F_{k_r}$$

chips for some j with $1 \leq j \leq r$, provided $j = 1$ or

$$F_{k_{j-1}} > 2(F_{k_j} + \cdots + F_{k_r}).$$

Thus, for $n = 1000$, the first player should take 13 chips – the *only* lucky number in this case: only for $m_1 = 13$ can he force a victory by leaving his opponent a Fibonacci number of chips, making it impossible for the second player to force a win.

5.6 Fibonacci, Rabbits and Computers

Fibonacci numbers abound in nature. They govern the number of leaves, petals and seed grains of many plants (see Fig. 5.7 [5.9, 10]), and among the bees the number of ancestors of a drone n generations back equals F_{n+1} (Fig. 5.8).

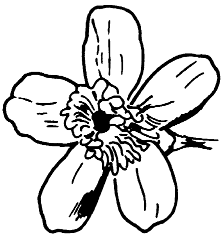
Rabbits, not to be outdone, also multiply in Fibonacci rhythm if the rules are right: offspring beget offspring every "season" except the first after birth – and they never die (Fig. 5.9). As already mentioned, this was the original Fibonacci problem [5.11] considered in 1202 by Fibonacci himself.

Leonardo da Pisa, as Fibonacci was also known, was a lone star of the first magnitude in the dark mathematical sky of the Middle Ages. He travelled widely in Arabia and, through his book *Liber Abaci*, brought the Hindu-Arabic number system and other superior methods of the East to Europe. He is portrayed in Fig. 5.10.

Fibonacci numbers also tell us in how many ways a row of n squares can be covered by squares or "dominoes" (two squares side-by-side). Obviously for $n = 2$, there are two ways: either 2 squares or 1 domino. For $n = 11$ there are 144 ways. What is the general rule?

n squares can be covered by first covering $n - 1$ squares and then adding another square or by first covering $n - 2$ squares and then adding 1 domino. Thus, calling the number of different coverings of n squares f_n, we have

$$f_n = f_{n-1} + f_{n-2},$$



Buttercups :	5 petals
Lilies and irises :	3 petals
Some delphiniums :	8 petals
Corn marigolds :	13 petals
Some asters :	21 petals
Daisies :	34, 55 and 89 petals

Fig. 5.7 Flowers have petals equal to Fibonacci numbers

i.e., the familiar recursion for the Fibonacci numbers F_n. With the initial values $f_1 = 1$ and $f_2 = 2$, we thus see that $f_n = F_{n+1}$.

Fibonacci numbers also crop up in computer science and artificial languages. Suppose there is a “language” with variables $A, B, C, \dots$ and functions of one or two variables $A(B)$ or $A(B,C)$. If we leave out the parentheses, how many ways can a string of n letters be parsed, i. e., grammatically decomposed without repeated multiplication? For a string of three letters, there are obviously two possibilities: $A \cdot B(C)$ and $A(B,C)$. In general, the answer is F_n ways, or so says Andrew Koenig of Bell Laboratories [5.12].

Another area in which Fibonacci numbers have found useful application is that of efficient sequential search algorithms for unimodal functions. Here the kth interval for searching is divided in the ratio of Fibonacci numbers F_{n-k}/F_n, so that after the $(n-1)$st step, the fraction of the original interval (or remaining uncertainty)

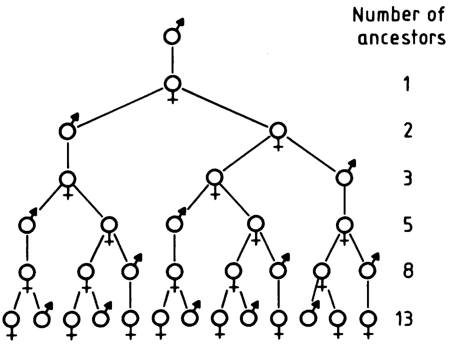


Fig. 5.8 Bees have Fibonacci-number ancestors

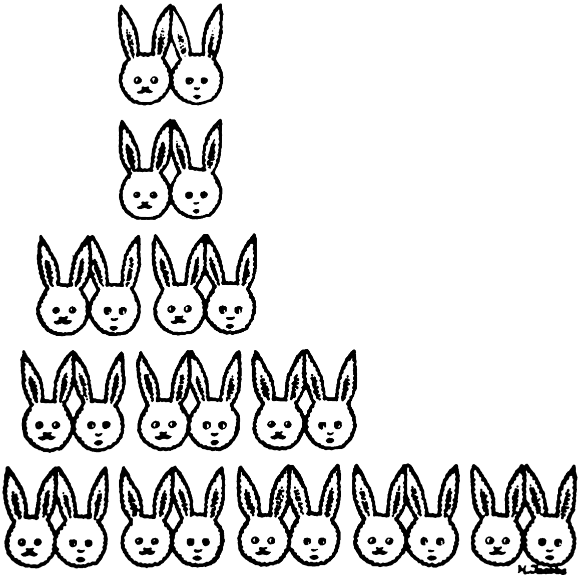


Fig. 5.9 Pairs of rabbits under certain (artifical) birth constraints, multiply according to a Fibonacci series

is $1/F_n \simeq \sqrt{5}/[(1+\sqrt{5})/2]^n$ as opposed to $(1/\sqrt{2})^n$ for "dichotomic" sequential search. After 20 steps, the precision of the Fibonacci-guided search is 6.6 times higher than the dichotomic one [5.13]. For an extensive treatment of applications of number theory in numerical analysis, see [5.14].

Fig. 5.10 Leonardo da Pisa, widely known as Fibonacci ("blockhead"), the great mathematical genius of the Middle Ages – a mathematical dark age outside the Middle East (and the Middle Kingdom!)

5.7 Fibonacci and Divisibility

It can be proved by induction that

$$F_{n+m} = F_m F_{n+1} + F_{m-1} F_n, \tag{5.48}$$

which for $m = 2$ is our fundamental recursion (5.24).

By choosing m as a multiple of n, one can further infer that F_{nk} is a multiple of F_n (and of F_k).

Example: $F_{30} = 832040$, which is divisible by $F_{15} = 610$, $F_{10} = 55$, $F_6 = 8$, $F_5 = 5$, etc.

In other words, every third F_n is even, every fourth F_n is divisible by $F_4 = 3$, every fifth F_n by $F_5 = 5$, etc. As a consequence, all F_n for composite n (except $n = 4$) are composite. However, not all F_p are prime. For example, $F_{53} = 953 \cdot 55945741$.

In 1876 Lucas showed even more, namely that, magically, the two operations "take GCD" and "compute Fibonacci" commute:

$$(F_m, F_n) = F_{(m,n)}, \tag{5.49}$$

a "magic" that can be proved with the help of Euclid's algorithm.

Example: $(F_{45}, F_{30}) = (1134903170, 832040) = 610 = F_{15}$.

One of the most interesting divisibility properties of the Fibonacci numbers is that for *each* prime p, there is an F_n such that p divides F_n. More specifically, $p \neq 5$ divides either F_{p-1} [for $p \equiv \pm 1 \pmod 5$] or F_{p+1} [for $p \equiv \pm 2 \pmod 5$]. And of course, for $p = 5$ we have $p = F_p$. In fact, *every integer* divides some F_n (and therefore infinitely many).

Also, for odd prime p,

$$F_p \equiv 5^{\frac{p-1}{2}} \pmod p \tag{5.50}$$

holds.

Many intriguing identities involve powers of Fibonacci numbers, e.g.

$$(F_n^2 + F_{n+1}^2 + F_{n+2}^2)^2 = 2[F_n^4 + F_{n+1}^4 + F_{n+2}^4]$$

5.8 Generalized Fibonacci and Lucas Numbers

By starting with initial conditions different from $F_1 = F_2 = 1$, but keeping the recursion (5.24), one obtains the generalized Fibonacci sequences, which share many properties with the Fibonacci sequences proper. The recursion for the generalized

Fibonacci sequence G_n, in terms of its initial values G_1 and G_2 and the Fibonacci numbers, is

$$G_{n+2} = G_2 F_{n+1} + G_1 F_n. \tag{5.51}$$

Of course, for $G_1 = G_2 = 1$ the original Fibonacci sequence is obtained. For some initial conditions, there is only a shift in the index, as for example with $G_1 = 1$ and $G_2 = 2$.

However, for $G_1 = 2$ and $G_2 = 1$, one obtains a different sequence:

$$2,\ 1,\ 3,\ 4,\ 7,\ 11,\ 18,\ 29, \ldots, \tag{5.52}$$

the so-called Lucas sequence [5.8]. Of course, obeying the same recursion as the Fibonacci numbers, the ratio of successive Lucas numbers also approaches the Golden ratio. However, they have "somewhat" different divisibility properties.

For example, statistically only *two out of three primes* divide some Lucas number. This result is deeper than those on the divisibility of Fibonacci numbers that we mentioned; it was observed in 1982 by Jeffrey Lagarias [5.15].

A closed form for the Lucas numbers is

$$L_n = g^n + \left(-\frac{1}{g}\right)^n, \quad \text{with} \quad g = \frac{1}{2}\left(1+\sqrt{5}\right), \tag{5.53}$$

where the second term is again alternating and geometrically decaying. This suggests the simpler formula obtained by rounding to the nearest integer:

$$L_n = \left\lfloor g^n + \frac{1}{2} \right\rfloor, \qquad n \geq 2. \tag{5.54}$$

Like the ratios of Fibonacci numbers, the ratios of successive Lucas numbers have very simple continued fractions:

$$\frac{L_n}{L_{n+1}} = [1, 1, \ldots, 1, 3],$$

where the number of 1's equals $n-2$.

Equation (5.53) leads to an intriguing law for the continued fractions of the odd powers of the Golden ratio g. With

$$g^{2n+1} = L_{2n+1} + g^{-2n-1},$$

we get

$$g^{2n+1} = \left[L_{2n+1}; \overline{L_{2n+1}}\right],$$

i. e., a periodic continued fraction of period length 1 with all partial quotients equal to the corresponding Lucas number.

What are the continued fractions of the *even* powers of g? What other irrationals have similarly simple continued fractions? What is the eighth root of the infinite continued fraction:

$$2207 - \cfrac{1}{2207 - \frac{1}{2207} - \dots}?$$

(The latter question was asked as problem B-4 in the 56th Annual William Lowell Putnam Mathematical Competition; see *Mathematics Magazine* **69**, 159 (April 1996), where a somewhat tortuous solution was given.)

Lucas numbers can be used to advantage in the calculation of large even Fibonacci numbers by using the simple relation

$$F_{2n} = F_n L_n \tag{5.55}$$

to extend the accuracy range of limited-precision (noninteger arithmetic) calculators. Similarly, we have for the even Lucas numbers

$$L_{2n} = L_n^2 - 2(-1)^n. \tag{5.56}$$

For odd-index F_n, one can use

$$F_{2n+1} = F_n^2 + F_{n+1}^2 \tag{5.57}$$

to reach higher indices.

The “decimated” Lucas sequence

$$\tilde{L}_n = L_{2^n}, \quad \text{i.e.,}$$

$$3,\ 7,\ 47,\ 2207,\ 4870847, \dots,$$

for which the simple recursion (5.56) $\tilde{L}_{n+1} = \tilde{L}_n^2 - 2$ holds, plays an important role in the primality testing of Mersenne numbers M_p with $p = 4k+3$ (see Chap. 3 for the more general test).

It is not known whether the Fibonacci or Lucas sequences contain infinitely many primes. However, straining credulity, R. L. Graham [5.16] has shown that the generalized Fibonacci sequence with

$$G_1 = 1786\ 772701\ 928802\ 632268\ 715130\ 455793$$
$$G_2 = 1059\ 683225\ 053915\ 111058\ 165141\ 686995$$

contains no primes at all!

Interesting results are obtained by introducing random signs into the Fibonacci recursion:

$$f_n = f_{n-1} \pm f_{n-2},$$

or a “growth factor” b:

$$f_n = f_{n-1} + b f_{n-2},$$

or combining both random signs and $b \neq 1$. For $b = \frac{1}{2}$ and random signs (± 1 with equal probabilities) the series doesn't grow but converges on 0. For which value of b does the series neither grow indefinitely nor decay to 0?

Another generalization of Fibonacci numbers allows more than two terms in the recursion (5.24). In this manner kth order Fibonacci numbers $F_n^{(k)}$ are defined that are the sum of the k preceding numbers with the initial conditions $F_0^{(k)} = 1$ and $F_n^{(k)} = 0$ for $n < 0$. For $k = 3$, the 3rd order Fibonacci numbers sequence starts as follows (beginning with $n = -2$): 0, 0, 1, 1, 2, 4, 7, 13, 24, 44, ...

Generalized Fibonacci numbers have recently made their appearance in an intriguing railroad switch yard problem solved by Ma Chung-Fan of the Institute of Mathematics in Beijing (H. O. Pollak, personal communication). In a freight classification yard, a train arrives with its cars in more or less random order, and before the train leaves the yard the cars must be recoupled in the order of destination. Thus, the cars with the nearest destination should be at the front of the train so they can simply be pulled off when that destination is reached, those with the second stop as destination should be next, etc. Recoupling is accomplished with the aid of k spur tracks, where usually $4 \leq k \leq 8$. The initial sequence of cars is decomposed into $\leq k$ subsequences by backing successive cars onto the various spurs, and the subsequences can then be recombined in an arbitrary order as segments in a new sequence. For any initial sequence, the desired rearrangement should be accomplished with a minimum number of times that a collection of cars is pulled from one of the spurs. These are called "pulls". For example, if 10 cars with possible destinations 1 to 7 are given in the order 6324135726, we wish to get them into the order 1223345667. On two tracks, this can be done by first backing

341526 onto the first track,
6237 onto the second track.

Pull both into the order 3415266237 (that's two pulls); then back them onto the two tracks in order

1223 onto the first track,
345667 onto the second track.

Then pull the first track's content onto the second, and pull out the whole train in the right order. Thus, it takes 4 pulls on 2 tracks to get the train together.

Define the *index* $m(\sigma)$ of the sequence $\sigma = 6324135726$ as follows: Start at the leftmost (in this case the only) 1, put down all 1's, all 2's to the right of the last 1, 3's to the right of the last 2 if you have covered *all* the 2's, etc. In this case, the first subset defined in this way is 12 (positions 5 and 9). The next subset takes the other 2 and the second 3 (positions 3 and 6); it can't get to the first 3. The next subset takes the first 3, the 4, the 5, and the second 6; the last subset is 67. Thus 6324135726 has been decomposed into 4 nondecreasing, non-overlapping, non-descending sequences

12, 23, 3456, 67. The “4” is the index $m(\sigma)$ of the given sequence σ; the general definition is analogous; it is the number of times the ordering comes to the left end of the sequence.

Now Ma showed that the minimum number of pulls in which a sequence σ can be ordered on k tracks is the integer j such that

$$F_{j-1}^{(k)} < m(\sigma) \le F_j^{(k)}.$$

Fibonacci would be delighted!

5.9 Egyptian Fractions, Inheritance and Some Unsolved Problems

A rich sheik, shortly before his death (in one of his limousines; he probably wasn’t buckled up) bought 11 identical cars, half of which he willed to his eldest daughter, one quarter to his middle daughter, and one sixth to his youngest daughter. But the problem arose how to divide the 11 cars in strict accordance with the will of the (literally) departed, without smashing any more cars. A new-car dealer offered help by lending the heirs a brand-new identical vehicle so that each daughter could now receive a whole car: the eldest 6, the middle 3 and the youngest 2. And lo and behold, after the girls (and their retinues) had driven off, one car remained for the dealer to reclaim!

The problem really solved here was to express $n/(n+1)$ as a sum of 3 *Egyptian fractions*, also called *unit fractions*:

$$\frac{n}{n+1} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}. \tag{5.58}$$

In the above story, $n = 11$ and $a = 2$, $b = 4$, $c = 6$.

Interestingly, for $n = 11$, there is another solution (and potential story) with $a = 2$, $b = 3$, $c = 12$ because *two* subsets of the divisors of 12 $(1, 2, 3, 4, 6, 12)$ add to 11. Check: $2 + 3 + 6 = 1 + 4 + 6 = 11$. Check! The inheritance problem is related to *pseudoperfect* numbers, defined as numbers equal to a sum of a subset of their divisors [5.17].

For 3 heirs and 1 borrowed car there are only 5 more possible puzzles, the number of cars being $n = 7$, 17, 19, 23 and 41 [5.18].

As opposed to continued fractions, unit fractions are of relatively little use (other than in tall tales of inheritance, perhaps). In fact, they probably set back the development of Egyptian mathematics incalculably. However, they do provide fertile ground for numerous unsolved problems in Diophantine analysis [5.17].

Of special interest are sums of unit fractions that add up to 1. Thus, for example, it is not known what is the smallest possible value of x_n, called $m(n)$, in

$$\sum_{k=1}^{n} \frac{1}{x_k} = 1, \qquad x_1 < x_2 < \cdots < x_n. \tag{5.59}$$

It is "easy" to check that $m(3) = 6$, $m(4) = 12$ and $m(12) = 30$. But what is the general law? Is $m(n) < cn$ for some constant c? Unknown!

Is $x_{k+1} - x_k \leq 2$ ever possible for all k? Erdös in [5.17] thinks not and offers ten (1971?) dollars for the solution.

Graham [5.19] was able to show that for $n > 77$, a partition of n into distinct positive integers x_k can always be found so that $\sum 1/x_k = 1$.

5.10 Farey Fractions

Another kind of fraction, the *Farey Fractions* have recently shown great usefulness in number theory [5.20].

For a fixed $n > 0$, let all the reduced fractions with nonnegative denominator $\leq n$ be arranged in increasing order of magnitude. The resulting sequence is called the Farey sequence of order n or belonging to n.

Example: for $n = 5$, in the interval [0,1] we have:

$$\frac{0}{1}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{1}{1}. \tag{5.60}$$

For other intervals, the Farey fractions are congruent modulo 1 to the Farey fractions in (5.60). In the interval $(c, c+1]$ there are exactly $\sum_{b=1}^{n} \phi(b) \approx 3n^2/\pi^2$ Farey fractions [see Chap. 8 for the definition of $\phi(n)$].

Calling two successive Farey fractions a/b and c/d, then

$$b + d \geq n + 1, \quad \text{and} \tag{5.61}$$

$$cb - ad = 1, \quad \textit{for} \quad \frac{a}{b} < \frac{c}{d}. \tag{5.62}$$

One of the outstanding properties of the Farey fractions is that given any real number x, there is always a "nearby" Farey fraction a/b belonging to n such that

$$\left| x - \frac{a}{b} \right| \leq \frac{1}{b(n+1)}. \tag{5.63}$$

Thus, if $b > n/2$ the approximating error (5.63) is bounded by $2/n^2$. This compares well with the approximate approximating error $\pi^2/12n^2$ which would result if the Farey fractions were completely uniformly distributed.

What is the spectrum (Fourier transform) of the process defined by (5.63) when x goes uniformly from 0 to 1?

The following recursion provides a convenient method of generating the Farey fractions x_i/y_i of order n: Set $x_0 = 0$, $y_0 = x_1 = 1$ and $y_1 = n$. Then

$$\begin{aligned} x_{k+2} &= \left\lfloor \frac{y_k+n}{y_{k+1}} \right\rfloor x_{k+1} - x_k, \\ y_{k+2} &= \left\lfloor \frac{y_k+n}{y_{k+1}} \right\rfloor y_{k+1} - y_k. \end{aligned} \tag{5.64}$$

The *mediant* of two fractions $\frac{a}{b}$ and $\frac{e}{f}$ is defined by

$$\mathit{mediant}\left(\frac{a}{b},\frac{e}{f}\right) := \frac{a+e}{b+f}, \tag{5.65}$$

which lies in the interval $\left(\frac{a}{b},\frac{e}{f}\right)$. Each term in a Farey series $\ldots\ \frac{a}{b}, \frac{c}{d}, \frac{e}{f}\ \ldots$ is the mediant of its two neighbours:

$$\frac{c}{d} = \frac{a+e}{b+f}. \tag{5.66}$$

In fact, the mediant of any two terms is contained in the Farey series, unless the sum of their (reduced) denominators exceeds the order n of the series.

There is also an interesting geometrical interpretation of Farey fractions in terms of point lattices, especially the fundamental point lattice consisting of all integer pairs (x,y). The Farey fractions a/b belonging to n are precisely all those lattice points $(x=a,\ y=b)$ in the triangle defined by $y=0$, $y=x$, $y=n$ which can be "seen" from the origin $x=y=0$, or, equivalently, which can "see" the origin with no other "Farey points" lying on the line of sight (see also Fig. 4.8).

Farey fractions are useful in rational approximations. Continued fractions give the excellent approximation

$$\frac{1}{\pi} \approx \frac{113}{355}.$$

But suppose we want to construct mechanical gears in the approximate ratio $\pi:1$ using fewer than 100 teeth on the smaller of the two gears. Continued fractions would then give us

$$\frac{1}{\pi} \approx \frac{7}{22},$$

but we can do better with Farey fractions. In a table published by the London Royal Society [5.21] of the Farey series of order 1025 we find near 113/355 the entries

$$\frac{99}{311}, \frac{92}{289}, \frac{85}{267}, \frac{78}{245}, \frac{71}{223} \quad \text{and} \quad \frac{64}{201},$$

any one of which is a better approximation than $7/22$.

Or suppose we want one of the gears to have 2^n teeth. We find in the table

$$\frac{1}{\pi} \approx \frac{163}{512},$$

with an error of $1.5 \cdot 10^{-4}$. This table is of course quite voluminous, having a total of 319765 entries (and a guide to locate the fraction nearest to any given number in the interval (0,1) quickly). (With $\sum_{b=1}^{n} \phi(b) \approx (3/\pi^2)n^2$, (see Chap. 8) we expect about 320000 Farey fractions of order 1025.)

Another important practical application of Farey fractions implied by (5.62) is the solution of Diophantine equations (see Chap. 7). Suppose we are looking for a solution of

$$243b - 256a = 1 \tag{5.67}$$

in integer a and b. By locating the Farey fraction just below $243/256$, namely $785/827$, we find $a = 785$ and $b = 827$. Check: $243 \cdot 827 = 200961$ and $256 \cdot 785 = 200960$. Check!

Of course, we can reduce the above solution for a modulo 243 (Chap. 6) giving the smallest positive solution $a = 56$ and $b = 59$. Thus, a table of Farey fractions of a given order n contains *all* integer solutions to equations like (5.67) with coefficients smaller than n.

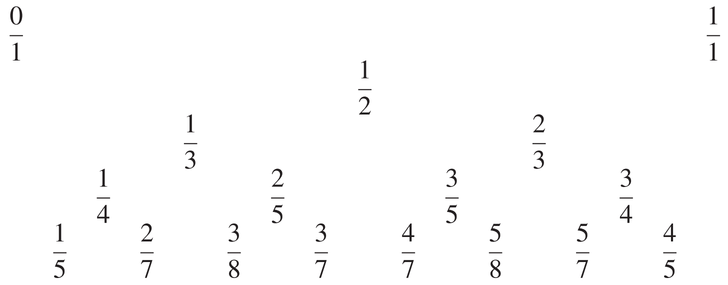
Another, and quite recent, application of Farey series is the recovery of undersampled periodic (or nearly periodic) waveforms [5.22]. If we think of “nearly periodic waveforms” as a line-scanned television film, for example, then for most pictorial scenes there are similarities between adjacent picture elements (“pixels”), between adjacent scan lines, and between successive image frames. In other words, the images and their temporal sequence carry redundant information (exception: the proverbial “snowstorm”).

Because of this redundancy, such images can, in general, be reconstructed even if the image is severely “undersampled”, i. e., if only every nth pixel $(n \gg 1)$ is preserved and the others are discarded. The main problem in the reconstruction is the close approximation of the ratio of the sampling period to one of the quasi periods in the sampled information by a rational number with a given maximal denominator – precisely the problem for which Farey fractions were invented!

5.10.1 Farey Trees

While Farey sequences have many useful applications, such as classifying the rational numbers according to the magnitudes of their denominators, they suffer from a great irregularity: the number of additional fractions in going from Farey sequences of order $n-1$ to those of order n equals the highly fluctuating Euler function $\phi(n)$. A much more regular order is infused into the rational numbers by *Farey trees*, in which the number of fractions added with each generation is simply a power of 2.

Starting with two fractions, we can construct a Farey tree by repeatedly taking the mediants of all numerically adjacent fractions. For the interval $[0,1]$, we start with $0/1$ and $1/1$ as the initial fractions, or “seeds”. The first five generations of the Farey tree then appear as follows:



Each rational number between 0 and 1 occurs exactly once somewhere in the infinite Farey tree.

The location of each fraction within the tree can be specified by a binary address, in which 0 stands for moving to the left in going from level n to level $n+1$ and 1 stands for moving to the right. Thus, starting at $1/2$, the rational number $3/7$ has the binary address 011. The complement of $3/7$ with respect to 1 (i. e., $4/7$) has the complementary binary address: 100. This binary code for the rational numbers is useful in describing frequency locking in coupled oscillators.

Note that any two *numerically* adjacent fractions of the tree are unimodular. For example, for $4/7$ and $1/2$, we get $2 \cdot 4 - 1 \cdot 7 = 1$.

Some properties of the Farey tree are particularly easy to comprehend in terms of continued fractions w in the interval [0, 1]:

$$w = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3} \cdots}}$$

or more conveniently $w = [a_1, a_2, a_3, \ldots]$, where the "partial quotients" a_k are positive integers. Irrational w have nonterminating continued fractions. For quadratic irrational numbers the a_k will (eventually) repeat periodically. For example, $1/\sqrt{3} = [1, 1, 2, 1, 2, 1, 2, \ldots] = [1, \overline{1, 2}]$ is preperiodic and has a period of length 2; $1/\sqrt{17} = [\overline{8}]$ has period length 1 and $1/\sqrt{61}$ has period length 11. (It is tantalizing that no simple rule is known that predicts period lengths in general.) Interestingly, for any fraction on level n of the Farey tree, the sum over all its a_k equals n:

$$\sum_k a_k = n \qquad n = 2, 3, 4, \ldots.$$

We leave it to the reader to prove this equation (by a simple combinatorial argument, for example).

There is also a direct way of calculating, from each fraction on level $n-1$, its two neighbours or direct descendants on level n. First write the original fraction as a continued fraction in two different ways, which is always possible by splitting off a 1 from the final a_k. Thus, for example, $2/5 = [2,2] = [2,1,1]$. Then add 1 to the last term of each continued fraction; this yields $[2,3] = 3/7$ and $[2,1,2] = 3/8$, which are indeed the two descendants of $2/5$.

Conversely, the close parent of any fraction (the one on the adjacent level) is found by subtracting 1 from its last term (in the form where the last term exceeds 1, because $a_k = 0$ is an illegal entry in a continued fraction). The other (distant) parent is found by simply *omitting* the last term. Thus, the two parents of $3/7 = [2,3]$ are the close parent $[2,2] = 2/5$ and the distant parent $[2] = 1/2$. (But which parent is greater, in general – the close or the distant one? And how are mediants calculated using only continued fraction?)

Interestingly, if we zigzag down the Farey tree from its upper right ($1/1 \to 1/2 \to 2/3 \to 3/5 \to 5/8$, and so on), we land on fractions whose numerators and denominators are given by the Fibonacci numbers F_n, defined by $F_n = F_{n-1} + F_{n-2}$; $F_0 = 0$, $F_1 = 1$. In fact, on the nth zig or zag, starting at $1/1$, we reach the fraction F_{n+1}/F_{n+2}, which approaches the golden mean $\gamma = (\sqrt{5} - 1)/2 = 0.618\ldots$ as $n \to \infty$. (Starting with $0/1$ we land on the fractions F_n/F_{n+2}, which converge on $\gamma^2 = 1 - \gamma$.) The binary address of γ in the Farey tree is 101010....

As already noted, the continued fraction expansions of the ratios F_n/F_{n+1} have a particularly simple form:

$$\frac{F_n}{F_{n+1}} = [1,1,\ldots,1] \quad (\text{with } n\ 1'\text{s}).$$

Obviously, continued fractions with small a_k converge relatively slowly to their final values, and continued fractions with only 1's are the most slowly converging of all. Since

$$\gamma = \lim_{n\to\infty} \frac{F_n}{F_{n+1}} = [1,1,1,\ldots] = [\overline{1}],$$

where the bar over the 1 indicates infinitely many 1's, the golden mean γ has the most slowly converging continued fraction expansion of all irrational numbers. The golden mean γ is therefore sometimes called (by physicists and their ilk) "the most irrational of all irrational numbers" – a property of γ with momentous consequences in a wide selection of problems in nonlinear physics, from the double swing to the three-body problem.

Roughly speaking, if the frequency ratio of two coupled oscillators is a rational number P/Q, then the coupling between the driving force and the "slaved" oscillator is particularly effective because of a kind of a resonance: every Q cycles of the driver, the same physical situation prevails so that energy transfer effects have a chance to build up in a resonancelike manner. This resonance effect is particularly strong if Q is a *small* integer. This is precisely what happened with our moon: resonant energy transfer between the Moon and the Earth by tidal forces slowed the Moon's spinning motion until the spin period around its own axis locked into the 28-day cycle of its revolution around the Earth. As a consequence the Moon always shows us the same face, although it wiggles ("librates") a little.

Similarly, the frequency of Mercury's spin has locked into its orbital frequency at the rational number 3/2. As a consequence, one day on Mercury lasts two Mercury *years*. (And one day – in the distant future, one hopes – something strange like that may happen to Mother Earth!)

The rings of Saturn, or rather the gaps between them, are another consequence of this resonance mechanism. The orbital periods of any material (flocks of ice and rocks) in these gaps would be in a rational resonance with some periodic force (such as the gravitational pull from one of Saturn's "shepherding" moons). As a consequence, even relatively weak forces have a cumulatively significant effect over long time intervals, accelerating any material out of the gaps.

For rational frequency ratios with large denominators Q, such a resonance effect would, of course, be relatively weak, and for *irrational* frequency ratios, resonance would be weaker still or absent.

For strong enough coupling, however, even irrational frequency ratios might be affected. But there is always one irrational frequency ratio that would be least disturbed: the golden mean, because, in a rational approximation to within a certain accuracy, it requires the largest denominators Q. This property is also reflected in the Farey tree: on each level n the two fractions with the largest denominators are the ones that equal F_{n-1}/F_{n+1} and F_n/F_{n+1}, which for $n \to \infty$ approach $\gamma^2 = 0.382\ldots$ and $\gamma = 0.618\ldots$, respectively. (Conversely, the fractions with the smallest Q on a given level of the Farey tree are from the harmonic series $1/Q$ and $1 - 1/Q$.)

Another way to demonstrate the unique position of the golden mean among all the irrational numbers is based on the theory of rational approximation, an important part of number theory. For a good rational approximation, one expands an irrational number w into a continued fraction and terminates it after n terms to yield a rational number $[a_1, a_2, \ldots, a_n] = p_n/q_n$. This rational approximation to w is in fact the best for a given maximum denominator q_n. For example, for $w = 1/\pi = [3, 7, 15, 1, 293, \ldots]$ and $n = 2$, we get $p_n/q_n = 7/22$, and there is no closer approximation to $1/\pi$ with a denominator smaller than 22.

Now, even with such an optimal approximation as afforded by continued fractions, the differences for the golden mean γ

$$\left|\gamma - \frac{p_n}{q_n}\right|$$

exceed c/q_n^2 (where c is a constant that is smaller than but arbitrarily close to $1/\sqrt{5}$) for *all* values of n above some n_0. And this is true only for the golden mean γ and the "noble numbers" (defined as irrational numbers whose continued fractions end in all 1's). Thus, in this precise sense, the golden mean (and the noble numbers) keep a greater distance from the rational numbers than does any other irrational number. Small wonder that the golden mean plays such an important role in synchronization problems.

5.10.2 Locked Pallas

On 5 May, 1812, Gauss communicated to his friend Friedrich Wilhelm Bessel (1784–1846) a strange discovery he had just made: the periods of revolution around the sun of Jupiter and Pallas are exactly in the ratio 18/7. Gauss asked Bessel not to tell anyone else because he was afraid of being accused of *Zahlen-Mystik* (number mysticism), then, as now, rampant in astronomy. Instead (to preserve priority) he "published" his result in the *Göttingschen Gelehrten Anzeigen* (No. 67, 25 April 1812) as a cryptogram, a string of 16 0's and 1's:

$$1\,1\,1\,1\,0\,0\,0\,1\,0\,0\,1\,0\,1\,0\,0\,1.$$

He would divulge the key to unlock his discovery at an appropriate time, Gauss added. But he never did "divulge" and the great Gauss was rather peeved (although he did not perish) when later other astronomers did publish and claimed priority.

Knowing the encrypted message (18/7) my student Inga Holube showed the way to read Gauss's cryptogram. By appropriate segmentation,

$$1\,1\,1 \quad 1\,0\,0\,0 \quad 1\,0\,0\,1\,0 \quad 1\,0\,0\,1,$$

and by interpreting the resulting snippets as binary numbers she obtained the numbers

$$7 \qquad 8 \qquad 18 \qquad 9.$$

Thus Gauss was probably saying that the period 7 belongs to the 8th planet (as Pallas was considered at the time) and the period 18 belongs to the 9th planet (Jupiter) [5.23].

Incidentally, 7 and 18 are not "any old" numbers; they are close *Lucas* numbers (see Sect. 5.8). The series of Lucas numbers are constructed like the Fibonacci series: each number is the sum of its two predecessors, but the series begins with 1 and 3 and continues 4, 7, 11, 18, ... As mentioned before (Sect. 5.8), the ratios of successive Lucas numbers have simple continued fractions and approach the golden mean. Like the Fibonacci numbers they play an important role in nonlinear dynamics and synchronization problems.

What ratio of periods would the orbits of Pallas and Jupiter lock into if the gravitational coupling between them was increased (for example by increasing the mass of Jupiter). The real experiment cannot of course be done but such questions can conveniently be studied by computer simulation.

Analytically, too, we can venture a guess. For increased coupling strengths, the period ratios are typically represented by fractions lying higher in the Farey tree (see Sect. 5.10.1), such as one of their "parent" fractions. For $7/18 = [2,1,1,3]$, the parent fractions are $[2,1,1,2] = 5/13$ and $[2,1,1] = 2/5$. Thus, the period ratio of Jupiter and Pallas might lock into $5/13$ or $2/5$.

For an even stronger gravitational coupling we have to consult the "grandparents" of $7/18$, i. e., the parents of $5/13$ and $2/5$. With our familiar algorithm for Farey families this yields $3/8$ and $2/5$ and $1/3$ and $1/2$, respectively. Curiously, the ratio $2/5$ is both parent and *grand*parent to $7/18$. Unusual relationships between humans but perhaps par for the course among the Greek gods – especially if Jupiter is part of the party.

5.11 Fibonacci and the Problem of Bank Deposits

There is an interesting family of problems, appearing in many guises, to which Fibonacci numbers provide a quick solution. Suppose Bill, a wealthy Texan chemist, opens a new bank account with x_1 dollars. The next (business) day he deposits x_2 dollars, both integer dollar amounts. Thereafter his daily deposits are always the sum of the previous two deposits. On the nth day Bill is known to have deposited x_n dollars. What were the original deposits?

A solution of this problem, posed by L. A. Monzert (cf. Martin Gardner [5.24]), argues that, for sufficiently large n, successive deposits should be in the golden ratio. This reasoning permits one to find the $(n-1)$th deposit x_{n-1} and, together with x_n, by *backward* recursion, all prior ones.

However, with the knowledge gained in this chapter, we can find a *direct* answer to this financial problem, one that is valid even for small n. Since the recursion rule for x_n is like that for the Fibonacci numbers, the x_n must be expressible as a linear combination of Fibonacci numbers. In fact, two such terms suffice:

$$x_n = aF_{n+k} + bF_{n+m}. \tag{5.68}$$

With $F_0 = 0$ and $F_{-1} = F_1 = 1$, the initial conditions are satisfied by

$$x_n = x_1 F_{n-2} + x_2 F_{n-1}. \tag{5.69}$$

Now, because of (5.27), an integer solution to (5.69) is given by

$$x_1 = (-1)^n x_n F_{n-3}, \tag{5.70}$$

$$x_2 = -(-1)^n x_n F_{n-4}.$$

Solutions (not only to mathematical problems) become that simple if one knows and *uses* the proper relations!

However, we are not quite done yet. According to (5.70) one or the other initial deposit is negative; but we want all deposits to be positive of course. Looking at (5.69) we notice that we can add to x_1 any multiple of F_{n-1} as long as we subtract the same multiple of F_{n-2} from x_2. Thus, the general solution is

$$\begin{aligned} x_1 &= (-1)^n x_n F_{n-3} + mF_{n-1}, \\ x_2 &= -(-1)^n x_n F_{n-4} - mF_{n-2}. \end{aligned} \tag{5.71}$$

We can now ask for what values (if any!) of m *both* x_1 and x_2 are positive. Or, perhaps, for what value of m x_2 is positive and as small as possible. The answer, which leads to the longest chain of deposits to reach a given x_n, is

$$m = \left\lfloor -(-1)^n x_n \frac{F_{n-4}}{F_{n-2}} \right\rfloor. \tag{5.72}$$

For $x_{20} = 1\,000\,000$ (dollars), (5.72) yields $m = -381966$ and (5.71) gives $x_1 = 154$ and $x_2 = 144$.

lf we had asked that the twenty-*first* deposit be one million dollars, (5.72) would have given the same absolute value of m, and with (5.71), $x_2 = 154$ and $x_1 = -10$. In other words, we would have posed an illicit problem.

Here we have, unwittingly, solved a Diophantine equation, of which more in Chaps. 6 and 7.

5.12 Error-Free Computing

One of the overriding problems in contemporary computing is the accumulation of *rounding errors* to such a degree as to make the final result all but useless. This is particularly true if results depend on the input data in a *discontinuous* manner. Think of matrix inversion.

The inverse of the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1+\varepsilon \end{pmatrix} \tag{5.73}$$

for $\varepsilon \neq 0$ equals

$$A^{-1} = \begin{pmatrix} 1+1/\varepsilon & -1/\varepsilon \\ -1/\varepsilon & 1/\varepsilon \end{pmatrix}. \tag{5.74}$$

An important generalization of a matrix inverse, applicable also to singular matrices, is the *Moore-Penrose inverse* A^+ [5.25]. For nonsingular matrices, the Moore-Penrose inverse equals the ordinary inverse:

$$A^+ = A^{-1}. \tag{5.75}$$

As $\varepsilon \to 0$ in (5.73), the matrix A becomes singular and the Moore-Penrose inverse no longer equals A^{-1} but can be shown to be

$$A^+ = \tfrac{1}{4}A. \tag{5.76}$$

In other words, as $\varepsilon \to 0$, the elements of A^+ become larger and larger only to drop discontinuously to $1/4$ for $\varepsilon = 0$.

Examples of this kind of sensitivity to small errors abound in numerical analysis. For many computations the only legal results are integers, for example, the coefficients in chemical reaction equations. If the computation gives noninteger coefficients, their values are rounded to near integers, often suggesting impossible chemical reactions.

In some applications of this kind, double-precision arithmetic is a convenient remedy. (The author once had to invoke double precision in a very early (ca. 1959) digital filter, designed to simulate concert hall reverberation, because the sound would refuse to die away when the music stopped.) In other situations, number-theoretic transforms can be used, of which the Hadamard transform (see Chap. 18) is only one example.

However, quite general methods for error-free computing have become available in the recent past, and it is on these that we shall focus attention in this section.

Specifically, we want to sketch a strategy for computing that will not introduce *any* rounding errors whatsoever, no matter how long or complex the computation. How is this possible? Of the four basic mathematical operations, three (addition, subtraction and multiplication) are harmless: if we start with integers, we stay with integers – no rounding problems there. But *division* is a real bugbear. Computers can never represent the fraction $1/7$, for example, in the binary (or decimal system) without error, no matter how many digits are allowed. If we could only do away with division in our computations! Surprisingly, this is in fact possible, as we shall see.

Of course, computers cannot deal with continuous data – both input and output are, by necessity, rational numbers, and the rational numbers we select here to represent both input data and final results are *Farey fractions* of a given order N (see Sect. 5.10). Once we have chosen a large enough value of N to describe adequately the input data of a problem and all of the answers to that problem, then within this precision, *no* errors will be generated or accumulated.

In this application, we shall generalize our definition of Farey fractions a/b of order N, where a and b are coprime, to include negative and improper fractions:

$$0 \leq |a| \leq N, \quad 0 < |b| \leq N. \tag{5.77}$$

The error-free strategy, in its simplest form, [5.25] then proceeds as follows. A prime modulus m is selected such that

$$m \geq 2N^2 + 1, \tag{5.78}$$

and each Farey fraction a/b, with $(b,m) = 1$, is mapped into an *integer* k modulo m:

$$k = \langle ab^{-1} \rangle_m, \tag{5.79}$$

where the *integer* b^{-1} is the inverse of b modulo m and the acute brackets signify the smallest nonnegative remainder modulo m (see Sect. 1.5). It is in this manner that we have abolished division! The inverse b^{-1} can be calculated by solving the Diophantine equation

$$bx + my = 1, \tag{5.80}$$

using the Euclidean algorithm (see Sect. 7.2). The desired inverse b^{-1} is then congruent modulo m to a solution x of (5.80).

After this conversion to integers, all calculations are performed in the integers modulo m. For example, for $N = 3$ and $m = 19$, and with $3^{-1} = 13$, the fraction $2/3$ is mapped into $26 \equiv 7$, and the fraction $-1/3$ is mapped into $-13 \equiv 6$. The operation $(2/3)+(-1/3)$ is then performed as $7+6=13$, which is mapped back into $1/3$, the correct answer.

It is essential for the practical application of this method that fast algorithms be available for both the forward and backward mappings. Such algorithms, based on the Euclidean algorithm, were described by *Gregory* and *Krishnamurthy* [5.25], thereby reclaiming error-free computing from the land of pious promise for the real world.

Sometimes *intermediate* results may be in error, but with no consequence for the final result, as long as it is an order-N Farey fraction. For example, for $m=19$, $2^{-1}=10$, so that $1/2$ maps into 10, and $(1/3)-(1/2)$ maps into $13-10=3$, which is the image of 3 – an erroneous result because $-1/6$ is the correct answer! But 3 is still useful as an intermediate result. For example, multiplying 3 by 2 produces 6, which is the image of $-1/3$, the correct result.

For the large values of N that are needed in practical applications, the prime m has to be correspondingly large. Since calculating modulo very large primes is not very convenient, a multiple-modulus residue (or Chinese remainder) system, see Chap. 16, is often adopted. For example, for $N=4$, the smallest prime not smaller than $2N^2+1=33$ is 37. Instead, one can calculate with the residues modulo the *two* primes $m_1=5$ and $m_2=7$, whose product $m=m_1\cdot m_2=35$ exceeds $2N^2+1=33$. Such calculations, described in Chap. 17, are much more efficient than the corresponding operation in single-modulus systems, the savings factor being proportional to $m/\sum m_i$. For decomposition of large m into many small prime factors, the savings can be so large as to make many otherwise impossible calculations feasible.

Another preferred number system for carrying out the calculations is based on the integers modulo a prime *power*: $m=p^r$. For example, for $N=17$, the modulus m must exceed 578 and a convenient choice would be $p=5$ and $r=4$, so that m equals $5^4=625$. There is only one problem with this approach: all fractions whose denominators contain the factor 5 cannot be represented because 5 has no inverse modulo 625. However, an ingenious application of p-adic algebra and finite-length *Hensel codes* has solved the problem and looks like the wave of the future in error-free computing. We shall attempt a brief description; for details and practical applications the reader is referred to [5.25].

Essentially, what the Hensel codes do is to remove bothersome factors p in the denominators, so that the "purified" fractions do have unique inverses.

For *integers*, the p-adic Hensel codes are simply obtained by "mirroring" the p-ary expansion. With the 5-ary expansion of 14, for example,

$$14=2\cdot 5^1+4,$$

the Hensel code for 14 becomes

$$H(5,4,14)=.4200. \tag{5.81}$$

In general, $H(p,r,\alpha)$ is the Hensel code of α to the (prime) base p, having precisely r digits.

A *fraction* a/b whose denominator b does *not* contain the factor p is converted to an integer modulo p^r, which is then expressed as a Hensel code. For example, with $p^r = 5^4 = 625$, we get

$$\tfrac{1}{16} \mathrel{\overset{\wedge}{=}} \left\langle 16^{-1} \right\rangle_{625} = 586 = 4 \cdot 5^3 + 3 \cdot 5^2 + 2 \cdot 5 + 1,$$

or in Hensel code:

$$H\left(5, 4, \tfrac{1}{16}\right) = .1234. \tag{5.82}$$

Similarly, with $\langle 3/16 \rangle_{625} = 508 = 4 \cdot 5^3 + 1 \cdot 5 + 3$, becomes

$$H\left(5, 4, \tfrac{3}{16}\right) = .3104. \tag{5.83}$$

Of course, the Hensel code for $3/16$ can be obtained directly by multiplying (5.82) with the code for 3:

$$H(5, 4, 3) = .3000, \tag{5.84}$$

where the multiplication proceeds from *left to right*. (Remember, Hensel codes are based on a *mirrored* p-ray notation.) Thus, $H(5, 4, 1/16) \times H(5, 4, 3)$ equals

$$\begin{array}{rr} & .1234 \\ \times & .3000 \\ \cline{2-2} & .3142 \\ \text{carries} & 112 \\ \cline{2-2} = & .3104 \end{array}$$

which agrees with $H(5, 4, 3/16)$, see (5.83). Note that any carries beyond four digits (the digit 2 in the above example) are simply dropped. It is ironic that such a “slipshod” code is the basis of *error-free* computation!

If the numerator contains powers of p, the corresponding Hensel code is simply right-shifted, always maintaining precisely r digits. For example

$$H\left(5, 4, \tfrac{5}{16}\right) = 0.0123. \tag{5.85}$$

Powers of p in the denominator are represented by a *left*-shift. Thus, with $H(5, 4, 1/3) = .2313$,

$$H\left(5, 4, \tfrac{1}{15}\right) = 2.313. \tag{5.86}$$

To expand the range of the Hensel codes to arbitrary powers of p in the denominator or numerator, a floating-point notion, $\hat{H}(p, r, \alpha)$, is introduced. For example,

$$\hat{H}\left(5, 4, \tfrac{1}{15}\right) = (.2313, -1), \quad \text{and} \tag{5.87}$$

$$\hat{H}(5,4,375) = (.3000,3), \tag{5.88}$$

where the first number on the right is the mantissa and the second number the exponent.

When multiplying floating-point Hensel codes, their mantissas are multiplied and their exponents are added. For example,

$$\tfrac{1}{3} \times \tfrac{6}{5} \overset{\wedge}{=} (.2313,0) \times (.1100,-1) = (.2000,-1),$$

and with

$$.2313 \times .1100 = .2000,$$

we obtain

$$(.2313,0) \times (.1100,-1) = (.2000,-1),$$

which corresponds to $2/5$, the correct answer. (Remember, all operations proceed from left to right and Hensel code .1000 corresponds to 1 and not $1/10$.)

Strange and artificial as they are, Hensel codes perform numerical stunts and never slip a single digit.

Part III
Congruences and the Like

Chapter 6
Linear Congruences

Suppose a certain airline is consistently 25 hours late in departure and arrival (this has happened, but no names will be mentioned) while another one, flying the same route, is only 2 to 3 hours late. If you were in a hurry, which airline would you fly – food, lack of leg room and all else being equal? Obviously, being 25 hours late is as good (or bad) as being only 1 hour late. In other words, in a *daily recurring* event an extra day, or even several, makes no difference. The mathematics that deals with this kind of situation is called *modular arithmetic*, because only remainders *modulo* a given integer matter.

Another application of modular arithmetic occurs in wave interference phenomena such as ripples on a lake or patterns of light and dark in a hologram. In all these cases a path difference of, say, half a wavelength is indistinguishable from a path difference of one and a half or two and a half, etc., wavelengths.

And of course, there are many applications in mathematics proper. For example, few people care what n^{560} is, but the remainder of n^{560} when divided by 561 for all n coprime to 561 is a question of some actuality. (As it happens, all these remainders are 1 – a terrible thing to happen, as we shall see.) But first we have to know the ground rules of modular arithmetic, so that n^{560} for, say, $n = 500$, a 1512-digit monster of a number, cannot frighten us.

6.1 Residues

When c divided by m leaves the remainder b (not necessarily positive) we write (following Gauss)

$$c \equiv b \pmod{m}; \tag{6.1}$$

read: c is congruent b modulo m.

More generally, we define the above congruence as meaning

$$m \,|\, (c - b); \tag{6.2}$$

read: m divides c minus b (without remainder).

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_6, © Springer-Verlag Berlin Heidelberg 2009

Example: $16 \equiv -2(\bmod 9)$ implies $9|(16+2)$.

Together with c, all

$$b = mq + c; \qquad q = \text{integer} \tag{6.3}$$

belong to the same *residue class* modulo m [6.1].

A *complete residue system* modulo m consists of m integers, one representative each from each residue class. The most common residue systems are the *least nonnegative* residue system modulo m, consisting of the integers $0, 1, 2, \ldots, m-1$, and the *least absolute* residue system, consisting of the integers $0, \pm 1, \pm 2, \ldots, \pm(m-1)/2$ (for odd m).

For many purposes, one calculates with the congruence sign for a given modulus as if it were an equal sign.

Addition:

$$\textit{Example:} \qquad \begin{array}{l} 13 \equiv 4 \ (\bmod 9) \\ \underline{16 \equiv -2 \ (\bmod 9)} \\ 29 \equiv 2 \ (\bmod 9) \qquad \text{Check!} \end{array} \tag{6.4}$$

Multiplication of the two upper congruences results in

$$208 = -8 \ (\bmod 9) \qquad \text{Check!} \tag{6.5}$$

The congruence

$$c \equiv 5 \ (\bmod m) \tag{6.6}$$

can be “cancelled” by the GCD of c, b and m. With $(c, b, m) = d$, we may write

$$\frac{c}{d} \equiv \frac{b}{d} \left(\bmod \frac{m}{d}\right). \tag{6.7}$$

Example: $28 \equiv 4 \ (\bmod 6)$ can be converted to $14 \equiv 2 \ (\bmod 3)$.

Another useful rule is the following. If

$$mc \equiv mb \ (\bmod n) \tag{6.8}$$

and the GCD $(m, n) = d$, then

$$c \equiv b \left(\bmod \frac{n}{d}\right). \tag{6.9}$$

Example: $28 \equiv 4 \ (\bmod 6)$ implies $7 \equiv 1 \ (\bmod 3)$.

Among the many useful applications of linear congruences is the ancient error-detecting algorithm sailing under the name of “casting out 9’s” [6.2]. If we add two decimal numbers column by column, then if in any one column the sum exceeds

9, we reduce the result modulo 10 and add 1 (or 2 or 3, etc.) to the next column. Thus, in terms of the sum of the decimal digits, we have added 1 (or 2 or 3, etc.) and subtracted 10 (or 20 or 30, etc.). We have therefore changed the sum of the digits by a multiple of 9; in other words, the sum of the digits has not changed modulo 9.

Example:

$$\begin{array}{rll} 86 & \text{sum of digits} = 14, & \text{sum of sum of digits} = 5 \\ +\ \ 57 & \text{sum of digits} = 12, & \text{sum of sum of digits} = 3 \\ \hline = 143 & \text{sum of digits} = 8, & \end{array}$$

Check: $14 + 12 = 26 \equiv 8 (\operatorname{mod} 9)$. Check! Of course in the check we can apply the same rule and consider sums of sums of digits: Check: $5 + 3 = 8 \equiv 8 (\operatorname{mod} 9)$. Check!

The same rule also holds for multiplication.

Example:

$$\begin{array}{rl} 15 & \text{sum of digits} = 6, \\ \times \quad 17 & \text{sum of digits} = 8, \\ \hline = \ \ 255 & \text{sum of digits} = 12, \text{ sum of sum of digits} = 3\,. \end{array}$$

Check: $6 \times 8 = 48$, sum of digits $= 12$, sum of sum digits $= 3$. Check!

The reason why sums of digits when multiplied give the same result modulo 9 as the numbers themselves is that, trivially, any power of 10 is congruent 1 modulo 9:

$$10 \equiv 1 \pmod 9 \quad \text{and therefore} \quad 10^k \equiv 1 \pmod 9 \quad \text{for any} \quad k \geq 0.$$

The only problem with this ancient error-detecting code is that it can fail to signal an error. In fact, for random errors, about 10 % of the errors go undetected. Fortunately, though, the casting-out-9's is not restricted to the decimal system; it works for any base b, casting out $(b-1)$'s. Specifically, it also works for a much neglected (and very simple) number system: the base-100 or "hectic" system. One of the advantages of the hectic system is that it needs no new digits. For example, the year of Gauss's birth in hectic notation looks like this:

$$17\ 77, \quad \text{sum of digits} = 94,$$

and that of his death

$$18\ 55, \quad \text{sum of digits} = 73,$$

and the difference of these two hectic numbers (his age when he died) is

$$00\ 78, \quad \text{sum of digits} = 78.$$

Check: $73 - 94 = -21 \equiv 78 \pmod{99}$. Check!

This example is perhaps too simple, but with the hectic error-detecting algorithm the undetected error rate has dropped to about 1 %.

A simple rule exists also for divisibility by 11, used in the International Standard Book Numbers (ISBN). It follows from $10 \equiv -1 \bmod 11$ and $100 \equiv 1 \bmod 11$ that divisibility by 11 of an integer and its digital sum, taken with *alternating* signs, are equivalent. Thus 517, for example, is immediately seen to be divisible by 11 because $5 - 1 + 7 = 11$. If the result of this operation is itself a large number, the operation can of course be repeated until manageable numbers, like 0 or 11, are reached.

Divisibility checks also exist for 7, 13, 17, 19 etc., but they are not as efficient. Thus, to check n for a factor 7, one writes

$$n = 10a + b$$

and

$$m = a - 2b\,.$$

Now if $m \equiv 0 \bmod 7$, it follows that $a \equiv 2b$ and $10a \equiv 20b \equiv 6b$. Thus $n \equiv 6b + b \equiv 7b \equiv 0 \bmod 7$. Of course the rule can be iterated.

The same approach tells us that, for divisibility by 13, we should check $a + 4b$. Thus, for example, for $n = 91$, we have $9 + 4 = 13$, implying that 91 is divisible by 13.

Similarly, for $n = 17$, we should look at $a - 5b$ and for $n = 19$, the test is $a + 2b$.

6.2 Some Simple Fields

Complete residue systems modulo a prime form a *field*, i. e., a set of numbers for which addition, subtraction, multiplication and division (except by 0) are defined and for which the usual commutative, associative and distributive laws apply [6.3].

For the least nonnegative residue system modulo 2, consisting of 0 and 1 (perhaps the most important one in this computer age) we have the addition table:

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \tag{6.10}$$

which is isomorphic both to the logical operation "exclusive or" and to multiplication of signs (if we identify 0 with + and 1 with −).

The multiplication table for 0 and 1:

$$\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array} \tag{6.11}$$

is isomorphic to the logical "and" and the set-theoretic "intersection".

Multiplication for a complete residue system modulo a composite number has no inverse for some of its members, as can be seen from the multiplication table modulo 4:

	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

(6.12)

Thus, there is no number which when multiplied by 2 gives 1, i. e., 2 has no inverse. Also, division by 2 is not unique: 2 divided by 2 could be either 1 or 3.

This grave defect is rectified by *prime* residue systems which consist only of those residue classes that are coprime to the modulus. Thus, the least nonnegative prime residue system modulo 10 consists of the integers 1, 3, 7 and 9, and their multiplication table is well behaved:

	1	3	7	9
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

(6.13)

If f is a polynomial with integer coefficients, then $a \equiv b \pmod{m}$ implies

$$f(a) \equiv f(b) \pmod{m}, \tag{6.14}$$

For $(m,n) = 1$ if x and y run through complete residue systems modulo n and m, respectively, then $mx + ny$ runs through a complete residue system modulo mn.

Example: $n = 5$, $m = 3$:

	$x = 0$	1	2	3	4
$y = 0$	0	3	6	9	12
1	5	8	11	14	2
2	10	13	1	4	7

We will encounter this kind of decomposition of a residue class modulo a product of coprimes again when we discuss simultaneous congruences and the Chinese Remainder Theorem with its numerous applications in Chap. 17.

6.3 Powers and Congruences

What is $2^{340} \pmod{341}$? Obviously, the ancient Chinese did not know or they would not have formulated their primality test (Sect. 2.3). With the aid of the congruence notation we may rewrite the Chinese test as

$$2^n \equiv 2 \pmod{n}, \tag{6.15}$$

if *and only if n* is prime. Here the first "if" is all right (see Fermat's Theorem, Chap. 8) but the "and only if" is wrong, as we shall presently demonstrate with the composite $n = 341 = 11 \cdot 31$. For odd n we may write, because of (6.9),

$$2^{n-1} \equiv 1 \pmod{n}.$$

Of course, it is foolish actually to calculate 2^{340}, a 103-digit number, if we are interested only in the remainder modulo 341. Instead we will decompose 340 into $10 \cdot 34$ and first raise 2 to the 10th power, giving 1024. Dividing 1024 by 341 leaves the remainder 1:

$$2^{10} \equiv 1 \pmod{341}. \tag{6.16}$$

Now raising the result to the 34th power is easy because $1^{34} = 1$. Thus,

$$2^{340} \equiv 1 \pmod{341}, \quad \text{or} \tag{6.17}$$

$$341 \mid (2^{341} - 2), \tag{6.18}$$

in spite of the fact that 341 is composite. Woe to the Chinese and three cheers for the congruence notation! We could even have done this in our heads, without recourse to pencil and paper, and thereby demolished a false "theorem" which had stood undisputed for so many centuries.

Composite numbers which masquerade as primes vis-a-vis Fermat's theorem are called *pseudoprimes*. 341 is actually the smallest pseudoprime to the base 2. In a certain sense, pseudoprimes have become almost as important as actual primes in modern digital encryption. We will hear more about pseudoprimes and even *absolute* and *strong pseudoprimes* in Chap. 20.

Of course, in calculating $2^{340} \pmod{341}$ we were lucky, because 2^{10} already gave the remainder 1. In general we will not be so lucky, and we need a universal algorithm that will see us through any base b or exponent n.

More formally, to calculate $b^n \pmod{m}$, we first find the binary decomposition of n:

$$n = \sum_{k=0}^{\lfloor \log_2 n \rfloor} n_k 2^k \quad \text{with} \quad n_k = 0 \text{ or } 1, \tag{6.19}$$

where $\lfloor\ \rfloor$ is the Gauss bracket, or "floor function", signifying the integer part.

The binary expansion coefficients n_k can be found by any of a variety of "analog"-to-digital conversion algorithms.

Omitting all terms with $n_k = 0$ in the sum, we may write

$$n = \sum_{m=1}^{M} 2^{c_m}. \tag{6.20}$$

In other words we write, for example, $6 = 2 + 4$ or $13 = 1 + 4 + 8$, etc.

Using this decomposition into a sum of powers of 2, we write b^n as follows:

$$b^n = \prod_{m=1}^{M} b^{c_m} = \underbrace{(\ldots(b^2)\ldots)^2}_{c_1 \text{ squarings}} \cdots \underbrace{(\ldots(b^2)\ldots)^2}_{c_M \text{ squarings}}. \tag{6.21}$$

In words: we calculate b^n by squaring b c_1 times in succession. Then we square b c_2 times. (Of course, we make use of the previous result, i. e., we need square only $c_2 - c_1$ more times, etc.) Then we multiply the results of all these squarings together to obtain b^n.

Apart from the gain in computational efficiency (if n is a power of 2, then only about $\log_2 n$ squarings are required, instead of n multiplications), the main raison d'être for the repeated squaring algorithm is that if we want the result modulo m, then after each squaring we can reduce the intermediate result modulo m without running the risk of calculator "overflow" (as long as m^2 is smaller than the largest number the machine can handle).

Here is another rule that makes working with powers and congruences easier:

$$(x+y)^n \equiv x^n + y^n \pmod{n}, \tag{6.22}$$

n prime (or absolute pseudoprime), i. e., of the $n+1$ terms obtained upon expanding $(x+y)^n$ binomially only two remain, because all others, being multiplied by $\binom{n}{k}$, with $k \neq 0$ or n, are divisible by n and therefore do not contribute to the end result modulo n.

Incidentally, the condition $k \neq 0$ or n can be expressed with the following more widely applicable and succinct notation:

$$k \not\equiv 0 \pmod{n}.$$

Read: k is *not* congruent to zero modulo n.

Chapter 7
Diophantine Equations

Diophantine equations, i. e., equations with integer coefficients for which integer solutions are sought, are among the oldest subjects in mathematics. Early historical occurrences often appeared in the guise of *puzzles*, and perhaps for that reason, Diophantine equations have been largely neglected in our mathematical schooling. Ironically, though, Diophantine equations play an ever-increasing role in modern applications, not to mention the fact that some Diophantine problems, especially the unsolvable ones, have stimulated an enormous amount of mathematical thinking, advancing the subject of number theory in a way that few other stimuli have.

Here we shall deal with some of the basic facts and rules and get to know *triangular* and *Pythagorean numbers, Fermat's Last Theorem*, an unsolved conjecture by *Goldbach*, and another conjecture by *Euler* – one that was refuted, although it looked quite convincing while it lasted.

7.1 Relation with Congruences

The congruence

$$ax \equiv c \pmod{m} \tag{7.1}$$

has a solution *iff* $(a,m) \mid c$. In fact, there are then (a,m) solutions that are incongruent modulo m [7.1].

Example:

$$3x \equiv 9 \pmod{6}. \tag{7.2}$$

With $(3,6) = 3$ and $3 \mid 9$, there are exactly three incongruent solutions: $x = 1$, 3 and 5. Adding or subtracting multiples of 6 to these three solutions gives additional solutions *congruent* to the three already found (7, 9, 11, 13, etc.).

For $(a,m) = 1$, the solution is *unique* modulo m.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_7, © Springer-Verlag Berlin Heidelberg 2009

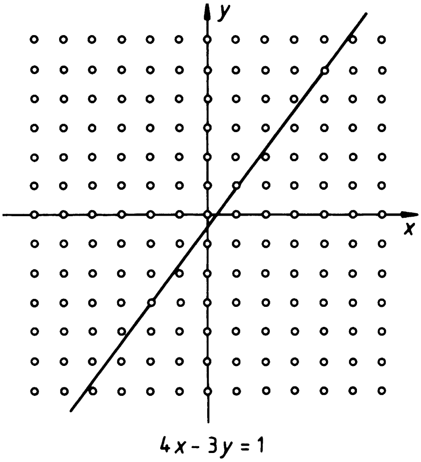


Fig. 7.1 Linear Diophantine equations have simple geometric interpretations: the straight line $3y = 4x - 1$ cuts through points $\{-2,-3\}$ and $\{4,5\}$, two solutions of the equation

The solution of the congruence $ax \equiv c \pmod{m}$ is identical with solutions in integers x and y of the *Diophantine equation* [7.1]:

$$ax = my + c, \tag{7.3}$$

so named after Diophantus of Alexandria (ca. A. D. 150) [7.2].

For $(a,m) = d$ and $m = m'd$, $a = a'd$ and $c = c'd$ [note that (a,m) must divide c for a solution to exist], we can write instead of the above equation

$$a'x = m'y + c', \tag{7.4}$$

whose solution is unique modulo m', because $(a',m') = 1$. Additional solutions that are incongruent modulo m are obtained by adding km', where $k = 1,2,\ldots,d-1$.

Diophantine equations also have a geometric interpretation, which is illustrated for $3y = 4x - 1$ in Fig. 7.1. The straight line representing this equation goes through only those points of the two-dimensional integer lattice shown in Fig. 7.1 for which x and y are solutions. In the illustration this is the case for the points $\{x,y\} = \{-2,-3\}$, $\{1,1\}$ and $\{4,5\}$. Additional solutions are obviously given by linear extrapolation with multiples of the difference $\{4,5\} - \{1,1\} = \{3,4\}$.

7.2 A Gaussian Trick

For $(b,m) = 1$, *Gauss* [7.3] suggested writing the congruence $bx \equiv c \pmod{m}$ as

$$x \equiv \frac{c}{b} \pmod{m} \tag{7.5}$$

and adding or subtracting multiples of m to c and b so that cancellation becomes possible *as if* c/b were, in fact, a fraction.

Example: $27x \equiv 1 \pmod{100}$.

Solution:

$$x \equiv \frac{1}{27} \equiv \frac{-99}{27} \equiv \frac{-11}{3} \equiv \frac{-111}{3} \equiv -37 \equiv 63 \pmod{100}.$$

Check: $27 \cdot 63 = 1701 \equiv 1 \pmod{100}$. Check!

Another method uses Euclid's algorithm (Sect. 2.7) for solving congruences or Diophantine equations. A congruence is first converted into a Diophantine equation. For example, the congruence

$$15x \equiv 1 \pmod{11} \tag{7.6}$$

has a solution, because $(15,11) = 1$ divides 1. The corresponding Diophantine equation is

$$15x = 11y + 1. \tag{7.7}$$

Solving for y, we obtain

$$y = x + \frac{4x-1}{11}. \tag{7.8a}$$

For (7.8a) to have an integer solution $4x - 1$ must be a multiple of 11:

$$4x - 1 = 11w.$$

Now solving for x, we get

$$x = 2w + \frac{3w+1}{4}. \tag{7.8b}$$

By now, the denominator has become so small that a solution is obvious: for x to be an integer, $3w + 1$ must be a multiple of 4, for example $w = 1$. Or, more formally:

$$3w + 1 = 4v,$$

whence

$$w = v + \frac{v-1}{3}. \tag{7.8c}$$

Here an integer solution is even more obvious: $v = 1$.

Solutions for x and y are now obtained by *backward* substitutions: with $v = 1$, (7.8c) gives $w = 1$ (as we noted before); and with $w = 1$, (7.8b) gives $x = 3$ and finally, if we so desire, (7.8a) gives $y = 4$.

Check: $15 \cdot 3 = 11 \cdot 4 + 1 = 45 \equiv 1 \pmod{11}$. Check!

The trick of this method of solution is that, in going from (7.8a) to (7.8b) and (7.8c), we have made the denominators smaller and smaller – just as in the Euclidean algorithm. In fact, the Euclidean algorithm applied to $15/11$, the original factors in

(7.7), gives precisely the emainders and denominators as appear in the equations (7.8a–c):

$$\begin{aligned}15:11 &= 1+\frac{4}{11}\\ 11:4 &= 2+\frac{3}{4}\\ 4:3 &= 1+\frac{1}{3}.\end{aligned}$$

For numerous calculations with the same modulus, it is most convenient to calculate a *table* of inverses once and for all. Such a table for the modulus 11 would contain, for example, the entry $4^{-1} = 3$. Check: $4 \cdot 3 = 12 \equiv 1 \pmod{11}$. Check!

Thus, the solution of $4x \equiv 7 \pmod{11}$ becomes $x \equiv 4^{-1} \cdot 7 = 21 \equiv 10 \pmod{11}$. Check: $4 \cdot 10 = 40 \equiv 7 \pmod{11}$. Check!

7.3 A Stamp Problem

Suppose you have an unlimited supply of 26-cent and 41-cent postage stamps. Which postage amounts can you cover exactly? Obviously, the smallest amounts that are covered (26, 41, 52, 78, 82 cents etc) have sizeable gaps that cannot be covered (1 to 25 cents, 27 to 40 cents etc.) But for larger amounts, these gaps get smaller and smaller. Thus, the question arises whether there is an amount after which *all* postages can be covered. An answer to this question (not in this guise, of course) was provided by the English mathematician J. Sylvester in 1884, who showed that for two kinds of stamps whose different values, a and b, are relatively prime, the critical amount is $(a-1)(b-1)$, or 1000 cents for $a = 26$ and $b = 41$ cents. Let us see whether there is a solution for $26\mathrm{x} + 41\mathrm{y} \equiv 1000$ as promised by Sylvester. First we solve the standard linear Diophantine equation.

$$26\mathrm{x} + 41\mathrm{y} = 1$$

Using the (Euclidean) algorithm described in this section, we find $\mathrm{x} = -11$ and $\mathrm{y} = 7$ Check:

$$-11 \cdot 26 \div 7 \cdot 41 = 1$$

(Alternatively, the solution can be found with the Mathematica$^{\mathrm{TM}}$ command “Extended GCD” which does Euclid’s algorithm for you.)

To get the solution for 1000, we multiply the solution for 1 by 1000 and add (subtract) k 26 41 to the first (second) term:

$(-11000 + \mathrm{k}41)26 + (7000 - \mathrm{k} - 26)41 = 1000$ For the factor of 26 to become positive k has to exceed 268. For k = 269 we get

$$29 \cdot 26 + 6 \cdot 41 = 1000$$

for the basic solution.

For the 1001 we simply add the basic solution $-11\cdot 26+7\cdot 41=1$ to yield

$$18\cdot 26+13\cdot 41=1001$$

and similarly for 1002:

$$7\cdot 26+20\cdot 41=1002.$$

Proceeding in the same manner we would get for 1003

$$-4\cdot 26+27\cdot 41=1003$$

Of course, there are no negative postages, but by adding (subtracting) $26\cdot 41$ to the first (second) term we get

$$37\cdot 26+1\cdot 41=1003$$

This procedure can be continued indefinitely for all higher postages. But what about 999? By subtracting our basic solution ($-11\cdot 26+7\cdot 41=1$) from the one for 1000 we obtain

$$40\cdot 26-1\cdot 41=999$$

Again we run into a negative postage, but this time there is no way out. We are stuck By adding $26\cdot 41$ to the second term to make it positive while subtracting $26\cdot 41$ from the first summand, we get

$$-1\cdot 26+25\cdot 41=999$$

But now we need a negative 26-cent stamp. In other words the postage 999 cents cannot be covered by 26-cent and 41-cent stamps. In fact, 1000 cents is the smallest postage after which all higher postages can be realized. According to Sylvester this crossover should occur at (a-1)(b-1) or, for $a=26$ and $b=41$, at $25\times 40=1000$. The interested reader may want to prove this result for general coprime a and b.

Unfortunately, there is no simple formula for the largest impossible amount for three or more postages. This was shown to be a NP-hard problem (see B. Cipra: "Exact Postage Poser Still Not Licked" in *Science* Vol. **319**, page. 899).

With an unlimited supply of 26-cent and 41-cent stamps 500 postages (between 1 cent and 999 cents) cannot be covered. Suppose the post office decides to issue a third stamp, which value (above 41 cents) should it be given to minimize the number of uncovered postages? The surprising answer is 42 cents, which leaves only 152 postages uncovered. If values above 26 cents are considered, the optimum stamp value is 32 cents, which leaves just 134 postages uncovered. Of course, some of those choices require a lot of stamp licking.

The stamp problem described here is related to the work (on linear forms with integer coefficients) of the German mathematician F.G. Frobenius (1849–1917), a student of Kronecker, Kummer and Weierstrass in Berlin. He was the teacher of Edmund Landau and Issai Schur to mention only the two most famous alummi.

7.4 Nonlinear Diophantine Equations

A simple example of a nonlinear Diophantine equation is

$$x^2 - Ny^2 = \pm 1, \tag{7.9}$$

where x and y are integers. Two solutions for $+1$ on the right-hand side (the so-called Pell equation) are obviously $x = \pm 1$, $y = 0$. Are there others? For $N = 2$,

$$x^2 - 2y^2 = \pm 1 \tag{7.10}$$

has the solution $x = \pm 3$, $y = \pm 2$. Are there more? Yes, and they are obtained by the continued-fraction (CF) expansion of $\sqrt{N}$. For $n = 2$, we have

$$\sqrt{2} = 1 + \sqrt{2} - 1, \tag{7.11}$$

$$\frac{1}{\sqrt{2}-1} = 2 + \sqrt{2} - 1, \tag{7.12}$$

i. e., as we already know, we obtain the periodic CF:

$$\sqrt{2} = \left[1;\overline{2}\right]. \tag{7.13}$$

The approximants A_k and B_k are thus obtained recursively from

$$A_k = 2A_{k-1} + A_{k-2}, \tag{7.14}$$

and similarly for the B_k. With the initial conditions $A_0 = A_1 = 1$ and $B_0 = 0$, $B_1 = 1$, we obtain

$$\begin{aligned} A_k &= 1,\ 1,\ 3,\ 7,\ 17,\ 41,\ 99, \ldots \\ B_k &= 0,\ 1,\ 2,\ 5,\ 12,\ 29,\ 70, \ldots . \end{aligned} \tag{7.15}$$

Here each pair of values (A_k, B_k) corresponds alternately to a solution of $x^2 - 2y^2 = 1$ and $x^2 - 2y^2 = -1$. This is not really too surprising, because we already know that A_k/B_k will tend to $\sqrt{2}$ alternately from above and below. However, the general proof is a bit tedious [7.1].

The first two solutions $\{1,0\}$ and $[1,1]$ we already know. We will check the third and fourth solutions: $7^2 - 2 \cdot 5^2 = -1$. $17^2 - 2 \cdot 12^2 = 1$. Check!

The CF's of the squareroots of integers are not only periodic; they are also palindromic, i. e., the periods are symmetric about their centres, except for the last number of the period, which equals twice the very first number (left of the semicolon). For example, $\sqrt{29} = [5;\overline{2,1,1,2,10}]$ or $\sqrt{19} = [4;\overline{2,1,3,1,2,8}]$.

Some periods are very short, such as those of numbers of the form $n^2 + 1$, which have period length 1. For example, $\sqrt{10} = [3;\overline{6}]$ or $\sqrt{101} = [10;\overline{20}]$. But the squareroots of other integers can have rather long periods. For example, $\sqrt{61} = [7;\overline{1,4,3,1,2,2,1,3,4,1,14}]$ has a period length of 11, and $\sqrt{109}$ has a period length of 15. There seems to be no simple formula that predicts long-period lengths.

Solving Pell's equation for such integers is quite tedious, a fact that Fermat exploited when he wrote his friend Frénicle in 1657, rather mischievously, to try $N = 61$ and $N = 109$ "pour ne vous donner pas trop de peine". (The *smallest* solution of the latter problem has 14 and 15 decimal places, respectively. Poor Frénicle!)

7.5 Triangular Numbers

The *k*th *triangular number* is defined as

$$\Delta_k = 1 + 2 + \ldots + k = \tfrac{1}{2}k(k+1). \tag{7.16}$$

Δ_k is the number of unordered pairs of $k+1$ objects or, more tangibly, the number of handshakings when $k+1$ persons meet (no self-congratulations, please!). The smallest Δ_k are 0, 1, 3, 6, 10, 15, ... Their first differences form a linear progression: 1, 2, 3, 4, 5, ...

On July 10, 1796, Gauss wrote in his still very fresh diary (then in its 103rd day):

$$\text{Eureka!} \quad n = \Delta + \Delta + \Delta, \tag{7.17}$$

by which he meant that every integer can be represented by the sum of 3 triangular numbers. For example, $7 = 3+3+1 = 6+1+0$; $8 = 6+1+1$; $9 = 3+3+3 = 6+3+0$; $10 = 6+3+1$, etc. What this means is that the Δ_k, although they grow like $k^2/2$, are still distributed densely enough among the integers that three of them suffice to reach any (nonnegative) whole number.

Gauss's discovery implies that every integer of the form $8n+3$ as a sum of three odd squares, an interesting *nonlinear* Diophantine equation, is always solvable. First we note that the square of an odd number, $2k+1$, equals 8 times a triangular number plus 1:

$$(2k+1)^2 = 4k^2 + 4k + 1 = 8\Delta_k + 1. \tag{7.18}$$

Hence if

$$n = \Delta_{k_1} + \Delta_{k_2} + \Delta_{k_3},$$

then we obtain, in view of Gauss's Eureka discovery, a solution to the following nonlinear Diophantine equation:

$$\sum_{m=1}^{3} (2k_m + 1)^2 = 8n + 3. \tag{7.19}$$

Example: $35 = 4 \cdot 8 + 3 = 1 + 9 + 25$.

Finally, there is a connection with the perfect numbers P_p: every even perfect number is also a triangular number:

$$P_p = (2^p - 1)2^{p-1} = (2^p - 1)2^p\tfrac{1}{2} = \Delta_{2^p - 1}. \tag{7.20}$$

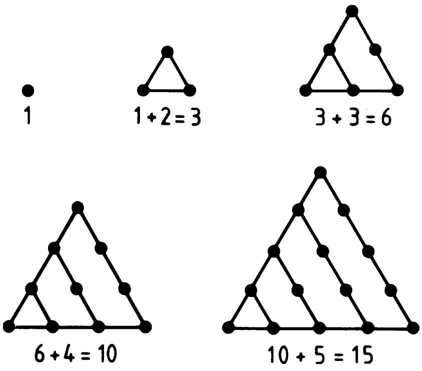


Fig. 7.2 Geometrical interpretation of the triangular numbers (and the reason for their name). Note the simple recursion

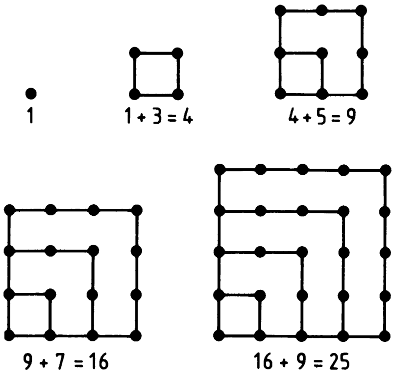


Fig. 7.3 The square numbers, their geometrical interpretation and their recursion

Triangular numbers can also be illustrated *geometrically* as the number of equidistant points in triangles of different sizes (Fig. 7.2). They were defined this way in antiquity (by the Pythagoreans). These points form a triangular lattice.

In a generalization of this concept, *square numbers* are defined by the number of points in square lattices of increasing size, as illustrated in Fig. 7.3.

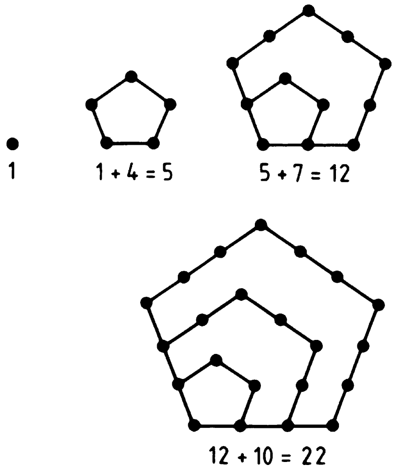


Fig. 7.4 The pentagonal numbers, their geometrical interpretation and their recursion. The pentagonal numbers, $n(3n-1)/2$, also play a role in partitioning problems (see Chap. 22)

Higher n-gonal or *figurate* numbers, such as pentagonal numbers (Fig. 7.4) and hexagonal numbers, are defined similarly. Can the reader derive the general formula for figurate numbers? Calling the kth (beginning with $k = 0$) n-gonal number $g_n(k)$, the answer is

$$g_n(k) = \tfrac{1}{2}(n-2)k^2 + \tfrac{1}{2}nk + 1, \tag{7.21}$$

which for $n = 4$ gives the square numbers $g_4(k) = (k+1)^2$. Check!

7.6 Pythagorean Numbers

Another nonlinear Diophantine equation is the well-known

$$x^2 + y^2 = z^2, \tag{7.22}$$

expressing Pythagoras's theorem for right triangles. Solutions in integers are called Pythagorean *triplets*, the smallest positive one being 4, 3, 5:

$$4^2 + 3^2 = 5^2.$$

To avoid redundancy, we shall require x to be even and y to be odd. (If both x and y were even, then z would also be even and the equation could be divided by 4.) All *basic* solutions, i. e., those for which x, y and z do not have a common divisor and x is even, are obtained from the two coprime integers m and n, $m > n > 0$, at least one of which must be even, as follows:

$$\begin{aligned} x &= 2mn \\ y &= m^2 - n^2 \\ z &= m^2 + n^2. \end{aligned} \tag{7.23}$$

It is easy to verify that $x^2 + y^2 = z^2$ and that x is even and $y > 0$ is odd. Of course, z is also odd.

With these conventions, the first case is $m = 2$, $n = 1$, yielding the triplet $(4,3;5)$. The next basic case is $m = 3$, $n = 2$, yielding $(12,5;13)$. The third basic case, $m = 4$, $n = 1$, yields $(8,15;17)$.

Incidentally, at least one pair of basic Pythagorean triples has the same product xy, meaning that there are (at least) two incongruent right triangles with integer sides and equal areas:

$$\begin{aligned} m &= 5, \quad n = 2 : (20,21;29) \\ m &= 6, \quad n = 1 : (12,35;37) \end{aligned} \tag{7.24}$$

both of which have area 210. Are there other equal-area pairs?

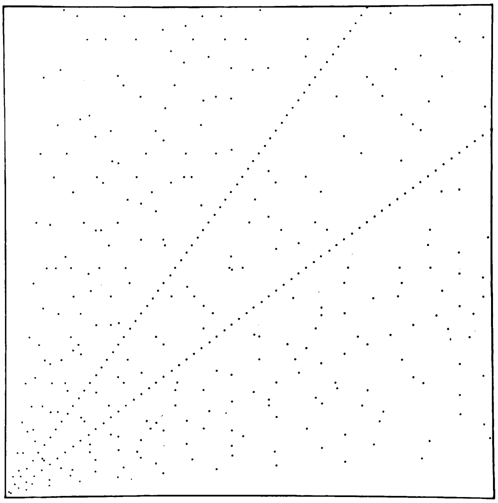


Fig. 7.5 The Pythagorean numbers $x^2 + y^2 = z^2$ in the range $1 \le x \le 52$, $1 \le y \le 52$. The two prominent straight "lines" correspond to the basic triplet $4^2 + 3^2 = 5^2$, the triplet $3^2 + 4^2 = 5^2$ and their multiples

Figure 7.5, prepared by Suzanne Hanauer of Bell Laboratories, shows the x and y values of all (not just the basic) Pythagorean triplets up to $x, y = 52$. The two pronounced straight "lines" are the solutions obtained from the basic triplet $(3, 4; 5)$. Each plotted point in Fig. 7.5 is one corner of an integer right triangle obtained by connecting it with the origin $\{0, 0\}$ and drawing the normal to the abscissa through it. It is apparent that there is a certain "thinning out" as x and y get larger.

7.7 Exponential Diophantine Equations

Another type of Diophantine equation, in which the unknowns appear in the exponent is exemplified by

$$2^n = 3^m - 1. \tag{7.25}$$

The background against which the author encountered this equation is the following. The Fast Fourier Transform (FFT) works most efficiently for data whose length is a power of 2. On the other hand, pseudorandom "maximum-length" sequences, which are ideal for precision measurements (see Chap. 27), have period lengths $p^m - 1$, where p is a prime. In most applications, the preferred sequences are binary, i. e., $p = 2$. Of course, there is no integer solution to the equation $2^n = 2^m - 1$, except the uninteresting one $n = 0$, $m = 1$.

For some physical applications, ternary-valued sequences would also be acceptable. Thus, the question arises whether the equation $2^n = 3^m - 1$ has integer solutions other than the two obvious ones $n = 1$, $m = 1$ and $n = 3$, $m = 2$.[1] Lewi Ben Gerson (1288–1344) proved that these are indeed the only solutions.

[1] In fact, in 1844 *Catalan* [7.4] posed a more general question and conjectured that 2^3 and 3^2 are the only perfect powers that differ by 1.

7.8 Fermat's Last "Theorem"

Perhaps the most famous Diophantine equation is

$$x^n + y^n = z^n, \qquad (n > 2) \tag{7.26}$$

for which Fermat asserted that no nontrivial $(xyz \neq 0)$ solution in integers exists. Fermat thought he had a proof, but this seems more than doubtful after centuries of vain efforts by some of the greatest mathematicians who came after Fermat.

Some special cases for n are relatively easy to prove, for example the quartic case, $n = 4$. The cubic case, $n = 3$, was solved by Euler. Sophie Germain in Paris, who mailed her proof to Gauss in Göttingen under the male pseudonym of Monsieur Le Blanc [7.5], showed Fermat's Last Theorem to be likely true for all odd primes p such that $2p+1$ is also prime. These primes are now called Sophie Germain primes, the smallest being $p = 3$.

The greatest breakthrough (as some would say today) was made in 1851 by the German mathematician Ernst Eduard Kummer. He showed that Fermat's Last Theorem (FLT) was true for what he called regular primes p_r, defined as those primes which do not divide any of the numerators of the Bernoulli numbers B_k up to B_{p_r-3} [7.6].

The only irregular primes below 100 are 37, 59, and 67, i. e., three primes out of 25. It is believed that the asymptotic fraction of the irregular primes tends toward $1 - 1/\sqrt{\mathrm{e}} = 0.393469\ldots$. Thus, there would be an infinity of irregular primes, roughly 40 % of all primes. This is in stark contrast to the Fermat primes, of which only 5 are presently known.

In 1908 the Göttingen Academy of Sciences established the Wolfskehl Prize to the tune of 100,000 (gold!) marks for proving FLT. In 1958, two World Wars later, this was reduced to 7,600 Deutsche Mark, but FLT had still not been proved or disproved. By 1976, it had been shown to be correct for all prime exponents smaller than 125,000.

If the proof of FLT has proved so difficult, perhaps the theorem is just not true, and one should look for a counterexample. (One counterexample, of course, would suffice to demolish FLT once and for all.) But unfortunately, since the exponent n has to be larger than 125000, and it can also be shown that x must be greater than 10^5, any counterexample would involve numbers *millions* of decimal digits long [7.7]. Hence the counterexample route seems closed, even if FLT was false. Consequently, we will have to conclude, more than 300 years after its somewhat offhand assertion, that FLT will probably never be disproved. However work on the FLT has led to some profound mathematical insights and innovations.

In the meantime, the "impossible" *has* happened and Fermat's Last Theorem has been proved. Although the proof, presented by Andrew Wiles on 23 June 1993, was still defective, all holes, even the most recalcitrant ones, have been closed with the collaboration of Richard Taylor. The proposition proved by Wiles and Taylor is the main part of the Shimura-Taniyama-Weil conjecture related to elliptic curves $y^2 = x^3 + ax + b$ [7.8].

In the summer of 1997 the Göttingen Academy of Sciences, after due consideration of the published proof, is expected to award Wiles the Wolfskehl Prize, which now amounts to about 70,000 Marks.

FLT, although for long the most famous unsolved case, is not an isolated quirk. Another seemingly unprovable (but probably false) conjecture is the following one by Georg Cantor: All numbers generated by the recursion

$$p_{n+1} = 2^{p_n} - 1, \quad \text{with} \quad p_0 = 2 \tag{7.27}$$

are prime. The first Cantor numbers after 2 are $p_1 = 3$, $p_2 = 7$, $p_3 = 127$ and $p_4 = 2^{127} - 1$, all of which are (Mersenne) primes. Unfortunately, little is known about the next Cantor number, p_5, other than that it has more than 5×10^{37} decimal digits! (This should not be confounded with the large – but "infinitely" smaller – number 5×10^{37}, which has only 38 digits.) Nevertheless, with the latest advances in primality testing (see Chap. 12) perhaps the primality of p_5 can be confirmed – or refuted, thereby demolishing another conjecture. (Note: if p_m is composite, then *all* p_n, $n \geq m$ are composite – see Sect. 3.5).

One of the perennial conjectures is the famous *Goldbach conjecture*, asserting that every even number > 2 is the sum of two primes. Some progress has been made on related weaker assertions, and the Goldbach conjecture itself has been *numerically* confirmed up to very large numbers. But alas, even if it had been shown to hold up to $10^{10^{10}}$, there would be no guarantee that it would not fail for $10^{10^{10}} + 2$. See Sect. 4.13 for some numerical results and an heuristic estimate.

7.9 The Demise of a Conjecture by Euler

Euler conjectured that (excepting trivial cases)

$$x_1^n + x_2^n + \ldots + x_k^n = z^n \tag{7.28}$$

has nontrivial integer solutions *iff* $k \geq n$. For $n = 3$ and $k = 2$, Euler's conjecture corresponds to the proven case $n = 3$ of FLT: the sum of 2 cubes cannot be another cube. For $n = 3$ and $k = 3$, (7.28) asserts that the sum of *three* cubes can be another cube. Euler's conjecture stood for two centuries but fell in 1966 to the joint effort of Lander and Parkin, who found a counterexample for $n = 5$:

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5, \tag{7.29}$$

which can be verified with a good pocket calculator (and which Euler himself could have done in his head).

Thus Fermat was avenged, whose fifth "prime" Euler had shown to be composite.

In fact, in 1987 Noam Elkies, a Harvard graduate student, found a first case where only *three* fourth powers are needed to get another fourth power. A year later, R. Frye found a solution in which the sum of three fourth powers equals 422481^4.

Another, more difficult problem that has also been successfully tackled, is the Diophantine equation

$$\mathrm{a}^4+\mathrm{b}^4+\mathrm{c}^4+\mathrm{d}^4=(\mathrm{a}+\mathrm{b}+\mathrm{c}+\mathrm{d})^4.$$

(See *The American Mathematical Monthly*, March 2008).

7.10 A Nonlinear Diophantine Equation in Physics and the Geometry of Numbers

In 1770, Lagrange proved that "the set of squares is a basis of order 4". This means that *every* positive integer can be represented as the sum of 4 squares. If we allow ourselves just 3 squares, then some integers cannot be so represented.

Which integers n can be expressed as the sum of 3 squares? The author first encountered this problem in the formula for the resonant frequencies of a cube-shaped resonator, which in units of the lowest resonant frequency are:

$$f_{x,y,z}^2=x^2+y^2+z^2. \tag{7.30}$$

An important question in some areas of physics is whether $f_{x,y,z}^2$ can take on all positive integer values or whether there are gaps. (This problem occurred in the author's Ph.D. thesis on normal-mode statistics.) To answer this question let us consider the complete residue system modulo 8:

$$r=0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7. \tag{7.31}$$

Hence,

$$r^2\equiv 0,\ 1\quad\text{or}\quad 4\pmod{8}.$$

Thus, the sum of 3 squares modulo 8 is precisely those numbers that can be generated by adding 3 of the integers 0, 1, 4 (with repetition allowed). This is possible for integers 0 through 6, but *not* 7. Thus, certainly,

$$x^2+y^2+z^2\neq 8m+7. \tag{7.32}$$

However, it can be shown that these "forbidden" numbers, when multiplied by a nonnegative power of 4, are also not possible as the sum of 3 squares [7.1]. In fact,

$$x^2+y^2+z^2=n \tag{7.33}$$

has a solution in integers *iff*

$$n\neq 4^k(8m+7),\qquad k\geq 0. \tag{7.34}$$

This means that on average, the fraction

$$\frac{1}{8}+\frac{1}{4\cdot 8}+\frac{1}{16\cdot 8}+\ldots=\frac{1}{6} \tag{7.35}$$

of all integers cannot be represented as the sum of 3 squares.

Since any positive integer n can be represented by a sum of 4 squared integers, it is interesting to ask how many ways $r_2(n)$ a given n can be so represented. The answer (due to Jacobi), including permutation, signs and 0's, is

$$r_2(n)=8\sum_{d|n,4\not|d} d, \tag{7.36}$$

i. e., 8 times the sum of the divisors of n that are not divisible by 4.

Example: $r_2(4)=8(1+2)=24$. Check:

$$\begin{aligned}4&=(\pm 1)^2+(\pm 1)^2+(\pm 1)^2+(\pm 1)^2 \quad &&(16 \text{ cases}) \qquad \text{and}\\ 4&=(\pm 2)^2+0^2+0^2+0^2 &&(+8 \text{ cases}) \quad \text{Check!}\end{aligned}$$

For a recent proof see *Hirschhorn* [7.9].

On the other hand, Fermat proved that a certain class of integers can be represented by the sum of just 2 squares, and in a *unique* way at that. This class consists of all primes of the form $4k+1$.

This result has an enticing geometrical interpretation. Consider the *integer lattice* in the plane, i. e., all the points (x,y) in the plane with integer coordinates. Draw a circle around the origin $(0,0)$ with radius $p^{1/2}$, where p is a prime with $p\equiv 1$ $(\operatorname{mod} 4)$. Then there are exactly eight lattice points on the circle.

No solutions exist for the primes $p\equiv -1\,(\operatorname{mod} 4)$. For composite n, we have to distinguish between the factor 2 and the two kinds of primes $p_i\equiv 1\,(\operatorname{mod} 4)$ and $q_i\equiv -1\,(\operatorname{mod} 4)$:

$$n=2^{\alpha}\prod_{p_i} p_i^{\beta_i}\prod_{q_i} q_i^{\gamma_i}.$$

Solutions for $n=x^2+y^2$ exist only if all γ_i are even. *Hardy* and *Wright* [7.1, p. 299] give four different proofs, no less, one going back to Fermat and his "method of descent".

If all γ_i are even, the number of solutions is $\prod_i(\beta_i+1)$, including permutations. ("Trivial" solutions with x or y equal to 0 are counted only once.) Example: $325=5^2\cdot 13$ has 6 solutions. $325=1^2+18^2=6^2+17^2=10^2+15^2=15^2+10^2=17^2+6^2=18^2+1^2$ and no others.

The question *which* squared integers sum to a given prime p is a little trickier to answer. It requires the *Legendre symbol* (a/p), see Eq. (16.17), which equals $+1$ if a is a quadratic residue modulo p, and -1 for a a nonresidue, and 0 for $a\equiv 0 \operatorname{mod} p$. Thus, for example, for $p=5$, beginning with $a=0$

$$(a/5)=0 \quad +1 \quad -1 \quad -1 \quad +1 \quad 0 \quad \ldots$$

Now taking the product of 3 consecutive terms, beginning with $0\ +1\ -1$, yields

$$0 \quad +1 \quad +1 \quad 0$$

summing to $+2$. Thus, $5 = 2^2 + y^2$ where, obviously, $y = 1$.

For $p = 173$, the same procedure yields $233 = 13^2 + y^2$ with $y = 8$.

The *two-squares theorem*, as it has been called, has an unexpected application in the generation of circulary polarized sound waves. Ordinarily, sound waves in air are thought of as *longitudinal* waves with no transverse field components that could give rise to elliptic or circular polarizations. But in a laterally restricted medium, such as an air-filled duct, sound waves do exhibit lateral motions and are thus polarizable.

Figure 7.6 shows a square box, a short piece of duct with a square cross-section, covered with plexiglas. This cavity resonator supports acoustic modes at frequencies

$$f_n = \frac{c}{2L}\sqrt{n},$$

where c is the velocity of sound in air, L is the side length of the box and n is a positive integer equal to the sum of two squares, the number of half-wavelengths along the side of the box:

$$n = k^2 + m^2.$$

To excite an elliptic wave, the mode must be degenerate, i. e., k and m must be different. Also, to avoid a mode "salad", we require precisely *two* modes at the same resonance frequency. Thus all exponents β_i in the above factorization of n must be 0 except one, which must equal 1, or n must be a square, in which case all β_i must

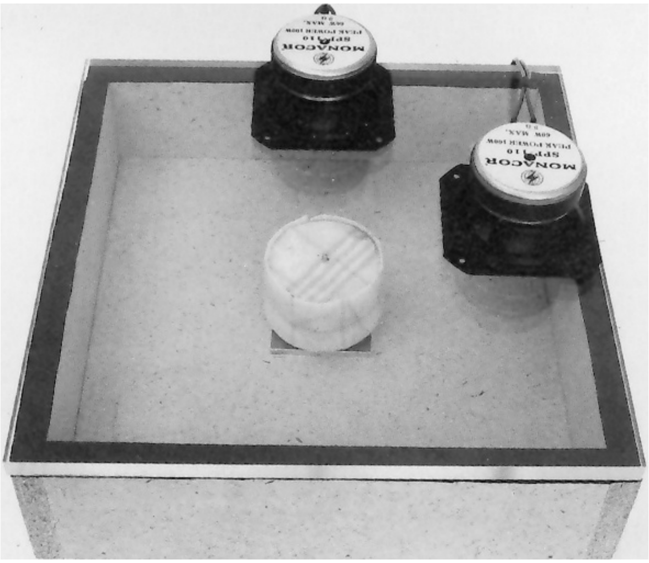


Fig. 7.6 Acoustic resonator with a square cross-section for circularly polarized sound waves

equal 0. Values of n which fulfill these conditions are $1 = 1^2 + 0^2$, $4 = 2^2 + 0^2$, $5 = 2^2 + 1^2$, $9 = 3^2 + 0^2$, $10 = 3^2 + 1^2$, $13 = 3^2 + 2^2$, etc. If the cavity shown in Fig. 7.6 is excited at one of these resonances by means of the two loudspeakers with a phase difference of 90°, a circularly polarized acoustic field will fill the cavity. The circularity can be made visible by a piece of sound-absorbing material supported by a sharp needle in the centre of the cavity: turn on the loudspeakers and the "carousel" will start rotating because the absorber absorbs sound energy, which, in a circularly polarized field, has an angular momentum. Reverse the electrical connections to one of the loudspeakers and the carousel will slow down and start rotating the other way. Replace the absorber by an empty yoghurt cup, which does not absorb much sound, and nothing will happen [7.10].

Geometrically speaking, the resonance frequencies of our square cavity form a two-dimensional square lattice. According to a well-known result by H. Weyl the asymptotic number of resonances ("eigenvalues") up to a given frequency f equals $\pi L^2 f^2/c^2$, which agrees with a result already obtained by Gauss concerning the asymptotic number of representations of the integers by the sum of two squares.

The connection between geometry and number theory was forged into a strong link by Hermann Minkowski. In his *Geometrie der Zahlen* (published in 1896) Minkowski established and proved many beautiful relationships at the interface of geometry and number theory. His most famous result, known as *Minkowski's theorem*, says that *any* convex region symmetrical about $(0,0)$ having an area greater than 4 contains integer lattice points other than $(0,0)$. This theorem, and its generalization to higher dimensional spaces, is particularly useful in proofs concerning the representation of numbers by quadratic forms, such as the above result on the decomposition of certain primes into sums of squares.

In an address before the Göttingen Mathematical Society commemorating the 100th anniversary of Dirichlet's birth, Minkowski hypothesized that some day soon, number theory would triumph in physics and chemistry and that, for example, the decomposition of primes into the sum of two squares would be seen to be related to important properties of matter.[2]

Another intriguing geometrical concept by Minkowski is that of a *Strahlkörper* (literally: ray body) defined as a region in n-dimensional Euclidean space containing the origin and whose surface, as seen from the origin, exhibits only one point in any direction. In other words, if the inner region was made of transparent glass and only the surface was opaque, then the origin would be visible from each surface point of the Strahlkörper (i. e., there are no intervening surface points). Minkowski proved that if the volume of such a Strahlkörper does not exceed $\zeta(n)$, a volume preserving linear transformation exists such that the Strahlkörper has no points in common with the integer lattice (other than the origin). Here $\zeta(n)$ is Riemann's zetafunction which we encountered already in Chap. 4 in connection with the distribution of primes.

[2] "In letzterer Hinsicht bin ich übrigens für die Zahlentheorie Optimist und hege still die Hoffnung, dass wir vielleicht gar nicht weit von dem Zeitpunkt entfernt sind, wo die unverfälschteste Arithmetik gleichfalls in Physik und Chemie Triumphe feiern wird, und sagen wir z. B., wo wesentliche Eigenschaften der Materie als mit der Zerlegung der Primzahlen in zwei Quadrate im Zusammenhang stehend erkannt werden." [7.11].

The fact that $\zeta(n)$ should determine a Strahlkörper property is not totally surprising. If we look at our “coprimality function”, Fig. 4.8, it consists of precisely all those points of the lattice of positive integers from which the origin is visible, i. e., the white dots in Fig. 4.8 define the surface of a (maximal) Strahlkörper. And the asymptotic density of dots is $1/\zeta(2) = 6/\pi^2$ (see Sect. 4.4) or, in n dimensions, $1/\zeta(n)$.

The reader may wants to explore the more general decomposition of primes p

$$p = x^2 + cy^2$$

with $c = \pm 2, \pm 3, \ldots$.
For $c = 2$, Lagrange showed that all primes

$$p \equiv 3 \bmod 8$$

are *uniquely* representable, e. g. $19 = 1 + 2 \cdot 3^2$.

For $c = -2$ and $p \equiv 7 \bmod 8$ there are infinitely many solutions obtainable by a simple linear recursion, which the reader may want to discover.

7.11 Normal-Mode Degeneracy in Room Acoustics (A Number-Theoretic Application)

The minimum frequency spacing of two nondegenerate normal modes of a cubical room, in the units used above in (7.30), is

$$\Delta f_{\min} = \frac{1}{2f_{x,y,z}}. \tag{7.37}$$

Because of the gaps in the numbers representable by the sum of 3 squares, the *average* nondegenerate frequency spacing becomes $7/6$ of this value:

$$\overline{\Delta f} = \frac{7}{12f}. \tag{7.38}$$

The asymptotic density of normal modes per unit frequency (using a famous formula on the distribution of eigenvalues, proved in its most general form by Hermann Weyl) is

$$\Delta Z = \frac{\pi}{2} \cdot f^2. \tag{7.39}$$

Thus the average degree of degeneracy D (i. e., the number of modes having the *same* resonance frequency) becomes [7.12]:

$$D = \Delta Z \cdot \overline{\Delta f} = \frac{7\pi}{24} \cdot f \simeq 0.92 \cdot f. \tag{7.40}$$

The degree of degeneracy is important because a high degeneracy can be detrimental to good room acoustics: normal modes that coincide in frequency are missing elsewhere and leave gaps in the frequency scale. Consequently, musical notes generated at those frequencies are not well transmitted to the attending audience (or microphones). This problem is most significant for small enclosures such as recording studios, where the mode density, especially at the lower end of the audiofrequency range, is already small and any unnecessary degeneracy impairs the acoustic responsiveness [7.12].

7.12 Waring's Problem

A problem that has stimulated much mathematical thought, by Hilbert among others, is *Waring's problem* [7.1]: given a positive integer $n > 0$, what is the least number of terms $G(k)$ in the sum:

$$n = \sum_j m_j^k \tag{7.41}$$

for all *sufficiently* large n?

Another question is how many terms $g(k)$ are needed so that *all* n can be represented as in (7.41). Of course:

$$g(k) \geq G(k). \tag{7.42}$$

As we saw in Sect. 7.9, $g(2) = 4$. It is also known that $g(3) = 9$. In fact, there are only finitely many n for which 9 third powers are required. Probably the only two cases are

$$\begin{aligned} 23 &= 2 \cdot 2^3 + 7 \cdot 1^3 \quad \text{and} \\ 239 &= 2 \cdot 4^3 + 4 \cdot 3^3 + 3 \cdot 1^3. \end{aligned}$$

Thus, by definition:

$$G(3) \leq 8.$$

Also $G(3) \geq 4$, but the actual value of $G(3)$ is still not known – another example of how a seemingly innocent question can lead to mathematically most intractable problems!

If we ask *how many different* ways $r_k(n)$ an integer n can be represented as the sum of kth powers (including different sign choices and permutations), then for example by (7.36), $r_2(5) = 8$. Of particular interest are the following asymptotic averages [7.1]:

$$\sum_{n=1}^{N} r_2(n) = \pi N + 0\left(\sqrt{N}\right) \quad \text{and} \tag{7.43}$$

$$\sum_{n=1}^{N} r_3(n) = \frac{4\pi}{3}N^{3/2} + 0(N). \tag{7.44}$$

Both (7.43) and (7.44) are intuitively obvious because they count the number of integer lattice points in a circle and, respectively, a sphere of radius $\sqrt{N}$.

Chapter 8
The Theorems of Fermat, Wilson and Euler

These three theorems, especially the one by *Euler*, play a central role in many modern applications, such as digital *encryption*. They are deeply related to the theory of groups, and indeed, their most elegant proofs are group theoretic. Here, however, we shall stress the purely arithmetic viewpoint. We also introduce the important *Euler ϕ function* (or totient function) which reaches into every corner of number theory and which, by way of illustration, tells us how many ways an n-pointed star can be drawn by n straight lines without lifting the pen.

8.1 Fermat's Theorem

For prime p and integer b not divisible by p, the following congruence holds:

$$b^{p-1} \equiv 1 \pmod{p}, \qquad p \nmid b. \tag{8.1}$$

Let us check this for $p = 5$: $1^4 = 1, 2^4 = 16 \equiv 1, 3^4 = 81 \equiv 1, 4^4 = 256 \equiv 1 \pmod{5}$.

One of several proofs of Fermat's theorem relies on the fact that a complete residue system modulo a prime, excluding 0, i. e., the numbers $1, 2, \ldots, p-1$ (if we decide to use least positive residues) form a multiplicative group of order $p-1$. It is well known from elementary group theory that *any* element of such a group, when raised to the power equal to the order of the group, will yield the unit element.

Another proof starts from the product

$$P = 1b \cdot 2b \cdot 3b \ldots (p-2)b \cdot (p-1)b = b^{p-1}(p-1)\,! \tag{8.2}$$

Here multiplication modulo p of the numbers $1, 2, 3, \ldots, p-2, p-1$ by any factor b which is coprime to p will only change their sequence, but not the value of the product. Thus, $P \equiv (p-1)!\pmod{p}$. Since the factor $(p-1)!$ is coprime to p, we can cancel it and conclude that b^{p-1} must be congruent 1 modulo p.

Example: for $p = 5$ and $b = 2$,

$$P = 2 \cdot 4 \cdot 6 \cdot 8 = 2^4 \cdot 4!, \tag{8.3}$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_8, © Springer-Verlag Berlin Heidelberg 2009

but taking each factor in P modulo 5 gives

$$P \equiv 2 \cdot 4 \cdot 1 \cdot 3 = 4! \pmod{5}.$$

Hence $2^4 \equiv 1 \pmod{5}$.

8.2 Wilson's Theorem

If and only if p is prime, then

$$(p-1)\,! \equiv -1 \pmod{p}. \tag{8.4}$$

For a proof when p is a prime, not relying on group theory, consider the product

$$2 \cdot 3 \cdot 4 \ldots (p-3) \cdot (p-2). \tag{8.5}$$

Here each factor has its own inverse (modulo p) somewhere among the other factors. For example, for $p = 7$, in the product $2 \cdot 3 \cdot 4 \cdot 5$ the factors 2 and 4 form a pair of inverses modulo 7, and so do 3 and 5. The product of such a pair is by definition congruent 1 modulo p, and so is the product of all the $(p-3)/2$ pairs in the above product. Thus,

$$2 \cdot 3 \cdot 4 \ldots (p-3) \cdot (p-2) \equiv 1 \pmod{p}, \tag{8.6}$$

and multiplying by $p-1$ yields:

$$(p-1)\,! \equiv p-1 \equiv -1 \pmod{p}, \quad \text{Q.E.D.} \tag{8.7}$$

Another way to state Wilson's theorem is as follows:

$$p \mid (p-1)\,! + 1. \tag{8.8}$$

Example: for $p = 5$: 5 divides $24 + 1$. Check!

Using this form of Wilson's theorem, we construct the function

$$f(n) = \sin\left[\pi \frac{(n-1)\,! + 1}{n}\right], \tag{8.9}$$

which has a truly remarkable property: $f(n)$ is zero if *and only* if n is prime. Here we seem to have stumbled across the ultimate test for primality!

By contrast, Fermat's theorem, always true for primes, *can* also be true for composite numbers. Not so for Wilson's theorem, which holds if *and only* if p is prime.

Unfortunately, this primality test is of *no* practical advantage, because calculating $f(n)$ takes much longer than even a slow sieve method. For example, to test whether 101 is prime, we would have to compute 101! first – a 160 digit tapeworm!

Another proof of Wilson's theorem uses Fermat's theorem in the following form:

$$x^{p-1} - 1 \equiv 0 \pmod{p}, \tag{8.10}$$

which holds for $x = 1, 2, \ldots, p-1$. According to the fundamental theorem of algebra these $p-1$ roots must be *all* the roots of the above equation. We can therefore write

$$x^{p-n} - 1 \equiv (x-1) \cdot (x-2) \ldots (x-p+1) \pmod{p}. \tag{8.11}$$

Now letting $x = p$ yields

$$p^{p-1} - 1 \equiv (p-1) \cdot (p-2) \ldots 1 = (p-1)\,! \pmod{p}. \tag{8.12}$$

Here p^{p-1} is of course congruent 0 modulo p and Wilson's theorem is proved.

8.3 Euler's Theorem

For b and m coprime, i. e., $(b,m) = 1$, Euler asserted that

$$b^{\phi(m)} \equiv 1 \pmod{m}. \tag{8.13}$$

Here $\phi(m)$ is Euler's ϕ function (also called totient function) defined as the number of positive integers r smaller than m that are coprime to m, i. e., for which $1 \leq r < m$ and $(r,m) = 1$ holds.

Example: $m = 10$: $r = 1,3,7,9$. Thus $\phi(10) = 4$. Note: $\phi(1) = 1$.

For *prime* moduli, each of the numbers $r = 1, 2, \ldots, p-1$ is coprime to p and therefore $\phi(p) = p-1$.

For prime powers p^{α}, one obtains in a similar fashion

$$\phi(p^{\alpha}) = (p-1)p^{\alpha-1} = p^{\alpha}\left(1 - \frac{1}{p}\right). \tag{8.14}$$

Example: $\phi(9) = 6$. Check! (There are exactly six positive integers below 9 coprime to 9, namely, 1, 2, 4, 5, 7, and 8.)

Euler's ϕ function is a so-called *multiplicative function*, which in number theory is defined as a function for which

$$\phi(n \cdot m) = \phi(n) \cdot \phi(m) \quad \text{for all } (n,m) = 1 \tag{8.15}$$

holds. Thus, we have for any

$$m = \prod_i p_i^{e_i} \tag{8.16}$$

$$\phi(m) = \prod_i (p_i - 1) p_i^{e_i - 1} \qquad (e_i > 0), \tag{8.17}$$

or in a form without the restriction on the exponents:

$$\phi(m) = m \prod_i \left(1 - \frac{1}{p_i}\right). \tag{8.18}$$

As can be seen from (8.17), the totient function $\phi(m)$ is always even, except $\phi(2) = \phi(1) = 1$.

Example: for $m = 10$, $\phi(10) = 10(1 - \frac{1}{2})(1 - \frac{1}{5}) = 4$; and $1^4 = 1$, $3^4 = 81 \equiv 1$, $7^4 = 49^2 \equiv (-1)^2 = 1$, $9^4 = 81^2 \equiv 1^2 = 1 \pmod{10}$.

One proof of Euler's theorem parallels that of Fermat's: Consider the prime residue system modulo m:

$$r_1, r_2, \ldots, r_{\phi(m)} \tag{8.19}$$

and multiply each r_k with b, where $(b, m) = 1$. This multiplication changes the *sequence* of the above residues, but their total product is not affected. For example, for $m = 10$,

$$r = 1,\ 3,\ 7,\ 9,$$

which upon multiplication with $b = 7$ becomes

$$rb = 7,\ 21,\ 49,\ 61 \equiv 7,\ 1,\ 9,\ 3 \pmod{10}.$$

This is, of course, a consequence of the fact that prime residue systems form a multiplicative group.

Hence,

$$b^{\phi(m)} r_1 r_2 \ldots r_{\phi(m)} \equiv r_1 r_2 \ldots r_{\phi(m)} \pmod{m}. \tag{8.20}$$

Since the residues are by definition prime to m, we can cancel these factors and are left with

$$b^{\phi(m)} \equiv 1 \pmod{m}. \tag{8.21}$$

The reader may wish to show the following curious but important property of the totient function:

$$m \mid \phi(p^m - 1), \tag{8.22}$$

which holds for any m.

8.4 The Impossible Star of David

How many ways can an n-pointed star be drawn from n straight lines without lifting the pen? An n-pointed star has of course n outer corners. For simplicity we assume that these corners lie equidistantly on a circle. Connecting corners that are adjacent on the circle produces a polygon, not a star. To obtain a star we must connect each corner with one of its $n-2$ non-neighbors. For $n = 6$, we get the six-pointed Star of David: two superimposed triangles.

But suppose we are asked to draw a six-pointed star in six straight consecutive strokes. The Star of David then becomes impossible (the reader is invited to try the following himself). Suppose the points are numbered $1,2,3,4,5,6$ and we connect 1 with 3 with 5 with 1. But that is not a six-pointed star; that is a triangle and points $2,4,6$ have been left out. Suppose we skip *two* points. Then we connect 1 with 4 with 1, and we have missed four points. Suppose we permit irregular skippings and connect, for example, 1 with 3 with 6 with 2 with 4 – but now we cannot skip any more because the only missing point is the adjacent point 5.

No matter how hard we try, a six-pointed star cannot be completed in this way. Yet it is easy to draw a five-pointed star by five straight lines without lifting the pen: connect point 1 with 3 with 5 with 2 with 4 with 1. And there are even *two* seven-pointed stars: skipping *one* point at a time, connect point 1 with 3 with 5 with 7 with 2 with 4 with 6 with 1 (Fig. 8.1); or, skipping *two* points at a time, connect point 1 with 4 with 7 with 3 with 6 with 2 with 5 with 1 (Fig. 8.2).

When is a star possible? If $k-1$ is the number of points skipped, then k must be greater than 1 and smaller than $n-1$ (or we would get a polygon) and k must be coprime with n, the number of corners. Because $\phi(5) = 4$, there are four values of k such that $(k,5) = 1$, of which only two differ from 1 or $n-1$: $k = 2$ and $k = 3$.

The two stars drawn with $k = 2$ and $k = 3$ are indistinguishable once the drawing is completed. Thus, there is only one possible five-pointed star.

From the discussion it should be clear that for general n, the following holds:

$$\text{Number of possible stars} = \frac{\phi(n)-2}{2}. \tag{8.23}$$

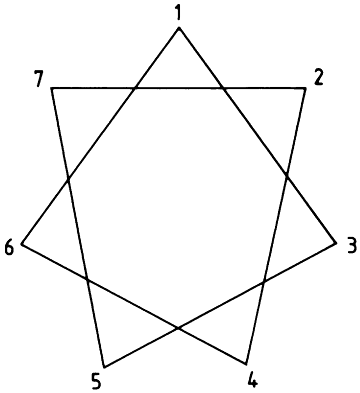


Fig. 8.1 One of two possible regular seven-pointed stars. It is based on the fact that 7 is coprime with 2 (and 5)

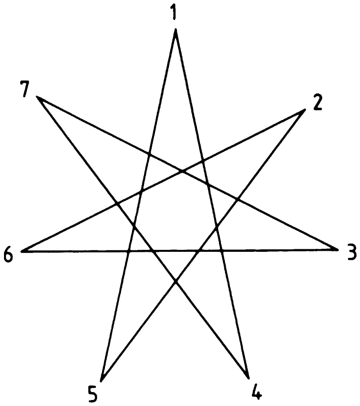


Fig. 8.2 The other possible seven-pointed stars based on the coprimality of 7 with 3 (and 4)

Let us try this formula for $n = 6$. With $\phi(6) = 2$, we get *no* stars – as we have already discovered by trial and error.

For $n = 7$ we get, with $\phi(7) = 6$, $(6-2)/2 = 2$ stars. Check! (See Figs. 8.1, 8.2.)

Question: for which n do we get 6 different stars? Answer: for *no* n – the number $2 \cdot 6 + 2 = 14$ is a so-called *nontotient*: $\phi(n)$ can never equal 14. (Can the reader show this?)

The number 6 seems to have an unlucky star: not only is a singly connected six-pointed star impossible, but there is no n for which exactly six different stars are possible. Other nontotients are 26 and 34. What is the general condition for a number to be a nontotient?

8.5 Dirichlet and Linear Progression

Consider the linear progression

$$a_k = mk + c, \quad (c,m) = 1, \qquad k = 0,1,2,\dots. \tag{8.24}$$

There are exactly $\phi(m)$ incongruent choices for c and therefore $\phi(m)$ nonoverlapping progressions. In 1837 Dirichlet showed, in one of the most spectacular early applications of analytic number theory, that the primes are equally distributed over these $\phi(m)$ different progressions, i. e., each progression contains asymptotically the fraction $1/\phi(m)$ of all the primes [8.1].

For $m = 3$, for example, $\phi(3) = 2$, and all primes (except 3 itself) belong to one of two residue classes: $3k + 1$ and $3k - 1$. And Dirichlet asserted that these two kinds of primes are equally frequent, each class (on average) laying claim to one-half of all primes.

For $m = 4$, $\phi(4) = 2$, there are likewise two kinds of primes: $4k \pm 1$, and again they are equally frequent (but see below!).

For $m = 6$, $\phi(6) = 2$, and we conclude that primes of the form $6k \pm 1$ are equally frequent.

For $m = 10$, $\phi(10) = 4$, and all the primes (except 2 and 5) have 1, 3, 7 or 9 as their last digits, and we find that each class asymptotically gets a quarter of the total.

For $m = 100$, $\phi(100) = 40$, the conclusion from Dirichlet's theorem is that the second-to-last digit (the "tens" digit) will be 0, 1, 2, 3, 4, 5, 6, 7, 8 or 9 with equal probability. In fact, in the interval from 7200 to 8800, where we would expect to find a total of $16600/\ln 8000 = 178$ primes, there are actually 176 primes with the following distribution over the tens digits:

0:	15	5:	15
1:	20	6:	18
2:	20	7:	14
3:	17	8:	20
4:	20	9:	17 ,

with a mean value of 17.6 primes and a standard deviation of $\sigma = 2.2$ – considerably less than for a Poisson process, which would have $\sigma = 4.2$.

The author once thought that nothing but equipartition of the primes between different prime residue classes of a given modulus k was even thinkable. After all, what is the difference between primes of the form $4k+1$ and $4k-1$? Well, for one, $+1$ is a quadratic residue of 4 and -1 is not. As a result, there is a subtle "quadratic effect" on the distribution of primes taken modulo 4. As *Chebyshev* [8.2] noted in 1853, there are more primes of the form $4k-1$ than $4k+1$, although not as *many* as some people believed. In fact, the difference in number between these two kinds of primes changes sign infinitely often! The number of primes below $x = 2.5 \cdot 10^9$ is about $1.2 \cdot 10^8$, and the absolute excess of primes $4k-1$ is roughly 2000 – subtle but finite.

There are even subtler cubic residue effects. However, in *relative* terms, all these effects vanish as x approaches infinity, and Dirichlet is ultimately vindicated.

Chapter 9
Permutations, Cycles and Derangements

Permutations pervade much of mathematics including number theory. Besides innumerable peaceful uses, permutations were crucial in classical cryptography, such as the German *Geheimschreiber* (secret writer) and *Enigma* enciphering machines—and their demise. The *Geheimschreiber* was broken during World War II by the Swedish mathematician Arne Beurling—with the occasional help from the leading Swedish statistician Harald Cramér (see B. Beckman: *Codebreakers*).

Polish and British cryptanalysts were able to break the Enigma code by observing—among other factors—the cycle structure of the code. Cycles of permutations and their distributions are therefore considered in Section 9.4 of this chapter.[1]

9.1 Permutations

The number of arrangements ("permutations") of n distinct objects equals the *factorial* of n:

$$n! := 1 \cdot 2 \cdot 3 \cdots n, \tag{9.1}$$

a formula easily proved by induction. Factorials grow very fast: while 5! equals just 120, 10! is already equal to 3 628 800. A good and relatively simple approximation is Stirling's famous formula:

$$n! \approx \sqrt{(2\pi n)}n^n \, \mathrm{e}^{-n}, \tag{9.2}$$

which yields 3 598 696 for $n = 10$.

A better approximation multiplies the Stirling result by $\mathrm{e}^{1/12n}$, yielding $10! \approx 3\,628\,810$ (for an error of less than 0.0003%!).

Factorials are also related to the "Euler" integral. Repeated partial integration shows that

[1] While the Germans, after a few years, became aware of the *Geheimschreiber's* vulnerability and curtailed its use, the fact that the Allies had broken *Enigma* was one of the best-kept secrets of the war. The *Enigma* decrypts therefore continued to provide the Allies with invaluable information during the entire war.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_9, © Springer-Verlag Berlin Heidelberg 2009

$$\int_0^\infty t^n \mathrm{e}^{-t}\,dt = n! \tag{9.3}$$

The related *gamma-function*

$$\Gamma(z) := \int_0^\infty t^{z-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$$

is single-valued and analytic in the entire complex plane, except for the points $z = -n$ $(n = 0, 1, 2, \ldots)$ where it possesses simple poles with residues $(-1)^n/n!$. $\Gamma(z)$ obeys the recurrence formula

$$\Gamma(z+1) = \mathrm{z}\Gamma(z) \tag{9.4}$$

and the curious "reflection" formula

$$\Gamma(z)\Gamma(1-z) = -z\Gamma(-z)\Gamma(z) = \pi\csc(\pi z), \tag{9.5}$$

which for $z = 1/2$ yields

$$\Gamma(1/2) = \sqrt{\pi}.$$

9.2 Binomial Coefficients

As we learn in high school (?), the "binomial" $(1+x)^n$ can be expanded (multiplied out) as follows:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \tag{9.6}$$

where the $\binom{n}{k}$ (read n choose k) are the binomial coefficients—

$$\text{With 0! defined as 1, } \binom{n}{0} = \binom{n}{n} = 1, \tag{9.7}$$

$\binom{n}{1}$ equals n and $\binom{n}{2}$ equals $n(n-1)/2 = 0, 1, 3, 6, 10, 15, \ldots$ the "triangular" numbers (see Sec. 7.4). The binomial coefficient $\binom{n}{2} = 1/2\ n(n-1)$ is (by definition) the number of *pairs* that can be selected from n distinct objects. Thus, at a party of n people, each guest clinking his glass with everyone else, produces a total of $1/2\ n(n-1)$ clinkings. (Of course, for n = 1, the number of possible clinkings is zero, just as there is no applause with just one hand clapping. For two people ($n = 2$), there is just one clinking.)

Permutations when just two objects change places are called *transpositions*. Every permutation can be decomposed into a unique (modulo 2) number of transpositions. If this number is odd, the permutation is called *odd*. Otherwise it is called an *even* permutation. The identity permutation is even because the number of transpositions is 0 (an even number). For example for $n = 5$, there are a total of $n! = 120$

permutations of which 60 are odd and 60 are even, the latter forming the famous *symmetrical group* S_5 which was crucial in Galois' historic proof that the general quintic equation has no solution in radicals. This put to rest a problem that had baffled mathematicians for centuries.

Other special permutations with unique properties are those generated by primitive roots, the number-theoretic logarithm (index), and the Zech logarithm (see Chaps. 14 and 27).

If the n objects consist of m groups containing k_1, $k_2, \ldots, k_m$ elements, respectively, the binomial coefficients can be generalized to the *multinomials*. The multinomial coefficients are defined by

$$\frac{n!}{k_1!, k_2! \ldots k_m!}, \quad \text{where} \quad \sum_{r=1}^{m} k_r = n \tag{9.8}$$

The differences between consecutive triangular numbers equal 1, 2, 3, 4, 5, ..., i.e. a set covering *all* positive integers. Thus, as the young Gauss discovered, they are sufficiently dense so that every positive integer can be represented by the sum of just 3 triangular numbers Δ. Or, as Gauss wrote in 1796 in his still new notebook:

$$\text{Eureka! n} = \Delta + \Delta + \Delta.$$

Note that already $n = 5$ requires 3 triangular numbers $(5 = 3 + 1 + 1)$.

By contrast, the *square* numbers, 0, 1, 4. 9, 16, 25 ... have differences equal to 1, 3, 5, 7, 9 ..., i.e. they cover only the odd numbers. They are therefore less dense and up to 4 squares are required to represent all positive integers. For example, $7 = 4 + 1 + 1 + 1$ cannot be represented by just 3 squares. The same is true for $n = 15, 23, 31, \ldots 28, \ldots$ (see Sect. 7.9 for more on the sum of 3 squares).

9.3 The Binomial and Related Distributions

If p is the probability that one of n possible events occurs, the probability of k events occurring in n independent trials is proportional to the binomial coefficient "n choose k". With k ranging from 0 to n, the (discrete) probability *distribution* is the so-called binomial distribution

$$p(k) = c \binom{n}{k}, 0 \leq k \leq n \tag{9.9}$$

where the constant c must be chosen so that

$$\sum_{k=0}^{n} p(k) = c \cdot 2^n \tag{9.10}$$

equals 1, i.e. c must equal $1/2^n$.

The mean value of k equals np and its variance is $np(1-p)$, which for fixed n achieves its maximum for $p = 1/2$.

For large n, the binomial distribution looks like samples from a Gaussian (normal) distribution. And in fact, for large n, the binomial distribution can be approximated by a normal distribution with mean np and variance $np(1-p)$.

For n $\to \infty$, but np fixed: $np = m$, the binomial distribution turns into the important Poisson distribution

$$p_m(k) = \frac{m^k}{k!}\mathrm{e}^{-m}, \quad k = 0, 1, 2 \ldots \tag{9.11}$$

with mean *and* variance equal to m.

The Poisson distribution gives the number of "clicks" per second of a Geiger-counter near a radioactive source with an average click rate equal to m clicks per second. The Poisson distribution also describes the occurrence of other "rare" events, i.e. events for which p is so small that, even as n $\to \infty$, np stays finite.

9.4 Permutation Cycles

One important subject in the study of permutations is their *cycle structure*. It was by the analysis of cycles, and particularly an invariance property of the cycle structure that the Polish mathematician Marian Rejewski, before the outbreak of World War II, was able to crack the *Enigma* enigma (see C. Christensen, *Mathematics Magazine*, Vol. **80**, No. 4 (October 2007), pp. 247–273).

For $n = 2$, the two possible permutations are the "identity" permutation (1, 2) and the transposition (2, 1). Note that the notation (2, 1) means that the first object is now in the second position and the second object appears in the first position.

Here (1, 2) has two cycles of period length 1 each:

$$1 \to 1 \text{ and } 2 \to 2.$$

Whereas the transposition (2, 1) has only *one* cycle of period length 2:

$$1 \to 2 \to 1.$$

Thus, the total number of cycles equals 3, two of which have period length 1 and one ($1 \to 2 \to 1$) has length 2.

Now let us study the cycle structure for the case of $n = 3$.

For the 6 permutations of 3 objects, cycle-analysis yields for the identity permutation (1, 2, 3): $1 \to 1$, $2 \to 2$, $3 \to 3$, i.e. 3 cycles of length 1.

For the permutation (1, 3, 2) we have 2 cycles: $1 \to 1$ and $2 \to 3 \to 2$, one of which has length 1 and the other cycle has length 2.

For (2, 1, 3), we find again 2 cycles, $2 \to 1 \to 2$ and $3 \to 3$, with lengths 1 and 2, respectively.

For (2, 3, 1), a "cyclic" permutation, we have just 1 cycle, $1 \to 2 \to 3 \to 1$, with length 3.

For (3, 1, 2), the other cyclic permutation, we find again just 1 cycle, $1 \to 3 \to 2 \to 1$, with length 3.

And, finally, for the permutation (3, 2, 1), we find 2 cycles, $1 \to 3 \to 1$ and $2 \to 2$, with lengths 1 and 2, respectively.

Thus, for $n = 3$, we have found a total number of 11 cycles, namely 2 permutations with 1 cycle, 3 permutations with 2 cycles and 1 permutation with 3 cycles. In general, the number of cycles is given by the Stirling numbers of the first kind $\mathrm{S}_n{}^{(m)}$, with the generating function

$$x(x-1)\cdots(x-n+1) = \sum_{m=0}^{n} S_n{}^{(m)} x^m \tag{9.12}$$

and the recurrence relation

$$\mathrm{S}_{n+1}{}^{(m)} = \mathrm{S}_n{}^{(m-1)} - n\mathrm{S}_n{}^{(m)}.$$

In fact, the number of permutations of n symbols which have exactly m cycles equals

$$\#_n(m) = (-1)^{(n-m)}\mathrm{S}_n{}^{(m)}, \tag{9.13}$$

which for $n = 3$ and for $m = 1, 2, 3$ yields the values 2, 3, 1, respectively.

The *total* number of cycles is given by the simple formula

$$\sum_{m=1}^{n} m\#_n(m) = \#_{n+1}(2), \tag{9.14}$$

which, for $n = 3$, yields 11 (as we already found by enumerating all 6 permutations for $n = 3$).

For more on the Stirling numbers, see Graham, Knuth and Patashnik: *Concrete Mathematics*, a veritable treasure trove of discrete (discreet?) mathematics.

Now let us look at the total number of different cycle *lengths*. By summing up the above results for $n = 3$, we find that there are a total of six cycles of period length 1, three cycles of length 2 and two cycles of length 3. Bell Labs mathematician S.P. Lloyd has shown that, in general, for the $n!$ permutations of n distinct objects there are $n!$ cycles of length 1, $n!/2$ cycles of length 2 and, generally, $n!/k$ cycles of length k, $1 \leq k \leq n$.

Thus, the total number of cycles equals $n!(1 + 1/2 + \quad +1/n)$. Here, the sum is the *harmonic number* H_n, which can be approximated by a definite integral from $x = 1/2$ to $x = n + 1/2$, over $1/x$, yielding

$$\mathrm{H}_n \approx \ln(2(n+1/2)). \tag{9.15}$$

However, considering that $1/x$ is concave (i.e. a "sagging" function), the factor 2 in the above formula overestimates H_n. Taking a cue from Euler, we replace the factor 2 by $\mathrm{e}^{\gamma} \approx 1.781$, where $\gamma = 0.57721\ldots$ is Euler's constant. This yields the astonishingly accurate approximation

$$\mathrm{H}_n \approx \ln(1.781(n+1/2)), \tag{9.16}$$

giving (exact values in parentheses) $H_1 \approx 0.98$ (1.0), $H_2 \approx 1.49$ (1.5), $H_3 \approx 1.83$ (1.83...) and $H_{50} \approx 4.49915$ (4.49921) with an error of less than 0.007%. (Of course, $n!H_n$ must be an integer, so that multiplying the approximation for H_n by $n!$ and rounding to the nearest integer gives an even better approximation for H_n).

In a computer simulation using the Random Permutation routine of Mathematica^{TM}, I found for 10^5 random permutations of 50 objects, 100 098 cycles of length 1, 49 970 cycles of length 2 etc. down to 2 005 cycles of length 50 and a total number of cycles (450 836) in close agreement (0.2%) with the expected number of cycles of $10^5 \cdot H_{50} = 449\,921$. These agreements are remarkable, given that the investigated 10^5 permutations are a very small fraction of the total of $50! \approx 3 \cdot 10^{64}$ permutations.

9.5 Derangements

A (complete) *derangement* is a permutation that leaves no object in its original place. A well-known derangement problem is that of n envelopes and n letters: What is the probability that not a single letter will end up in its proper envelope if the assignment of letters to envelopes is random? (For large n, the probability tends to $1/e \approx 0.37$ or 37%.) For two objects, there is exactly one derangement, the transposition (2, 1). For 3 objects, there are two complete derangements, the two cyclic permutations (2, 3, 1) and (3, 1, 2).

No matter what n is there is always exactly one permutation, with no derangements, namely the identity permutation where *all* objects are in their original place. There are never any permutations with just a single derangement because if one object is "deranged", there must be another one that is also "out of place". From these facts (plus a few other "insights") I once guessed the proper formula for the number $D(\mathrm{n})$ of complete derangements of n symbols:

$$D(n) = n! \sum_{\mathrm{k}=0}^{n} \frac{(-1)^k}{k!} \tag{9.17}$$

where the sum converges to 1/e for $n \to \infty$.

This is not to be confused with the birthday problem: how many people must be present at a party so that the probability that at least two persons have the same birthday exceeds 1/2? (Answer: 23).

9.6 Ascents and Descents

One aspect of permutations that has taken on considerable significance in recent times is the question of the longest rising (or falling) subsequence. Thus, in the permutation (1, 3, 5, 2, 4) of the first five positive integers, the longest rising subsequence is 1, 3, 5 while the longest *falling* subsequence is 5, 2. According to a theorem by the celebrated Paul Erdös and Gyorgy Szekeres, (proved by the pigeonhole principle) any list of k^2+1 distinct numbers contains an increasing (or decreasing) subsequence of length $k+1$.

Thus, for $k=2$, any list of $k^2+1=5$ distinct numbers contains a subsequence of length 3.

Here is a list of three random permutations of 5 numbers 1, 2, 3, 4, 5:

$$(3,4,5,1,2); \quad (2,5,3,1,4); \quad (2,3,4,1,5).$$

The longest monotone subsequence of the first permutation is an increasing subsequence, namely (3, 4, 5). The longest subsequence of the second permutation is a decreasing one: (5, 3, 1). Both have length 3.

The third permutation has a longest increasing subsequence of length 4, (2, 3, 4, 5), i.e. greater than the guaranteed minimum of 3. Note that non-monotonic members (like the 1 in the third permutation) can intervene. The members of such subsequences don't have to be contiguous. They are therefore sometimes called *scattered* subsequences.

It is interesting to observe that the distribution of increasing (or decreasing) subsequences of random permutations is related to the distribution of the eigenvalues of certain chaotic dynamical systems. Such distributions are therefore of great contemporary concern.

9.7 Quantum Decrypting

The days of the RSA[2] public-key encryption scheme may be numbered. The reader will recall that the difficulty of breaking RSA encrypted messages hinges on the difficulty of factoring large numbers. While the ever-advancing speeds of number-theoretic factoring[3] can be easily held at bay by using ever larger key numbers, going from, say, 300 digits to 400 digits, a new paradigm is arising on the cryptographic horizon that will thoroughly undo RSA: *quantum factoring.* In 1994 Peter Shor[4] then working at Bell Laboratories in Murray Hill, New Jersey, proposed a quantum algorithm for very fast factoring large composite numbers. Shor's algorithm is based on finding the order ("period length") of certain number-theoretic sequences.

As is well known (no, this is not translated from Russian), to compute the *de*crypting exponent t from the (public) *en*crypting exponent s, the following

[2] The RSA algorithm was named after Donald Rivest, Adi Shamir and Leonhard Adleman who published it in 1977. It was actually invented by Clifford Cocks three years earlier in a project classified TOP SECRET by British Intelligence.

[3] In May 2007 the largest number factored was 2^{1039}-1—which has over 300 decimal digits—with the help of some 500 computers running "in parallel" for 6 months (see *Discover Magazine* (January 2008), pp. 17–30).

[4] P. Shor: *Proc. 35th Annual Symposium of the Foundations of Computer Science*, p. 124. See also *SIAM Journal on Computing* **26** (1997), p. 1484, for a full version of Shor's paper.

Diophantine equation must be solved:

$$s \cdot t \equiv 1 \bmod \phi(m)$$

where ϕ is Euler's ϕ-function (also called totient-function) and m is the (public) encrypting modulus. Now, if m is the product of two primes, p and q, then $\phi(m) = (p-1)(q-1)$. Thus, to obtain the value of ϕ, the (secret) factors of m, i.e. the individual primes, must be known—not just their product!

How does Shor get these factors? Take a look at Euler's generalization of Fermat's "Little Theorem":

$$a^{\phi(m)} \equiv 1 \bmod m$$

for coprime a and m. It follows that the sequence $a^k \bmod m$, $k = 1, 2, 3 \ldots$ has a period length that divides $\phi(m)$. For example, for $m = 10$, and $a = 3$, we get the sequence

$$3, 9, 7, 1, 3, 9 \ldots,$$

which has a period length of 4. (Remember: $9 \cdot 3 = 27 \equiv 7$ mod 10.) And 4 does indeed divide $\phi(10) = \phi(2 \cdot 5) = (2-1)(5-1) = 4$.

The same period length is obtained for $a = 7$: to wit 7, 9, 3, 1, 7, 9 .... But for $a = 9$ we get 9, 1, 9, 1 ..., i.e. a period length of 2, but still a divisor of $\phi(10) = 4$.

The period lengths of 3^k and 7^k are therefore as long as possible. Such numbers a are called primitive roots. 3 and 7 are thus two (the only) primitive roots of 10. (Only integers 1, 2, 4, p^k and $2 \cdot p^k$, where p is an odd prime, have primitive roots. The number of primitive roots equals $\phi(\phi(m))$, or 2 for $m = 10$.)

Now, if the primes p and q are different (and larger than 2), then $m = p \cdot q$ has no primitive roots and the period length L can never attain the value $\phi(m) = (p-1)$ $(q-1)$. But, L is of course still a divisor of $(p-1)(q-1)$. In fact, in a large number of all (legal) choices of a, L equals $(p-1)(q-r)/2$, so that the calculation of p and q from L and $m = p \cdot q$ is easily accomplished.

One interesting relation between the period lengths of a^k modulo p, q, and pq, respectively, is

$L_{pq} =$ LCM (L_p, L_q) where the L are period lengths and LCM stands for "least common multiple".

For example, for $p = 5$ and $q = 7$ and $a = 3$ we get $L_5 = 4$, $L_7 = 6$ and $L_{35} = 12$, which is indeed the least common multiple of 4 and 6 and which also a divisor of $\phi(35) = 4 \cdot 6 = 24$. Here is a somewhat larger (randomly generated) example: $p = 229$, $q = 349$, $a = 7$, for which $L_p = 228$. $L_q = 348$ and $L_{pq} = 6612$—which divides $(p-1)(q-1) = 79344$ and is divisible by both 228 and 348, as it should as the least common multiple.

This is not surprising, because—as every physicist knows (but a mathematician still has to prove)—the period length L_{ab} of two added oscillations (periodic sequences) with period lengths L_a and L_b, respectively, is simply LCM (L_a, L_b). In physics and musicology this is known as a beat note and its frequency equals the largest common divisor of the two "beating" tones or, what is the same, the beat period is the least common multiple of the two (or more) beating periods. But there is a big difference between telling the fundamental frequency (or period) of a musical note and determining the period length of a number-theoretic sequence a^k. For example, for $a = 97$ and the above ($p = 229$, and $q = 349$), the sequence a^k mod (pq) starts as 97, 9409, 33542, 56734, ... and continues in a seemingly random fashion showing no periodicity. Only after $6612 = 228 \cdot 348/12$ steps does it start over again: ..., 11535, 1, 97, 94909, ... Such long periods are difficult to discern on a graphical printout. However, if converted to an audible tone, then even for fundamental period lengths as long as one or several seconds, the periodicity can be heard. (Typically, depending on the sampling rate, it sounds like the idling engine of a motorboat.)

9.8 Decrypting without Factoring

While the need of factoring the encryption modulus m into its prime factors was considered an article of faith for breaking RSA, some number theorists have come up with a method of decrypting RSA that does *not* require factoring. In fact, with a sprinkling of (elementary) group theory and Euler's Theorem, it can be shown that the Diophantine equation

$$s \cdot t \equiv 1 \bmod \phi(m)$$

can be solved for t *without* factoring m. Surprisingly, $\phi(m)$ in the above equation can be replaced by the *order* of the (publicly transmitted) encrypted message modulo m, the (public) encryption modulus. For the still necessary period finding one could use the Shor algorithm. So, while RSA has not yet been cracked, it is good to know that factoring is not a *sine qua non*. Also, *the order* is usually smaller than $\phi(m)$.

As an example, let us take $p = 617$ and $q = 2273$, i.e. $p \cdot q = 1402441$. For an encrypting exponent $s = 101$ and message $n =$ **31415**, the encrypted message is 31415^{101} mod 1402441 which equals 81679. Now the decrypting exponent t, as usually obtained, is given by

$$s \cdot t \equiv 1 \bmod \phi(m),$$

which requires factoring of m (a 7-digit number in the example).

In the alternative method, $\phi(m)$ in the above equation is replaced by the order of the cryptogram, 81697, modulo m. This yields for the decrypting exponent $t = 1122413$ and it is easy to confirm the correctness of this result by calculating, modulo m, 81697^t which equals, wonder of wonders, the original message: **31415**—and we still don't know (or care) what p and q are.

But where is the connection with quantum mechanics (QM) and its calculating speed? Well, QM is good at Fourier transforming or spectral analysis. And Shor finds the period lengths of a^k by Fourier analysis on a "quantum computer". I put quantum computer between quotation marks because Shor's algorithm isn't really a full-blown quantum computer—it's just a super fast period-length finder relying on quantum mechanics.

Of course, as a final step, the Shor algorithm calls for a measurement leading, in a quantum system, to a "collapse" of the wave function. However, the system collapses with high probability to the desired state, namely the spectral peak whose frequency is to be determined.

9.9 Quantum Cryptography

In an early realization of quantum cryptography, Anton Zeilinger and co-workers[5] transmitted an image (of a prehistoric statuette of a woman—the "Venus von Willendorf"), making use of *entangled photon states* originally called "verschränkte Zustände" by Erwin Schrödinger.

Entangled states are at the core of the Einstein, Podolsky, Rosen (EPR) paradox. Einstein, for one, never believed in the *spukhafte Fernwirkungen* (spooky actions at a distance) that are implied by EPR. But he was wrong and the seeming paradoxon invented by him, Podolsky and Rosen is now an experimentally verified foundation of quantum physics.

The use of entangled states allows *single* photons to be used in the quantum encryption scheme—a breathtaking achievement, especially in view of the fact that in the original interpretation of quantum mechanics, its laws were considered to be applicable only to large ensembles (Of course, because of photon loss during transmission and detection, the "single-photon" schemes usually employ several photons).

In the quantum cryptography scheme, invented by Charles Bennet and Gilles Brossard[6]; (see Sect. 9.11), the encrypted data is transmitted via an open (public) channel. But the data is made unintelligible by a secret key, a *one-time-pad*. And it is the one-time pad key, a sequence of random bits, that is transmitted via a secure, unbreakable, quantum channel.

9.10 One-Time Pads

One-time pads are considered the only really secure method of encryption because the key bits are used only a single time and then never used again so that no statistical information can be exploited. (The clever use of statistical dependencies is of course the root of most decrypting schemes.—C.E. Shannon derived a mathematical requirement for a key to be secure involving its entropy and the entropy of the message to be encrypted.)

In the world's navies, secret keys are often printed with water-soluble ink on blotting paper. But sometimes the ship doesn't sink and the key is recovered—as in

[5] T. Jennewein *et al*: Quantum Cryptography with Entangled Photon Phys Rev. Lett. **84**, 4729–4732 (15 May 2000) See also Bouwmeester, Ekert, Zeilinger (Eds): *The Physics of Quantum Information* (Springer, 2000).

[6] See C.H Bennet and G. Brossard, in *Proc. IEEE Int. Conference on Computers*, Bangalore (1984).

the case of the German submarine *U 505* that was "sunk" off West Africa on June 4th, 1944, by the U.S. Navy and then dragged to Bermuda across the Atlantic below the water's surface (to hide the fact that the key was captured). Three weeks later the captured code books were at Bletchley Park, the British deciphering center (with Alan Turing in residence).[7]

An ingenious variant of the one-time pad was used by the famous Soviet spy Richard Sorge. Sorge (who, on KGB orders, joined the Nazi party as a camouflage) memorized which page of the German *Statistisches Jahrbuch* for 1937 (publicly available at German embassies around the world) he had to consult on any particular day to extract the key. He was able to tell Stalin that the Japanese would "go south" and not attack the Soviet Union, which allowed the Russians to transfer their crack Siberian divisions to Moscow in November 1941—with the well-known result: Hitler's first major defeat.[8]

9.11 The Bennet-Brossard Key Distribution Scheme (BB84)

The seed idea for the BB84 scheme was the (totally impractical) proposal for "quantum money" by Stephen Wiesner in which each dollar bill, in addition to its serial number, carries 20 different polarized photons known only to the issuing bank. Because of the rules of quantum mechanics, such a bill could never be copied because the secret polarizations where (randomly) choses from *two* possible "channels": either horizontal/vertical or $\pm 45^\circ$. Here a "1" might be encoded by a vertical or a $+45^\circ$ polarization. A "0" would be encoded by a horizontal or -45° polarization.

If the potential money faker measured, for example, a $\pm 45^\circ$ photon with a horizontal/vertical photon counter, he would get a random result—without knowing that it was random! The bank, on the other hand, knowing all 20 polarizations, would, of course, have no problem reading the secret code and verifying the validity by comparing it to the serial number of the bill. The only problem with this lovely, totally secure, scheme: how do you store 20 photons for any length of time in a paper bill?

While the idea of using two different polarization channels was impractical for the creation of quantum money, it was resurrected for a quantum-mechanical secret-key distribution scheme called BB84. Here Alice, who wants to send secret messages to Bob (in the presence of an eavesdropper Eve) first constructs a one-time key pad shared with Bob. For this purpose she generates a sequence of random 1s

[7] *U 505* is now at the Museum of Science and Industry in Chicago. Earlier submarines whose codebooks (and equipment) were captured include *U 110* (May 1941) and *U 559* (October 1942).

[8] Sorge was made a posthumous Hero of the Soviet Union under Khrushchev and a street in *East* Berlin was named after him. The German Democratic Republic also issued a postage stamp with a portrait of Sorge—but not until after Stalin's death, who loathed Sorge—a "thorn in his side"—for being privy to his greatest blunder: ignoring the massive warnings of the imminent Nazi attack in June 1941.

and 0s[9]. Alice also generates a sequence of polarization channels chosen randomly from horizontal/vertical and ±45°.

Next she transmits to Bob the random bits, each one over one of the randomly chosen polarization channels. Bob detects the photons he receives with randomly chosen polarization channels, which of course agrees only 50% of the time with the channel used by Alice. In a subsequent public communication Alice tells Bob which channels she has used and Bob discards all results measured by his using the wrong channel. Alice can do this publicly because she only communicates the channels and not the actual bits she transmitted. So evil Eve is none the wiser.

Quantum mechanics also guarantees that any eavesdropping can be easily detected. As a consequence of Heisenberg's indeterminacy principle, Eve's observing the photon stream from Alice to Bob will necessarily change some of the photon polarization states. To know that this is happening, Alice and Bob only have to compare some, say 50, bits and, if they all agree, they can safely assume that their photon link was undisturbed. If they do this publicly, they cannot of course use the check bits for encrypting.

One of the remaining principal difficulties is the inability of the polarized photon channel to work over large distances in the atmosphere, (which is apt to change polarizations) thereby precluding—for the time being—worldwide key distribution via satellites. However, optical glass fibers are sufficiently stable and such systems have in fact been successfully implemented.

Another open—political—problem is whether states should allow the free use of quantum cryptography because it would allow criminals unfettered communication.

[9] based on some *physical* source of randomness, such as a radioactive decay. If she were to use an algorithmic (pseudo) random number generated instead, she would of course, according to John von Neumann, live in a state of sin—there is no way to generate truly random numbers on a computer.

Part IV
Cryptography and Divisors

Chapter 10
Euler Trap Doors and Public-Key Encryption

Here we describe a *public-key digital encryption* scheme that does not require secret-key distribution – one of the Achilles' heels of secret-key cryptographic systems because once the secret encrypting key becomes known, transmissions are no longer secret. Figure 10.1 illustrates the situation for a generic encryption/decryption system. Both transmitter and receiver need to be in possession of the same key, a long string of "random" binary digits. At the transmitter, on the left, these bits (0 or 1) are added modulo 2 to the message bits to form the encrypted message. A "bad guy" tapping into the output cannot decrypt the transmitted cryptogram unless he knows the key. In fact, he does not even know whether there is traffic on the tapped line: even if there is no traffic (a "message" consisting only of 0's) the line still carries random 0's and 1's. (This disguise of traffic volume is itself an important asset of the illustrated encryption method – many impending enemy actions have been betrayed simply by a rise in traffic.)

The problem implicit in the illustrated method is of course the secure transportation or transmission of the key between sender and receiver. Many ingenious schemes have been invented to solve the key problem. In addition the keys have to be protected from enemy ogling. To forestall illicit snooping, the navies of the world print sensitive keys on water-soluble paper. In other words, the key will dissolve into thin water when the ship sinks. A great idea – but not all vessels will sink when attacked by the enemy. Figure 10.2 shows a case in point: His Majesty's Ship, the German light cruiser *Magdeburg*, run aground off the Baltic Coast in August 1914 while trying to outmaneuver a Russian man-of-war. The poor *Magdeburg* refused to sink and the Russians got hold of the undissolved key book of the imperial German Navy (see Fig. 10.3). After examining it for several weeks, they shipped it to the British Admiralty (First Lord none other than Winston Churchill) who could make excellent use of it.

This kind of mishap is completely avoided if there *is* no secret key to be kept secret. In fact, in public-key encryption, every potential recipient of secret messages *publishes* his encrypting key (hence the name *public* key). But knowledge of the *en*crypting key is of no practical help in *de*cryption. The public-key is the key to a "trap door" through which messages can vanish, not to be recovered – except through a different route to which only the legitimate receiver holds the required key, the decrypting key. (Trap door means easily *in*, but out only with the proper

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_10, © Springer-Verlag Berlin Heidelberg 2009

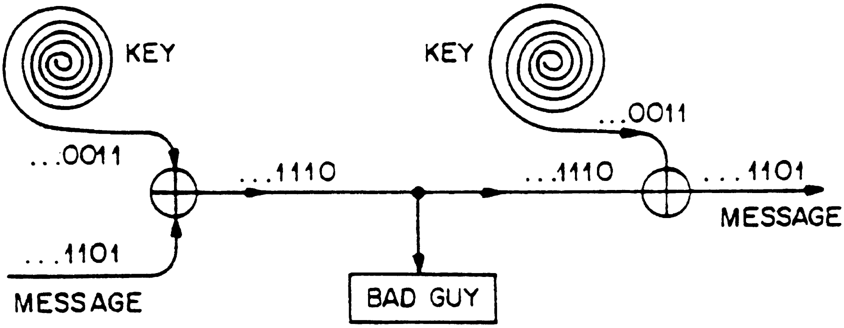


Fig. 10.1 Encrypting a binary message by adding modulo 2 bits from a key. The identical key must be available at the receiver on the right

tools.) It is as if everyone had a (chained!) box in front of his house into which secret messages could be stuffed and to which only the owner had the proper opening key.

The required trap-door behaviour is realized for digital messages (represented by long strings of digits) by the fact that it is easy to multiply two large numbers, but impossible to factor sufficiently large numbers in a reasonable amount of time.

The public-key cryptographic systems described here are based on modular arithmetic, requiring knowledge of the Euler ϕ function of the modulus n to calculate inverses, i.e., to decrypt. And of course, to determine $\phi(n)$, we have to know the factors of n.

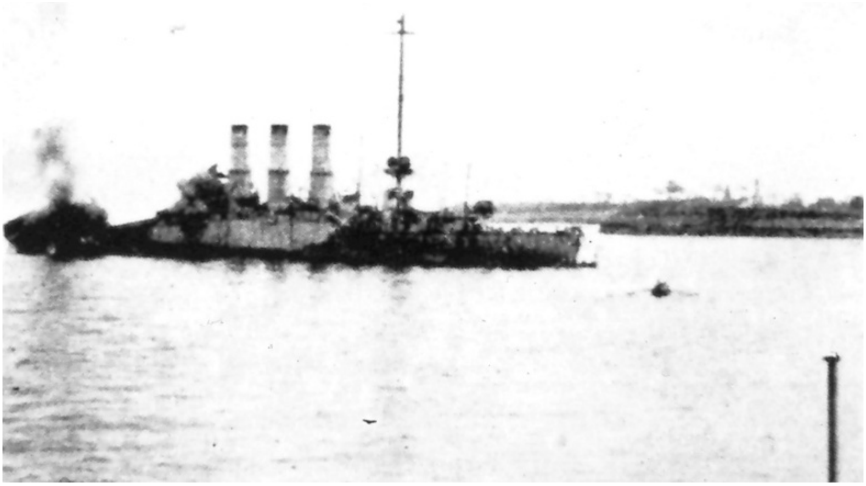
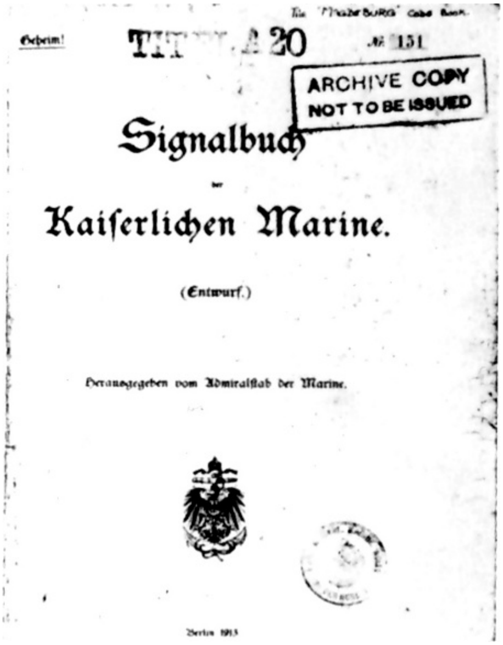


Fig. 10.2 The German cruiser *Magdeburg* run aground off the Baltic coast in August 1914



The "MAGDEBURG" Code Book.

Geheim! TIT[illegible] 20 № 151

ARCHIVE COPY
NOT TO BE ISSUED

Signalbuch
der
Kaiſerlichen Marine.

(Entwurf.)

Herausgegeben vom Admiralſtab der Marine.

Berlin 1913

Fig. 10.3 The key book of the German Navy captured by the Russians

Some exceedingly interesting interrelations between cryptography and *error-correcting codes* (Sect. 28.2) are discussed by *Sloane*, and the reader seeking deeper insights into these seemingly conflicting goals is referred to [10.1].

10.1 A Numerical Trap Door

An important practical application of Euler's theorem today is *public-key encryption* [10.2], and specifically the construction of *trap-door functions* for such secret activities [10.3].

What is a trap-door function? It is, as the name implies, a (mathematical) function that is easy to calculate in one direction but *very* hard to calculate in the opposite direction. For example, it takes a modern computer only microseconds to multiply two 500-digit numbers. By contrast, to take a 1000-digit number having two 500-digit factors and decompose it into its factors can take *ages*, even on the very fastest computers available and using the most efficient factoring algorithms known today.

Although no completely satisfactory theoretical definition of trap-door functions has been given, or a proof that they exist, reasonable *candidates* for such functions have been proposed, for example the Euler trap-door function to be discussed here. These are not *proved* secure, but are considered *empirically* (at present) secure trap doors.

10.2 Digital Encryption

Before encryption, a written message is usually first converted into a positive integer $1 < M < r$, i.e., a string of (decimal or binary) digits. In addition, we will assume that M and r are coprime.

The encryption we want to discuss here consists in raising M to a certain power s and keeping only the remainder modulo r. Thus, the encrypted message E is represented by the number

$$E \equiv M^s \pmod{r}, \qquad 1 < E < r. \tag{10.1}$$

Now, if we had chosen a prime for our modulus r, decrypting E, i.e., recovering M from E, would be easy. First we find the inverse, t, of the encrypting exponent s modulo $r-1$ by solving the congruence

$$st \equiv 1 \pmod{r-1}. \tag{10.2}$$

For this congruence to have a solution, s and $r-1$ must be coprime: $(s, r-1) = 1$. One method of obtaining t from s is by an application of Euler's theorem:

$$t \equiv s^{\phi(r-1)-1} \pmod{r-1}, \tag{10.3}$$

because for such a t, according to Euler,

$$st \equiv s^{\phi(r-1)} \equiv 1 \pmod{r-1}, \tag{10.4}$$

provided that $(s, r-1) = 1$.

Now, if we raise the encrypted message E to the power t and reduce the result modulo r, we obtain the original message M back. First,

$$E^t \equiv M^{st} \pmod{r}, \tag{10.5}$$

and since $st \equiv 1 \pmod{r-1}$ we can write

$$st = (r-1)k + 1 \quad \text{(for some integer } k\text{)}. \tag{10.6}$$

Thus,

$$E^t \equiv M^{(r-1)k+1} \pmod{r}. \tag{10.7}$$

Now Euler intervenes again (or rather Fermat, as long as r is prime), telling us that

$$M^{r-1} \equiv 1 \pmod{r} \tag{10.8}$$

if $(M, r) = 1$, which we initially assumed.

With this intermediate result, we obtain

$$E^t \equiv 1^k M^1 = M \pmod{r}, \tag{10.9}$$

i.e., we have recovered the original message M from the encrypted message E, which is what we wanted to show.

Example: $s = 3$, $r = 17$; $\phi(r-1) = 8$.

With this choice of s and r, for which $(s, \phi(r)) = 1$ as required, the decrypting exponent t becomes

$$t \equiv s^{\phi(r-1)-1} = 3^7 \equiv 11 \pmod{16}.$$

Check: $st = 3 \cdot 11 = 33 \equiv 1 \pmod{16}$. Check!

Now take a message, say $M = 4$. The encrypted message becomes

$$E \equiv M^s = 4^3 = 64 \equiv 13 \pmod{17}.$$

The decrypted message is obtained by raising $E = 13$ to the decrypting exponent $t = 11$:

$$\begin{aligned} E^t &= 13^{11} = (13^5)^2 \cdot 13 = (371\,293)^2 \cdot 13 \equiv 13^2 \cdot 13 \\ &\equiv 4 = M \pmod{17}. \qquad \text{Check!} \end{aligned}$$

We have now set the stage for our central theme.

10.3 Public-Key Encryption

We shall try to understand one of the greatest advances in the safeguarding and secure transmission of secrets: *public-key encryption*. This story has both “logistic” and mathematical aspects. I shall attempt to illuminate both.

One form of public-key encryption makes use of a very large *composite* number, the encrypting modulus r. Specifically, we shall assume that r has exactly two prime factors, p and q, both of which are also very large

$$r = pq. \tag{10.10}$$

More specifically, we may think of p and q as being each about 100 decimals long, so that r has roughly 200 decimal places.

In addition, the method makes use of an encrypting exponent s, which is chosen so that s and $\phi(r)$ are coprime:

$$(s, \phi(r)) = 1. \tag{10.11}$$

Everyone who ever wishes to receive a secret message selects such a triplet of numbers p, q and s and *publishes* s and the product r of p and q. The publishing may take the form of a kind of “telephone” book, where behind the name and address of each prospective secret-message recipient R, two numbers appear, an encrypting exponent s and an encrypting modulus r – selected in the manner just described. Thus, anyone – say a person named S – wishing to send a secret message raises its digital form M to the power s and reduces it modulo r.

Now notice something very important here: a key never had to be sent from R to S; R simply published it in a “telephone” book for everyone to see and use if he or she so desires. Thus, the grave problems of secure key transmission, which fill volumes in the literature of espionage, have been circumvented.

But cannot everyone now decrypt the encrypted message? The answer is *no*, if the encryption has made use of a trap-door (or almost one-way) function.

But what if the encryption was by a true one-way process like, let us say, the randomizing of air molecules in a bicycle tire: if the tire was initially half filled with air – also known among cyclists as (half) flat – and then pumped up full, all the molecules just pumped in could *never* be extracted again while leaving all the original ones in the tire. Here we have an everyday example of a true one-way function. In public-key encryption the function is not totally one-way – or the sender might as well burn the message and scatter it to the winds. No, it is only a trap door, i.e., a door that can still be opened from one side – the side of the legitimate recipient. Here is how.

As in the previous example, the recipient needs a decrypting exponent t, given again by the congruence

$$ts \equiv 1 \pmod{\phi(r)}. \tag{10.12}$$

Now, *if the factors of r are known*, then $\phi(r)$ is also known. For $r = pq$, we have

$$\phi(r) = (p-1)(q-1), \tag{10.13}$$

and t can be calculated from the following congruence:

$$t \equiv s^{\phi(\phi(r))-1} \pmod{\phi(r)}, \tag{10.14}$$

because then

$$ts \equiv s^{\phi(\phi(r))} \pmod{\phi(r)}, \tag{10.15}$$

which, according to Euler, is congruent to 1 modulo $\phi(r)$, provided $(s, \phi(r)) = 1$, as initially required of s.

Having demonstrated that t is the proper (and, incidentally, essentially unique) decrypting exponent, decryption now proceeds as before. The received encrypted message is

$$E \equiv M^s \pmod{r}, \tag{10.16}$$

and decrypting with t yields

$$E^t \equiv M^{st} = M^{\phi(r)k+1} \pmod{r}. \tag{10.17}$$

Now, with Euler,

$$M^{\phi(r)} \equiv 1 \pmod{r}. \tag{10.18}$$

Thus,

$$E^t \equiv M \pmod{r}, \tag{10.19}$$

as desired. Note that for the encryption to work, M has to be coprime to r, but the chances of this not being the case are of order 10^{-100}.

The main point here is that the decrypting exponent t cannot be derived from the publicly known s and r, but only from s and the *factors* of r, namely p and q. Thus, as long as the receiver keeps knowledge of p and q to himself and publishes only their product r, he is safe, i.e., the messages he receives cannot be read by unauthorized third parties – *provided that r cannot be factored.*

At present, a 200-digit number that is the product of two 100-digit numbers cannot be factored in any reasonable time – say a couple of minutes. In fact, not so long ago, the most efficient factoring algorithms on a very fast computer were estimated to take 40 trillion years, or 2000 times the present age of the universe. That *sounds* very safe indeed. But algorithms get more efficient by the month and computers become faster and faster every year, and there is no *guarantee* that one day a so-called "polynomial-time" algorithm[1] will not emerge that will allow fast factoring of even 1000-digit numbers [10.4]. Few mathematicians believe that a true polynomial-time algorithm is just around the corner, but there also seems to be no prospect of proving that this will not occur [10.5].

10.4 A Simple Example

Next we will illustrate public-key encryption with a simple example: $s = 7$, $r = 187$. Say the message is $M = 3$; then the encrypted message is

$$E \equiv M^7 = 3^7 = 2187 \equiv 130 \pmod{187}.$$

Now for the receiver to decrypt E, he needs to know the decrypting exponent t, which he can obtain from the factors of $r = 187 = 11 \cdot 17$, which only *he* knows (or that only he would know if r were a much larger number).

With $r = 11 \cdot 17$, he has $\phi(r) = 10 \cdot 16 = 160$ and $\phi(\phi(r)) = 64$, and

$$\begin{aligned} t &\equiv s^{\phi(\phi(r))-1} = 7^{63} = (7^9)^7 = (40\,353\,607)^7 \\ &\equiv 7^7 = 823\,543 \equiv 23 \pmod{160}. \end{aligned}$$

[1] A polynomial-time algorithm is one in which computing time increases like some power or polynomial of the "size" of the problem. Thus, if the polynomial were of the third degree, for example, every doubling of size would increase computation time eightfold.

Knowing that $t = 23$, the receiver can now proceed with this decryption:

$$E^t = 130^{23} \equiv 3 = M \pmod{187}. \qquad \text{Check!}$$

10.5 Repeated Encryption

Here are two more examples which illustrate an important point:

$$\begin{array}{llll} M_1 = 123, & E_1 = 183, & E_1^t \equiv 123 \pmod{187}. & \text{Check!} \\ M_2 = E_1 = 183, & E_2 = 72, & E_2^t \equiv 183 \pmod{187}. & \text{Check!} \end{array}$$

Thus, for the above choice of parameters the encrypting operation is *not* its own inverse. Symbolically:

$$E\{E\{M\}\} \neq M, \quad \text{or} \tag{10.20}$$

$$E^{-1}\{\,\} \neq E\{\,\}. \tag{10.21}$$

This is important, because otherwise an eavesdropper who tried, perhaps accidentally, to double-encrypt with the publicly known key would recover the message.

Because of the uniqueness of the decrypting exponent t, a necessary and sufficient condition for $E^{-1} \neq E$ is that the exponents for encryption and decryption be different: $s \neq t$.

If we continue encrypting $M_1 = 123$, the third round will look like this:

$$M_3 = E_2 = 72, \quad E_3 = 30, \quad E_3^t = 72. \qquad \text{Check!}$$

Thus, for this choice of parameters,

$$E\{E\{E\{M\}\}\} \neq M, \tag{10.22}$$

or $E^3 \neq 1$, for short. Of course, for some number n of repetitions of the encryption process the original message must reappear. Symbolically: $E^n = 1$. The smallest n for which this is true is called the *order* of E. For practical security considerations the order of E should be fairly high. However, it is easy to see that it cannot be larger than $\phi(r)$, because there are only $\phi(r)$ different possible messages (remember: $(M,r) = 1$).

Whether the maximal order $\phi(r)$ can in fact be obtained is another question for which we require the concept of a "primitive element", to be introduced in Chap. 14. In our example, starting with $M = 123$, the fourth round of encrypting yields

$$M_4 = E_3 = 30, \quad E_4 = 123, \quad E_4^t = 30. \qquad \text{Check!}$$

So after 4 successive encryptions, we get the original message back, i.e., the order of our encrypting exponent is 4. And, indeed $7^4 = 2401 \equiv 1 \pmod{\phi(r) = 160}$.

From a practical point of view, such a small order would be completely unacceptable, but then so would our choice of a three-digit integer as encryption modulus.

The encrypting exponent $s = 3$ and its inverse modulo 160, $s^{-1} = 107$, have order 8. Other triplets of exponent, inverse and order are: $(9,89,4)$, $(11,131,8)$, $(13,37,8)$, $(17,113,4)$, $(19,59,8)$, $(21,61,8)$, $(23,7,4)$, $(27,83,8)$, $(29,149,8)$, $(31,31,1)$. As we shall see in Chap. 14, all orders must be some divisor of $\phi(r) = 160$. In the above examples we found the divisors 1, 4 and 8.

Large orders for repeated encryption can be achieved by a judicious choice of prime numbers, p and q, for the encryption modulus $n = pq$. Specifically, the prime numbers p and q should be of the following form

$$p = 2p_1 + 1 \qquad \text{with} \qquad p_1 = 2p_2 + 1,$$

$$\text{and} \tag{10.23}$$

$$q = 2q_1 + 1 \qquad \text{with} \qquad q_1 = 2q_2 + 1,$$

where, p_1, p_2, q_1 and q_2 are also prime numbers.

With such p and q, $\phi(n) = 4p_1q_1$ and $\phi(\phi(n))$, the number of primitive roots (see Sect. 14.1), equals $8p_2q_2$, assuring large orders for almost all encryption exponents. In fact, most exponents will have the *maximal* order of $\phi(\phi(n))/4 = 2p_2q_2$. With p_2 and q_2 of the order of 10^{100} (i.e. n is of the order 10^{200}), the maximal order would thus be about $2 \cdot 10^{200}$ and the probability of an exponent having an order smaller than p_2q_2 would be roughly $1/p_2 + 1/q_2 \approx 2 \cdot 10^{-100}$. In other words, “bad” exponents would be highly unlikely [10.6].

Luckily, there is no dearth of prime numbers of the desirable forms (10.23). The probability of an integer near p_2 being prime is approximately $1/\ln p_2$, and the probabilities that in addition both $p_1 = 2p_2 + 1$ and $p = 2p_1 + 1$ are prime are about $2/\ln 2p_1$ and $2/\ln 4p_1$, respectively. (The two factors of 2 in the numerators stem from the fact that p_1 and p_2 are already by definition odd.) Assuming independence, the total probability of finding a “special” prime p_2 such that p_1 and p are also prime is roughly $4/(\ln 2p_1)^3$. Indeed the range 100 000 to 110 000 contains 861 primes of which 17 are of the desired special form. The theoretical expectations are 865 primes and roughly 22 special primes.

10.6 Summary and Encryption Requirements

For public-key encryption according to the method just outlined, proceed as follows:

(1) Pick two *large* primes p and q and publish their *product* the modulus $r = pq$. This might make factoring difficult and there will be few “forbidden” messages for which $(M, r) \neq 1$.

(2) Select an encrypting exponent s which is coprime to $\phi(r) = (p-1)(q-1)$: $(s, \phi(r)) = 1$. Make sure that the order of s is large, to prevent accidental decryption by repeated encryption.

(3) Calculate the decrypting exponent t and *do not* publish it.

Here the following problem occurs: To determine t from s and p and q we have to calculate

$$t \equiv s^{\phi(\phi(r))-1} \pmod{\phi(r)},$$

where $\phi(r) = (p-1)(q-1)$. To calculate $\phi(\phi(r))$, we have to factor $p-1$ and $q-1$. Now since p and q are large numbers, so are $p-1$ and $q-1$ and the required factoring might be difficult or impossible.[2]

This problem can be avoided by *constructing* $p-1$ and $q-1$ from *known* primes and then testing the resulting p and q for primality. For example, we might try $p-1 = 2 \cdot 3 \cdot 11 = 66$, and 67 is indeed prime. The subject of primality testing, which becomes very important here, will be discussed in Chap. 12.

On the other hand, factoring $p-1$ and $q-1$, even though they may be 100-digit numbers, is in general not as difficult as factoring a 200-digit number, whose smallest prime factor has 100 digits. We will also consider the factoring problem in greater detail in Chap. 12 when we discuss prime divisor functions. We will see there that the median value $p_{0.5}$ of the prime factors of n is given approximately by

$$p_{0.5} \approx \mathrm{e}^{(\ln n/2.81)^{0.5}}.$$

Thus, for $n \approx 10^{100}$, half the prime factors can be expected to be smaller than roughly 9000.

The application of cryptography to computer data security is extensively discussed in [10.8]. Other up-to-date information in this fast-moving field can be found in a new journal, *Cryptologica*, published by Albion College (Albion, MI $2^3 \cdot 3 \cdot 7 \cdot 293$) and in the proceedings of the recurring international symposia on information theory sponsored by the Institute of Electrical and Electronics Engineers (New York, N.Y. 10017).

[2] The is an implied logical contradiction here: if we consider subtraction of 1 not to change the attribute “large”, then by induction we can show that 0 is a large number. However, it is clear here that we only mean that if 10^{50} is large, then so is $10^{50}-1$. A similar quandary, related to the expression “close enough”, is illustrated by the following situation: Suppose all the young gentlemen in a class were to line up on one side of the room, and all the young ladies on the other. At a given signal, the two lines move toward each other, halving the distance between them. At a second signal, they move forward again, halving the remaining distance; and so on at each succeeding signal. Theoretically the boys would never reach the girls; but actually, after a relatively small number of moves, they would be close enough for all practical purposes [10.7].

Chapter 11
The Divisor Functions

Some numbers have few divisors, such as primes, and some numbers have many divisors, such as powers of 2. The number of divisors, although it fluctuates wildly from one integer to the next, obeys some interesting rules, and averages are quite predictable, as we shall see below.

11.1 The Number of Divisors

By definition, $\phi(n)$ is the number of integers in the range from 1 to n which have 1 as greatest common divisor (GCD) with n (see Chap. 8). To codify this statement we introduce the following notation, making use of the number sign #:

$$\phi(n) := \#\{m : 1 \leq m < n,\ (m,n) = 1\}. \tag{11.1}$$

This notation is, of course, equivalent to the more frequently encountered notation using the sum sign:

$$\phi(n) = \sum_{(m,n)=1, m<n} 1.$$

Let (cf. Chap. 2)

$$n = \prod_{p_i|n} p_i^{e_i}. \tag{11.2}$$

Then all divisors of n are of the form

$$d_k = \prod_{p_i|n} p_i^{f_i} \quad \text{with } 0 \leq f_i \leq e_i. \tag{11.3}$$

Here each exponent f_i can take on $e_i + 1$ different values. Thus the number of distinct divisors of n is given by the *divisor function* defined as

$$d(n) := \#\{d_k : d_k|n\},$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_11, © Springer-Verlag Berlin Heidelberg 2009

and equal to

$$d(n) = \prod_i (e_i + 1). \tag{11.4}$$

Example: $n = 12 = 2^2 \cdot 3^1$, $d(12) = 3 \cdot 2 = 6$.

Check: $d_k = 1, 2, 3, 4, 6, 12$, i. e., there are indeed 6 distinct divisors of 12. Check!

In some applications (11.3) is used more than once. Consider a small (elite) university with $N = 3174$ (or was it only 1734) students, subdivided into groups of equal size, each group being cared for by one of x tutors. The tutors in turn are supervised by y professors, each professor looking after the same number of tutors. It is clear that $y|x|N$. Repeated applications of (11.4) show that there are a total of 54 solutions for x and y. Even requiring that there will be more than one professor and more tutors than professors, leaves 31 possibilities. The university president, unhappy with this extravagant freedom of choice, fixed the number z of students in each tutor's group in such a (unique) way that the solution became unique. Not surprisingly, the president, a former mathematician, called this the perfect solution. (But what *is* his unique "perfect" z?)

Now take some m, $1 \leq m \leq n$. Its GCD with n must be one of the divisors of n:

$$(m,n) = d_k \tag{11.5}$$

for some k. How many numbers are there that share the same GCD? We shall denote the size of this family by N_k:

$$N_k := \#\{m : 1 \leq m \leq n,\ (m,n) = d_k\}. \tag{11.6}$$

Of course, by definition of Euler's function, for $d_k = 1$, $N_k = \phi(n)$. We can rewrite the above definition of N_k as follows:

$$N_k = \#\left\{m : 1 \leq m \leq n, \left(\frac{m}{d_k}, \frac{n}{d_k}\right) = 1\right\}, \tag{11.7}$$

and now we see that

$$N_k = \phi\left(\frac{n}{d_k}\right). \tag{11.8}$$

Example: for $n = 12$ and $d_k = 4$, $N_k = \phi(3) = 2$, and there are indeed precisely 2 integers in the range 1 to 12 that share the GCD 2 with 12, namely 2 and 10. Check!

Now each m, $1 \leq m \leq n$, must have one of the $d(n)$ distinct divisors of n as the GCD with n. Hence,

$$\sum_{k=1}^{d(n)} \phi\left(\frac{n}{d_k}\right) = n. \tag{11.9}$$

Reverting to our old notation of summing over all divisors, we may write instead

$$\sum_{d|n} \phi\left(\frac{n}{d}\right) = \sum_{d|n} \phi(d) = n, \tag{11.10}$$

an interesting and important result.

The sum over all divisors of the argument of a number-theoretic function is called its *summatory* function. Thus, the summatory function of Euler's ϕ function is its argument!

Example: $n = 18$:

	Integers m for which $(m,n) = d$						Number of such integers
Divisors $d = 1$	1	5	7	11	13	17	$6 = \phi(18)$
2	2	4	8	10	14	16	$6 = \phi(18/2)$
3	3	15					$2 = \phi(18/3)$
6	6	12					$2 = \phi(18/6)$
9	9						$1 = \phi(18/9)$
18	18						$1 = \phi(18/18)$

The divisor function $d(n)$ is *multiplicative*, i. e., for coprime n and m:

$$d(nm) = d(n) \cdot d(m) \quad \text{for} \quad (n,m) = 1, \tag{11.11}$$

which follows immediately from the formula (11.2) for $d(n)$ in terms of prime exponents.

For the special case that n is the product of k distinct primes, none of which is repeated,

$$d(n = p_1 p_2 \dots p_k) = 2^k. \tag{11.12}$$

Such n are also called *squarefree*, for obvious reasons. For example, 18 is not squarefree, but 30 is, being the product of 3 distinct primes. Thus, 30 has $2^3 = 8$ divisors.

As we saw in Sect. 4.4, the probability of a large integer being squarefree is about $6/\pi^2 \approx 0.61$. Thus, a (narrow) majority of integers are squarefree. In fact, of the 100 integers from 2 to 101, exactly 61 are squarefree. And even among the first 20 integers above 1, the proportion (0.65) is already very close to the asymptotic value. Thus, in this particular area of number theory, 20 is already a large number. [But the reader should be reminded that in other areas (cf. Chap. 4) even $10^{10\,000}$, for example, is not so terribly large.]

Using the notation for the sum over all divisors, we could have introduced $d(n)$ in the following way:

$$d(n) := \sum_{d|n} 1. \tag{11.13}$$

Thus, $d(n)$ is the summary function of n^0.

11.2 The Average of the Divisor Function

Using the Gauss bracket, there is still another way of expressing $d(n)$ which is especially suited for estimating an asymptotic average:

$$d(n) = \sum_{k=1}^{n} \left(\left\lfloor \frac{n}{k} \right\rfloor - \left\lfloor \frac{n-1}{k} \right\rfloor \right). \tag{11.14}$$

Here, if n is divisible by k, then the difference in the parentheses will be 1; otherwise it will be zero. The important point in the above expression is that it is extended over *all* k, not just the divisors of n, as in the definition of $d(n)$.

Now if we sum, we obtain

$$\sum_{n=1}^{N} d(n) = \sum_{k=1}^{N} \left\lfloor \frac{N}{k} \right\rfloor \approx N \sum_{k=1}^{N} \frac{1}{k}. \tag{11.15}$$

Of course, the estimate on the right is an upper limit, because by dropping the Gauss bracket we have increased (by less than 1) all summands for which N is not divisible by k. However, for large N, this increase should be relatively small. Hence we expect the average value to go with the sum of the reciprocal integers, i. e., the logarithm

$$\frac{1}{N} \sum_{n=1}^{N} d(n) \approx \ln N. \tag{11.16}$$

The exact results is [10.1]

$$\frac{1}{N} \sum_{n=1}^{N} d(n) = \ln N + 2\gamma - 1 + 0\left(\frac{1}{\sqrt{N}} \right). \tag{11.17}$$

Here γ is Euler's constant:

$$\gamma := \lim_{n \to \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} - \ln n \right) = 0.57721 \ldots, \tag{11.18}$$

which makes the additive constant in (11.17) about 0.15442. The notation $0(1/\sqrt{N})$ means that the absolute error in (11.17) is smaller than $c/\sqrt{N}$, where c is some constant.

11.3 The Geometric Mean of the Divisors

There is also a nice formula for the *product* of all divisors of a given integer n. With

$$n = \prod_{p_i | n} p_i^{e_i}, \tag{11.19}$$

we have

$$\prod_{d|n} d = n^{\frac{1}{2}d(n)}. \tag{11.20}$$

Example: $n = 12$. The product of divisors equals $1 \cdot 2 \cdot 3 \cdot 4 \cdot 6 \cdot 12 = 1728$. The number of divisors $d(12) = 6$ and $1728 = 12^3$. Check!

To obtain the geometric mean of the divisors of n, we have to take the $d(n)$-th root of their product, giving $\sqrt{n}$ according to (11.20).

A curious result? Not really! Divisors come in *pairs*: if d divides n, so does its distinct "mate" n/d (exception: for $n = d^2$, the mate is not distinct). And the geometric mean of each of these pairs equals $\sqrt{n}$, and so does the overall geometric mean.

11.4 The Summatory Function of the Divisor Function

The summary function

$$\sigma(n) := \sum_{d|n} d, \tag{11.21}$$

like all summatory functions of multiplicative functions, is multiplicative. Thus, it suffices to consider the problem first only for n that are powers of a single prime and then to multiply the individual results. This yields with (11.3):

$$\sigma(n) = \prod_{p_i|n} \frac{p_i^{1+e_i} - 1}{p_i - 1}. \tag{11.22}$$

The asymptotic behaviour of $\sigma(n)$ is given by [11.1]:

$$\frac{1}{N^2} \sum_{n=1}^{N} \sigma(n) = \frac{\pi^2}{12} + 0\left(\frac{\ln N}{N}\right), \tag{11.23}$$

a result that can be understood by dropping the -1 in the denominator in each term of the product (11.23) and converting it into a product over *all* primes with the proper probabilistic factors, and proceeding as in the case of $\phi(n)/n$ (see Sect. 11.6).

The above expression converges quite rapidly. For example, for $N = 5$, the result is 0.84 compared to $\pi^n/12 \approx 0.82$.

11.5 The Generalized Divisor Functions

Generalized divisor functions are defined as follows:

$$\sigma_k(n) := \sum_{d|n} d^k. \tag{11.24}$$

Of course, $\sigma_o(n) = d(n)$ and $\sigma_1(n) = \sigma(n)$. These generalized divisor functions obey a simple symmetry with respect to their index:

$$\sigma_k(n) = \sum_{d|n} d^k = \sum_{d|n} \left(\frac{d}{n}\right)^{-k} = n^k \sigma_{-k}(n). \tag{11.25}$$

Example: $n = 6$, $k = 3$. Divisors $d = 1, 2, 3, 6$.

$$1 + 8 + 27 + 216 = 216\left(1 + \frac{1}{8} + \frac{1}{27} + \frac{1}{216}\right).$$

The above symmetry relation, for $k = 1$, leads to the following relation between divisor means of n: Arithmetic mean $\overline{d}$ times harmonic mean $\hat{d}$ equals geometric mean $\tilde{d}$ squared equals n:

$$\overline{d}\hat{d} = \tilde{d}^2 = n. \tag{11.26}$$

Example: $n = 6$; $d = 1, 2, 3, 6$. $\overline{d} = 3$, $\hat{d} = 2$, $\tilde{d} = \sqrt{6}$, $3 \cdot 2 = (\sqrt{6})^2 = 6$. Check!

11.6 The Average Value of Euler's Function

Euler's ϕ function is a pretty "wild" function. For example,

$$\phi(29) = 28, \quad \phi(30) = 8, \quad \phi(31) = 30, \quad \phi(32) = 16.$$

If we are interested in the asymptotic behaviour of $\phi(n)$, we had better consider some *average* value. The following probabilistic argument will give such an average automatically, because our probabilities ignore fine-grain fluctuations such as those in the above numerical example.

Consider first

$$\frac{\phi(n)}{n} = \prod_{p_i|n} \left(1 - \frac{1}{p_i}\right). \tag{11.27}$$

Using our by now customary (but of course unproved) probabilistic argument, we convert the above product over primes that divide n into a product over *all* primes. The probability that "any old" prime will divide n equals $1/p_i$, and the probability that it will *not* equals $1 - 1/p_i$. In that case, the prime p_i "contributes" the factor 1 to the product. Thus, we may write

$$\begin{aligned}\prod_{p_i|n} \left(1 - \frac{1}{p_i}\right) &\approx \prod_{p_i} \left[\left(1 - \frac{1}{p_i}\right)\frac{1}{p_i} + 1\left(1 - \frac{1}{p_i}\right)\right] \qquad (11.28)\\ &= \prod_{p_i} \left(1 - \frac{1}{p_i^2}\right),\end{aligned}$$

an infinite product that we have encountered before and which we calculated by converting the reciprocal of each factor into an infinite geometric series and then multiplying everything out. This produces every squared integer exactly once. Thus, we find

$$\text{average of } \frac{\phi(n)}{n} \approx \left(\sum_{n=1}^{\infty}\frac{1}{n^2}\right)^{-1} = \frac{6}{\pi^2}. \tag{11.29}$$

The result corresponds to the asymptotic probability that two arbitrarily selected integers are coprime – as it should, if we remember the definition of $\phi(n)$. In fact, the number of white dots in Fig. 4.8 in a vertical line up to the 45° diagonal equals Euler's function for that coordinate.

The formal results for the asymptotic behaviour of $\phi(n)$ are as follows [11.1]:

$$\frac{1}{n}\sum_{k=1}^{n}\frac{\phi(k)}{k} = \frac{6}{\pi^2} + 0\left(\frac{\ln n}{n}\right), \quad \text{and} \tag{11.30}$$

$$\frac{1}{n^2}\sum_{k=1}^{n}\phi(k) = \frac{3}{\pi^2} + 0\left(\frac{\ln n}{n^2}\right). \tag{11.31}$$

The result (11.30) corresponds of course to our probabilistic estimate, and the formula (11.31) is likewise unsurprising because the *average* factor inside the sum in (11.31) compared to the sum (11.30) is $n/2$.

Example: $n = 4$

$$\frac{1}{n}\sum_{k=1}^{n}\frac{\phi(k)}{k} = \frac{1}{4}\left(1+\frac{1}{2}+\frac{2}{3}+\frac{1}{2}\right) = 0.667,$$

as compared to the asymptotic value 0.608. And

$$\frac{1}{n^2}\sum_{k=1}^{n}\phi(k) = \frac{1}{16}(1+1+2+2) = 0.375,$$

which also compares well with the asymptotic value (0.304).

What are the probabilities that *three* integers will not have a common divisor? And what is the probability that each of the three *pairs* that can be formed with three integers will be made up of coprime integers? The reader can find the (simple) answers or look them up in Sect. 4.4.

Chapter 12
The Prime Divisor Functions

Here we consider only *prime* divisors of n and ask, for given order of magnitude of n, "how many prime divisors are there typically?" and "how many *different* ones are there?" Some of the answers will be rather counterintuitive. Thus, a 50-digit number (10^{21} times the age of our universe measured in picoseconds) has only about 5 different prime factors on average and – even more surprisingly – 50-digit numbers have typically fewer than 6 prime factors in all, even counting repeated occurrences of the same prime factor as separate factors.

We will also learn something about the distribution of the number of prime factors and its implications for the important factoring problem. Thus, we discover that even for numbers as large as 10^{50}, the two smallest primes, 2 and 3, account for about 25% of all prime factors!

12.1 The Number of Different Prime Divisors

In connection with encrypting messages by means of Euler's theorem, the number of distinct *prime* divisors of a given integer n, $\omega(n)$, is of prime importance. Its definition is similar to that of the divisor function $d(n)$, except that the sum is extended – as the name implies – only over the prime divisors of n:

$$\omega(n) := \sum_{p_i|n} 1. \tag{12.1}$$

It is easily seen that $\omega(n)$ is additive, i. e., for $(n,m) = 1$,

$$\omega(nm) = \sum_{p_i|nm} 1 = \sum_{p_i|n} 1 + \sum_{p_i|m} 1 = \omega(n) + \omega(m). \tag{12.2}$$

Of particular interest to our encrypting desires will be the behaviour of $\omega(n)$ for large n, i.e., its asymptotic behaviour. We shall try to get an idea of this behaviour by means of our usual "dirty tricks". First, we will convert the sum of those primes that divide n into a sum over *all* primes up to n, using the "probability" factor $1/p_i$:

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_12, © Springer-Verlag Berlin Heidelberg 2009

$$\omega(n) = \sum_{p_i|n} 1 \approx \sum_{p_i \le n} \frac{1}{p_i}. \tag{12.3}$$

This, in turn, we will convert into a sum over all *integers* up to n, using the probability factor for primality $1/\ln x$:

$$\overline{\omega}(n) \approx \sum_{x \le n} \frac{1}{x \ln x},$$

which we will approximate by an integral:

$$\overline{\omega}(n) \approx \int_2^n \frac{\mathrm{d}x}{x \ln x} = \ln(\ln n) + 0.367\ldots. \tag{12.4}$$

Of course $\omega(n)$ is a wildly fluctuating function and exact results [12.1] are available only for asymptotic averages, just as in the case of $\phi(n)$ and $d(n)$:

$$\frac{1}{n} \sum_{k=1}^{n} \omega(k) = \ln(\ln n) + 0(1), \tag{12.5}$$

where $0(1)$ is a fancy way of writing a bounded quantity.

To get a better grip on this constant, we calculate the sum over the reciprocal primes in (12.3) out to some p_m and convert only the remaining sum to a sum over all integers using the probability factor $\ln x$:

$$\overline{\omega}(n) \approx \sum_{p_i=2}^{m} \frac{1}{p_i} + \sum_{x=p_m+1}^{n} \frac{1}{x \ln x}. \tag{12.6}$$

Approximating the second sum by an integral, we have

$$\overline{\omega}(n) \approx \sum_{p_i=2}^{p_m} \frac{1}{p_i} + \ln\ln n - \ln\ln p_m. \tag{12.7}$$

In other words, our estimate tells us that the difference between $\overline{\omega}(n)$ and $\ln\ln n$, i.e., the constant in (12.5), is given by

$$\overline{\omega}(n) - \ln\ln n \approx \lim_{p_m \to \infty} \sum_{p_i=2}^{p_m} \frac{1}{p_i} - \ln\ln p_m. \tag{12.8}$$

In the last century Kronecker, assuming that the limiting average of $\omega(n)$ existed, obtained

$$\overline{\omega}(n) = \ln(\ln n) + b_1, \tag{12.9}$$

with

$$b_1 = \gamma + \sum_{p_i=2}^{\infty} \left[\ln\left(1 - \frac{1}{p_i}\right) + \frac{1}{p_i} \right], \tag{12.10}$$

where γ is again Euler's constant.

To compare Kronecker's constant b_1 with ours, we make use of the following asymptotic result (Merten's theorem [12.1]):

$$\lim_{p_m \to \infty} \mathrm{e}^{\gamma} \ln p_m \prod_{p_i=2}^{p_m} \left(1 - \frac{1}{p_i}\right) = 1, \tag{12.11}$$

which yields for Kronecker's constant

$$b_1 = \lim_{p_m \to \infty} \sum_{p_i=2}^{p_m} \frac{1}{p_i} - \ln(\ln p_m), \tag{12.12}$$

which is identical with our "crude" estimate (12.8)!

Equation (12.12) is not very suitable to obtain a numerical value for b_1, because it converges rather slowly. (In fact, even for p_m as large as 104759, the relative error is still larger than 10^{-3}.) A faster converging series is obtained by expanding the logarithm in (12.10), which yields

$$\gamma - b_1 = \sum_{p_i=2}^{\infty} \left(\frac{1}{2p_i^2} + \frac{1}{3p_i^3} + \dots \right). \tag{12.13}$$

Now, if we remember the Riemann zetafunction (Chap. 4), we have

$$\zeta(k) = \sum_{n=1}^{\infty} \frac{1}{n^k} = \prod_{p_i=2}^{\infty} \left(1 - \frac{1}{p_i^k}\right)^{-1}, \tag{12.14}$$

or

$$\ln \zeta(k) = -\sum_{p_i=2}^{\infty} \ln\left(1 - \frac{1}{p_i^k}\right). \tag{12.15}$$

Expanding the logarithm, we obtain

$$\ln \zeta(k) = \sum_{p_i=2}^{\infty} \left(\frac{1}{p_i^k} + \frac{1}{2p_i^{2k}} + \dots \right). \tag{12.16}$$

Introducing this result into (12.13) yields

$$\gamma - b_1 = \tfrac{1}{2} \ln \zeta(2) + \tfrac{1}{3} \ln \zeta(3) + \tfrac{1}{5} \ln \zeta(5) - \tfrac{1}{6} \ln \zeta(6) + \dots. \tag{12.17}$$

This sum written in terms of the Möbius function $\mu(m)$ (Chap. 21) is:

$$\gamma - b_1 = -\sum_{m=2}^{\infty} \frac{\mu(m)}{m} \ln \zeta(m). \tag{12.18}$$

This sum converges very quickly and, for just 7 terms yields a relative accuracy of about 10^{-5}. The result is

$$b_1 = 0.2614\ldots. \tag{12.19}$$

How do Milton Abramowitz and Irene Stegun feel about this? On page 862 of their *Handbook of Mathematical Functions* [12.2] they list the prime factors of the integers from 9000 to 9499 (see Fig. 12.1). I have counted a total of 1260 distinct prime factors for these 500 integers. Thus, $\overline{\omega} = 2.52$, which should be compared to our $\ln(\ln 9250) + 0.26 = 2.47$. Close enough? Certainly, because as we said before, $\omega(n)$ fluctuates and an average, even over 500 consecutive integers, is not completely smooth. (More about the fluctuations of $\omega(n)$ in a moment.)

12.2 The Distribution of $\boldsymbol{\omega(n)}$

The probability that the prime factor p_i does not occur in the prime factor decomposition of $n > p_i$ is given by

$$1 - \frac{1}{p_i}.$$

The probability that it *does* occur (at least once) is therefore

$$\frac{1}{p_i}.$$

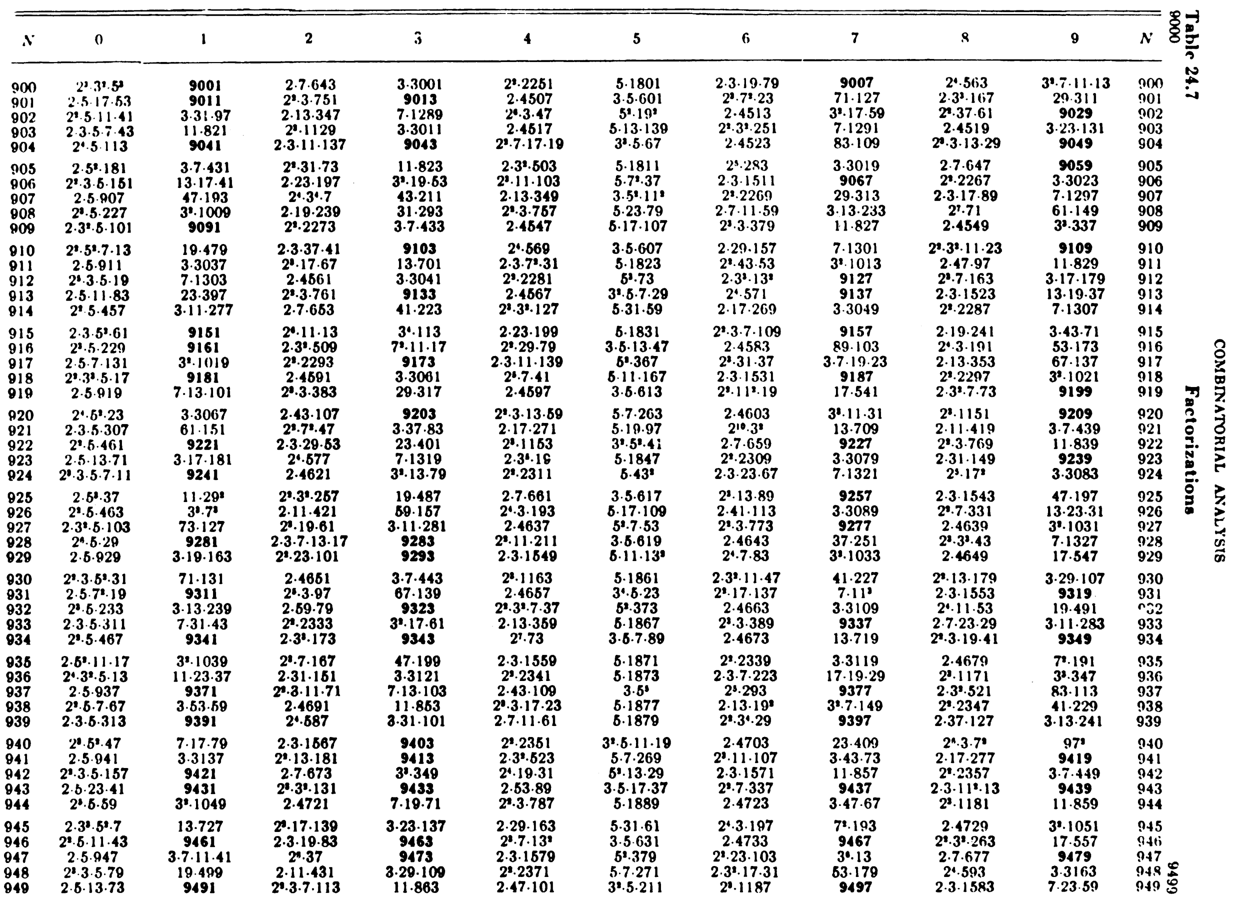
The mean occurrence is therefore

$$m_i = \frac{1}{p_i}, \tag{12.20}$$

and its variance, according to the formula for the binomial distribution for two possible outcomes, equals

$$\sigma_i^2 = \frac{1}{p_i}\left(1 - \frac{1}{p_i}\right) = m_i - \frac{1}{p_i^2}. \tag{12.21}$$

Assuming divisibility by different primes to be independent, we get for the overall mean



COMBINATORIAL ANALYSIS

Table 24.7 **Factorizations**

9000 9499

N	0	1	2	3	4	5	6	7	8	9	*N*
900	$2^3\cdot3^2\cdot5^3$	**9001**	2·7·643	3·3001	$2^2\cdot2251$	5·1801	2·3·19·79	**9007**	$2^4\cdot563$	$3^2\cdot7\cdot11\cdot13$	900
901	2·5·17·53	**9011**	$2^2\cdot3\cdot751$	**9013**	2·4507	3·5·601	$2^3\cdot7^2\cdot23$	71·127	$2\cdot3^3\cdot167$	29·311	901
902	$2^2\cdot5\cdot11\cdot41$	3·31·97	2·13·347	7·1289	$2^6\cdot3\cdot47$	$5^2\cdot19^2$	2·4513	$3^2\cdot17\cdot59$	$2^2\cdot37\cdot61$	**9029**	902
903	2·3·5·7·43	11·821	$2^3\cdot1129$	3·3011	2·4517	5·13·139	$2^2\cdot3^2\cdot251$	7·1291	2·4519	3·23·131	903
904	$2^4\cdot5\cdot113$	**9041**	2·3·11·137	**9043**	$2^2\cdot7\cdot17\cdot19$	$3^3\cdot5\cdot67$	2·4523	83·109	$2^3\cdot3\cdot13\cdot29$	**9049**	904
905	$2\cdot5^2\cdot181$	3·7·431	$2^2\cdot31\cdot73$	11·823	$2\cdot3^2\cdot503$	5·1811	$2^5\cdot283$	3·3019	2·7·647	**9059**	905
906	$2^2\cdot3\cdot5\cdot151$	13·17·41	2·23·197	$3^2\cdot19\cdot53$	$2^3\cdot11\cdot103$	$5\cdot7^2\cdot37$	2·3·1511	**9067**	$2^2\cdot2267$	3·3023	906
907	2·5·907	47·193	$2^4\cdot3^4\cdot7$	43·211	2·13·349	$3\cdot5^2\cdot11^2$	$2^2\cdot2269$	29·313	2·3·17·89	7·1297	907
908	$2^3\cdot5\cdot227$	$3^2\cdot1009$	2·19·239	31·293	$2^2\cdot3\cdot757$	5·23·79	2·7·11·59	3·13·233	$2^7\cdot71$	61·149	908
909	$2\cdot3^2\cdot5\cdot101$	**9091**	$2^2\cdot2273$	3·7·433	2·4547	5·17·107	$2^3\cdot3\cdot379$	11·827	2·4549	$3^3\cdot337$	909
910	$2^2\cdot5^2\cdot7\cdot13$	19·479	2·3·37·41	**9103**	$2^4\cdot569$	3·5·607	2·29·157	7·1301	$2^2\cdot3^2\cdot11\cdot23$	**9109**	910
911	2·5·911	3·3037	$2^3\cdot17\cdot67$	13·701	$2\cdot3\cdot7^2\cdot31$	5·1823	$2^2\cdot43\cdot53$	$3^2\cdot1013$	2·47·97	11·829	911
912	$2^5\cdot3\cdot5\cdot19$	7·1303	2·4561	3·3041	$2^2\cdot2281$	$5^3\cdot73$	$2\cdot3^3\cdot13^2$	**9127**	$2^3\cdot7\cdot163$	3·17·179	912
913	2·5·11·83	23·397	$2^2\cdot3\cdot761$	**9133**	2·4567	$3^2\cdot5\cdot7\cdot29$	$2^4\cdot571$	**9137**	2·3·1523	13·19·37	913
914	$2^2\cdot5\cdot457$	3·11·277	2·7·653	41·223	$2^3\cdot3^2\cdot127$	5·31·59	2·17·269	3·3049	$2^2\cdot2287$	7·1307	914
915	$2\cdot3\cdot5^2\cdot61$	**9151**	$2^6\cdot11\cdot13$	$3^4\cdot113$	2·23·199	5·1831	$2^2\cdot3\cdot7\cdot109$	**9157**	2·19·241	3·43·71	915
916	$2^3\cdot5\cdot229$	**9161**	$2\cdot3^2\cdot509$	$7^2\cdot11\cdot17$	$2^2\cdot29\cdot79$	3·5·13·47	2·4583	89·103	$2^4\cdot3\cdot191$	53·173	916
917	2·5·7·131	$3^2\cdot1019$	$2^2\cdot2293$	**9173**	2·3·11·139	$5^2\cdot367$	$2^3\cdot31\cdot37$	3·7·19·23	2·13·353	67·137	917
918	$2^2\cdot3^3\cdot5\cdot17$	**9181**	2·4591	3·3061	$2^5\cdot7\cdot41$	5·11·167	2·3·1531	**9187**	$2^2\cdot2297$	$3^2\cdot1021$	918
919	2·5·919	7·13·101	$2^3\cdot3\cdot383$	29·317	2·4597	3·5·613	$2^2\cdot11^2\cdot19$	17·541	$2\cdot3^2\cdot7\cdot73$	**9199**	919
920	$2^4\cdot5^2\cdot23$	3·3067	2·43·107	**9203**	$2^2\cdot3\cdot13\cdot59$	5·7·263	2·4603	$3^3\cdot11\cdot31$	$2^3\cdot1151$	**9209**	920
921	2·3·5·307	61·151	$2^2\cdot7^2\cdot47$	3·37·83	2·17·271	5·19·97	$2^{10}\cdot3^2$	13·709	2·11·419	3·7·439	921
922	$2^2\cdot5\cdot461$	**9221**	2·3·29·53	23·401	$2^3\cdot1153$	$3^2\cdot5^2\cdot41$	2·7·659	**9227**	$2^2\cdot3\cdot769$	11·839	922
923	2·5·13·71	3·17·181	$2^4\cdot577$	7·1319	$2\cdot3^5\cdot19$	5·1847	$2^2\cdot2309$	3·3079	2·31·149	**9239**	923
924	$2^3\cdot3\cdot5\cdot7\cdot11$	**9241**	2·4621	$3^2\cdot13\cdot79$	$2^2\cdot2311$	$5\cdot43^2$	2·3·23·67	7·1321	$2^5\cdot17^2$	3·3083	924
925	$2\cdot5^3\cdot37$	$11\cdot29^2$	$2^2\cdot3^2\cdot257$	19·487	2·7·661	3·5·617	$2^3\cdot13\cdot89$	**9257**	2·3·1543	47·197	925
926	$2^2\cdot5\cdot463$	$3^3\cdot7^3$	2·11·421	59·157	$2^4\cdot3\cdot193$	5·17·109	2·41·113	3·3089	$2^2\cdot7\cdot331$	13·23·31	926
927	$2\cdot3^2\cdot5\cdot103$	73·127	$2^3\cdot19\cdot61$	3·11·281	2·4637	$5^2\cdot7\cdot53$	$2^2\cdot3\cdot773$	**9277**	2·4639	$3^2\cdot1031$	927
928	$2^6\cdot5\cdot29$	**9281**	2·3·7·13·17	**9283**	$2^2\cdot11\cdot211$	3·5·619	2·4643	37·251	$2^3\cdot3^3\cdot43$	7·1327	928
929	2·5·929	3·19·163	$2^2\cdot23\cdot101$	**9293**	2·3·1549	$5\cdot11\cdot13^2$	$2^4\cdot7\cdot83$	$3^2\cdot1033$	2·4649	17·547	929
930	$2^2\cdot3\cdot5^2\cdot31$	71·131	2·4651	3·7·443	$2^3\cdot1163$	5·1861	$2\cdot3^2\cdot11\cdot47$	41·227	$2^2\cdot13\cdot179$	3·29·107	930
931	$2\cdot5\cdot7^2\cdot19$	**9311**	$2^5\cdot3\cdot97$	67·139	2·4657	$3^4\cdot5\cdot23$	$2^2\cdot17\cdot137$	$7\cdot11^3$	2·3·1553	**9319**	931
932	$2^3\cdot5\cdot233$	3·13·239	2·59·79	**9323**	$2^2\cdot3^2\cdot7\cdot37$	$5^2\cdot373$	2·4663	3·3109	$2^4\cdot11\cdot53$	19·491	932
933	2·3·5·311	7·31·43	$2^2\cdot2333$	$3^2\cdot17\cdot61$	2·13·359	5·1867	$2^3\cdot3\cdot389$	**9337**	2·7·23·29	3·11·283	933
934	$2^2\cdot5\cdot467$	**9341**	$2\cdot3^3\cdot173$	**9343**	$2^7\cdot73$	3·5·7·89	2·4673	13·719	$2^2\cdot3\cdot19\cdot41$	**9349**	934
935	$2\cdot5^2\cdot11\cdot17$	$3^2\cdot1039$	$2^3\cdot7\cdot167$	47·199	2·3·1559	5·1871	$2^2\cdot2339$	3·3119	2·4679	$7^2\cdot191$	935
936	$2^4\cdot3^2\cdot5\cdot13$	11·23·37	2·31·151	3·3121	$2^2\cdot2341$	5·1873	2·3·7·223	17·19·29	$2^3\cdot1171$	$3^3\cdot347$	936
937	2·5·937	**9371**	$2^2\cdot3\cdot11\cdot71$	7·13·103	2·43·109	$3\cdot5^5$	$2^5\cdot293$	**9377**	$2\cdot3^2\cdot521$	83·113	937
938	$2^2\cdot5\cdot7\cdot67$	3·53·59	2·4691	11·853	$2^3\cdot3\cdot17\cdot23$	5·1877	$2\cdot13\cdot19^2$	$3^2\cdot7\cdot149$	$2^2\cdot2347$	41·229	938
939	2·3·5·313	**9391**	$2^4\cdot587$	3·31·101	2·7·11·61	5·1879	$2^2\cdot3^4\cdot29$	**9397**	2·37·127	3·13·241	939
940	$2^3\cdot5^2\cdot47$	7·17·79	2·3·1567	**9403**	$2^2\cdot2351$	$3^2\cdot5\cdot11\cdot19$	2·4703	23·409	$2^6\cdot3\cdot7^2$	97^2	940
941	2·5·941	3·3137	$2^2\cdot13\cdot181$	**9413**	$2\cdot3^2\cdot523$	5·7·269	$2^3\cdot11\cdot107$	3·43·73	2·17·277	**9419**	941
942	$2^2\cdot3\cdot5\cdot157$	**9421**	2·7·673	$3^3\cdot349$	$2^4\cdot19\cdot31$	$5^2\cdot13\cdot29$	2·3·1571	11·857	$2^2\cdot2357$	3·7·449	942
943	2·5·23·41	**9431**	$2^3\cdot3^2\cdot131$	**9433**	2·53·89	3·5·17·37	$2^2\cdot7\cdot337$	**9437**	$2\cdot3\cdot11^2\cdot13$	**9439**	943
944	$2^5\cdot5\cdot59$	$3^2\cdot1049$	2·4721	7·19·71	$2^2\cdot3\cdot787$	5·1889	2·4723	3·47·67	$2^3\cdot1181$	11·859	944
945	$2\cdot3^3\cdot5^2\cdot7$	13·727	$2^2\cdot17\cdot139$	3·23·137	2·29·163	5·31·61	$2^4\cdot3\cdot197$	$7^2\cdot193$	2·4729	$3^2\cdot1051$	945
946	$2^2\cdot5\cdot11\cdot43$	**9461**	2·3·19·83	**9463**	$2^3\cdot7\cdot13^2$	3·5·631	2·4733	**9467**	$2^2\cdot3^2\cdot263$	17·557	946
947	2·5·947	3·7·11·41	$2^8\cdot37$	**9473**	2·3·1579	$5^2\cdot379$	$2^2\cdot23\cdot103$	$3^6\cdot13$	2·7·677	**9479**	947
948	$2^3\cdot3\cdot5\cdot79$	19·499	2·11·431	3·29·109	$2^2\cdot2371$	5·7·271	$2\cdot3^2\cdot17\cdot31$	53·179	$2^4\cdot593$	3·3163	948
949	2·5·13·73	**9491**	$2^2\cdot3\cdot7\cdot113$	11·863	2·47·101	$3^2\cdot5\cdot211$	$2^3\cdot1187$	**9497**	2·3·1583	7·23·59	949

Fig. 12.1 The prime factors of n in the range $9000 \leq n \leq 9499$. The number of distinct prime factors in this range is 1260; the corresponding theoretical expectation equals 1237 ± 32. The number of prime factors, including multiple occurrences, is 1650, compared to a theoretical expectation of 1632 ± 31

$$\overline{\varpi}(n) \approx \sum_{p_i<n} \frac{1}{p_i} \approx \ln(\ln n) + 0.2614, \tag{12.22}$$

as before [see (12.9) and (12.19)]. The overall variance becomes, with (12.21),

$$\sigma_\omega^2 \approx \overline{\varpi}(n) - \sum_{p_i=2}^{\infty} \frac{1}{p_i^2}, \tag{12.23}$$

where we have extended the sum out to infinity because it converges quite rapidly.

The numerical value of the sum can be obtained most efficiently with the help of Riemann's zetafunction, expanded as in (12.16). This yields

$$\begin{aligned} \sum_{p_i=2}^{\infty} \frac{1}{p_i^2} &= \ln \zeta(2) - \frac{1}{2} \ln \zeta(4) - \ldots \\ &= \sum_{m=1}^{\infty} \frac{\mu(m)}{m} \ln \zeta(2m) \approx 0.452248, \end{aligned} \tag{12.24}$$

where $\mu(m)$ is again the Möbius function (see Chap. 21).

Thus,

$$\sigma_\omega^2 \approx \overline{\varpi}(n) - 0.45 \tag{12.25}$$

and, because $\sigma_\omega^2 \approx \overline{\varpi}$, we expect ω to be approximately *Poisson* distributed [12.3]. Of course, each number has at least *one* prime factor (itself, if it is prime), so that the Poisson distribution must be shifted by 1:

$$\mathrm{Prob}\{\omega(n) = k\} \approx \frac{(\overline{\varpi} - 1)^{k-1}}{(k-1)!} \mathrm{e}^{-\overline{\varpi}+1}, \quad k = 1, 2, \ldots, \overline{\varpi} > 1, \tag{12.26}$$

with $\overline{\varpi}$ from (12.22).

The mode (most probable value) of this distribution occurs for

$$\check{k} = \lfloor \overline{\varpi} \rfloor + 1, \tag{12.27}$$

where $\check{k}$ is read "kay check". Although intended for large n, (12.27) seems to work very well even for small n. Equation (12.27) predicts that the most probable number $\check{k}$ of different prime factors of n is as follows:

$$\begin{aligned} &\check{k} = 1 \text{ for} && n < 9 \\ &\check{k} = 2 \text{ for} && 9 \leq n < 296 \\ &\check{k} = 3 \text{ for} && 296 \leq n < 5 \cdot 10^6 \\ &\check{k} = 4 \text{ for} && 5 \cdot 10^6 \leq n < 2 \cdot 10^{18} \\ &\check{k} = 5 \text{ for} && 2 \cdot 10^{18} \leq n < 4 \cdot 10^{49} \\ &\check{k} = 6 \text{ for} && 4 \cdot 10^{49} \leq n < 8 \cdot 10^{134} \text{ etc.} \end{aligned}$$

Thus, up to almost 10^{135} the most likely number of different prime factors is 6 or less!

According to (12.26), the probability that n has exactly one prime factor, i.e., that n is either a prime or a prime power, equals about $2/\ln n$. This value is somewhat larger than the one we would expect from the distribution of primes. But then, we should not expect the Poisson distribution for $\omega(n)$ to be exact. For example, σ_ω^2 should equal $\overline{\omega}-1$ for the shifted Poisson distribution and not $\overline{\omega}-0.45$ as in (12.23) and (12.24).

12.3 The Number of Prime Divisors

Apart from the "little" $\overline{\omega}(n)$ we need a "big" $\Omega(n)$, the number of prime divisors of n, counted with multiplicity. For

$$n = \prod_{p_i|n} p_i^{e_i}, \tag{12.28}$$

we have the definition

$$\Omega(n) := \sum_{p_i|n} e_i. \tag{12.29}$$

The divisor function $\Omega(n)$ is *completely* additive, i.e.,

$$\Omega(mn) = \Omega(m) + \Omega(n), \tag{12.30}$$

whether m and n are coprime or not.

To estimate an average value of $\Omega(n)$, we convert the sum appearing in its definition into a sum over all primes up to n:

$$\Omega(n) \approx \sum_{p_i \le n} e_i \frac{1}{p_i^{e_i}} \left(1 - \frac{1}{p_i}\right), \tag{12.31}$$

recognizing that the probability that p_i occurs e_i times equals $(1-1/p_i)/p_i^{e_i}$. Averaging over these values of e_i yields

$$\overline{\Omega}(n) \approx \sum_{p_i \le n} \frac{1}{p_i - 1}. \tag{12.32}$$

Note the closeness of our estimates of $\overline{\Omega}(n)$ and $\overline{\omega}(n)$ according to (12.3)! The difference (which some friends did not even think converged) is given by

$$\overline{\Omega}(n) - \overline{\omega}(n) \approx \sum_{p_i \le n} \frac{1}{p_i(p_i - 1)}, \tag{12.33}$$

in agreement with a result by Kronecker. (This sum is upperbounded by the sum over all integers out to infinity, which equals 1.)

Since the sum does not only converge, but converges quite rapidly, we will only bother about its value taken out to infinity. First we write

$$\sum_{p_i=2}^{\infty} \frac{1}{p_i(p_i-1)} = \sum_{p_i=2}^{\infty} \left(\frac{1}{p_i^2} + \frac{1}{p_i^3} + \frac{1}{p_i^4} + \dots \right) \tag{12.34}$$

and then introduce the zetafunction again, making use of (12.16). This yields

$$\begin{aligned} \sum_{p_i=2}^{\infty} \frac{1}{p_i(p_i-1)} &= \ln\zeta(2) + \ln\zeta(3) \\ &+ \frac{1}{2}\ln\zeta(4) + \ln\zeta(5) \\ &+ \frac{1}{6}\ln\zeta(6) + \dots \approx 0.77317 \quad \text{or} \end{aligned} \tag{12.35}$$

$$\overline{\Omega}(n) \approx \overline{\omega}(n) + 0.77317 \approx \ln(\ln n) + 1.0346. \tag{12.36}$$

What do Abramowitz and Stegun have to say? In their table of prime factors for n in the range 9000 to 9499 [12.2, p. 862], I counted a total of 1650 prime factors, including multiplicity, yielding $\overline{\Omega} = 3.30$. Our theoretical value $\ln(\ln 9250) + 1.0346 \approx 3.25$, which is as similar as could be expected.

Incidentally, sums taken over all primes, with primes appearing in the denominator as in (12.13) and (12.34), need not always lead to irrational results. A noteworthy counterexample (from an entire family of like-fashioned expressions) is

$$\prod_{p=2}^{\infty} \frac{p^2+1}{p^2-1} = \frac{5}{2}. \tag{12.37}$$

This seems preposterous, but a quick numerical check indicates that the product certainly could not deviate much from $5/2$, and in fact, the infinite product *does* equal $5/2$. This is actually not too difficult to see, because

$$\prod \frac{p^2+1}{p^2-1} = \prod \frac{p^4-1}{(p^2-1)^2} = \prod \frac{1-\dfrac{1}{p^4}}{\left(1-\dfrac{1}{p^2}\right)^2},$$

or, expanding into geometric series:

$$\prod \frac{p^2+1}{p^2-1} = \frac{\prod \left(1 + \dfrac{1}{p^2} + \dfrac{1}{p^4} + \dots\right)^2}{\prod \left(1 + \dfrac{1}{p^4} + \dfrac{1}{p^8} + \dots\right)} = \frac{\left(\sum\limits_{n=1}^{\infty} \dfrac{1}{n^2}\right)^2}{\sum\limits_{n=1}^{\infty} \dfrac{1}{n^4}}. \tag{12.38}$$

We have encountered the sum in the numerator several times before (Chaps. 4, 8), and found it to equal $\pi^2/6$. The sum in the denominator equals $\zeta(4) = \pi^4/90$, and if we had not heard of the zetafunction, we could find out by calculating a certain definite integral over the Fourier series

$$\sin x - \sin 3x + \sin 5x - \ldots.$$

(The reader may want to try this.) The result is

$$\prod_{p=2}^{\infty} \frac{p^2+1}{p^2-1} = \frac{\frac{\pi^4}{36}}{\frac{\pi^4}{90}} = \frac{5}{2}.$$

Consideration of this product also leads to some rather unexpected relations for $\Omega(n)$. Expanding

$$\prod \frac{p^2+1}{p^2-1} = \frac{\Pi\left(1+\frac{1}{p^2}+\frac{1}{p^4}+\ldots\right)}{\Pi\left(1-\frac{1}{p^2}+\frac{1}{p^4}-\ldots\right)} \tag{12.39}$$

and multiplying out, one obtains, in the denominator, a sum of each reciprocal square $1/n^2$ exactly once, with a sign that depends on the parity (odd or even) of the total numbers of prime factors of n. Thus, with (12.37), remembering that $\Omega(1) = 0$:

$$\sum_{n=1}^{\infty} \frac{(-1)^{\Omega(n)}}{n^2} = \frac{\pi^2}{15}, \tag{12.40}$$

or

$$\sum_{\Omega(n)\,\text{odd}} \frac{1}{n^2} = \frac{\pi^2}{20}, \tag{12.41}$$

two noteworthy results.

Similar procedures give the equally remarkable

$$\sum_{n=1}^{\infty} \frac{(\pm 1)^{\Omega(n)}}{n^2} 2^{\omega(n)} = \left(\frac{5}{2}\right)^{\pm 1}, \tag{12.42}$$

or

$$\sum_{\Omega(n)\,\text{odd}} \frac{2^{\omega(n)}}{n^2} = \frac{21}{20}. \tag{12.43}$$

12.4 The Harmonic Mean of $\Omega(n)$

In order to estimate, as we would like to, the geometric mean of the prime factors of n, we need the *harmonic* mean of $\Omega(n)$. If we designate geometric means by a tilde, then the desired mean is given by

$$\tilde{p}(n) := n^{1/\Omega(n)}. \tag{12.44}$$

Now if we average over several (similar) values of n, we are led to the harmonic mean of $\Omega(n)$, which we identify by a "hat":

$$\hat{\Omega}(n) := \left(\overline{1/\Omega(n)}\right)^{-1}. \tag{12.45}$$

With this notation, we have

$$\tilde{p}(n) \approx n^{1/\hat{\Omega}(n)}. \tag{12.46}$$

Of course, like any harmonic mean of a fluctuating quantity, $\hat{\Omega}(n)$ is smaller than the previously computed arithmetic mean $\overline{\Omega}(n) \approx \ln(\ln n) + 1.035$. By how much? to answer this question, we have to find out about the *distribution* of $\Omega(n)$. Reverting to our earlier "unaveraged" estimate of $\Omega(n)$:

$$\Omega(n) \approx \sum_{p_i \le n} e_i \frac{1}{p_i^{e_i}} \left(1 - \frac{1}{p_i}\right), \tag{12.47}$$

we recognize geometric distributions[1] in the exponents e_i. The mean value m_i for each term of the sum is

$$m_i = \frac{1}{p_i - 1}, \tag{12.48}$$

a result we used before in estimating $\overline{\Omega}(n)$.

Now we also want the *variance* σ_i^2 of each term, which for a geometric distribution is given in terms of the mean m_i by the following well-known formula:[2]

$$\sigma_i^2 = m_i + m_i^2. \tag{12.49}$$

[1] Physicists call a related distribution "Bose-Einstein" in honor of Bose, the Indian scientist who discovered its significance for photons and other "bosons", and Einstein, who publicized it when people would not believe it.

[2] This formula played a role in physics that can hardly be overestimated. According to Maxwell's equation, the intensity fluctuations σ_i^2 in "black-body" radiation should equal the squared intensity m_i^2. It was Einstein who discovered, from deep considerations of entropy, that the actual fluctuations exceeded m_i^2 by m_i, recognizing the additional term m_i as stemming from a non-Maxwellian "granularity" of the field. This observation led him to the *photon* concept for electromagnetic radiation on much more persuasive grounds than Planck's inherently contradictory discretization of the energies of harmonic oscillators. As a result, Einstein believed in the reality of the photons from 1905 on (and he received his Nobel prize in physics for this work and not for his theory of relativity), while Planck continued to doubt the meaningfulness of his "ad hoc" trick.

By summing over the index i, assuming independence of the p_i, we obtain the variance of $\Omega(n)$:

$$\sigma_\Omega^2 = \overline{\Omega}(n) + \sum_{p_i=2} \frac{1}{(p_i-1)^2}. \tag{12.50}$$

Using the expansion (12.16) again, we can write the sum here as

$$\sum_{p_i}\left(\frac{1}{p_i^2}+\frac{2}{p_i^3}+\frac{3}{p_i^4}+\ldots\right) = \ln\zeta(2)+2\ln\zeta(3)$$
$$+\frac{5}{2}\ln\zeta(4)+\ldots \approx 1.3751. \tag{12.51}$$

Again, $\sigma_\Omega^2 \approx \overline{\Omega}$, and we also expect a shifted Poisson distribution for Ω:

$$\text{Prob}\{\Omega(n)=k\} \approx \frac{(\overline{\Omega}-1)^{k-1}}{(k-1)!}\,\mathrm{e}^{-\overline{\Omega}+1}\ , \quad k=1,2,\ldots,\ \overline{\Omega}>1\ , \tag{12.52}$$

with $\overline{\Omega}$ from (12.36).

This theoretical distribution is shown by dots in Fig. 12.2 for $\overline{\Omega}=3.25$ ($n \approx 9500$). The shaded bars are from actual prime factor counts between $n=9000$ and 9499. The agreement is remarkably good.

For the shifted Poisson distribution, the harmonic mean $\hat{\Omega}$ is easily evaluated:

$$\hat{\Omega} = \frac{\overline{\Omega}-1}{1-\mathrm{e}^{-\overline{\Omega}+1}}, \quad \text{or} \tag{12.53}$$

$$\hat{\Omega} \approx \frac{\ln(1.035\ln n)}{1-1/1.035\ln n}. \tag{12.54}$$

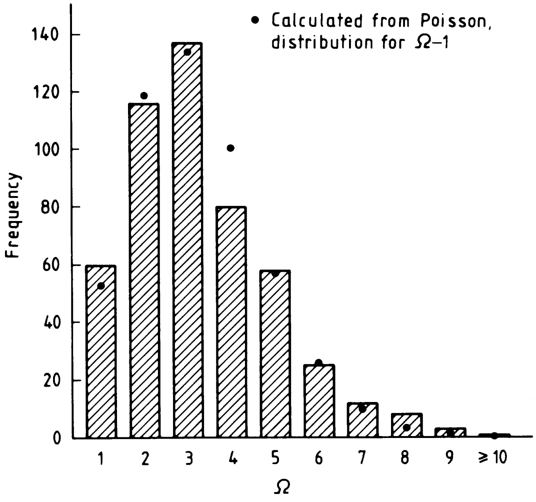


Fig. 12.2 The distribution of the number of prime factors (*bars*) in the interval $9000 \leq n \leq 9499$ and the Poisson distribution (*dots*) for the theoretical mean

For $n = 9250$, we obtain $\hat{\Omega} \approx 2.51$. The “experimental” value for the range $n = 9000$ to 9499 is $\hat{\Omega} = 2.47$ – as close as one can hope.

The geometric mean of the prime factors, calculated with the theoretical value of $\hat{\Omega}$, becomes

$$\tilde{p}(9250) \approx 38,$$

while the actual value in the range $n = 9000$ to 9499 is $\tilde{p} = 40$.

For $n = 10^{50}$, a range of interest for public-key encryption, $\hat{\Omega} \approx 4.8$, and the geometric mean $\tilde{p} \approx 2.4 \cdot 10^{10}$ – 40 orders of magnitude smaller than n.

12.5 Medians and Percentiles of $\boldsymbol{\Omega}(\boldsymbol{n})$

With (12.36), the probability that the integer n equals a prime that divides N is given approximately by

$$w(n = p|N) = \frac{1}{n \ln n(\ln(\ln N) + 1.035)}. \tag{12.55}$$

Thus, the cumulative distribution for a prime divisor of N to be smaller than n is approximated by

$$W(n;N) = \frac{\ln(\ln n) + 1.035}{\ln(\ln N) + 1.035}. \tag{12.56}$$

From this expression the median value $n_{0.5}$ follows directly:

$$n_{0.5} = \mathrm{e}^{\sqrt{\ln N/2.81}}. \tag{12.57}$$

Example: $N = 9250$, $n_{0.5} = 6$. Thus, the primes 2, 3, and 5 should account for roughly half the prime factors around $N = 9250$. The actual count in the interval 9000 to 9499 is as follows (with the theoretical value, $500/(p_i - 1)$, in parenthesis):

$$\begin{aligned} p_i = 2 &: 500 \text{ times} \quad (500) \\ p_i = 3 &: 250 \text{ times} \quad (250) \\ p_i = 5 &: 126 \text{ times} \quad (125) \end{aligned}$$

Thus, the total number of occurrences of 2, 3, and 5 is 876 times, or 53% of the total of 1650 prime factors in that interval – in very good agreement with our theoretical prediction.

For $N = 10^{50}$, $n_{0.5} = 600$, a remarkably small value.

The above distribution formula gives the following percentile values n_f, defined by $W(n_f) = f$:

$$n_f(N) = \exp\left[(\ln N)^f 2.81^{f-1}\right]. \tag{12.58}$$

The lower-quartile value $n_{0.25}$ (for $N = 9250$) becomes 2.2, which compares well with the count of 30% (500 out of 1650) for the factor 2 in the interval 9000 to 9499. In fact, according to (12.56), 29% of the prime factors should be below 2.5.

The theoretical upper-quartile value $n_{0.75} = 57.8$ is in very good agreement with the count of 75%. (1231 out of 1650) prime factors up to and including 59. But the median of the largest prime factor of N is about $N^{0.6}$ (!).

12.6 Implications for Public-Key Encryption

For $N = 10^{50}$, the theoretical lower- and upper-quartile values for the prime factors are 4.5 and $6 \cdot 10^{11}$, respectively. Thus, in three out of four cases of integers around 10^{50}, one will encounter prime factors not exceeding $6 \cdot 10^{11}$. If one assumes that *rapid* factoring of such integers is no problem, then 75% of such large, *randomly* selected integers can be easily factored.

This conclusion is in stark contrast to the (correct) assertion that sufficiently large integers *constructed so as to contain only two very large prime factors* cannot be easily factored.

Additional results on large prime factors in a given interval can be found in [12.4]. As an introduction to the art of generating large primes, see [12.5].

Chapter 13
Certified Signatures

Here we learn how certified signatures can be attached to secret messages in the context of public-key encryption. The degree of certitude (in the sense of avoiding random confusions) achievable by this method, which is based on modular arithmetic, appears to exceed by far that of notarized signatures, fingerprinting or, conceivably, even genetic analysis.

Certified signatures are also important in protecting computer systems against illicit entry and manipulation, and safeguarding data files from unauthorized "readers", falsification or destruction.

13.1 A Story of Creative Financing

Baron von Münchhausen, a close relative of the fabulous liar of the same name, and founder of the Georg-August University at Göttingen under the auspices of his King in Hanover, Georg August[1], received a secret message in (say) 1743 saying (in part):

"SPEND ALL EXCESS FUNDS OF KINGDOM ON NEW UNIVERSITY IN GOTTINGEN." signed "GEORGE".

How does von Münchhausen know that it was really King George who sent that generous but unlikely message? George is about to establish two more institutions of higher learning in his American colonies: King's College on an island called Manhattan (later to be known as Columbia University) and the College of New Jersey (now Princeton University) and the royal treasure has few, if any, "excess funds" to throw in the direction of Göttingen. The signature looks fine, but it could have been faked.

13.2 Certified Signature for Public-Key Encryption

In one of the great advances of modern secure and reliable communication (apart from public-key encryption itself), *certified signatures* can now be attached to public-key encryption messages in such a manner as to remove any doubt about

[1] Also known in London as George II, King of England, etc., etc.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_13, © Springer-Verlag Berlin Heidelberg 2009

the sender [13.1]. This works as follows: The sender, call him N, encrypts his name, address, etc., by his *de*crypting key t_0 (which only he knows!). Thus, he forms

$$S \equiv N^{t_0} \pmod{r_0}, \tag{13.1}$$

which he appends to his message M (which includes his name) and encrypts both M and S by the (public!) encrypting key of the receiver s_1, r_1.

The receiver decrypts using his secret key t_1, and reads the message M followed by a string of "garbled" symbols S, which must be the certified signature, because the message was identified as carrying such a signature. The message also purports to have been sent by N. Thus, knowing the protocol, the receiver applies the publicly known encrypting parameters of N, namely s_0 and r_0, to S and obtains

$$S^{s_0} \equiv N^{t_0 s_0} \equiv N \pmod{r_0}, \tag{13.2}$$

i. e., the name and address, etc., of the sender. And no one, but no one, who did not know t_0, could have constructed S so that with the above operation it would yield N. A certified signature to put all other "certified" signatures – including fingerprinting and (present-day) genetic analysis – to shame!

The reader can find further information on digital signatures and authentications to counteract potential threats[2] in financial, diplomatic and military "transactions" in [13.2–4].

With the spread of the Internet and electronic banking, data security and guaranteed signatures have taken on a wholly new dimension, see [13.5]. Proof of purchase and the important concept of "oblivious transfer" on the Internet are discussed in [13.6]. See also [13.7].

[2]

reneging	the *originator* subsequently disowns a transaction
forgery	the *recipient* fabricates a transaction
alteration	the *recipient* alters a previous valid transaction
masquerading	an *originator* attempts to masquerade as another

Chapter 14
Primitive Roots

In this chapter we introduce the concepts of *order* and the *primitive root*, two of the more fascinating and useful ideas in number theory. On the fundamental side, they helped the young Gauss to reduce the equation $x^{16} + x^{15} + \ldots + x + 1 = 0$ to several quadratic equations leading to the construction of the regular 17-gon. These same concepts also allow us to see why the decimal fraction of $1/7$ has a period of length 6, while the decimal fraction for $1/11$ has a period of only 2. And why does $1/99007599$, written as a binary fraction, have a period of nearly 50 million 0's and 1's? We shall see!

Closely related to the primitive root is the concept of *index*, a kind of number-theoretic logarithm that permits us to solve exponential Diophantine equations and even show that $2^n = 3^m - 1$ has only two, and precisely two solutions ($n = 1$, $m = 1$ and $n = 3$, $m = 2$).

Periodic sequences constructed from primitive roots also have an interesting Fourier-transform property that permits the construction of wave-scattering surfaces with very broad scatter and little specular reflection. Such surfaces can be useful in improving concert hall acoustics, in noise abatement measures, and in making ships and planes more difficult to see by sonar or radar. And, of course, there are applications to our main theme: digital encryption and electronic contracting (Sect. 20.3).

14.1 Orders

Some of the things we want to accomplish by electronic mail – other than public key encryption and certified signatures – have to do with certifiable "coin tossing", *registered mail* with or without *receipt*, and *signed contracts*. For these options we need the number-theoretic concepts of a *primitive root* and a *quadratic residue*, both delightful entities in their own right.

Let us look at increasing powers of 2 modulo 7:

$n = 1$	2	3	4	5	6
$2^n \equiv 2$	4	1	2	4	1

etc.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_14, © Springer-Verlag Berlin Heidelberg 2009

Here the period after which the sequence repeats for the first time is obviously 3. One therefore says that the integer 2 has *order* 3 modulo 7:

$$\mathrm{ord}_7 2 = 3. \tag{14.1}$$

Of course, the order of any integer must be divisor of $p-1$, where p is the prime modulus (7 in our example). This is so because of Fermat's theorem, which requires of any integer b coprime to the modulus p that the congruence $b^{p-1} \equiv 1 \pmod{p}$ must hold. Obviously, for $p = 7$, the order could never be 4, for example, because then the sequence of powers would repeat after 4 and 8 steps, etc., and not, as required by Fermat, after $p-1 = 6$ steps.

What is the largest order of any integer modulo a prime p? Certainly it cannot be larger than $p-1$, because there are only $p-1$ values of least positive residues such that $(m, p) = 1$, and once all residues have appeared they *must* repeat.

What is the order of 3 modulo 7? The following table will tell us.

$n =$	1	2	3	4	5	6	7	8
$3^n \equiv$	3	2	6	4	5	1	3	2

etc.

Thus, the order of 3 modulo 7 is 6, the highest possible value. Therefore 3 is called a *primitive root* modulo 7. A primitive root is also called a generating element, or generator, because it generates a complete residue system (in our example the integers from 1 to 6) in some permutation.

Once we have found a primitive root g, we can immediately find another one, its inverse modulo p:

$$g_2 \equiv g_1^{\phi(p)-1} \pmod{p}, \tag{14.2}$$

or, since $\phi(p) = p-1$ for a prime,

$$g_2 \equiv g_1^{p-2}. \tag{14.3}$$

In our example, with $g_1 = 3$, we get $g_2 \equiv 3^5 = 243 \equiv 5 \pmod{7}$. Check: $5 \cdot 3 = 15 \equiv 1 \pmod{7}$. Check! And 5 raised to successive powers yields

$n =$	1	2	3	4	5	6	7	8
$5^n \equiv$	5	4	6	2	3	1	5	4

etc.

Thus 5, too, has order $p-1 = 6$ and is therefore another primitive root.

How many primitive roots are there? If we raise a given primitive root g to the power $m > 1$, where $(m, \phi(p)) = 1$, then g^m must be another primitive root. Thus there are $\phi(\phi(p))$ primitive roots. (For $p = 7$, the number is $\phi(6) = 2$, both of which we have already found: 3 and 5.)

If, by contrast, the greatest common divisor (GCD) d of m and $\phi(p)$ is greater than 1, $(m, \phi(p)) = d > 1$, then the order of $g = g_1^m$ is only $\phi(p)/d$. To show this, we observe first that $\phi(p)/d$ is a period of g:

$$g^{\phi(p)/d} = g_1^{\phi(p)m/d} \equiv 1^{m/d} = 1 \pmod{p}, \tag{14.4}$$

and second that it is the *shortest* period, because by introducing the least common multiple $[\phi(p), m]$, we can write

$$g^{\phi(p)/d} = g_1^{[\phi(p),m]} = g_1^{\phi(p)\cdot k} \equiv 1^k = 1 \pmod{p}. \tag{14.5}$$

Thus, $\phi(p)/d$ is the smallest exponent for which $g^{\phi(p)/d}$ is congruent 1 modulo p.

How many positive $m < p$ are there that have order $T = \phi(p)/d$? As we saw in Chap. 7, there are exactly $\phi[\phi(p)/d]$ values of m that have d as the GCD with $\phi(d)$. Hence, the number of residue classes that have order T equals $\phi(T)$. This is illustrated by the following table for $p = 7$, $\phi(p) = 6$:

m	m^2	m^3	m^4	m^5	m^6	T
1	1	1	1	1	1	1
2	4	1	2	4	1	3
3	2	6	4	5	1	6
4	2	1	4	2	1	3
5	4	6	2	3	1	6
6	1	6	1	6	1	2

Indeed, there are exactly $\phi(1) = 1$ order $T = 1$, $\phi(2) = 1$ order $T = 2$, $\phi(3) = 2$ orders $T = 3$ and $\phi(6) = 2$ orders $T = 6$. Further, all T divide $\phi(7) = 6$.

Primitive roots are possessed by the integers 1, 2, 4, p^k and $2 \cdot p^k$ (where p is an *odd* prime and $k > 0$). All that has been said about primitive roots for a prime modulus transfers, *mutatis mutandi*, to these other cases.

The smallest integer not having a primitive root is 8. A prime residue system modulo 8 is given by 1, 3, 5, and 7, and all of these have order 1 or 2: $1^1 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod{8}$. There is no residue that has order $\phi(8) = 4$.

Why are 3 and 5 primitive roots modulo 7 and not, say, 4? How are the primitive roots distributed within a residue system? For example, 71 and 73 both have 24 primitive roots, of which they share exactly one half, namely

$$11,\ 13,\ 28,\ 31,\ 33,\ 42,\ 44,\ 47,\ 53,\ 59,\ 62,\ 68.$$

What distinguishes these numbers?

Gauss said in his *Disquisitiones* [14.1] that the distribution of primitive roots is a deep mystery; there is no way to predict where they will occur – only their total number is known. But Gauss *did* give some fast algorithms for ferreting them out.

14.2 Periods of Decimal and Binary Fractions

As every high-school student knows, $1/2$ written as a decimal fraction is 0.5 and $1/50$ becomes 0.02. Both 0.5 and 0.02 are *terminating* decimal fractions. By contrast $1/3$ becomes a nonterminating decimal fraction, namely $0.3333\ldots$, and so does $1/7$:

$$0.142857142857\ldots.$$

Both $1/3$ and $1/7$ lead to *periodic* decimal fractions. By contrast $\sqrt{2} = 1.41421356\ldots$ and $\pi = 3.14159265\ldots$ are irrational and have nonterminating aperiodic decimal representation.

What reduced rational fractions m/n, where $(m,n) = 1$, have terminating decimal representation? The answer is very simple and devolves directly from the prime factor decomposition of the denominator n:

$$n = \prod_{p_i|n} p_i^{e_i}, \tag{14.6}$$

where the product is over all prime p_i that divide n. Now, if the only p_i in (14.6) are 2 and 5, then the fraction terminates because 2 and 5 are the only prime factors of 10.

Specifically, if $n = 2^a 5^b$ and, for example, $a > b$, then $n = 2^{a-b}10^b$ and for $a - b = 2$, say,

$$\frac{1}{n} = 0.0\ldots025,$$

where the number of zeros to the right of the decimal point equals b.

Example: $n = 80 = 2^4 \cdot 5 = 2^3 \cdot 10$. Thus, $1/80 = 2^{-3} \times 10^{-1} = 0.0125$.

The numerator m of the fraction m/n simply converts a terminating decimal fraction into another terminating decimal fraction.

Things become more interesting if the denominator n is divisible by prime factors other than 2 or 5. Let us begin with the prime factor 3 and write

$$\frac{1}{3} = \frac{3}{9} = \frac{3}{10-1} = \frac{3}{10} \cdot \frac{1}{1 - \frac{1}{10}} = \frac{3}{10} \cdot \left[1 + \frac{1}{10} + \frac{1}{100} \ldots\right].$$

This brings the periodic nature of the decimal fraction for $1/3$ into direct evidence. The fraction $3/10$ is, of course, 0.3 and the 3 has to be repeated over and over again with increasing right shifts:

$$\frac{1}{3} = 0.3 + 0.03 + 0.003 + \ldots = 0.333\ldots.$$

To save ink, periodic decimal fractions are usually written with a bar over a single period. Thus, $1/3 = 0.\overline{3}$ and $1/7 = 0.\overline{142857}$.

But why does $1/7$ have a period length of 6? Modeling $1/7$ on what we did to $1/3$, we might try to express $1/7$ as a rational fraction with a denominator that is one less than a power of 10. Thus, we are looking for the smallest positive factor f such that

$$7f = 10^k - 1, \tag{14.7}$$

or, equivalently, we want to know the smallest k for which

$$10^k \equiv 1 \pmod{7}. \tag{14.8}$$

That is, of course, just the definition of *order* (in the arithmetic sense!) that we encountered in Sect. 14.1. Thus,

$$k = \operatorname{ord}_7 10 = 6. \tag{14.9}$$

Check: $10^6 = 142857 \cdot 7 + 1$, and no lower power of 10 exceeds a multiple of 7 by 1. Check!

Hence, $1/7$ has a decimal period of length 6 with the digits

$$f = \frac{10^6 - 1}{7} = 142857, \quad \text{or} \tag{14.10}$$
$$\frac{1}{7} = 0.\overline{142857}.$$

More generally, $1/p$, $p \neq 2$ or 5, has a period length

$$k = \operatorname{ord}_p 10. \tag{14.11}$$

Example: For $p = 11$, $k = 2$, $f = 99/11 = 9$; hence $1/11 = 0.\overline{09}$. For $p = 13$, $k = 6$, $f = 999999/13 = 76923$; hence $1/13 = 0.\overline{076923}$.

It is also clear that the period cannot be longer than $p - 1$, because in carrying out the long division $1/p$, there are at most $p - 1$ possible remainders, namely $1, 2, \ldots, p - 1$, after which the remainders and therefore also the decimal digits *must* repeat.

In fact, $\operatorname{ord}_p 10$ is always less than p, because according to Fermat's theorem, for $(p, 10) = 1$:

$$10^{p-1} \equiv 1 \pmod{p}. \tag{14.12}$$

Thus $\operatorname{ord}_p 10$ is either $p - 1$ (as in the case of $p = 7$) or a proper divisor of $p - 1$ (as in the cases $p = 11$ and $p = 13$).

The longest possible period $p - 1$ occurs whenever 10 is a primitive root of p. According to *Abramowitz* and *Stegun* [14.2], 10 is a primitive root of $p = 7$, 17, 19, 23, 29, 47, 59, 61, 97, etc.

Example: $\dfrac{1}{17} = 0.\overline{0588235294117647}$, which indeed has period length 16.

Of course, pocket calculators are not accurate enough to determine the 96 digits of the decimal period of $1/97$ directly. However, there is a trick that allows us to get the desired digits nevertheless. We shall illustrate this with the 16 digits of the period of $1/17$. A 10-digit pocket calculator shows that

$$\frac{100}{17} = 5.88235294(1),$$

where the last digit may have been rounded off and is therefore uncertain. We have thus found 9 of the 16 digits. The next digits are obtained by calculating, say,

$$\frac{160}{17} = 9.41176470(6).$$

Thus we have found all 16 digits of $1/17$. By merging the two digit strings we obtain

$$\frac{1}{17} = 0.\overline{0588235294117647}.$$

We leave it to the reader to discover a general and efficient algorithm for generating the digits of periodic fractions with a calculator of limited accuracy. In [14.3] all primes with period lengths less than 17 are listed. Surprisingly, 37 is the only prime with period length 3.

Without derivation or proof we also state that for

$$n = \prod_{p_i \neq 2,5} p_i^{n_i}, \tag{14.13}$$

the decimal expression has a period length T equal to the least common multiple of the orders of 10 with respect to the different p_i. Thus, with

$$k_i := \operatorname{ord}_{p_i^{n_i}} 10, \tag{14.14}$$

$$T = [k_1, k_2, \ldots]. \tag{14.15}$$

Proving this is a nice exercise. If n also contains factors 2 or 5, the decimal fractions are mixed, meaning they have a nonperiodic “head”.

Example: $1/119 = 1/(7 \cdot 17)$ has period length $T = [6,16] = 48$. And for $1/2737 = 1/(7 \cdot 17 \cdot 23)$, $T = [6,16,22] = 528$.

Nonunitary fractions have a cyclically shifted period with respect to the corresponding unitary fractions, provided $T = \phi(n)$. Otherwise there are $\phi(n)/T$ different cycles, all of length T.

Example: $1/7 = 0.\overline{142857}$ and $6/7 = 0.\overline{857142}$. But for 13 we have $\phi(13) = 12$ and $\operatorname{ord}_{13} 10 = 6$; thus there are $12/6 = 2$ different cycles: $1/13 = 0.\overline{076923}$ and $2/13 = 0.\overline{153846}$.

Everything that has been said here about decimal fractions carries over to other number bases. For example, $1/3$ in binary notation has period length $T = \operatorname{ord}_3 2 = 2$. In fact, $1/3 = 0.\overline{01}$. For $1/5$ in binary, $T = \operatorname{ord}_5 2 = 4$, and indeed, $1/5 = 0.\overline{0011}$. The prime 9949 has 2 as a primitive root [14.2]. Therefore, $\operatorname{ord}_{9949} 2 = 9948$ and $1/9949$ will generate a sequence of 0’s and 1’s with period lengths $9948 = 2^2 \cdot 3 \cdot 829$.

Another prime in the same range having 2 as a primitive root is 9851. With $9850 = 2 \cdot 5^2 \cdot 197$, the binary expansion of $1/99007599 = 1/(9949 \cdot 9851)$ has a period length $T = [2^2 \cdot 3 \cdot 829, 2 \cdot 5^2 \cdot 197] = 48993900$. Here is a method of generating long pseudorandom binary sequences! What are their spectral properties?

14.3 A Primitive Proof of Wilson's Theorem

Because the sequence of least positive residues $g^k \pmod{p}$, $k = 1, 2, \ldots, p-1$ (where g is a primitive root of the prime p), is a permutation of the integers $1, 2, \ldots, p-1$, one has

$$(p-1)! = 1 \cdot 2 \ldots (p-1) \equiv g \cdot g^2 \ldots g^{p-1} = g^{p\frac{p-1}{2}} \equiv g^{\frac{p-1}{2}} \pmod{p}. \qquad (14.16)$$

Now according to Fermat,

$$g^{p-1} \equiv 1 \pmod{p}. \qquad (14.17)$$

Therefore

$$g^{\frac{p-1}{2}} \equiv \pm 1 \pmod{p}. \qquad (14.18)$$

However the plus sign is impossible, because g is a primitive root and $p-1$ is the smallest exponent m for which g^m is congruent to 1. Thus,

$$g^{\frac{p-1}{2}} \equiv -1 \pmod{p}, \qquad \text{i.e.,} \qquad (14.19)$$

$$(p-1)! \equiv -1 \pmod{p}, \qquad (14.20)$$

which is Wilson's theorem. (However, note that we had to assume the *existence* of a primitive root!)

14.4 The Index – A Number-Theoretic Logarithm

Let m have the primitive root g. For the prime residue system $(k, m) = 1$, one defines the *index* of k modulo m as the smallest positive t for which

$$g^t \equiv k \pmod{m}, \qquad (14.21)$$

and writes

$$t = \operatorname{ind}_g k.$$

Read: t equals the index to the base g of k.

Example: for $m = 5$ and $g = 2$:

$$\operatorname{ind}_2 1 = 0, \quad \operatorname{ind}_2 2 = 1, \quad \operatorname{ind}_2 3 = 3, \quad \operatorname{ind}_2 4 = 2.$$

It is easy to see that

$$\operatorname{ind}_g(ab) \equiv \operatorname{ind}_g a + \operatorname{ind}_g b \pmod{\phi(m)}, \qquad (14.22)$$

a property the index shares with the logarithm. And in fact, the index is used much like a logarithm in numerical calculations in a prime residue system. For example, the congruence

$$3x \equiv 2 \pmod{5}$$

is converted to

$$\operatorname{ind}_2 3 + \operatorname{ind}_2 x \equiv \operatorname{ind}_2 2 \pmod{4},$$

or with the above "index table"

$$\operatorname{ind}_2 x \equiv 1 - 3 = -2 \equiv 2 \pmod{4}.$$

Thus,

$$x = 4.$$

Check: $3 \cdot 4 = 12 \equiv 2 \pmod{5}$. Check!

A rule that is handy for base conversion is

$$\operatorname{ind}_a b \cdot \operatorname{ind}_b a \equiv 1 \pmod{\phi(m)}, \tag{14.23}$$

which is reminiscent of $\log_a b \cdot \log_b a = 1$ for logarithms.

In preparing index tables, it is only necessary to list values for primes, because index values of composites are obtained by addition. Also, one-half of a complete index table is redundant on account of the following symmetry relation:

$$\operatorname{ind}(m-a) \equiv \operatorname{ind} a + \tfrac{1}{2}\phi(m) \pmod{\phi(m)}, \tag{14.24}$$

which is a consequence of

$$g^{\frac{1}{2}\phi(m)} \equiv -1 \pmod{m}. \tag{14.25}$$

14.5 Solution of Exponential Congruences

The exponential congruence

$$a^x \equiv b \pmod{m}, \tag{14.26}$$

if m has a primitive root, can be solved by index-taking

$$x \cdot \operatorname{ind} a \equiv \operatorname{ind} b \pmod{\phi(m)}, \tag{14.27}$$

which has a solution *iff*

$$(\operatorname{ind} a, \phi(m)) \mid \operatorname{ind} b. \tag{14.28}$$

In fact, in that case, there are $(\operatorname{ind} a, \phi(m))$ incongruent solutions.

Example: $7^x \equiv 5 \pmod{17}$; with the primitive root $g = 3$ as a base, we have

$$x \cdot 11 \equiv 5 \pmod{16}.$$

Since $(11,16) = 1$ divides 5, there is one (and only one) incongruent solution. Using Gauss's recipe,

$$x \equiv \frac{5}{11} \equiv \frac{5}{-5} = -1 \equiv 15 \pmod{16}.$$

Check: $7^{15} = 7^{16}/7 \equiv 1/7 \equiv 5 \pmod{17}$. Check!

In Sect. 7.6 on exponential Diophantine equations we considered the equation

$$2^n = 3^m - 1 \tag{14.29}$$

and asked whether there were solutions other than $2 = 3 - 1$ and $8 = 9 - 1$. Unfortunately, the answer was negative, otherwise we could have used ternary maximum-length sequences for precision measurements whose period was a power of 2, making them amenable to Fast Fourier Transformation (FFT) algorithms.

Now we consider another equation and ask: does

$$3^n = 2^m - 1$$

have any solutions other than $n = 1$, $m = 2$? If so, we could use binary maximum-length sequences whose period is a power of 3, making only slightly less efficient FFT algorithms based on the factor 3 (rather than 2) applicable.

We shall answer the above question using the concept of the *order* of an integer [14.4]. We ask: is there a solution of the above equation for $n > 1$? If there were, then $3^n = 9 \cdot k$ for some integer $k \geq 1$. Thus,

$$2^m \equiv 1 \pmod{9}.$$

Now the order of 2 modulo 9 is 6:

$$\operatorname{ord}_9 2 = 6.$$

Check: $2^r \equiv 2, 4, 8, 7, 5, 1 \pmod{9}$. Check! This means that 6 must divide the exponent m in the above congruence:

$$m = 6b \quad \text{for some integer } b.$$

Hence,

$$2^m = 2^{6b} = (2^3)^{2b} = 8^{2b} \equiv 1^{2b} = 1 \pmod{7}.$$

In other words, 7 divides $2^m - 1 = 3^n$, a contradiction because by the fundamental theorem, 3^n cannot be divisible by 7. Consequently, $n \leq 1$ and $3^1 = 2^2 - 1$ is the only solution. Too bad for our intended application!

In a similar vein [14.4], we prove that there are no solutions to

$$2^a = 3^b - 1 \tag{14.30}$$

for $a > 3$ or $2^a = 16 \cdot k$ for some integer $k \geq 1$. Thus, for $a > 3$:

$$3^b \equiv 1 \pmod{16}.$$

Since $\operatorname{ord}_{16} 3 = 4$, b must be some multiple of 4:

$$b = 4 \cdot r$$

for some integer $r \geq 0$. Thus,

$$3^{4r} \equiv 1 \pmod{16}.$$

Now, note that $\operatorname{ord}_5 3 = 4$, i. e.,

$$3^4 \equiv 1 \pmod{5},$$

and therefore also

$$3^{4r} = 3^b \equiv 1 \pmod{5},$$

or, equivalently,

$$5 | (3^b - 1) = 2^a,$$

a contradiction because 5 cannot divide a power of 2! Thus, $a = 1$, $b = 1$ and $a = 3$, $b = 2$ are the sole solutions of $2^a = 3^b - 1$.

14.6 What is the Order T_m of an Integer m Modulo a Prime p?

As another example of solving exponential congruences we shall consider the congruence

$$m^{T_m} \equiv 1 \pmod{p} \quad \text{or} \tag{14.31}$$
$$T_m \cdot \operatorname{ind} m \equiv 0 \pmod{\phi(p)}, \qquad \text{i.e.,} \tag{14.32}$$
$$T_m \cdot \operatorname{ind} m = k\phi(p). \tag{14.33}$$

Here the left-hand side must be both a multiple of $\phi(p)$ and $\operatorname{ind} m$ and, because of the definition of T_m as the *smallest* solution, $T_m \cdot \operatorname{ind} m$ must be the *least* common multiple $[\operatorname{ind} m, \phi(p)]$:

$$T_m \cdot \operatorname{ind} m = [\operatorname{ind} m, \phi(p)] = \frac{\operatorname{ind} m \cdot \phi(p)}{d}, \tag{14.34}$$

where d is the greatest common divisor of $\operatorname{ind} m$ and $\phi(p)$. Thus,

$$T_m = \frac{\phi(p)}{d}. \tag{14.35}$$

For example, for $p = 7$ and $m = 2$ and using 3 as the index base: $\operatorname{ind}_3 2 = 2$ and $T_2 = 6/(2,6) = 3$, i. e., the order of 2 modulo 7 is 3. Check: $2^3 = 8 \equiv 1 \pmod 7$ and $2^k \not\equiv 1 \pmod 7$ for $k < 3$. Check!

If we had taken 5 as the index base, the answer would have been the same: $\operatorname{ind}_5 2 = 4$ and $T_2 = 6/(4,6) = 3$. Check!

14.7 Index “Encryption”

The public-key encryption method described earlier is based on the fact that exponentiation modulo a large composite number whose factors are not known is apparently a “trap-door function”, i. e., it is easy to exponentiate with a known exponent and to calculate a remainder, but it is very difficult to go in the opposite direction, i. e., to determine which number has to be exponentiated to yield a known remainder.

Another way to describe this situation, for the case that the modulus has a primitive root, is to say that taking logarithms in number theory (i. e., determining an index) is a difficult operation. While the encrypted Message E is given by

$$E \equiv M^s \pmod r, \tag{14.36}$$

the original message M can be obtained, at least formally, by taking the index to the base g, where g is a primitive root of r:

$$\operatorname{ind}_g E \equiv s \cdot \operatorname{ind}_g M \pmod{\phi(r)} \quad \text{or} \tag{14.37}$$

$$M \equiv g^{(\operatorname{ind} E)/s} \pmod r. \tag{14.38}$$

Example: $r = 17$, $g = 3$, $s = 5$. Say the cryptogram is $E = 7$. Then, with

$$\operatorname{ind}_3 7 = 11 \pmod{16} \qquad \text{and}$$
$$\frac{\operatorname{ind} E}{s} = \frac{11}{5} \equiv \frac{-5}{5} = -1 \equiv 15 \pmod{16},$$

the original message is

$$M \equiv 3^{15} \equiv 6 \pmod{17}.$$

Check: $6^5 \equiv 7 \pmod{17}$. Check!

The disadvantage in serious applications of index encryption is that the modulus r is limited to integers that have primitive roots, i. e., primes, odd prime powers and twice odd prime powers (apart from 1, 2 and 4).

14.8 A Fourier Property of Primitive Roots and Concert Hall Acoustics

Consider the sequence

$$a_n = \exp\left(\frac{\mathrm{i}2\pi g^n}{p}\right), \tag{14.39}$$

where g is a primitive root of the prime p. This sequence is periodic, with period $\phi(p) = p-1$. Also, the a_n have magnitude 1.

The *periodic correlation sequence* is defined by

$$c_m := \sum_{n=0}^{p-2} a_n a_{n+m}^*, \tag{14.40}$$

where a^* stands for the complex conjugate of a. Obviously,

$$c_0 = p-1, \tag{14.41}$$

or, more generally, $c_m = p-1$ for $m \equiv 0 \pmod{p-1}$.

On the other hand, for $m \not\equiv 0 \pmod{p-1}$,

$$c_m = \sum_{n=0}^{p-2} \exp\left[\frac{\mathrm{i}2\pi g^n(1-g^m)}{p}\right]. \tag{14.42}$$

Here the factor $1-g^m \not\equiv 0 \pmod{p}$, and $g^n(1-g^m)$ therefore runs through a complete prime residue system except 0 as n goes from 0 to $p-2$. Thus, c_m is the sum over a complete set of pth roots of 1, except 1 itself. Since the "complete" sum equals 0, we have

$$c_m = -1, \quad \text{for } m \not\equiv 0 \pmod{p-1}. \tag{14.43}$$

Now, a periodic correlation function that has only two distinct values ($p-1$ and -1 in our case) has a *power spectrum* with only two distinct values [14.5]. By power spectrum we mean the absolute square of the Discrete Fourier Transform (DFT) defined by

$$A_m := \sum_{n=0}^{p-2} a_n \mathrm{e}^{-2\pi \mathrm{i} nm/(p-1)}. \tag{14.44}$$

It is easy to show that the power spectrum is given by

$$|A_m|^2 = \sum_{k=0}^{p-2} c_k \, \mathrm{e}^{-2\pi \mathrm{i} km/(p-1)}, \tag{14.45}$$

i. e., by the DFT of the correlation sequence. This is reminiscent of the well-known Wiener-Khinchin theorem [14.6].

For $m = 0$, or more generally $m \equiv 0 \pmod{p-1}$, we have, with the above two values for c_k,

$$|A_0|^2 = 1. \tag{14.46}$$

For $m \not\equiv 0 \pmod{p-1}$, with $c_0 = p-1$ and $c_k = -1$, we get

$$|A_m|^2 = p - 1 - \sum_{k=1}^{p-2} \mathrm{e}^{-2\pi \mathrm{i} km/(p-1)}, \tag{14.47}$$

where the sum is again over a complete set of roots of 1, except 1 itself. Thus,

$$|A_m|^2 = p \quad \text{for all } m \not\equiv 0 \pmod{p-1}. \tag{14.48}$$

Such a constant power spectrum is called “flat” or “white” (from “white light”, except that white light has a flat spectrum only in the *statistical* sense).

14.9 More Spacious-Sounding Sound

Flat power spectra are important in physics and other fields. (For example, a good loudspeaker is supposed to radiate a flat power spectrum when driven by a short electrical impulse.) Here, in addition, the original sequence a_n, whose spectrum is flat, has constant magnitude 1. This leads to an interesting application in concert hall acoustics.

It has been shown that concert halls with laterally traveling sound waves, all else being equal, have a superior sound [14.7] to those halls that furnish short-path sound arriving only from the front direction – as is the case in many modern halls with low ceilings (dictated by high building costs and made possible by modern air conditioning). To get more sound energy to arrive at the listeners’ ears from the sides (laterally), the author [14.8] proposed scattering, or diffusing, the sound which emanates from the stage and is reflected from the ceiling in all directions except the specular direction [14.9]. Also, the ceiling should not absorb any sound: in a large modern hall every “phonon”, so to speak, is valuable; otherwise the overall sound level (loudness) will be too low.

Thus, what is called for on the ceiling is something the physicist calls a *reflection phase-grating* that scatters equal sound intensities into all diffraction orders except the zero order. Here “order” is used not as defined in mathematics but as in physics. Zero-order diffraction corresponds to the specular direction, i. e., straight downward.

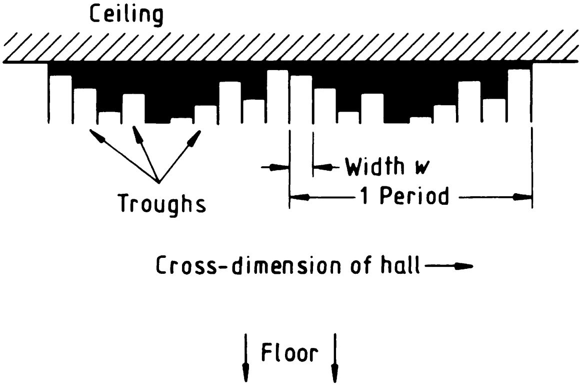


Fig. 14.1 Concert hall ceiling designed as a reflection phase-grating (based on the primitive root 2 of the prime 11)

Reflection phase-gratings can be realized by a hard surface with "wells" of different depths d_n, as shown in Fig. 14.1. Upon reflection, the phase of a normally incident wave is changed by $2d_n 2\pi/\lambda$, where λ is the wavelength. Now, if the depths d_n are chosen according to

$$d_n = \frac{1}{2}\frac{\lambda g^n}{p}, \tag{14.49}$$

where g is a primitive root of the prime p, and g^n can be the least residue modulo p, then the reflected wave has complex amplitudes[1] on its "wavefront" according to

$$a_n = \mathrm{e}^{2\pi \mathrm{i} g^n/p},$$

just like the periodic sequence that we considered above and that had a flat power spectrum.

Now, if the spatial distribution of wave amplitudes along a plane surface has a *flat* power spectrum, then the intensities of the wavelets scattered into the different diffraction orders will all be equal. Hence we expect a ceiling constructed according to this principle, as shown in Fig. 14.1 for $p = 11$ and $g = 2$, to scatter sound widely except in the specular direction (downward). That this is indeed so is illustrated in Fig. 14.2, which shows the result of actual measurements on a "primitive root" ceiling designed for improving concert hall acoustics. Such ceilings can be expected to increase the feeling of spaciousness, i. e., of being surrounded by or "bathed" in sound.

In order to form a *two*-dimensional array that scatters equal intensities into all diffraction orders (except the zeroth) over the *solid* angle, the prime p must be so chosen that $p-1$ has two coprime factors. For example, for $p = 11$, $p-1 = 10 = 2\cdot 5$

[1] This is taking an approximate ("Kirchhoff") view of diffraction. In reality, the complex amplitude cannot change abruptly. For an exact treatment, see [14.10].

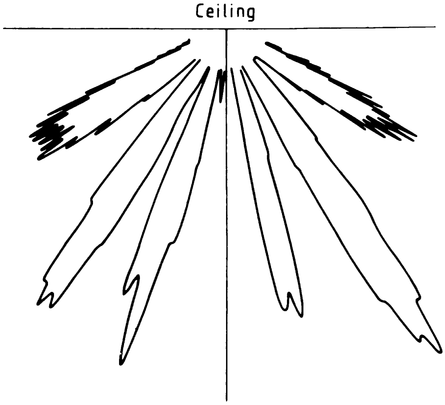


Fig. 14.2 Backscatter from primitive root ceiling. Note low specular reflection (vertically downward). $p = 7$, $g = 3$

and the ten numbers a_n can be used to fill a 2-by-5 array in "Chinese remainder" fashion (Chap. 17), for example as follows:

$$\begin{array}{ccccc} a_1 & a_7 & a_3 & a_9 & a_5 \\ a_6 & a_2 & a_8 & a_4 & a_{10}, \end{array} \qquad (14.50)$$

i. e., the horizontal (left-to-right) location of a_n in the array corresponds to $\langle n\rangle_5$ and the vertical (up-down) location of a_n is given by $\langle n\rangle_2$. More generally, the array locations can be given by $\langle k\cdot n\rangle_5$, with any k for which $(k,5) = 1$, and $\langle m\cdot n\rangle_2$, with $(m,2) = 1$ (i. e., m must be odd). Here the acute brackets signify least remainders (see Sect. 17.2 on Sino-representation).

In the most general case, an r-dimensional array with the desired r-dimensional correlation and Fourier properties can be formed if $p-1$ has r pairwise coprime factors $q_1, q_2, \ldots, q_r$ and the location in the array of a_n has the coordinates

$$\langle k_i\cdot n\rangle_{q_i}, \quad \text{with} \quad (k_i, q_i) = 1, \quad i = 1,2,\ldots,r. \qquad (14.51)$$

For three-dimensional arrays, the smallest prime p such that $p-1$ has three coprime factors is 31. Indeed, $30 = 2\cdot 3\cdot 5$, giving a 2-by-3-by-5 array.

Another important principle that can be employed to diffuse sound involves *quadratic residues* (Chap. 16), and an interesting application of primitive roots, to the splicing of telephone cables, is given in [14.11].

14.10 Galois Arrays for X-Ray Astronomy

X-rays are notoriously difficult to focus. For X-ray photons, the index of refraction of most earthly materials is so close to 1 that focussing lenses are all but impossible to construct. This is a pity because the skies abound with interesting emitters of X-rays. Fortunately, X-rays can be easily *blocked* by lead and other materials opaque

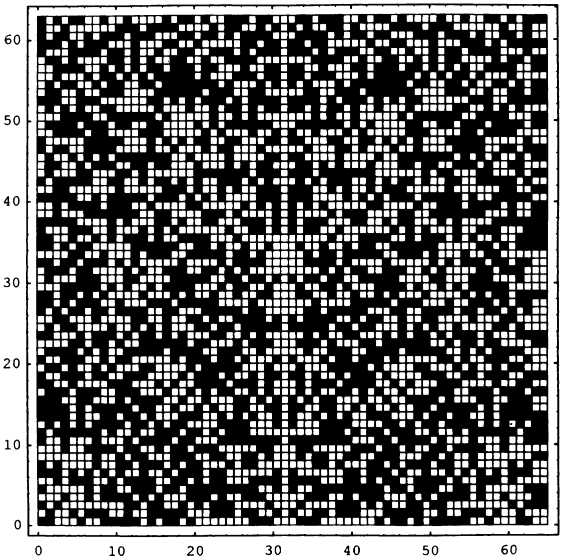


Fig. 14.3 Shadow mask for X-ray astronomy based on Galois sequence of length 4095

to X-rays. The question then is "can we image X-rays sources with the help of partially opaque masks?" And what patterns of opaqueness and translucence must such a "shadow mask" have to form useful images?

One solution is a two-dimensional pattern of opaqueness (0) and translucence (1) obtained from a Galois sequence (see Chap. 28) whose length $L = 2^m - 1$ can be factored into (at least) two coprime factors. For a nearly square-shaped mask, one chooses an even exponent, $m = 2k$, and obtains $L = 2^{2k} - 1 = (2^k + 1)$

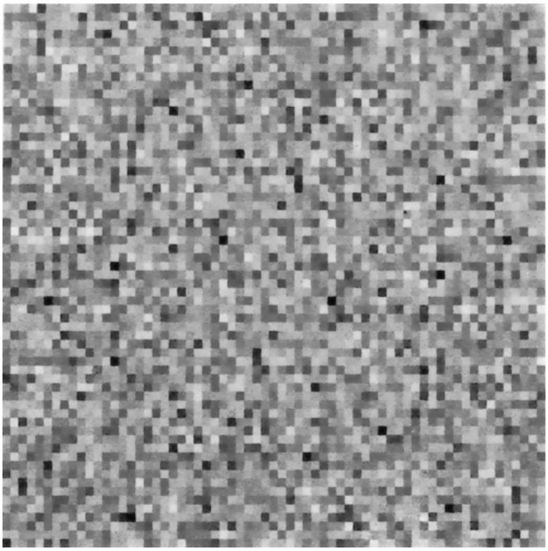


Fig. 14.4 Simulated X-ray shadow obtained with mask shown in Fig. 14.3

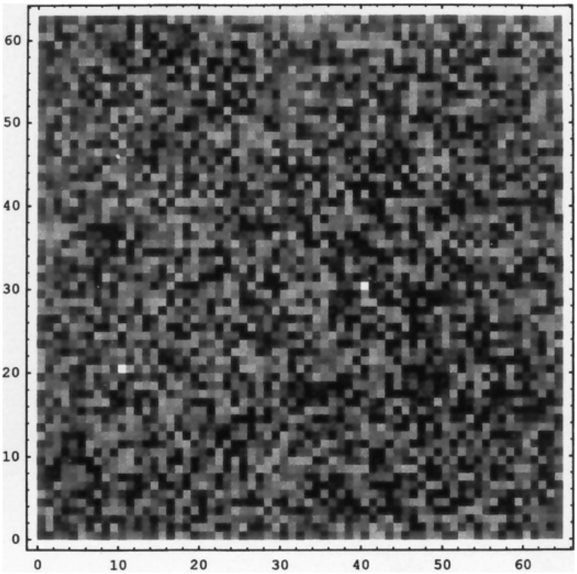


Fig. 14.5 X-ray image obtained by deconvolving data shown in Fig. 14.4. Note the two X-ray sources

$(2^k - 1)$. Figure 14.3 shows such a mask for $m = 12$ with side lengths 63 by 65. Note the bilateral symmetry around the vertical. Interestingly, the rows of length 63 are themselves Galois sequences. All this is no accident but can be proved rigorously, albeit by methods beyond the scope of this book. However, the interested reader might try to derive these results without the use of the relevant number-theoretic concepts (*trace* and *norm*).

Figure 14.4 shows the shadow of two distant X-ray sources cast by such a mask. Observing this data with photon counters and scanning it into a computer allows a reconstruction of the X-ray source by a two-dimensional deconvolution process. Figure 14.5 shows the result of this imaging method for *two* (incoherent) X-ray point sources [14.12].

14.11 A Negative Property of the Fermat Primes

The Fermat primes are, as Gauss discovered, precisely those primes p for which a "Euclidean" construction of the regular p-gon is possible (see Sect. 3.9). Thus, being a Fermat prime makes something possible.

Curiously, being a Fermat prime also makes something *impossible*, namely the construction of two- or higher-dimensional primitive-root arrays, as described in the preceding section. For such arrays, $p - 1$ must be factored into two or more coprime factors, but $p - 1$ has only *one* prime factor, namely 2. Thus, the same circumstance that allows the Euclidean construction forbids the construction of primitive-root arrays with more than one dimension (a form of mathematical justice?).

Chapter 15
Knapsack Encryption

As a diversion we return in this chapter to another (once) promising public-key encryption scheme using a trap-door function: *Knapsack encryption*. It, too, is based on residue arithmetic, but uses multiplication rather than exponentiation, making it easier to instrument and theoretically more transparent.

The required trap door is obtained from the ancient knapsack puzzle: given the total weight of a knapsack and the weight of individual objects, which objects are in the bag? As it turns out, this problem can be made quite difficult to solve for someone who doesn't have the proper "key".

15.1 An Easy Knapsack

Number-theoretic exponentiation is not the only trap-door function potentially suitable for public-key encryption. Another trap door is opened (and kept almost closed) by *knapsack problems*.

Suppose we have a set of six stones, weighing 1, 2, 4, 8, 16 and 32 kilograms, respectively. Now, if a knapsack containing some of these stones weighs 23 kilograms more than its empty weight, then which stones are in the knapsack? The answer is given by the binary decomposition of 23:

$$23 = 16 + 4 + 2 + 1,$$

i. e., the stones weighing 1, 2, 4 and 16 kilograms are in the knapsack and no others.

This is an example of an *easy* knapsack problem. The problem remains easy if each weight exceeds the sum of the lower weights by at least one measureable unit. The binary sequence of weights 1, 2, 4, 8, etc., fulfills this condition, because $2 > 1$, $4 > 2 + 1$, $8 > 4 + 2 + 1$, etc. Such sequences are called "superincreasing". In fact, the binary sequence is the smallest positive superincreasing sequence if 1 is the just discriminatable weight difference [15.1].

Table 15.1 shows such an easy knapsack embedded in leading and trailing digits. If the clear message in binary form is 110101, then the encrypted message is 21,853,232. It is obtained by multiplying each knapsack row by 0 or 1 in accordance

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_15, © Springer-Verlag Berlin Heidelberg 2009

Table 15.1 Encryption with an easy knapsack using binary "weights" 1, 2, 4, 8, 16, and 32 (see fourth and third columns in centre)

MESSAGE	EASY KNAPSACK	ENCRYPT
1	8301051	8301051
0	202061	
1	7804090	7804090
0	3508049	
1	2416013	2416013
1	3332078	3332078
		21853232

53 = 110101 IN BINARY
WHICH IS THE MESSAGE

with the binary message and adding as shown in Table 15.1. To decrypt this number we only need the two digits 53 and its binary expansion

$$53 = 32 + 16 + 4 + 1,$$

corresponding to 110101 in binary notation, the original clear message.

15.2 A Hard Knapsack

Adding extra digits has not made the knapsack one bit harder to solve – so far. But now look at Table 15.2, in which each row of the simple knapsack is multiplied by $s = 324,358,647$ and reduced modulo $r = 786,053,315$ to form a *hard* knapsack.

Table 15.2 Encryption with a hard knapsack, obtained from the easy knapsack at left by multiplication and residue reduction (After *N. J. A. Sloane* [15.1])

MESSAGE	EASY KNAPSACK	HARD KNAPSACK	ENCRYPT
1	8301051	515862967	515862967
0	202061	679273397	
1	7804090	513438305	513438305
0	3508049	402161783	
1	2416013	427531791	427531791
1	3332078	376052641	376052641
			1832885704
			=
			CIPHER

MULTIPLY BY s = 324358647
AND REDUCE MOD r = 786053315

Now, if encryption proceeds as before by multiplying the binary message with the rows of the knapsack (the hard knapsack!), the result, after the multiplied rows have been summed, is 1,832,885,704, which is “a little” harder to decrypt if only the public encryption key is available.

To design a trap-door function based on a hard knapsack, proceed as follows [15.2]:

1) Pick a set of easy knapsack weights a_i, i. e., one that forms a superincreasing sequence:

$$a_{i+1} > \sum_{k=1}^{i} a_k. \tag{15.1}$$

2) Pick a modulus r and a coprime multiplier s, i. e., $(s,r) = 1$.
3) Calculate the hard knapsack

$$b_k \equiv s a_k \pmod{r} \tag{15.2}$$

and publish *only* the b_k.
4) Calculate a decrypting multiplier t such that

$$st \equiv 1 \pmod{r} \tag{15.3}$$

and *do not* publish it.

Suppose M is a message and its binary digits are m_k:

$$M = \sum_{k=0}^{K} m_k 2^k. \tag{15.4}$$

The encrypted message E is now calculated as follows:

$$E = \sum_{k=0}^{K} m_k b_k. \tag{15.5}$$

To decrypt, the legitimate receiver, knowing t, now forms

$$tE = \sum_k m_k t b k_k. \tag{15.6}$$

But

$$t b_k \equiv t s a_k \equiv a_k \pmod{r}. \tag{15.7}$$

Thus,

$$tE \equiv \sum_k m_k a_k \pmod{r}, \tag{15.8}$$

and to recover the message bits m_k is an easy knapsack problem, because the a_k form a superincreasing sequence. Specifically, for $a_k = 2^k$, tE *is* the original message:

$$tE \equiv M \pmod{r}. \tag{15.9}$$

One of the advantages of knapsack encryption is that it does not rely on the supposed difficulty of factoring specially constructed very large numbers. On the other hand, knapsacks are presently under attack because *Shamir* [15.3] and others [15.4] have shown the equivalence of the knapsack problem to a problem in integer programming for which a “fast” algorithm was recently invented by H. W. Lenstra of the University of Amsterdam. Further progress in knapsack ripping has been made by L. Adleman, and by J. C. Lagarias and A. M. Odlyzko [15.5]. Thus knapsacks, as described here, have developed holes through which our “secret” weights can fall for everyone to see. Who will darn the ripped knapsack?

Part V
Residues and Diffraction

Chapter 16
Quadratic Residues

Here we will acquaint ourselves with the fundamentals of *quadratic residues* and some of their applications, and learn how to solve *quadratic congruences* (or perhaps see when there is no solution).

Certain periodic sequences based on quadratic residues have useful Fourier-transform and correlation properties. In fact, from the related *Legendre symbol*, we can construct binary-valued sequences that equal their own Fourier transform, up to a constant factor – a property reminiscent of the Gauss function $\exp(-x^2/2)$ which reproduces itself upon Fourier transformation. These Fourier properties are intimately related to the fact that quadratic-residue sequences generate perfect difference sets.

To demonstrate these Fourier properties, we will introduce the so-called *Gauss sums*, which added up to so many headaches for Carl Friedrich. We will also see (literally) that incomplete Gauss sums make pretty pictures.

An interesting physical application of quadratic-residue *sequences*, stemming from their Fourier property, is the design of broadly scattering diffraction gratings.

Another application, based on the auto- and cross-correlation properties of families of quadratic-residue sequences, occurs in *spread-spectrum* communications (an important technique for combating jamming and other interference).

And, of course, we will mention *quadratic reciprocity* and Gauss's famous *Theorema Fundamentale*.

16.1 Quadratic Congruences

Consider the quadratic congruence [16.1]

$$Ax^2+Bx+C\equiv 0 \pmod{m}. \tag{16.1}$$

For m a prime p, there is an inverse A' of A:

$$A'A\equiv 1 \pmod{p}, \tag{16.2}$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_16, © Springer-Verlag Berlin Heidelberg 2009

and we can write

$$x^2 + A'Bx + A'C \equiv 0 \pmod{p}. \tag{16.3}$$

If $A'B$ is even, then with

$$y = x + \tfrac{1}{2}A'B, \tag{16.4}$$

the original congruence can be written

$$y^2 \equiv b \pmod{p}, \quad \text{where} \tag{16.5}$$

$$b = \left(\frac{1}{2}A'B\right)^2 - A'C. \tag{16.6}$$

(For $A'B$ odd and p odd, $A'B$ is replaced by $A'B + p$ in the above substitution of y for x.)

Thus, to solve quadratic congruences, all we have to learn is how to extract square roots. For example, for $p = 7$ and $b = 2$, the solutions are $y = 3$ or $y = 4$. Of course, there may be no solution at all. For example, there is no y for which $y^2 \equiv 3 \pmod{7}$. The same is true for $y^2 \equiv 5$ or $y^2 \equiv 6 \pmod{7}$.

For a solution to $y^2 \equiv b \pmod{p}$ to exist, we say that b has to be a *quadratic residue* (R) modulo p. If b is a quadratic *nonresidue* (N), then there is no solution. For $p = 7$, the least positive residues are

$$R = 1,\ 2,\ 4;$$

and the nonresidues are

$$N = 3,\ 5,\ 6.$$

Not counting 0, there are exactly $(p-1)/2$ residues and also $(p-1)/2$ nonresidues. The properties R and N obey the rules for the multiplication of signs in ordinary arithmetic with R corresponding to + and N to −:

$$R \cdot R \equiv R, \quad R \cdot N \equiv N, \quad N \cdot N \equiv R. \tag{16.7}$$

Example: for $N \cdot N \equiv R$:

$$5 \cdot 6 = 30 \equiv 2 \pmod{7}.$$

16.2 Euler's Criterion

The integer a, $(a, p) = 1$, is a quadratic residue modulo p, p odd, *iff*

$$a^{\frac{1}{2}(p-1)} \equiv 1 \pmod{p}; \tag{16.8}$$

it is a quadratic nonresidue *iff*

$$a^{\frac{1}{2}(p-1)} \equiv -1 \pmod{p}. \tag{16.9}$$

Proof: According to Fermat, for $(a,p)=1$:

$$a^{p-1} \equiv 1 \pmod{p}. \tag{16.10}$$

Thus, since in a prime residue system there are only two values for the square root:

$$a^{\frac{1}{2}(p-1)} \equiv \pm 1 \pmod{p}. \tag{16.11}$$

Now, if g is a primitive root modulo p and if a can be written as follows:

$$a \equiv g^{2n} \pmod{p}, \tag{16.12}$$

then

$$a^{\frac{1}{2}(p-1)} = g^{(p-1)n} \equiv 1^n = 1 \pmod{p}, \tag{16.13}$$

and the solution of

$$y^2 \equiv a \pmod{p} \tag{16.14}$$

is

$$y \equiv g^n (\mathrm{mod}\, p), \tag{16.15}$$

i. e., a is quadratic residue.

For a equal to an odd power of a primitive root,

$$a^{\frac{1}{2}(p-1)} \equiv g^{\frac{1}{2}(p-1)} \equiv -1 \pmod{p}, \tag{16.16}$$

and a is a nonresidue. The same conclusion also follows from the fact that the formula g^{2n}, $n=1,2,\ldots$ generates $(p-1)/2$ distinct quadratic residues, so that the remaining $(p-1)/2$ residue classes must be quadratic nonresidues.

In a sequence of successive powers of a primitive root, R's and N's must therefore alternate.

Example: $p=7$, $g=3$, $n=1,2,\ldots$; $g^n \equiv 3,\ \underline{2},\ 6,\ \underline{4},\ 5,\ \underline{1} \pmod{7}$. Here the R's appearing for n even, are underlined.

16.3 The Legendre Symbol

The Legendre symbol (a/p) is a shorthand notation for expressing whether a is or is not a quadratic residue modulo p:

$$(a/p) := \left\{\begin{array}{rl} 1 & \text{for } a = R \\ -1 & \text{for } a = N \\ 0 & \text{for } a \equiv 0 \end{array}\right\} \pmod{p}. \tag{16.17}$$

Because of the multiplication rule for residues, one has

$$(a/p) \cdot (b/p) = (ab/p).$$

Let us look at an interesting property of what may be called *Legendre sequences*:

$$b_n = (n/p).$$

Example: $p = 7$:

$$b_n = 0,\ 1,\ 1,\ -1,\ 1,\ -1,\ -1;\ 0, \ldots,$$

etc., repeated periodically with period p. The sequence b_n is three-valued and its average is 0. Also, it reproduces itself by "decimating" with a quadratic residue. If we take every ath term, where a equals some R, then, because of

$$b_{na} = (an/p) = (a/p) \cdot (n/p) = (n/p), \tag{16.18}$$

we get

$$b_{na} = b_n. \tag{16.19}$$

For example, in the above sequence, with $a = 2$:

$$b_{2n} = 0,\ 1,\ 1,\ -1,\ 1,\ -1,\ -1;\ 0, \ldots,$$

which exactly equals b_n.

For a equal to a *non*residue, we get $b_{an} = -b_n$. This follows from the above derivation with $(a/p) = -1$ for a nonresidue. For example, with $a = 6 \equiv -1$, we obtain the backward-running sequence:

$$b_{-n} = 0,\ -1,\ -1,\ 1,\ -1,\ 1,\ 1;\ 0, \ldots = -b_n,$$

which is also the original sequence multiplied by -1.

More generally, we can say that when $p-1$ is a quadratic nonresidue of p, as in our example for $p = 7$, the sequence b_n is *antisymmetric*, i. e., $b_{-n} = -b_n$. From this follows that its discrete Fourier transform (DFT) is purely imaginary. For $p \equiv 3 \pmod{4}$, $p-1$ is a quadratic nonresidue. On the other hand, for $p \equiv 1 \pmod{4}$, $p-1$ is a residue.

Examples: for $p = 5$: $p-1 = 4 = 2^2$; similarly for $p = 13$; $p-1 = 12 \equiv 25 = 5^2 \pmod{13}$.

It is easy to see that this is so, because if $p \equiv 1 \pmod{4}$, then $p-1 \equiv 0 \pmod{4}$ and the congruence $y^2 \equiv 0 \pmod{4}$ always has a solution. In fact, it has two incongruent solutions, namely $y = 0$ and $y = 2$.

By contrast, if $p \equiv 3 \pmod 4$, then $p-1 \equiv 2 \pmod 4$, and there are *no* solutions to $y^2 \equiv 2 \pmod 4$. In fact, $0^2 = 0$, $1^2 = 1$, $2^2 \equiv 0$, $3^2 \equiv 1 \pmod 4$.

A convenient way to generate a Legendre sequence is by means of Euler's criterion

$$b_n := (n/p) \equiv n^{\frac{1}{2}(p-1)} \pmod p, \tag{16.20}$$

making use of the binary expansion of $(p-1)/2$ and reducing to least residues after each multiplication, as discussed in Sect. 6.3.

16.4 A Fourier Property of Legendre Sequences

The DFT of a Legendre sequence is given by

$$B_m = \sum_{n=0}^{p-1} b_n \mathrm{e}^{-2\pi \mathrm{i} nm/p}. \tag{16.21}$$

For $m \equiv 0 \pmod p$, we get

$$B_0 = 0. \tag{16.22}$$

For any m, we may replace b_n by $b_{nm}(m/p) = b_{nm} b_m$ yielding

$$B_m = b_m \sum_{n=0}^{p-1} b_{nm} \mathrm{e}^{-2\pi \mathrm{i} nm/p}, \tag{16.23}$$

or with $nm = k$, $m \not\equiv 0 \pmod p$:

$$B_m = b_m \sum_{k=0}^{p-1} b_k \mathrm{e}^{-2\pi \mathrm{i} k/p} = b_m B_1, \tag{16.24}$$

which, because of $b_0 = 0$, also holds for $m \equiv 0 \pmod p$. Thus, B_m, the DFT of the Legendre sequence, equals b_m times a constant (B_1) – a remarkable property, reminiscent of the fact that the Fourier integral of the Gauss distribution function $\exp(-x^2/2)$ reproduces the function to within a constant factor.[1]

[1] More generally, this is true for the Gauss function multiplied by a Hermite polynomial, a circumstance that plays an important role in quantum mechanics, because of the manifest symmetry in the phase space of the harmonic oscillator.

16.5 Gauss Sums

To determine the multiplier B_1, we consider the sum

$$S(p) := \sum_{k=0}^{p-1} (b_k + 1)\, \mathrm{e}^{2\pi \mathrm{i} k/p}. \tag{16.25}$$

Here $b_k + 1 = 2$ if $k \neq 0$ is a quadratic residue modulo p and $b_k + 1 = 0$ if k is a nonresidue. Thus, only those terms contribute to $S(p)$ for which $k \equiv n^2 \pmod{p}$ for some n. As n goes from 1 to $p-1$, every such k is "touched" exactly twice. Hence

$$S(p) = \sum_{n=0}^{p-1} \mathrm{e}^{2\pi \mathrm{i} n^2/p}, \tag{16.26}$$

a type of sum studied by Gauss.

Now, it is easy to show that

$$\begin{aligned} S^2(p) &= |S(p)|^2 = p \quad \text{for } p \equiv 1 \pmod{4} \quad \text{and} \\ S^2(p) &= -|S(p)|^2 = -p \quad \text{for } p \equiv 3 \pmod{4}, \end{aligned} \tag{16.27}$$

but it took Gauss (somewhat inexplicably) several painful years to determine the sign of $S(p)$, i. e., to prove that

$$\begin{aligned} S(p) &= \sqrt{p} \quad \text{for } p \equiv 1 \pmod{4} \quad \text{and} \\ S(p) &= \mathrm{i}\sqrt{p} \quad \text{for } p \equiv 3 \pmod{4}, \end{aligned} \tag{16.28}$$

where in both cases the positive sign of the square root is to be taken.

For p being replaced by the product of two primes, p and q, one has

$$S(pq) = (-1)^{(p-1)(q-1)/4}\, S(p)\, S(q). \tag{16.29}$$

The sums $S(p)$, or more generally $S(n)$ for any n, are called Gauss sums, and Gauss was eventually able to show that

$$S(n) = \begin{Bmatrix} (1+\mathrm{i})\sqrt{n} & \text{for } n \equiv 0 \\ \sqrt{n} & \text{for } n \equiv 1 \\ 0 & \text{for } n \equiv 2 \\ \mathrm{i}\sqrt{n} & \text{for } n \equiv 3 \end{Bmatrix} \pmod{4}. \tag{16.30}$$

Applying the above results to the DFT of our Legendre sequences, $b_m = (m/p)$, we obtain:

$$\begin{aligned} B_m &= \sqrt{p}\, b_m \quad \text{for } p \equiv 1 \pmod{4} \quad \text{and} \\ B_m &= -\mathrm{i}\sqrt{p}\, b_m \quad \text{for } p \equiv 3 \pmod{4}. \end{aligned} \tag{16.31}$$

If we replace $b_m = 0$ by $\tilde{b}_m = 1$, i. e., if we make our sequence binary-valued, then for $p \equiv 3 \pmod 4$, the new DFT is

$$\tilde{B}_m = 1 - \mathrm{i}\sqrt{p}\, b_m. \tag{16.32}$$

Thus, we have succeeded in constructing a *binary* sequence ($\tilde{b}_m = \pm 1$) with a *binary*-valued power spectrum

$$|\tilde{B}_m|^2 = 1 + p \quad \text{for } m \not\equiv 0 \pmod p, \tag{16.33}$$

and

$$|B_0|^2 = 1. \tag{16.34}$$

Such binary sequences with a binary power spectrum are extremely useful in several applications requiring Fourier transforms. If we store b_m in a computer memory, we need not also store the DFT, because a simple "multiply-and-add" operation converts the b_m into the $\tilde{B}_m$.

In Chap. 27 we will encounter another type of binary-valued periodic sequences with a binary power spectrum, the so-called pseudorandom or *maximum-length sequences*, with periods $2^n - 1$. Their DFT, by contrast, has $\phi(2^n - 1)/n$ different phase angles. For a typical value (for concert hall measurements), $n = 16$ and $\phi(2^n - 1)/n = 2048$, a sizable number when memory is expensive.

16.6 Pretty Diffraction

"Incomplete" Gauss sums defined by

$$S(m,k) := \sum_{n=0}^{k} \mathrm{e}^{2\pi \mathrm{i} n^2/m}, \qquad k = 0, 1, \ldots, m-1 \tag{16.35}$$

make pretty pictures. Figure 16.1 shows $S(m,k)$ for $m \equiv 2 \pmod 4$, for which the complete sum equals 0, i. e., the incomplete sum "returns to the origin" (marked by 0) after first having gone into and out of two spirals. For $m \to \infty$, these spirals,

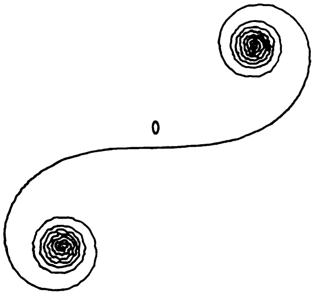


Fig. 16.1 Partial Gauss sums form spirals. Here the case $m \equiv 2 \pmod 4$, for which the complete sum returns to the origin (marked by 0)

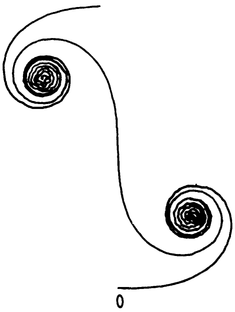


Fig. 16.2 Spiral representing partial Gauss sums for the case $m \equiv 3 \pmod 4$. The origin is marked by 0. The complete sum equals $+1$

properly scaled, will approach Cornu's spiral [16.2], the visual counterpart of the Fresnel integrals, which are important in diffraction theory. See also [16.3].

Figure 16.2 illustrates the case $m \equiv 3 \pmod 4$, where again the partial sum goes into and out of two spirals but "gets stuck" at $\sqrt{m}$ instead of returning to the origin (0), as predicted by Gauss.

16.7 Quadratic Reciprocity

If p and q are odd primes, then, using the Legendre symbol, the following law of *quadratic reciprocity* holds:

$$(p/q) \cdot (q/p) = (-1)^{\frac{(p-1)(q-1)}{4}}. \tag{16.36}$$

Another way to state this *Theorema Fundamentale*, so called by Gauss, who gave the first hole-free proof of this mathematical "gem" at age 18, is the following:

$$\text{for } \; p \equiv q \equiv 3 \pmod 4$$

$(p/q) \cdot (q/p) = -1$. Otherwise the product equals $+1$.

Example: $p = 3, q = 7$; $(3/7) = -1$, $(7/3) = (1/3) = 1$. Check!

The law of quadratic reciprocity can be very convenient when calculating quadratic residues.

Example: $(11/97) = (97/11) = (9/11) = (3/11)^2 = 1$.

Of particular interest are the following special cases:

$$(-1/p) = (-1)^{\frac{1}{2}(p-1)}, \tag{16.37}$$

which we have already encountered, in a more general form, as Euler's criterion, and

$$(2/p) = (-1)^{(p^2-1)/8}, \tag{16.38}$$

i. e., 2 is a quadratic residue *iff* $p \equiv \pm 1 \pmod 8$.

16.8 A Fourier Property of Quadratic-Residue Sequences

Consider the complex sequence

$$r_n = \mathrm{e}^{2\pi \mathrm{i} n^2/p}. \tag{16.39}$$

The magnitude of its individual terms equals 1, and it is periodic with period p. Its periodic correlation sequence is given by

$$c_k := \sum_{n=0}^{p-1} r_n r_{n+k}^* = \mathrm{e}^{-2\pi \mathrm{i} k^2/p} \sum_{n=0}^{p-1} \mathrm{e}^{-4\pi \mathrm{i} nk/p}. \tag{16.40}$$

Now, for any $k \not\equiv 0 \pmod p$, the sum, being over a complete set of pth roots of 1, vanishes. Thus,

$$c_k = 0 \quad \text{for } k \not\equiv 0 \pmod p. \tag{16.41}$$

On the other hand, for $k \equiv 0 \pmod p$, we obtain

$$c_0 = p. \tag{16.42}$$

Thus, we have found another constant-magnitude periodic sequence whose periodic correlation sequence has only two distinct values, one of which equals zero. As a direct consequence, the squared magnitude of the DFT, i. e., the power spectrum of the sequence r_n, has only *one* value:

$$|R_m|^2 = p, \quad \text{for } \textit{all } m. \tag{16.43}$$

Again there are applications to broad scattering of waves [16.4], such as diffusing sound waves from concert hall walls and ceilings to improve acoustic quality. Figure 16.3 shows a "quadratic-residue" ceiling for $p = 17$.

Such surface structures have the highly diffusing property not only at the "fundamental" frequency, but at a set of $p-1$ discrete frequencies. The reason is that for a frequency m times higher than the fundamental frequency, the phase shift suffered upon reflection from the "wells" is m times greater, i. e., the effective reflection coefficient is raised to the mth power. It is not difficult to show that the correlation property derived above holds not only for the sequence r_n, but also for the sequence r_n^m, as long as $m \not\equiv 0 \pmod p$. Thus, the ceiling shown in Fig. 16.3 is good for $p-1 = 16$ frequencies, spanning a range of 4 musical octaves; see Fig. 16.4 for $m = 3$.

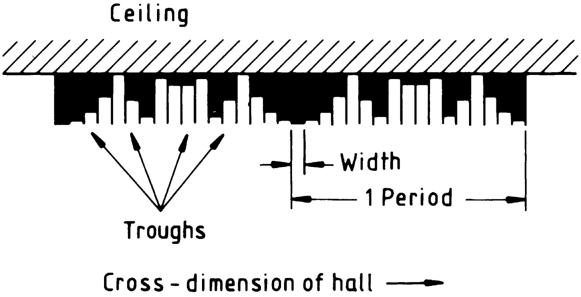


Fig. 16.3 Proposed concert hall ceiling design, a reflection phase-grating based on a quadratic-residue sequence for the prime 17

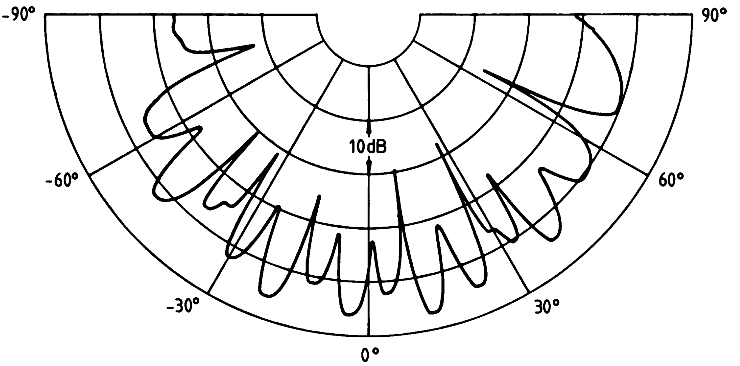


Fig. 16.4 Backscatter from model of concert hall ceiling illustrated in Fig. 16.3 for 3 times the fundamental design frequency. Note broad scatter, including strong specular reflection (vertically downward)

Fortunately, the good diffusing properties of such a quadratic-residue surface are also maintained at nonrational frequencies. Figure 16.5 shows a scatter diagram for $2\sqrt{2}$ times the fundamental frequency.

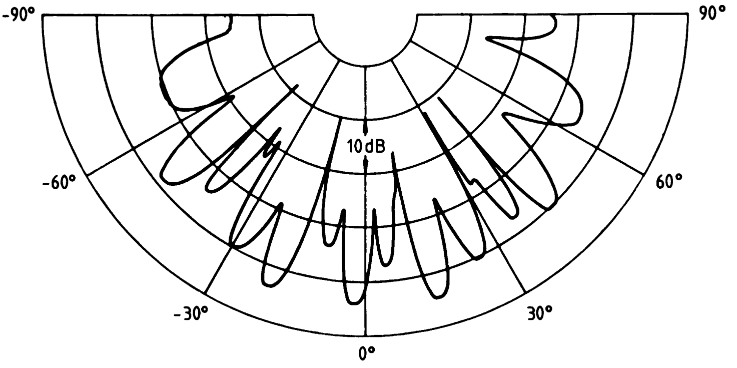


Fig. 16.5 Same as Fig. 16.4, for an irrational multiple of the fundamental design frequency

Such *reflection phase-gratings*, to use the physicist's term for such structures, also scatter well for oblique incidence. To guarantee good backscatter even at grazing incidence, the grating constant (width of the "wells") should be a *quarter*-wavelength, instead of the half-wavelength sufficient for normal incidence.

Quadratic-residue sequences are easily adapted to *two*-dimensional arrays a_{nm}, scattering equal intensities into all diffraction orders over the *solid* angle. We simply take *two* sequences r_n and r_m (based on the same prime p or different primes p_1 and p_2) and multiply them together:

$$a_{nm} = r_n \cdot r_m = \exp\left[2\pi\mathrm{i}\left(\frac{n^2}{p_1} + \frac{m^2}{p_2}\right)\right]. \tag{16.44}$$

Equation (16.44) is useful in the design of surfaces that scatter incident radiation widely with little of the total reflected energy going in any given direction. Potential applications are in radar and sonar camouflage and noise abatement.

Equation (16.44) obviously generalizes to arrays in three or more dimensions having the desired multidimensional correlation and Fourier properties.

Quadratic-residue sequences are also useful in the realization of broad (antenna) radiation patterns [16.5], which are discussed in Chap. 30.

16.9 Spread Spectrum Communication

In order to minimize interference, and especially *jamming*, in wireless ("radio") communication, the use of broad spectra is recommended [16.6]. Since frequency space in the "ether" is scarce, several users must therefore share the same frequency band. How should they signal among themselves to keep interference with other simultaneous users at a minimum?

Here again, quadratic residues offer a solution, and one that is especially transparent. Each user is assigned a particular quadratic-residue sequence as his "signature". By generalizing (16.39) we can construct $p-1$ such sequences:

$$r_n^{(m)} = \mathrm{e}^{2\pi\mathrm{i}mn^2/p}, \qquad m = 1, 2, \ldots, p-1. \tag{16.45}$$

Each of these sequences has, of course, the desired *auto*correlation property (16.41). But in addition, the cross-correlations with all other sequences is small. Thus, if each receiver is "matched" to his own sequence, he will only receive messages addressed to him – *without interference from any other simultaneously active user*, although they all share the same frequency band! (See Sect. 25.9 and [16.7, 8] for further details and additional references.)

16.10 Generalized Legendre Sequences Obtained Through Complexification of the Euler Criterion

Consider the Legendre sequence

$$b_n = \left(\frac{n}{p}\right), \tag{16.47}$$

where (n/p) is the Legendre symbol, see (16.17). Using Euler's criterion (16.8), (16.47) may be written as follows

$$b_n \equiv n^{(p-1)/2} \operatorname{mod} p, \tag{16.48}$$

or with a primitive root g of p:

$$b_0 = 0, \quad b_n = g^{(\operatorname{ind}_g n)(p-1)/2}, \tag{16.49}$$

where n is taken modulo p and $b_n = \pm 1$ for $n \not\equiv 0$. In (16.49) $\operatorname{ind}_g n$ is the index function (see Sect. 14.4) or "number-theoretic logarithm" defined by

$$g^{\operatorname{ind}_g n} \equiv n \operatorname{mod} p. \tag{16.50}$$

Since

$$g^{(p-1)/2} \equiv -1 \operatorname{mod} p, \tag{16.51}$$

(16.49) can also be written as

$$b_0 = 0, \quad b_n = \mathrm{e}^{\mathrm{i}\pi \operatorname{ind}_g n}. \tag{16.52}$$

Now, instead of admitting as possible values for b_n only two roots of unity (± 1), we can generalize (16.52) by allowing each b_n to equal one of the $p-1$ different $(p-1)$-th roots corresponding to the $p-1$ different values of $\operatorname{ind}_g n$. Thus, we define a *generalized Legendre sequence*, see also *Lerner* [16.9]:[2]

$$a_0 = 0, \quad a_n = \exp\left[\frac{2\pi\mathrm{i}(\operatorname{ind}_g n)}{p-1}\right], \qquad n \not\equiv 0 \operatorname{mod} p. \tag{16.53}$$

Note that $a_n \cdot a_m = a_{nm}$, as for the original Legendre sequence.

What are the spectral properties of the so defined periodic sequence? The (periodic) autocorrelation sequence

$$c_k = \sum_{n=1}^{p-1} a_n a_{n+k}^*, \tag{16.54}$$

[2] I am grateful to C. M. Rader for alerting me to Lerner's work.

equals $p-1$ for $k \equiv 0 \bmod p$. For $k \not\equiv 0$, (16.54) gives

$$c_k = \sum_{\substack{n=1 \\ n+k \not\equiv 0}}^{p-1} \exp\left\{\frac{2\pi\mathrm{i}\,[\mathrm{ind}_g n - \mathrm{ind}_g(n+k)]}{p-1}\right\}, \tag{16.55}$$

where the difference inside the brackets in the exponent assumes all integer values between 1 and $p-2$, modulo $p-1$, exactly once. Thus,

$$c_0 = p-1 \quad \text{and} \quad c_k = -1 \quad \text{for } k \not\equiv 0 \bmod p, \tag{16.56}$$

again as for the original Legendre sequence, see (30.22).

As a consequence of the correlation function being two-valued, the magnitude of the discrete Fourier transform (DFT), see Sect. 14.8 and [16.10], is also two-valued:

$$A_0 = 0 \quad \text{and} \quad |A_m|^2 = p-1 \quad \text{for } m \not\equiv 0 \bmod p. \tag{16.57}$$

In fact, the DFT A_m of the generalized Legendre sequence is *proportional* to the conjugate complex or reciprocal sequence:

$$A_m = A_1 a_m^{-1}, \tag{16.58}$$

a most remarkable property which also holds for the original Legendre sequence b_m, see (16.24).

In a further generalization, a scaling factor, $r = 1, 2, \ldots, p-2$, can be introduced in the definition (16.53), yielding

$$a_0 = 0 \quad \text{and} \quad a_n = \exp\left[\frac{2\pi\mathrm{i}r\,(\mathrm{ind}_g n)}{p-1}\right], \quad r \not\equiv 0 \bmod p-1\,, \; n \not\equiv 0 \bmod p\,. \tag{16.59}$$

For $r = (p-1)/2$ the original three-valued Legendre sequence is recovered. The generalized Legendre sequence can be used in the design of phase arrays and gratings and for spread-spectrum communication (see Sect. 16.9).

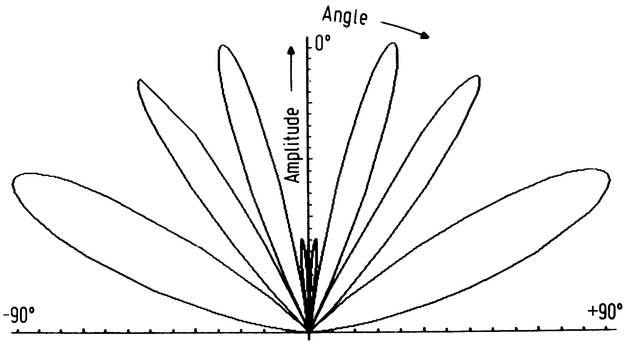


Fig. 16.6 Radiation pattern of index array based on prime number 7 and primitive root 3. Note the suppressed "broadside" lobe at 0°

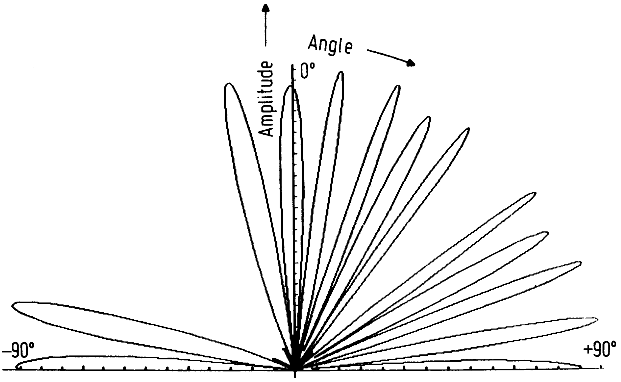


Fig. 16.7 Same as Fig. 16.6 except for an additional linear phase. Note the wide gap at $-45°$

In the latter application, because of the low crosscorrelation between sequences with different values of r, the possibility of choosing $p-2$ different values of r allows the specification of as many different "signature codes" for up to $p-2$ different users sharing the same frequency channel (see also Sect. 27.9).

Figure 16.6 shows the radiation pattern of an index array for the prime number 7 and the primitive root 3 at the design frequency. Note the suppressed broadside lobe. (The primitive root 5 gives the same result.) Figure 16.7 shows the radiation pattern of the same index array at *twice* the design frequency and an additional linearly increasing phase corresponding to a grating with an angle of incidence of 45°. Note the large gap in the pattern resulting from the missing lobe [16.11].

Part VI
Chinese and Other Fast Algorithms

Chapter 17
The Chinese Remainder Theorem and Simultaneous Congruences

One of the most useful and delightful entities in number theory is the *Chinese Remainder Theorem* (CRT). The CRT says that it is possible to reconstruct integers in a certain range from their *residues* modulo a set of *coprime* moduli. Thus, for example, the 10 integers in the range 0 to 9 can be reconstructed from their two residues modulo 2 and modulo 5 (the coprime factors of 10). Say the known residues of a decimal digit are $r_2 = 0$ and $r_5 = 3$; then the unknown digit is 8 (uniquely!).

It turns out that representing numbers in a notation based on the CRT has very useful applications, to wit: fast digital computations (convolutions and Fourier transforms), superfast *optical* transforms and, the solution of quadratic congruences, which are discussed in Chap. 19.

17.1 Simultaneous Congruences

So far, we have only talked about one congruence at a time. We shall now explore the inherent attractions of two or more *simultaneous* congruences:

$$x \equiv a_i \pmod{m_i}; \qquad i = 1, 2, \ldots, k. \tag{17.1}$$

Provided the moduli m_i are pairwise coprime:

$$(m_i, m_j) = 1 \quad \text{for } i \neq j, \tag{17.2}$$

a solution exists $\bmod(m_1, m_2 \ldots m_k)$ and is unique.

Next we shall construct the solution. Define:

$$M := \prod_{i=1}^{k} m_i \quad \text{and} \quad M_i = \frac{M}{m_i}. \tag{17.3}$$

Note that $(m_i, M_i) = 1$. Thus, there *are* solutions N_i of

$$N_i M_i \equiv 1 \pmod{m_i}. \tag{17.4}$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_17, © Springer-Verlag Berlin Heidelberg 2009

With these N_i, the solution x to be simultaneous congruences is [17.1]:

$$x \equiv a_1 N_1 M_1 + \ldots + a_k N_k M_k \pmod{M}. \tag{17.5}$$

To see this, we introduce as a convenient notation *acute brackets* with a subscript:

$$\langle x \rangle_m \tag{17.6}$$

by which we mean the least positive (or, occasionally, the least nonnegative) residue of x modulo m. With this notation we have

$$\langle x \rangle_{m_i} \equiv a_i N_i M_i \equiv a_i \pmod{m_i}, \tag{17.7}$$

because all other terms in the sum (17.5) making up x contain the factor m_i, and therefore do not contribute to the residue modulo m_i. Because $N_i M_i \equiv 1 \pmod{m_i}$, the solution is also unique modulo M. This is the *Chinese Remainder Theorem.*

Example: $m_1 = 3$, $m_2 = 5$; $M = 15$, $M_1 = 5$, $M_2 = 3$.

$$\begin{aligned} &\langle 5N_1 \rangle_3 = 1 \text{ yields } N_1 = 2 \quad \text{and} \\ &\langle 3N_2 \rangle_5 = 1 \text{ yields } N_2 = 2. \end{aligned}$$

Thus, the solution of the simultaneous congruences

$$\begin{aligned} &x \equiv a_1 \pmod{m_1}, \quad x \equiv a_2 \pmod{m_2} \qquad \text{is} \\ &x \equiv 10a_1 + 6a_2 \pmod{15}. \end{aligned}$$

Example: $a_1 = 2$ and $a_2 = 4$.

$$x \equiv 20 + 24 = 44 \equiv 14 \pmod{15}.$$

Check: $\langle 14 \rangle_3 = 2$, $\langle 14 \rangle_5 = 4$. Check!

17.2 The Sino-Representation: A Chinese Number System

Based on the Chinese Remainder Theorem (CRT), any positive integer x not exceeding M, where M can be written as a product of coprime integers:

$$0 < x \leq M = m_1 \ldots m_k, \quad \text{with} \tag{17.8}$$

$$(m_i, m_j) = 1 \quad \text{for } i \neq j, \tag{17.9}$$

can be represented uniquely by its least positive residues modulo the m_i. Thus, x is represented by a k-tuple of residues. In "vector" notation

$$x = (r_1, r_2, \dots, r_k), \quad \text{with} \tag{17.10}$$

$$r_i = \langle x \rangle_{m_i}, \qquad m_1 < m_2 < \dots < m_k. \tag{17.11}$$

This is the so-called *Sino-representation*. For only two moduli, thea Sino-representation admits of an illuminating graphic representation, in which r_1 and r_2 form the ordinate and abscissa, respectively. For $m_1 = 3$ and $m_2 = 5$, and using least nonnegative residues, we have

$r_2 =$	0	1	2	3	4
$r_1 = 0$	0	6	12	3	9
$r_1 = 1$	10	1	7	13	4
$r_1 = 2$	5	11	2	8	14

(17.12)

Here the position $(1,1)$ is occupied by 1 and the successive integers follow a 45°-degree downward diagonal jumping back to $r_1 = 0$ after $r_1 = 2$ and back to $r_2 = 0$ after $r_2 = 4$.

As can be seen from the above table, the integer 14, for example, has the representation $14 = (2,4)$. To go from the residues back to the integer, we use (17.5):

$$2 \cdot 2 \cdot 5 + 4 \cdot 2 \cdot 3 = 44 \equiv 14 \pmod{15}.$$

The Sino-representation is much more useful, in our digital age, than other ancient systems, for example the Roman number system (although *Shannon* [17.2] at Bell Laboratories did build a computer that truly calculated with Roman numerals – a masochistic numerical exercise if ever there was one).

Addition in the Sino-representation is "column by column": with

$$\begin{aligned} &a = (a_1, a_2) \quad \text{and} \quad b = (b_1, b_2), \\ &a + b = (\langle a_1 + b_1 \rangle_{m_1}, \langle a_2 + b_2 \rangle_{m_2}), \end{aligned} \tag{17.13}$$

which is easy enough to show.

Similarly, multiplication proceeds columnwise:

$$a \cdot b = (\langle a_1 b_1 \rangle_{m_1}, \langle a_2 b_2 \rangle_{m_2}). \tag{17.14}$$

These rules generalize to any finite number of moduli.

17.3 Applications of the Sino-Representation

Consider the two sequences, each of length M:

$$a(k) \quad \text{and} \quad b(k); \qquad k = 0, 1, \dots, M-1. \tag{17.15}$$

In numerous applications, the *convolution sequence* $c(k)$ of two such sequences is required [17.3]; it is defined by

$$\sum_{l=1}^{M} a(l)\cdot b(\langle k-l\rangle_M), \qquad k=0,1,\ldots,M-1. \tag{17.16}$$

Here arguments are taken modulo M, which makes this a *circular* convolution.

A *two*-dimensional circular convolution is defined as follows:

$$c(k_1,k_2) := \sum_{l_1=1}^{m_1}\sum_{l_2=1}^{m_2} a(l_1,l_2)\cdot b(\langle k_1-l_1\rangle_{m_1},\langle k_2-l_2\rangle_{m_2}). \tag{17.17}$$

If we introduce the Sino-representation for the index l in the *one*-dimensional convolution (17.17), we get

$$\tilde{c}(k_1,k_2) = \sum_{(l_1,l_2)} \tilde{a}(l_1,l_2)\cdot \tilde{b}(\langle k_1-l_1\rangle_{m_1},\langle k_2-l_2\rangle_{m_2}), \tag{17.18}$$

which is identical with the two-dimensional circular convolution (17.17). Thus, a one-dimensional circular convolution in Sino-representation is equivalent to a two- (or higher) dimensional circular convolution!

This has important consequences for the efficacy of calculating circular convolutions if a and b can be factored. In the one-dimensional case, we need M multiplications for each of the M values of the result $c(k)$, i. e., a total of M^2 multiplications. By contrast, in the two-dimensional Sino-representation, we need only m_1+m_2 multiplications for each value of $c(k_1,k_2)$, or a total of $(m_1+m_2)M$ multiplications. The reduction factor η_s in the number of multiplications is thus $(m_1+m_2)/M$ or, in the general case of decomposing M into r coprime factors,

$$\eta_s = \frac{m_1+m_2+\ldots+m_r}{M}. \tag{17.19}$$

Example: say $M \approx 10^6$. Make $M = 13\cdot 15\cdot 16\cdot 17\cdot 19 = 1{,}007{,}760$ (by appending 0's to the data if necessary). Then a five-dimensional Sino-representation yields a reduction in the number of multiplications by a factor $\eta_s^{-1} = 12597$, a very substantial saving!

In fact, the saving realized by the Sino-representation is exactly the same as that realizable by the famous Fast Fourier Transform (FFT) method [17.4]. In the FFT, too, the reduction factor equals

$$\eta_F = \frac{m_1+m_2+\ldots+m_r}{M}, \tag{17.20}$$

except that the m_i are no longer required to be coprime (at the price of some other complication, such as "twiddle" factors).

The greatest saving is of course realized if M is decomposed into as many small factors as possible. Thus, for $M \approx 10^6$, we can choose $M = 2^{20} = 1{,}048{,}576$ and get a reduction factor of $2\cdot 20/2^{20} = 3.81\cdot 10^{-5}$, which is, however, only twice as good as the Sino-decomposition.

In general for the FFT, choosing all $m_i = 2$, the optimum reduction factor is

$$\eta_F(m_i = 2) = \frac{2\log_2 M}{M}, \tag{17.21}$$

a well-known result [17.4].

Of course, if we wanted to apply FFT methods to circular convolution, 2 FFT's would be required [17.5] and the advantage of FFT over Sino-decomposition would be lost. And many more integers can be decomposed into, say, 5 coprime factors than factors of 2.

17.4 Discrete Fourier Transformation in Sino

The Discrete Fourier Transformation (DFT) of a sequence $a(n)$, $n = 0, 1, \ldots, M-1$, is defined by

$$A(k) := \sum_{n=1}^{M} a(n) W^{nk}, \qquad k = 0, 1, \ldots, M-1. \tag{17.22}$$

Here W is a (primitive) Mth root of unity, for example

$$W = \mathrm{e}^{-2\pi \mathrm{i}/M}. \tag{17.23}$$

Ordinarily, to compute all M values of $A(k)$ takes M^2 (complex) multiplications.

Now if M can be decomposed into coprime factors (we use only two factors in this exposition), $M = m_1 \cdot m_2$, and if we represent indices k and n in Sino:

$$\begin{aligned} k &= (k_1 = \langle k \rangle_{m_1}, k_2 = \langle k \rangle_{m_2}) \quad \text{and} \\ n &= (n_1 = \langle n \rangle_{m_1}, n_1 = \langle n \rangle_{m_2}), \end{aligned} \tag{17.24}$$

then the above DFT (17.22) can be rewritten

$$\tilde{A}(k_1, k_2) = \sum_{n_1=1}^{m_1} \sum_{n_2=1}^{m_2} \tilde{a}(n_1, n_2) W^{\langle n_1 k_1 \rangle_{m_1} N_1 M_1 + \langle n_2 k_2 \rangle_{m_2} N_2 M_2}, \tag{17.25}$$

where as before, $M_i = M/m_i$, and N_i is the (unique) solution of $N_i M_i \equiv 1 \pmod{m_i}$.

Because of the sum in the exponent in (17.25), W can be represented by the *product* of two factors, each one depending only on one n_i. Thus, the total number of multiplications is only $(m_1 + m_2)M$ instead of M^2. Again the reduction in the number of multiplications due to the Sino-representation is $\eta_s = (m_1 + m_2)/M$ or, in the general case, where M is decomposed into r coprime factors, $M = m_1 \cdot m_2 \ldots m_r$,

$$\eta_s = \frac{m_1 + m_2 + \ldots m_r}{M}.$$

Here, in the case of the DFT, too, the Sino-representation has converted the one-dimensional operation into a multidimensional one. In fact, a two-dimensional DFT is defined as follows:

$$A(k_1',k_2') := \sum_{n_1=1}^{m_1}\sum_{n_2=1}^{m_2} a(n_1,n_2)W^{M_1 n_1 k_1' + M_2 n_2 k_2'}, \tag{17.26}$$

where the factors $M_i = M/m_i$ in the exponent convert the W from an Mth root of unity to an m_ith root of unity, as required.

Comparing the two-dimensional DFT (17.26) with the Sino-representation, we see that they are identical if we set

$$k_i' = \langle N_i k_i \rangle_{m_i}. \tag{17.27}$$

In other words the Fourier array, indexed by k_i', differs from the Sino-array only by the indexing (k_i instead of k_i').

17.5 A Sino-Optical Fourier Transformer

The difference in indexing noted in the previous section becomes important in the following application of the Sino-representation to the DFT of one-dimensional data. Using Sino-representations, we can lay down the data in a two-dimensional array; see Fig. 17.1 for the case $M = 15$, $m_1 = 3$, $m_2 = 5$. As shown in this illustration, the "funny"-looking way of doing this becomes highly regular if we repeat the array periodically: consecutive numbers simply follow a −45° diagonal straight line.

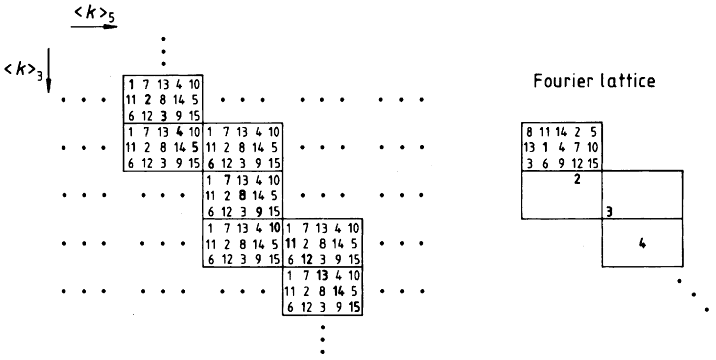


Fig. 17.1 One-dimensional data of period 15, put into two-dimensional array according to Sino-representation of index k modulo 3 and modulo 5 (the coprime factors of 15). On the right, the corresponding Fourier (DFT) array

If we Fourier-transform this data (one basic 3 × 5 array of it), we obtain the Fourier array[1] shown on the right of Fig. 17.1. Now, if we want to reconvert to one-dimensional data (the original data never was two-dimensional!), we have to know the sequence in which to "read" the data from the array. The proper "reading instruction" is embodied in the above index conversion equation $k_i' = \langle N_i k_i \rangle_{m_i}$. With $N_1 = N_2 = 2$, as we saw above, the proper reading "jumps" *two* units down and to the right to go from one sample to the next. These samples, of course, also fall on a straight line if the Fourier array is thought of as periodically repeated, as hinted in Fig. 17.1.

17.6 Generalized Sino-Representation

Instead of representing an integer k in the range 1 to M by its residues modulo the coprime factors m_i of M:

$$\begin{aligned} &k = (k_1, k_2, \ldots k_r) \quad \text{where} \\ &k_i = \langle k \rangle_{m_i}, \end{aligned} \tag{17.28}$$

we can, instead, choose the "components" k_i as follows:

$$k_i = \langle \alpha_i k \rangle_{m_i}, \tag{17.29}$$

where each α_i has to be chosen coprime to its modulus $m_i : (\alpha_i, m_i) = 1$. The inverse relation, i. e., the calculation of k from its components k_i, is the same as before:

$$k = \left\langle \sum_i k_i N_i M_i \right\rangle_M, \tag{17.30}$$

except that the congruences that determine the N_i look a little different:

$$\langle \alpha_i N_i M_i \rangle_{m_i} = 1. \tag{17.31}$$

The congruences have unique solutions, because $(\alpha_i M_i m_i) = 1$.

To check the above solution:

$$\langle a_j k \rangle_{m_j} = k_j \langle \alpha_j N_j M_j \rangle_{m_j} = k_j. \qquad \text{Check!}$$

Figure 17.2 illustrates the Sino-array for $\alpha_1 = 1$ and $\alpha_2 = 2$; in other words, the data is laid down in a "knight's-move" manner, as defined in chess. The solutions of $\langle \alpha_i N_i M_k \rangle_{m_i} = 1$ are $N_1 = 2$ and $N_2 = 1$. (Check: $\langle 1 \cdot 2 \cdot 5 \rangle_3 = 1$ and $\langle 2 \cdot 1 \cdot 3 \rangle_5 = 1$.

[1] If we do this optically by means of *Fourier optics* methods [17.6] using coherent light from a laser, the Fourier transform will be "calculated" (by interfering light waves) literally at the speed of light!

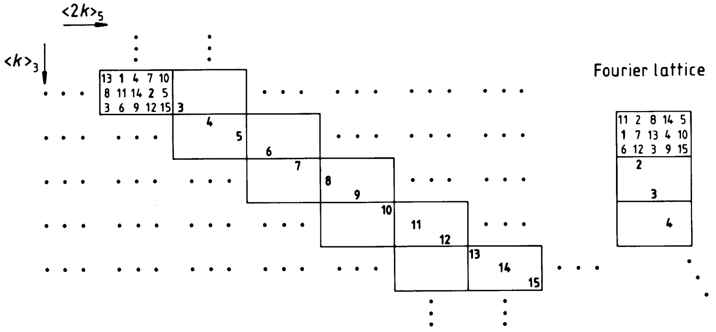


Fig. 17.2 Generalized conversion of one-dimensional data into two-dimensional array and corresponding Fourier array

Check!) Thus, the Fourier array also has to be read in a knight's-move fashion.

$$k_1' = \langle 2k_1 \rangle_3 \quad \text{and} \quad k_2' = \langle k_2 \rangle_5$$

as illustrated on the right of Fig. 17.2.

17.7 Fast Prime-Length Fourier Transform

In the preceding sections we have learned how to exploit compositeness of the number M of data samples to realize fast convolution and Fourier transform algorithms. But suppose M is *prime* and we cannot "pad" the data with zero-value samples to make the number of samples highly composite; suppose we are stuck with a prime M. What then?

Fortunately, *Rader* [17.7] comes to our rescue with a beautiful idea that bridges the prime-M gap. In fact, for Rader's algorithm to work, M *has* to be prime.

With (17.22) and (17.23) for the DFT, and assuming for the moment that $a(M) = 0$, we have

$$A(k) = \sum_{n=1}^{M-1} a(n) \exp\left(-\frac{2\pi \mathrm{i}}{M} nk\right). \tag{17.32}$$

Since M is prime, it has a primitive root g, which we put to work to *permute* our indices k and n as follows:

$$k =: \left\langle g^{k'} \right\rangle_M,$$

and

$$n =: \left\langle g^{n'} \right\rangle_M, \tag{17.33}$$

where k' and n' are the permuted indices; they cover the same range as k and n.

Defining further

$$a(g^{n'}) =: a'\left(n'\right),$$

$$A(g^{k'}) =: A'\left(k'\right),$$

we can write (17.32) as

$$A'(k') = \sum_{n'=1}^{M-1} a'(n') \exp\left(-\frac{2\pi \mathrm{i}}{M} g^{n'+k'}\right) \quad k' = 1,2,\ldots,M-1, \tag{17.34}$$

which we recognize as a circular convolution. Thus, the availability of a primitive root has allowed us to convert the *product nk* into the *sum* of the permuted indices $n'+k'$: we have turned a discrete Fourier transform into a circular convolution for which fast algorithms are available, even if $M-1$ is not highly composite [17.4]. (Note: the number of data samples is no longer the prime M, but the composite $M-1$.)

Since $k \equiv 0 \pmod{M}$ is not included in our permutation (17.33), we have to calculate $A(0)$ separately, which causes no sweat:

$$A(0) = \sum_{n=1}^{M-1} a(n). \tag{17.35}$$

If $a(M)$ [or $a(0)$] does not equal 0 as initially assumed, then all $A(k)$ have to be augmented by $a(0) = a(M)$.

But these are small technical details. The important thing to remember is that the existence of a primitive root has allowed us to convert a product in the exponent (Fourier transform) into a sum (convolution). At the same time, the number of data samples has been reduced by 1, thereby offering an escape from the unpleasant (for fast algorithms) primeness of M.

Converting products into sums smacks, of course, of logarithms, and indeed, although we never made it explicit, logarithms are implied by our permutation formulae (17.33). Instead of (17.33) we could have introduced the permuted indices as follows:

$$k' := \mathrm{ind}_g\, k \quad \text{and}$$

$$n' := \mathrm{ind}_g\, n, \tag{17.36}$$

where, as we recall (see Chap. 14), the index function is the number-theoretic equivalent of the usual logarithm.

Chapter 18
Fast Transformation and Kronecker Products

> *"It is important for him who wants to discover not to confine himself to one chapter of science, but to keep in touch with various others."*
>
> *– Jacques Hadamard*

While on the subject of fast computational algorithms based on the Chinese Remainder Theorem and primitive roots (discussed in the preceding chapter), we will now take time out for a glance at another basic principle of fast computation: decomposition into *direct* or *Kronecker* products. We illustrate this by showing how to factor Hadamard and Fourier matrices – leading to a Fast Hadamard Transform (FHT) and the well-known Fast Fourier Transform (FFT).

The Hadamard transformation is related to certain "Galois sequences" introduced in Chap. 26, and the FHT discussed here permits the computationally efficient utilization of these sequences in measurement processes of astounding precision.

18.1 A Fast Hadamard Transform

A Hadamard matrix is an $n \times n$ matrix of elements that equal $+1$ or -1 and whose rows and columns are mutually orthogonal [18.1]. Since the "norm" (sum of squared elements) of each row or column equals n, we have

$$H_n H_n^T = H_n^T H_n = nI_n, \tag{18.1}$$

where the raised T stands for "transpose" and I_n is the $n \times n$ unit matrix, having 1's in the main diagonal and 0's elsewhere.

The simplest nontrival Hadamard matrix H_n is

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{18.2}$$

All other Hadamard matrices necessarily have $n = 4k$.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_18, © Springer-Verlag Berlin Heidelberg 2009

Fig. 18.1 Construction of direct product of two matrices

$$V \otimes W = \begin{bmatrix} v_{11}W & v_{12}W & \cdots & v_{1m}W \\ \cdots & \cdots & \cdots & \cdots \\ v_{m1}W & v_{m2}W & \cdots & v_{mm}W \end{bmatrix}$$

Of special interest are the Hadamard matrices of Sylvester type with $n = 2^m$, which are obtained from H_2 by a “direct” or Kronecker product:

$$H_4 = H_2 \otimes H_2, \tag{18.3}$$

where $\otimes$ indicates the direct product, defined by replacing each element in the first factor by the product of that element with the second factor (see Fig. 18.1). Thus,

$$H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \tag{18.4}$$

All Sylvester-type Hadamard matrices are formed in the same manner:

$$H_{2^m} = \underbrace{H_2 \otimes H_2 \otimes \ldots \otimes H_2}_{m \text{ factors}}. \tag{18.5}$$

In a shorter notation, this is also expressed as a *Kronecker power* or exponentiation, indicated by a bracket in the exponent:

$$H_{2^m} =: H_n^{[m]}. \tag{18.6}$$

As an example H_8 is shown in Fig. 18.2, where a horizontal line stands for -1.

Various properties of the direct product are the following:

$$\begin{aligned} &(X+Y) \otimes W = X \otimes W + Y \otimes W \\ &(X \otimes Y)^T = X^T \otimes Y^T \\ &(X \otimes Y)(V \otimes W) = (XV) \otimes (YW) \\ &(X \otimes Y)^{-1} = X^{-1} \otimes Y^{-1} \\ &\mathrm{Trace}(X \otimes Y) = \mathrm{Trace}(X)\,\mathrm{Trace}(Y) \\ &\det(X \otimes Y) = (\det X)^n (\det Y)^m, \end{aligned} \tag{18.7}$$

where X is $n \times n$ and Y is $m \times m$.

$$H_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & - & 1 & - & 1 & - & 1 & - \\ 1 & 1 & - & - & 1 & 1 & - & - \\ 1 & - & - & 1 & 1 & - & - & 1 \\ 1 & 1 & 1 & 1 & - & - & - & - \\ 1 & - & 1 & - & - & 1 & - & 1 \\ 1 & 1 & - & - & - & - & 1 & 1 \\ 1 & - & - & 1 & - & 1 & 1 & - \end{bmatrix}$$

Fig. 18.2 The Sylvester-Hadamard matrix $H_8 = H_2^{[3]}$

The eigenvalues of $X \otimes Y$ are the nm products of the n eigenvalues of X and the m eigenvalues of Y. This last property is particularly revealing, because it shows that if we have a $k \times k$ matrix A, with $k = nm$, and if we can decompose A into two Kronecker factors

$$A = Y \otimes Y, \tag{18.8}$$

then instead of having to solve an nm-dimensional eigenvalue problem, we only have one n-dimensional and one m-dimensional problem to solve. The product nm is replaced by the sum $n+m$, just as in the Sino attack on convolutions and Fourier transforms, with a potentially enormous saving – especially if further Kronecker factors can be found, as in the case of the Sylvester-Hadamard matrices which completely decompose into the direct product of the 2×2 Hadamard matrices.

We shall demonstrate the attendant saving due to such decomposition for matrix-vector multiplications [18.2]. Let

$$z = H_4 \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} a+b+c+d \\ a-b+c-d \\ a+b-c-d \\ a-b-c+d \end{pmatrix}. \tag{18.9}$$

To calculate z, 12 additions/subtractions are required. In general, the number of operations is $n^2 - n$. [In (18.9) $n = 4$.]

However, by calculating intermediate results

$$\begin{aligned} a' &= a+b \\ b' &= a-b \\ c' &= c+d \\ d' &= c-d, \end{aligned} \tag{18.10}$$

we get the end result by only four more additions/subtractions:

$$z = \begin{pmatrix} a'+c' \\ b'+d' \\ a'-c' \\ b'-d' \end{pmatrix}, \tag{18.11}$$

a total of 8 operations. A saving of $12 - 8 = 4$ operations does not seem very impressive until we see the general formula: multiplication by a Sylvester-Hadamard matrix in the manner illustrated above for H_4 (possible because $H_4 = H_2 \otimes H_2$) requires

$$n \log_2 n \quad \text{instead of} \quad n^2 - n \tag{18.12}$$

additions/subtractions. For $n = 2^{16}$, the reduction factor equals 2.44×10^{-4}, a savings that may spell the difference between a problem being affordable (or even possible) to run on a given facility and impossible to run – perhaps on any machine. For further details and applications to spectroscopy see the excellent book by *Harwit* and *Sloane* [18.2].

Hadamard matrix operations owe their importance in part to the great technical significance of certain "maximum-length" sequences that are generated by primitive polynomials over finite number (Galois) fields, $\mathrm{GF}(2^m)$. We shall return to this fascinating topic when we discuss cyclotomic polynomials over Galois fields (Chaps. 26, 27).

18.2 The Basic Principle of the Fast Fourier Transforms

As we saw in Chap. 17, the Discrete Fourier Transform (DFT) is defined by [18.3]:

$$A_k := \sum_{n=1}^{M} a_n \mathrm{e}^{-2\pi \mathrm{i} nk/M}, \tag{18.13}$$

which can also be written as a matrix-vector product

$$A = F_M a, \tag{18.14}$$

where F_M is the $M \times M$ DFT matrix, the simplest nontrivial one being

$$F_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{18.15}$$

which is exactly like the Hadamard matrix H_2. Is F_4 perhaps the direct product of F_2 with itself? Not quite:

$$F_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -\mathrm{i} & -1 & \mathrm{i} \\ 1 & -1 & 1 & -1 \\ 1 & \mathrm{i} & -1 & -\mathrm{i} \end{pmatrix}. \tag{18.16}$$

However, if we permute rows 2 and 3 and multiply all imaginary terms by a "twiddle factor" i (not to be confused with the "fudge factor", so beloved among order-loving scientists), we obtain

$$\tilde{F}_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} = F_2 \otimes F_2 \ (!) \tag{18.17}$$

Thus, with a little twiddling and some permuting, we have succeeded in decomposing the 4×4 DFT matrix into the direct product of two (identical) 2×2 matrices, which – as an extra reward – are made up exclusively of ± 1's, requiring only additions or subtractions.

More generally, a properly twiddled and permuted $2^m \times 2^m$ DFT matrix can be written as a Kronecker power of F_2:

$$\tilde{F}_{2^m} = F_2^{[m]}, \tag{18.18}$$

which results in a fast algorithm as demonstrated for the Hadamard transform. This is the basic principle of all Fast Fourier Transforms (FFT). The savings in the number of operations is again governed by the ratio $\log_2 n/n$, where n is the number of data samples.

The rest is exhausting technical detail that is covered by a vast body of literature commensurate with the importance of the FFT (see [18.3] and the references therein).

Chapter 19
Quadratic Congruences

Here we shall see how the Chinese Remainder Theorem allows us to solve *quadratic congruences* for composite moduli. Quadratic congruences play a role in such digital communication tasks as certified receipts, remote signing of contracts, and coin tossing – or playing poker over the telephone (discussed in Chap. 20). Finally, quadratic congruences are needed in the definition of pseudoprimes, which were once almost as important as actual primes in digital encryption (see Chap. 20).

19.1 Application of the Chinese Remainder Theorem (CRT)

First consider quadratic congruences [19.1] modulo an odd prime p:

$$x^2 \equiv b \pmod{p}, \tag{19.1}$$

where p does not divide b, i. e., $p \not| b$. This congruence has precisely two solutions: x and $p - x$. The same is true for moduli that are powers of an odd prime: $p_i^{e_i}$.

Quadratic congruences modulo a general modulus

$$m = p_1^{e_1} \dots p_k^{e_k} \tag{19.2}$$

are solved by solving k *simultaneous* congruences modulo the coprime factors of m [19.2]:

$$x^2 \equiv b \pmod{p_i^{e_i}},$$

because owing to the CRT, solutions x to these simultaneous congruences will also solve the congruence

$$x^2 \equiv b \pmod{m = p_1^{e_1} \dots p_k^{e_k}}. \tag{19.3}$$

Example: $x^2 \equiv 9 \pmod{28}$. The coprime factors of 28 are 4 and 7. Thus, we solve

$$x^2 \equiv 9 \pmod{4} \qquad \text{and} \qquad x^2 \equiv 9 \pmod{7}$$

with the solutions

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_19, © Springer-Verlag Berlin Heidelberg 2009

$$\begin{aligned} \langle x\rangle_4 &= 3 & \langle x\rangle_7 &= 3 \quad \text{and} \\ \langle x\rangle_4 &= 4-3=1 & \langle x\rangle_7 &= 7-3=4. \end{aligned}$$

With these four values of x we can now construct four incongruent solutions modulo 28:

$$x = \big\langle \langle x\rangle_4 \cdot M_1 \cdot N_1 + \langle x\rangle_7 \cdot M_2 \cdot N_2 \big\rangle_{28},$$

where the N_i are given by

$$\langle N_i M_i \rangle_{m_i} = 1,$$

or with $M_1 = 7$ and $M_2 = 4$: $N_1 = 3$ and $N_2 = 2$ and

$$x = \big\langle \langle x\rangle_4 \cdot 21 + \langle x\rangle_7 \cdot 8 \big\rangle_{28}.$$

Substituting the above values for $\langle x\rangle_{m_i}$, we get the following four solutions

$$\begin{aligned} x_{++} &= \langle 3\cdot 21 + 3\cdot 8\rangle_{28} = 3 \\ x_{+-} &= \langle 3\cdot 21 + 4\cdot 8\rangle_{28} = 11 \\ x_{-+} &= \langle 1\cdot 21 + 3\cdot 8\rangle_{28} = 17 \\ x_{--} &= \langle 1\cdot 21 + 4\cdot 8\rangle_{28} = 25. \end{aligned}$$

Check: $3^2 \equiv 11^2 \equiv 17^2 \equiv 25^2 \equiv 9 \pmod{28}$. Check!

Of course, these solutions still come in pairs:

$$\begin{aligned} x_{++} &\equiv -x_{--} \pmod{28} \quad \text{and} \\ x_{+-} &\equiv -x_{-+} \pmod{28}. \end{aligned}$$

If we know one solution from each pair, we can find factors of the modulus by computing the greatest common divisor of their sum and the modulus. For example $(3+11, 28) = 14$ or $(3+17, 28) = 4$. We shall return to this point in Chap. 20 when we discuss “coin tossing” by telephone.

Part VII
Pseudoprimes, Möbius Transform, and Partitions

Chapter 20
Pseudoprimes, Poker and Remote Coin Tossing

In this chapter we take a closer look at numbers that are not primes, but are tantalizingly close to primes in some respects. Of course, a given number $n > 1$ is either prime or composite – in other words, n is either "pregnant" with factors or not; there is no third alternative. But nevertheless, it makes sense to define and, as we do in this chapter, discuss such odd entities as pseudoprimes, absolute (or universal) pseudoprimes and strong pseudoprimes. When talking about extremely large numbers, pseudoprimality is sometimes the only evidence we can go by.

20.1 Pulling Roots to Ferret Out Composites

Fermat says

$$b^{n-1} \equiv 1 \pmod{n}, \tag{20.1}$$

if b and n are coprimes and n is prime (cf. Chap. 8). However, as is well known, the above congruence can also hold for composite n.

Example:

$$2^{340} \equiv 1 \pmod{341}, \tag{20.2}$$

in spite of the fact that $341 = 11 \cdot 31$ is composite. Because of this primelike property, 341 is called a *pseudoprime to the base* 2: $341 = \mathrm{psp}(2)$.

More generally, a $\mathrm{psp}(a)$ is defined as an *odd composite* n for which

$$a^{n-1} \equiv 1 \pmod{n}$$

holds.

There are 245 pseudoprimes to the base 2 below one million, and the 6000th $\mathrm{psp}(2)$ is 1,187,235,193 [20.1]. The number of actual primes, by comparison, is about ten thousand times larger. Thus, numbers that are $\mathrm{psp}(2)$ are rather rare.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_20, © Springer-Verlag Berlin Heidelberg 2009

Nevertheless, there are infinitely many[1] [20.2] and, in fact, infinitely many in every prime residue class [20.3]. Numerically, the number of pseudoprimes less than n, $\Psi(n)$, seems asymptotically related to the prime-counting function $\pi(n)$ by

$$\Psi(n) = k\pi\left(\sqrt{n}\right),$$

with $k \approx 1.62$ (the Golden ratio??).

The exponent 340 in (20.2) is even, and now that we have learned to pull square roots, we might give it a try to see what happens. Of course, if 341 were prime, the only possiblee results would be $+1$ or -1.

$$2^{170} = (2^{10})^{17} \equiv 1^{17} = 1 \pmod{341}. \tag{20.3}$$

Again, 341 does not reveal its compositeness.

In general, composite numbers n for which

$$a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \pmod{n},$$

are called *Euler* pseudoprimes, abbreviated epsp(a) [20.4]. Here (a/n) is the Legendre symbol. Thus, because of (20.3), 341 is an epsp(2). Another epsp(2) is 1905. Check: $2^{952} \equiv (2/1905) = 1 \pmod{1905}$. Check!

Since 170 is also even, we can execute another square-root operation:

$$2^{85} = 32(2^{10})^8 \equiv 32 \pmod{341}, \tag{20.4}$$

since $2^{10} \equiv 1 \pmod{341}$.

Here we have trapped 341 and discovered its compositeness, because 32 is neither congruent 1 nor $-1 \pmod{341}$.

By repeated square rooting, we have, incidentally, also discovered the four solutions to the quadratic congruence

$$x^2 \equiv 1 \pmod{341}, \tag{20.5}$$

without inconveniencing the Chinese Remainder Theorem. Two of the solutions are obviously ± 1, and the other two are ± 32. In our repeated square rooting of 2^{340}, the solution $+1$ occurred twice and the solution $+32$ once.

Again (see Chap. 19), two absolutely different solutions tell us "something" about the factors of 341. Taking the greatest common divisors of the sum and difference of the two solutions with the modulus, we obtain

$$(32+1, 341) = 11$$

and

$$(32-1, 341) = 31, \tag{20.6}$$

[1] Indeed, every Fermat number, prime or not, obeys (20.1) (see Sect. 20.5).

20.2 Factors from a Square Root

This looks too good to be true: we have unearthed both factors of 341. Is this an accident? Or have we discovered a new approach to factoring the product of two primes? If so, we had better know more about it, even if the algorithm relies on the knowledge of two absolutely different square roots. To answer this question, we write down the solution to

$$x^2 \equiv 1 \pmod{341}$$

in terms of simultaneous congruences. Since both $\langle x \rangle_{11}$ and $\langle x \rangle_{31}$ equal ± 1, we have

$$x = \left\langle \pm \frac{341}{11} \left\langle \frac{341}{11} \right\rangle_{11}^{-1} \pm \frac{341}{11} \left\langle \frac{341}{31} \right\rangle_{31}^{-1} \right\rangle_{341} . \tag{20.7}$$

Here we have made use of a new but obvious symbol: if

$$N_i M_i \equiv 1 \pmod{m_i} \tag{20.8}$$

or, equivalently,

$$\langle N_i M_i \rangle_{m_i} = 1, \tag{20.9}$$

we write

$$N_i =: \langle M_i \rangle_{m_i}^{-1}, \tag{20.10}$$

and say N_i is the inverse of M_i modulo m_i. In the above numerical example, $N_{11} = 5$ and $N_{31} = 17$, and we have

$$x = \langle \pm 155 \pm 187 \rangle_{341} . \tag{20.11}$$

In terms of least residues, we have $x = \pm 1$ or $x = \pm 32$, as we saw before.

Now, looking at the general solutions of

$$x^2 \equiv 1 \pmod{pq}, \quad p \text{ and } q \text{ odd primes}, \tag{20.12}$$

we have

$$x = \pm \left\langle q \langle q \rangle_p^{-1} + p \langle p \rangle_q^{-1} \right\rangle_{pq} \tag{20.13}$$

and, using a different symbol, y, for the absolutely different solutions

$$y = \pm \left\langle q \langle q \rangle_p^{-1} - p \langle p \rangle_q^{-1} \right\rangle_{pq} . \tag{20.14}$$

Taking the upper signs, we see immediately that the sum equals

$$\langle x+y\rangle_{pq} = 2\left\langle q\langle q\rangle_p^{-1}\right\rangle_{pq} \tag{20.15}$$

and the difference

$$\langle x-y\rangle_{pq} = 2\left\langle p\langle p\rangle_q^{-1}\right\rangle_{pq}. \tag{20.16}$$

Consequently, since we have assumed that p and q are odd,

$$(x+y, pq) = q \quad \text{and} \tag{20.17}$$

$$(x-y, pq) = p. \tag{20.18}$$

Thus, calculating greatest common divisors of n and the sum and differences of quadratic congruences modulo n can indeed reveal the two factors of n.

20.3 Coin Tossing by Telephone

Alice and Bruno are divorcing (each other). Who gets the old jalopy, once the couple's precious chromium-embroidered self-propelled couch (rhymes with ouch)? Alice telephones Bruno and says: "Let's toss coins." Bruno: "Fine, I'll toss." But Alice, tossing her head (invisible to Bruno, because they are not connected by Picturephone®), objects: "Let *me* do the tossing – you know, you always liked my salads, if nothing else." So Bruno gives in: "O.K. I choose TAILS." Alice tosses (the coin) and up comes HEADS. Alice to Bruno: "Sorry, you lost – you'd better look around for another car."

How can we make this situation immune to cheating without resorting to video or witnesses? By extracting square roots! And here is how [20.5].

PROTOCOL

1) Alice selects two very large odd primes p and q and calculates $n = pq$.
2) Alice sends Bruno n.
3) Bruno picks a random number x smaller than n and calculates the least remainder of x^2 modulo n: $\langle x^2\rangle_n$.
4) Bruno sends Alice $\langle x^2\rangle_n$.
5) Alice, knowing p and q (and remembering the Chinese Remainder Theorem), calculates the four square roots, $\pm x$ and $\pm y$.
6) Alice sends Bruno one of these four values, not knowing which one Bruno had originally selected under (3).
7) Bruno squares the received number to check that Alice has not made a mistake (or cheated). If Alice sent $+x$ or $-x$, Bruno has lost, admits it, and Alice gets the remains of their car. However, if Alice sent $+y$ or $-y$, Bruno tells Alice that he has won and would she please have the car delivered to his new pad.
8) Alice does not believe that Bruno has won and challenges him to prove it.

9) Bruno adds his x to the y Alice sent him and, using Euclid's time-honored algorithm, determines the greatest common divisor with n, which is one of the two factors of n:

$$(x+y,n) = p \quad \text{or} \quad q.$$

10) Bruno sends Alice $(x+y,n)$ and $n/(x+y,n)$, i. e., the two factors p and q of n.

11) Alice knowing that Bruno has succeeded in factoring n, throws in the towel and hopes for better luck next time. Of course, Alice had picked p and q so large that Bruno (or anyone else, except perhaps the Government) could not factor n without knowledge of an absolutely different square root.

END OF PROTOCOL.

Example: (too small, except perhaps to wager on a set of seven-year-old snow tires): $p = 101$, $q = 113$, $n = 11413$.

$$\begin{aligned}
x &= 1234, & \langle x^2\rangle_{11413} &= 4827.\\
\langle x\rangle_{101} &= 22, & \langle x\rangle_{113} &= 104\\
\langle 113\rangle_{101}^{-1} &= 59 & \langle 101\rangle_{113}^{-1} &= 47\\
x,y &= \pm\langle 22\cdot 113\cdot 59 \pm 104\cdot 101\cdot 47\rangle_{11413} & &\\
x,y &= \pm 1234,\ \pm 4624; & \langle 4624^2\rangle_{1143} &= 4827.\ \text{Check!}\\
x+y &= 5858 & &\\
(5858, 11413) &= 101. \quad \text{Check!} & &\\
y-x &= 3390 & &\\
(3390, 11413) &= 113. \quad \text{Check!} & &
\end{aligned}$$

Suppose Bruno does not *want* to win. Then, even when Alice sends him $\pm y$ to his $\pm x$, he can pretend that what Alice sent him was the square root he already knew, and that therefore he has lost. If Alice *wants* Bruno to win, and if he did win the toss, then under the present protocol she has no way of telling whether Bruno lied when he said "I lost".

To take care of this peculiar situation, a different protocol has been advanced by *Silvio Micali* [20.6] of the University of California at Berkeley. It is called "tossing coins into wells". We will not describe Micali's protocol, but simply mention that the "well" into which Alice tosses the coin is so deep that Bruno cannot reach in and turn the coin over. But he can *see* whether "heads" or "tails" is up, and if Alice does not believe him, he can ask her to look at the "well" herself.

It is noteworthy, and perhaps surprising, that number theory, of all disciplines, provides the solution to such mundane problems.

It is obvious that the general idea exploited here, namely the exchange of partial information, can be exploited in other kinds of transactions such as

"I will tell you my secret only
if you tell me yours."

The problem, of course, arises when the two “clams” are remote from each other. If clam A releases her secret, clam B may say “thank you” and remain shut up. To avoid this kind of cheating, the secrets are represented by strings of bits, and one bit from each clam is exchanged at any one time following a procedure similar to the coin-tossing protocol. If either partner clams up in the middle of the exchange, then both have acquired the same number of bits (give or take one bit). Of course, the early bits of *B* (or *A*) could be relatively worthless. I shall leave it to the reader to find a way out of this semantic quandary.

Another potential application of the coin-tossing protocol is to play poker [20.7] over the wire. The author, not being a poker player himself, invites the interested reader to devise a safe protocol. But then, poker might lose much of its charm without the faces in a smoke-filled room.

20.4 Absolute and Strong Pseudoprimes

We showed above that 341 was a pseudoprime to the base 2 because it is composite and $341|(2^{341}-2)$. There are some numbers N called *Carmichael numbers* [20.8] or *absolute pseudoprimes*, composites defined by $N|(b^{N-1}-1)$ for all $(b,N)=1$. The smallest member of this dangerous species (dangerous for primality tests) is $561=3\cdot 11\cdot 17$.

How can it be possible for b^{N-1} to be congruent to 1 modulo N for *all b* coprime to N? For $N=561$ there are $\phi(561)=320$ such values b. Credulity is strained! Yet the mystery behind the Carmichael numbers is not very deep. All we need are three (or more) odd primes p_i such that each p_i-1 divides $N-1=\prod p_i-1$. This is obviously the case for $p_1=3$, $p_2=11$ and $p_3=17$ because 2, 10, and 16 all divide $3\cdot 11\cdot 17-1=560$.

Why is this sufficient? First note that according to Fermat,

$$b^{p_i-1}\equiv 1 \pmod{p_i}$$

for all $(b,p_i)=1$. These are simultaneous congruences (see Chap. 17), and by writing them

$$x\equiv 1-b^{p_i-1} \pmod{p_i}$$

we have brought them into the form of (17.1). These congruences have a *unique* solution modulo $N=\prod p_i$. (In Chap. 17 N is called M, and the coprime moduli m_i in (17.2) are in fact primes here, the primes p_i.) This unique solution is, of course, $x=0$. Thus, we have

$$\begin{aligned}&0\equiv 1-b^{p_i-1} \pmod{N}, \quad \text{or}\\ &b^{p_i-1}\equiv 1 \pmod{N}.\end{aligned}$$

Since $p_i - 1$ divides $N-1$:

$$N-1 = k_i(p_i - 1),$$

for some k_i, we have

$$b^{N-1} \equiv 1^{k_i} \equiv 1 \pmod{N}.$$

This congruence holds for all $(b, p_i) = 1$, $i = 1, 2, \ldots$, which implies $(b, N) = 1$. In other words, N is an *absolute* pseudoprime.

Another example is $N = 2821 = 7 \cdot 13 \cdot 31$, because 6, 12 and 30 all divide $N - 1 = 2820$. With $\phi(2820) = 2160$, there are 2160(!) bases b for which 2821 divides $b^{2820} - 1$.

The reader may wish to find other n-tuples of primes which give Carmichael numbers.

To get a better grip on pseudoprimes, the concept of a *strong pseudoprime* [20.4] has been introduced. Definition: N is a strong pseudoprime to the base b if it is composite and the following holds. With $N - 1 = K2^t$, K odd, consider

$$\begin{aligned} x_i &= \left\langle b^{K2^t} \right\rangle_N, \\ x_{t-1} &= \left\langle b^{K2^{t-1}} \right\rangle_N \\ &\vdots \\ &\vdots \\ x_0 &= \left\langle b^{K} \right\rangle_N, \end{aligned} \tag{20.19}$$

where $\langle\ \rangle_N$ means least absolute remainder modulo N. Now let x_i be the remainder with the largest index i that differs from $+1$. If $x_i = -1$, then N is called a *strong* pseudoprime to the base b, abbreviated spsp(b).

The strong pseudoprimality test is capable of unmasking even absolute pseudoprimes such as 561: While

$$2^{560} \equiv 2^{280} \equiv 1 \pmod{561},$$

the next halving of the exponent reveals 561 to be composite:

$$2^{140} \equiv 67 \not\equiv -1 \pmod{561}.$$

And 2821 is exposed as nonprime by the first halving of the exponent:

$$2^{1410} \equiv 1520 \not\equiv -1 \pmod{2821}.$$

According to *Rabin* [20.9], the probability is less than 0.25 that an odd N is in fact composite when N passes the above test. Thus, if the test for strong pseudoprimality is repeated, say, 50 times with 50 different bases, the probability that a composite number would have “slipped through the net” would be less than $(0.25)^{50} = 2^{-100} \approx 10^{-30}$.

As a consequence, if since the beginning of the universe (10^{17} seconds ago), one *million* primes had been tested every *second*, then the probability that a single composite number would be among the strong pseudoprimes so found would be is less than 10^{-7}. Of course, 10^{-7} is not zero, and with many extremely large numbers we shall probably *never* know with *certainty* whether they are composite or prime: they will keep their mystery forever.

Even so, strong pseudoprimality is a much more reliable indicator of primality than some "unambiguous" deterministic computations that take, say one hundred (or one trillion) years. Who will guarantee the fail-safety of the computer over such long running times?

And while we are talking of complexity and absolute provability, what about a proof that is more than 5000 pages long (or ten good-sized book volumes), like the recent proof that all finite simple groups are now known and that the so-called *monster group* is the largest sporadic simple group (i. e., has the largest order)?

Nevertheless, recent progress in *deterministic* primality testing promises reasonable speeds for numbers with more than a 1000 digits (Sect. 20.6).

20.5 Fermat and Strong Pseudoprimes

It is interesting to note that *all* Fermat numbers

$$F_n := 2^{2^n} + 1 \tag{20.20}$$

are either primes or strong pseudoprimes to the base 2. To see this, we write:

$$2^{F_n - 1} = 2^{2^{2^n}} = (2^{2^n})^{2^{2^n - n}} = (F_n - 1)^{2^{2^n - n}}. \tag{20.21}$$

Taking the remainder modulo F_n yields

$$\left\langle 2^{F_n - 1} \right\rangle_{F_n} = (-1)^{2^{2^n - n}} = 1. \tag{20.22}$$

Now, taking square roots will always give 1's until eventually, after $2^n - n$ square-rootings, a -1 will apear. Thus, all Fermat numbers are primes or strong pseudoprimes to the base 2 – a weak partial vindication of Fermat's erroneous conjecture ("quasi persuadé", to use Fermat's own words) that all F_n are in fact prime.

On the other hand

$$\left\langle 3^{F_5 - 1} \right\rangle_{F_5} = 3{,}816{,}461{,}520 \neq 1, \tag{20.23}$$

so that F_5 is not even an ordinary pseudoprime to the base 3.

20.6 Deterministic Primality Testing

Here we briefly review some background and the latest published advances in *deterministic* tests for primality [20.2]. Mersenne once remarked (in translation) that "to tell if a given number of 15 or 20 digits is prime or not, all time would not suffice for the test". How astonished he would be if he learned that the primality of 100-digit numbers can now, albeit 340 years later, be *ascertained* in a matter of minutes! And Gauss would be delighted too. In *Disquisitiones Arithmeticae* [20.10] he wrote (in Latin): "The dignity of the science itself [number theory, of course] seems to require that every possible means be explored for the solution of a problem so celebrated" (distinguishing prime numbers from composite numbers and resolving the latter into their prime factors).

However, the progress reported here is only one part of Gauss's "problem", namely *primality testing*. No comparable progress has been made on the second part of his problem, which is *factoring* [20.11]. We must always bear in mind the distinction between these two tasks. Primality testing has become very efficient in recent years, and that is good for public-key cryptography, because large primes are at a premium. Factoring, on the other hand, is still very difficult, but that is also good, nay vital, for the security of public-key cryptosystems.

The first deterministic primality test that we encountered was based on Wilson's theorem (Chap. 8): n is prime *iff* $n|(n-1)!+1$. Unfortunately, the factorial $(n-1)!$ destroys any attraction that Wilson's theorem might have evoked: the slowest sieve method will be lightning-fast compared to checking the divisibility of $(n-1)!+1$ for large n. If n has 100 digits, then $(n-1)!$ has roughly 10^{102} *digits* – not to be confused with the "relatively" small number 10^{102}; the utter futility is obvious.

Along came *Edouard A. Lucas* [20.12], who showed in 1876 that n is prime *iff* there is a b, coprime to n, such that

$$b^{n-1} \equiv 1 \pmod{n} \tag{20.24}$$

and

$$b^{\frac{n-1}{p}} \not\equiv 1 \pmod{n}$$

for *each* prime factor p of $(n-1)$.

Lucas's test is especially expeditious for $n = 2^m + 1$, because then $n-1$ has only one prime factor: 2.

Example: $n = 17$, $n-1 = 2^4 = 16$, $2^{16} \equiv 1 \pmod{17}$ and $2^8 = 256 \equiv 1 \pmod{17}$. Thus $b = 2$ does not tell us anything. But for $b = 3$, $3^{16} \equiv 1 \pmod{17}$ and $3^8 = 6561 \equiv 16 \not\equiv 1 \pmod{17}$. Thus, 17 *is* a prime.

The problem, of course, is that for large n not all the prime factors of $n-1$ may be known. Thus it came as a relief when in 1945, a primality test for n was formulated requiring knowledge of only *some* of the prime factors of $n-1$ and $n+1$. Later, a primality test was devised [20.13] for which partial factorization of

n^2+1 or n^2-n+1 or n^2+n+1 is required. However, if none of these numbers factor easily, the test does not fly.

Finally, in 1980, *Adleman* and *Rumely* [20.14] constructed a test that has drastically improved the efficacy of checking the primality of large numbers without restriction. With some improvements by H. Cohen and H. W. Lenstra, Jr., the test can now check a 100-digit number in less than a minute! And even a 1000-digit number can now be tested in about a week.

More generally, it is conjectured [20.2] that to test the primality of n requires a time not exceeding.

$$c(\ln n)^{b\ln(\ln(\ln n))},$$

where b and c are constants *independent of* n, and b is of order 1 and c depends, of course, on the calculating speed of the computer. With $c = 5\cdot 10^{-5}$ and $b \approx 1.5$ (empirical), the maximal testing time for a 2000-digit number is less than a year – only moderately unreasonable.

If we take present life expectancy at birth to be about 80 years, then a teenager can know the primality of any 5000-digit number during his expected lifetime – and sooner, if there is further progress in primality testing during the next decades, or perhaps the next few weeks. However, there will always be a "primality horizon" beyond which numbers will keep their prime secret "forever" – the half-life of a proton (ca. 10^{31} years), say.

Update: As a result of the Agrawal-Kayal-Saxena algorithm (2002), the situation has changed. Some of the recent developments are summarized at http://www.maths.anu.edu.au/∼brent/talks.html#Strachey

20.7 A Very Simple Factoring Algorithm

The simplest method for factoring N is to divide N by all primes up to $N^{1/2}$ and to watch for zero remainders. Divisibility by 2, 3, 5 and 11 is of course easily tested (see Sect. 6.1). But the main problem with this approach is the necessity to store many primes (more than 400 million primes for $N \le 10^{20}$).

The need to store any primes evaporates if we divide by *all* integers up to $N^{1/2}$. This method is of course very time consuming. A neat and simple algorithm, suitable for *pocket calculators*, follows from the following observation.

Suppose we *trial divide* N first by 2. If 2 is not a divisor of N, then trial dividing by any other even number is redundant. Thus, if suffices to trial divide only by 2 and odd numbers.

Similarly, after trial division with 3, trial division by any multiple of 3 is unnecessary. Thus, only trial divisors congruent 1 or 5 modulo 6 are still in the running.

Finally, after also trial dividing by 5, which is just as easy, only trial divisors congruent to 1, 7, 11, 13, 17, 19, 23 and 29 modulo 30 need be considered, i. e., 8 out of each 30 integers. I encourage the reader to program her or his calculator for this simple algorithm, which requires storing fewer than 10 integers.

Other factoring algorithms for calculators, home computers and large machines are described by H. C. Williams in “Factoring on a Computer” in *The Mathematical Intelligencer* (Volume 6, Number 3, 1984, pp. 29–36).

What are the factors (other than 2) of $11^{64}+1$, a 67-digit number? Is R_{1031} a “repunit” consisting of 1031 1’s (see Sect. 3.6) prime?

20.8 Factoring with Elliptic Curves

After decades of slow progress on factoring large integers, an unlikely candidate has given the factoring field a big boost: algebraic geometry, more specifically rational points on elliptic curves. Elliptic curves, defined by an equation of the form $y^2 = x^3 + ax + b$, have played a role also in the recent proof of Fermat’s Last Theorem (see Sect. 7.7).

The elliptic curve method (ECM) for factoring, is due to Hendrik Lenstra [Annals of Mathematics, 1987]. In 1997, Richard Brent showed that ECM could have been used to find the 49-digit factor of F9, but by then the factors of the 9th Fermat number F9 were already known by the Number Field Sieve (NFS) which doesn’t involve elliptic curves [Lenstra, Lenstra, Manasse & Pollard, Mathematics of Computation 61 (1993), 319–349].

F10 can be factored by ECM, and Brent did this in 1995. The possibility arises because the second-largest prime factor of F10 has “only” 40 decimal digits, so can be found by ECM.

Note that with ECM the work required depends mainly on the size of the factor, whereas for NFS the work depends on the size of the number being factored (so F10 is still beyond the range of NFS). Similarly, F11 can be factored by ECM, and again Brent succeed in this in 1988 soon after ECM was discovered, see http://www.maths.anu.edu.au/∼brent/F11.html

The smallest Fermat number whose complete factorisation is still unknown is F12. The reader can find a summary of the current status at http://www.prothsearch.net/fermat.html

20.9 Quantum Factoring

Enter quantum mechanics, the completely counter-intuitive field of physics of which Richard Feynman once said “I think it is safe to say that *nobody* understands quantum mechanics”. And the great Dane Niels Bohr thought that “If, in thinking about quantum mechanics, you don’t get *dizzy*, you really haven’t understood it”. (Einstein, however, didn’t go along with the general consensus and stuck to his “God doesn’t play dice” view of things.)

Quantum *computers*, still unrealized but looming on the horizon, can strikingly speed up the task of finding the prime factors of large numbers. Factors of a general n-digit number can be found in a time proportional to n^2 (as apposed to some exponential growth for classical computation). The still unsolved problem is to slow the decoherence (unavoidable in any quantum system with links to the classical world) sufficiently to extract the result of the computation [20.15].

Chapter 21
The Möbius Function and the Möbius Transform

After Euler's totient function, the Möbius function (named after the Möbius of strip fame) is one of the most important tools of number theory. It allows us to *invert* certain number-theoretic relations. In a sense, if we liken summation over divisors to integration, then taking the Möbius function is like differentiating. It is defined as follows:

$$\mu(n) := \begin{cases} 1 & \text{for } n = 1 \\ 0 & \text{if } n \text{ is divisible by a square} \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct } \textit{primes}. \end{cases} \tag{21.1}$$

Thus, $\mu(n) \neq 0$ only for 1 and squarefree integers. The Möbius function is *multiplicative*:

$$\mu(mn) = \mu(m)\mu(n) \quad \text{if } (m,n) = 1.$$

Also,

$$\mu(mn) = 0 \quad \text{if } (m,n) > 1, \tag{21.2}$$

because then mn is not squarefree.

The most important property of the Möbius function is that its summatory function (Sect. 11.4) equals zero, except for $n = 1$:

$$\sum_{d|n} \mu(d) = \delta_{n1}. \tag{21.3}$$

Here δ_{mn} is the *Kronecker symbol* defined by

$$\delta_{mn} := \begin{cases} 1 & \text{for } n = m \\ 0 & \text{for } n \neq m. \end{cases} \tag{21.4}$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_21, © Springer-Verlag Berlin Heidelberg 2009

21.1 The Möbius Transform and Its Inverse

Let F be the summatory function of f:

$$F(n) = \sum_{d|n} f(d). \tag{21.5}$$

Then

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) F(d), \tag{21.6}$$

and conversely, (21.6) implies (21.5). Thus, the Möbius function emerges as the one that permits us to determine a function from its summatory function.

$F(n)$ is also called the *Möbius transform* of $f(n)$ and $f(n)$ is called the *inverse* Möbius transform of $F(n)$ [21.1].

Example: As we have seen (and repeatedly used) in previous chapters, the following identity holds for Euler's function:

$$n = \sum_{d|n} \phi(d). \tag{21.7}$$

Thus, n is the Möbius transform of $\phi(n)$.

By applying the above inversion formula, we obtain the following expression for Euler's function:

$$\phi(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d = n \sum_{d|n} \mu(d) \frac{1}{d}, \tag{21.8}$$

a relation between Euler and Möbius that is good to know. How can we evaluate the sum on the right? We only need to consider squarefree divisors d^* of n because $\mu(d) = 0$ for nonsquarefree d. Writing

$$n = \prod_i p_i^{e_i}, \tag{21.9}$$

such divisors d^* are of the form p_i or $p_i p_j$ $(i \neq j)$, etc. Thus, with (21.1),

$$\begin{aligned} \sum_{d|n} \frac{1}{d} \mu(d) &= 1 - \sum_i \frac{1}{p_i} + \sum_{i \neq j} \frac{1}{p_i p_j} - \ldots \\ &= \prod_i \left(1 - \frac{1}{p_i}\right). \end{aligned} \tag{21.10}$$

In this manner, we have found another proof of

$$\frac{\phi(n)}{n} = \prod_i \left(1 - \frac{1}{p_i}\right), \tag{21.11}$$

the average of which, as we know, tends to $6/\pi^2$ (see Sect. 4.4).

The above proof is, of course, not complete without showing that (21.3) holds. Defining

$$\varepsilon(n) := \sum_{d|n} \mu(d), \tag{21.12}$$

we notice that trivially, $\varepsilon(1) = 1$. Next, we consider n to be a prime power p^α $(\alpha > 0)$. In that case, the divisors of n are $p, p^2, \ldots, p^\alpha$. Thus,

$$\varepsilon(p^\alpha) = 1 + \mu(p) + \mu(p^2) + \cdots + \mu(p^\alpha). \tag{21.13}$$

Here all terms vanish except the first two, which cancel. Hence

$$\varepsilon(p^\alpha) = 0, \quad (\alpha > 0). \tag{21.14}$$

The result for general n now follows from the fact that $\mu(n)$ is multiplicative, making its summatory function also multiplicative (which we do not prove here, but is easy to show).

21.2 Proof of the Inversion Formula

We want to show that for

$$F(n) = \sum_{d|n} f(d),$$

the inverse Möbius transform is

$$f(n) = \sum_{d|n} \mu(d) F\left(\frac{n}{d}\right),$$

and vice versa. We note that

$$F\left(\frac{n}{d}\right) = \sum_{d'|\frac{n}{d}} f(d') \tag{21.15}$$

and introduce this in the above equation, yielding

$$f(n) = \sum_{d|n} \mu(d) \sum_{d'|\frac{n}{d}} f(d'). \tag{21.16}$$

By inverting the order of summation over d and d', we obtain

$$f(n) = \sum_{d'|n} f(d') \sum_{d'|\frac{n}{d'}} \mu(d), \tag{21.17}$$

where the last sum equals zero, except for $n/d' = 1$, i. e., $d' = n$ – which proves (21.6). The converse is proved similarly.

According to the basic relation (21.3) the Möbius function $\mu(n)$ is the Möbius transform of the *Kronecker delta function* $\delta(n) = \delta_{n1}$ as defined in (21.4). The Möbius transform of $\delta(n)$ is the constant 1.

The Möbius transform of n^k is the generalized divisor function $\sigma_k(n)$ as defined in Chap. 11.

The *inverse* Möbius transform of n^k could be called a generalized Euler function $\phi_k(n)$, in view of the fact (21.8) that the inverse Möbius transform of n is $\phi(n) = \phi_1(n)$. The so defined generalized Euler function is given by

$$\phi_k(n) = n^k \prod_{p|n} \left(1 - \frac{1}{p^k}\right), \quad k = 1, 2, \ldots.$$

We mention one more function with interesting Möbius transform properties, the von Mangoldt function, which plays an important role in analytic number theory [21.2]:

$$\Delta(n) = \begin{cases} \ln p & \text{if } p \text{ is the only prime factor of } n \\ 0 & \text{if } n \neq p^m. \end{cases}$$

Its Möbius transform is the logarithm $\ln n$. The inverse Möbius transform of the von Mangoldt function is $-\mu(n)\ln n$ [21.3].

21.3 Second Inversion Formula

Let

$$G(n) = \prod_{d|n} g(d); \quad \text{then} \tag{21.18}$$

$$g(n) = \prod_{d|n} G(d)^{\mu(n/d)} \tag{21.19}$$

and vice versa. The proof follows from the one just given by setting $g = \mathrm{e}^f$ and $G = \mathrm{e}^F$.

Applying this formula to the one previously derived (Chap. 11)

$$\prod_{d|n} d = n^{1/2d(n)},$$

where $d(n)$ is the number of divisors of n, we obtain the following curious (and probably useless) representation of n^2:

$$n^2 = \prod_{d|n} d^{d(d)\mu(n/d)}. \tag{21.20}$$

21.4 Third Inversion Formula

For $x > 0$, let

$$H(x) = \sum_{n=1}^{\infty} h(nx). \quad \text{Then} \tag{21.21}$$

$$h(x) = \sum_{n=1}^{\infty} \mu(n) H(nx) \tag{21.22}$$

and vice versa, provided

$$\sum_{m,n} |f(mnx)|$$

converges.

21.5 Fourth Inversion Formula

For $x > 0$, let

$$H(x) = \sum_{n=1}^{|x|} h\left(\frac{x}{n}\right). \quad \text{Then} \tag{21.23}$$

$$h(x) = \sum_{n=1}^{|x|} \mu(n) H\left(\frac{x}{n}\right). \tag{21.24}$$

21.6 Riemann’s Hypothesis and the Disproof of the Mertens Conjecture

The Möbius function is involved in the following noteworthy sums:

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0, \tag{21.25}$$

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} \ln n = -1, \tag{21.26}$$

the former one being the result of E. Landau’s dissertation.

The partial sums over the Möbius function

$$M(N) := \sum_{n=1}^{N} \mu(n)$$

were conjectured by Mertens in 1897 to obey the bound $|M(N)| < N^{1/2}$ for all $N > 1$. The truth of this conjecture, which has attracted much attention over the last 100 years, would imply the Rieman hypothesis (see Sect. 4.2). But, alas, the Mertens conjecture was disproved in 1984 by A. Odlyzko and H. J. J. te Riele, although no counterexamples had been found for N up to 10^{10}. In fact, there may be no counterexamples below 10^{20} or even 10^{30}, but Odlyzko and te Riele proved that, for some N, $M(N)N^{-1/2}$ exceeds 1.06 (and may, in fact, be unbounded for $N \to \infty$). They did not actually give a counterexample but expect one to occur "near" $10^{10^{65}}$. Here, once more, *numerical* evidence was – and would continue to be! – very misleading. But, ironically, the disproof involved heavy numerical computation (of the zeros of the zetafunction).

21.7 Dirichlet Series and the Möbius Function

A Dirichlet series is of the form [21.1]:

$$F(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}, \tag{21.27}$$

where $F(s)$ is called the *generating function* of the function $f(n)$, which appears as the coefficient of n^{-s} in a Dirichlet series. In the following we shall ignore questions of convergence.

If $f(n)$ is multiplicative, then

$$F(s) = \prod_p \left[1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \ldots\right], \tag{21.28}$$

and if $f(n)$ is *completely* multiplicative then

$$F(s) = \prod_p \left[1 - \frac{f(p)}{p^s}\right]^{-1}, \tag{21.29}$$

For $f(n) = 1$ in (21.27) or (21.29), $F(s)$ equals the Riemann zetafunction $\zeta(s)$.

If we have a second Dirichlet series:

$$G(s) = \sum_{n=1}^{\infty} \frac{g(n)}{n^s}, \tag{21.30}$$

then the product of $F(s)$ and $G(s)$ is given by

$$F(s)G(s) = \sum_{n=1}^{\infty} \frac{h(n)}{n^s} \quad \text{where} \tag{21.31}$$

$$h(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right), \tag{21.32}$$

which is called a *number-theoretic convolution*. In other words, if two arithmetic functions are "convolved" as in (21.32), then the generating function is the product of their Dirichlet series.

Because, by definition (21.1):

$$\mu(p^k) = 0 \quad \text{for} \quad k > 1,$$

we have

$$1 - \frac{1}{p^s} = 1 + \frac{\mu(p)}{p^s} + \frac{\mu(p^2)}{p^{2s}} + \ldots, \tag{21.33}$$

and therefore

$$\frac{1}{\zeta(s)} = \prod_p \left[1 - \frac{1}{p^s}\right] = \prod_p \left[1 + \frac{\mu(p)}{p^s} + \frac{\mu(p^2)}{p^{2s}} + \ldots\right], \tag{21.34}$$

which, with (21.28) and (21.27), can be written:

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \zeta^{-1}(s). \tag{21.35}$$

For $s = 2$, we obtain the interesting result:

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} = \frac{6}{\pi^2} \tag{21.36}$$

In parting, we cite a few similarly derived Dirichlet series results, which the reader may wish to prove:

$$\sum_{n=1}^{\infty} \frac{|\mu(n)|}{n^s} = \frac{\zeta(s)}{\zeta(2s)}, \tag{21.37}$$

$$\sum_{n=1}^{\infty} \frac{d(n)}{n^s} = \zeta^2(s), \tag{21.38}$$

$$\sum_{n=2}^{\infty} \frac{\ln n}{n^s} = -\zeta'(s), \tag{21.39}$$

$$\sum_{n=1}^{\infty} \frac{(\ln n)^2}{n^s} = \zeta''(s), \tag{21.40}$$

$$\sum_{n=2}^{\infty} \frac{\Delta(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)}. \tag{21.41}$$

If $g(n)$ is the Möbius transform of $f(n)$, then their generating functions, $G(s)$ and $F(s)$, are related by

$$G(s) = \zeta(s)F(s). \tag{21.42}$$

Setting $f(n) = \phi(n)$, so that $g(n) = n$ and therefore

$$G(s) = \sum_{n=1}^{\infty} \frac{n}{n^s} = \zeta(s-1),$$

we obtain the relationship:

$$\sum_{n=1}^{\infty} \frac{\phi(n)}{n^s} = \frac{\zeta(s-1)}{\zeta(s)} \tag{21.43}$$

and, in a similar manner,

$$\sum_{n=1}^{\infty} \frac{\phi_k(n)}{n^s} = \zeta(s)^{-1}\zeta(s-k). \tag{21.44}$$

Relations between *prime* divisor functions and the zetafunction were already mentioned in Sect. 12.3 (12.40–12.43), although the relation with $\zeta(s)$ was not made explicit there, it is given by the following two identities:

$$\frac{\zeta(2s)}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{(-1)^{\Omega(n)}}{n^s} \quad \text{and} \tag{21.45}$$

$$\frac{\zeta^2(s)}{\zeta(2s)} = \sum_{n=1}^{\infty} \frac{2^{\omega(n)}}{n^s}. \tag{21.46}$$

Chapter 22
Generating Functions and Partitions

The generating functions introduced in Chap. 21 and defined by Dirichlet series are not the only kind of generating functions. Here we shall briefly get to know another type of generating function with many useful properties that are applicable in numerous fields of mathematics and other sciences. As an illustration of that utility, we shall acquaint ourselves with various partition problems such as the partitions of our main subject: the positive integers. For example, the integer 4 has 5 unrestricted partitions into integers:

$$4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1.$$

The subject of partitions is closely connected with Elliptic Modular functions [22.1], one of the more fertile fields of mathematics, which, however, is beyond the scope of this book.

Applications of partition functions abound throughout the natural sciences – we mention only (quantum) statistical physics, where questions such as how to partition a given number of energy quanta or, perhaps, a fluctuating number of photons among various "bins" or *eigenstates* are at the very root of our understanding of what radiation and matter are about.

22.1 Generating Functions

A *generating function* $F(s)$ for the function $a(n)$ can be defined quite generally by:

$$F(s) = \sum a(n)u(n,s), \tag{22.1}$$

where, again, we ignore questions of convergence [22.2].

With $u(n,s) = n^{-s}$, (22.1) is a Dirichlet series as discussed in Sect. 21.7. Another choice of $u(n,s)$ is:

$$u(n,s) = \mathrm{e}^{-sn},$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_22, © Springer-Verlag Berlin Heidelberg 2009

or, with

$$z = \mathrm{e}^{-n},$$
$$F(z) = \sum a(n) z^{-n}. \tag{22.2}$$

$F(z)$ is also known as the *z transform* of $a(n)$, especially in engineering applications – to time-discrete electrical circuits, for example [22.3].

While Dirichlet series, as a result of (21.31) and (21.32), play an important role in *multiplicative* problems of number theory (in particular the theory of primes and divisibility), the z transform is dominant in *additive* problems, because of the following all-important convolution property: If $F(z)$ is the z transform of $a(n)$ as defined in (22.2), and $G(z)$ is the z transform of $b(n)$, then the z transform of the *convolution* (ubiquitous in linear systems)

$$c(n) = \sum_k a(k) b(n-k) \tag{22.3}$$

equals the *product* of the z transforms:

$$\sum c(n) z^{-n} = F(z) G(z). \tag{22.4}$$

The decisive difference between Dirichlet series and z transforms is, of course, that in the former the variable (s) appears in the *exponent*, whereas in the latter the variable (z) appears as a base.

If we take z as a complex variable and consider it on the unit circle:

$$z = \mathrm{e}^{\mathrm{i}\theta},$$

then the z transform turns into a *Fourier series* – another important property of the power-series type generating function as defined by (22.2).

Generating functions of this type also play a great role in *probability theory*, where they are known as *characteristic functions*. They generate statistical *moments* or *cumulants* of probability distributions and permit their convenient manipulation [22.4]. Since summing two or more independent random variables results in a convolution of their distributions, the corresponding characteristic functions are simply multiplied to obtain the characteristic function of the summed random variables. (Thus, for example, the sum of several Poisson variables is immediately seen to be another Poisson variable.)

As a simple example of a generating function, we consider

$$F(z) = (1+z)^n,$$

which, of course, generates the binomial coefficients $\binom{N}{n}$:

$$(1+z)^N = \sum_{n=0}^{N} \binom{N}{n} z^n, \tag{22.5}$$

which tell us how many ways n objects can be selected from a set of N distinct objects without regard to order. (Often power series as in (22.5) are in z rather than z^{-1}, but that is a matter of choice depending on the field of application.)

We may also mention here the simple generating function of the Fibonacci numbers (Chap. 5):

$$\sum_{n=0}^{\infty} F_n z^n = \frac{z}{1-z-z^2}, \tag{22.6}$$

which is very useful in obtaining closed-form expressions for F_n as in (5.30).

22.2 Partitions of Integers

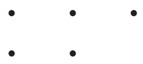
Let us ask ourselves how many ways $p(n)$ a positive integer n can be partitioned so that the sum of the parts equals n – without regard to order and with no restrictions (other than that the parts shall also be positive integers).

Example: $p(5) = 7$, because:

$$\begin{aligned} 5 &= 4+1 = 3+2 = 3+1+1 = 2+2+1 \\ &= 2+1+1+1 = 1+1+1+1+1, \end{aligned}$$

i. e., there are 7 ways to partition 5 in the prescribed manner.

Partitions obey an endearing symmetry. For example, consider the 5 dots:



This geometrical representation of a partition, called a Ferrer graph, illustrates *two* partitions of the integer 5. Reading column-wise (vertically) we see that $5 = 2+2+1$ (which we already knew), and reading row-wise (horizontally) we discover that $5 = 3+2$. Thus, each partition has a (not necessarily distinct) *conjugate*.[1] Note that the first partition ($5 = 2+2+1$) is into 3 integers of which the largest are 2, and the second partition ($5 = 3+2$) is into 2 integers of which 3 is the largest. Some partitions are self-conjugate, as for example $3+1+1$.

The following two theorems are an interesting consequence of conjugacy [22.2]:

1) *The number of partitions of n into m integers is equal to the number of partitions of n into integers the largest of which is m.*

For example, there are 2 partitions of 5 into 2 integers ($4+1$ and $3+2$) and there are equally 2 partitions of which 2 is the largest integer ($2+2+1$ and $2+1+1+1$).

[1] In quantum mechanics such geometric representations of partitions are known as Young tableaux, after the inventor who introduced them to the study of symmetric groups. They are important in the analysis of the symmetries of many-electron systems.

2) *The number of partitions of n into* at most *m integers is equal to the number of partitions of n into integers which do not exceed m.*

For example, there are 3 partitions of 5 into *at most* 2 integers (5, $4+1$ and $3+2$) and there are equally 3 partitions whose integers do not exceed 2, namely $2+2+1$, $2+1+1+1$ and $1+1+1+1+1$. Astounding consequences, considering the simplicity of the underlying symmetry!

22.3 Generating Functions of Partitions

The useful generating functions for *partitions* are of the power-series (or z-transform) type, because when we talk about partitioning, we are talking *additively*.

The generating function of $p(n)$ was found by Euler. Defining $p(0) = 1$, it is

$$\sum_{n=0}^{\infty} p(n)z^n = \frac{1}{(1-z)(1-z^2)(1-z^3)\dots}. \tag{22.7}$$

To see that Euler was right, we expand each factor into a geometric series:

$$\begin{array}{rl} \textit{integer } 1: & (\underline{z}^0 + \underline{\underline{z}}^1 + z^2 + z^{\overline{\overline{\overline{3}}}} + \dots) \\ \textit{integer } 2: & (\underline{z}^{\overline{\overline{\overline{0}}}} + \underline{\underline{z}}^2 + z^4 + z^6 + \dots) \\ \textit{integer } 3: & (\overline{\overline{\overline{\underline{\underline{z}}^0}}} + \underline{z}^3 + z^6 + z^9 + \dots) \\ \textit{taken} & \;\;0 \quad\;\; 1 \quad\;\; 2 \quad\;\; 3 \quad \dots \textit{ times}, \text{ etc.} \end{array} \tag{22.8}$$

In multiplying this out, we have to take one term from each factor. For example, what is the coefficient of z^3, i. e., how many ways are there to get z^3?

First, we take the term z^3 in the third factor (1-column) and the two 1's in the first two factors. (These 3 terms, one from each factor, are identified by underlining with a single horizontal stroke.) Next we take the term z^1 in the first factor, the term z^2 in the second factor, and the 1 in the third factor, indicated by underlining with a double stroke. This gives us another z^3 term in the product. Third, we take the factor z^3 from the first factor (3-column) and the 1's from the other two factors, as shown by overlining with a triple stroke. This gives us a third z^3 term. Since there are no more ways of making z^3 terms (the next factor – not shown in (22.8) – is $1+x^4+x^8+\dots$), we conclude, tentatively, that

$$p(3) = 3.$$

And indeed, there are precisely 3 partitions of 3: $3 = 2+1 = 1+1+1$.

Why is the Euler formula (22.7) correct? In what way did our way of picking z^3 terms correspond to the 3 partitions of 3? Let us identify the *first* factor in (22.8) with the partitioning integer 1. Having picked the term z^3 (3-column) from it (triple overlined) we say we have used the integer 1 exactly 3 times. In other words, we identify this z^3 term with the partition $1+1+1$: the integer 1 (first row) taken 3 times (3-column). Similarly, we identify the second and third factors in (22.8) with the partitioning integers 2 and 3, respectively. Thus, the double underlined contribution to the z^3 term in the product corresponds to the partition $1+2$. The single underlined terms correspond to the partition 3; in other words, the integer 3 is taken once (1-column) and the integers 1 and 2 are both taken 0 times (0-column). *This is how generating functions function.*

A more formal way to define partitions is to consider the expression

$$k_1 + 2k_2 + 3k_3 + \cdots = n, \tag{22.9}$$

and to ask how many sets of nonnegative integers k_m there are that satisfy (22.9). Here k_m is the number of times the integer m appears in the partition of n.

22.4 Restricted Partitions

The above discussion of the generating function (22.7) for unrestricted partitions immediately allows us to write down the generating functions $F_m(z)$ for the partition of n into integers the largest of which is m:

$$F_m(z) = \frac{1}{(1-z)(1-z^2)\dots(1-z^m)}. \tag{22.10}$$

Similarly, the generating function for partitions into *even* integers exceeding 4, for example, is

$${}_6F_{\text{even}}(z) = \frac{1}{(1-z^6)(1-z^8)(1-z^{10})\dots}.$$

If we are interested in partitions of n into *distinct* integers, a moment's thought will show that the generating function is

$$F_{\text{dist}}(z) = (1+z)(1+z^2)(1+z^3)\dots. \tag{22.11}$$

Since each factor in (22.11) has only one term other than 1, we can take each integer at *most* once, i. e., the integers in any partition are unequal or *distinct*. (Remember that the individual factors in these generating functions stand for the different integers.)

We shall illustrate the supreme usefulness of generating functions by the following example. $F_{\text{dist}}(z)$ according to (22.11) can be rewritten

$$\begin{aligned} F_{\text{dist}}(Z) &= \frac{1-z^2}{1-z}\,\frac{1-z^4}{1-z^2}\,\frac{1-z^6}{1-z^3}\cdots \\ &= \frac{1}{(1-z)(1-z^3)(1-z^5)\dots}. \end{aligned} \tag{22.12}$$

But the latter generating function is $F_o(z)$, corresponding to partitions into *odd* integers.

Thus, we have found the remarkable result that *distinct* partitions (that take each integer *at most once* without repetition) and partitions into *odd* integers are equally numerous:

$$p_{\text{dist}}(n) = p_o(n). \tag{22.13}$$

Example: There are three partitions of 5 into distinct integers, namely 5, $4+1$ and $3+2$. And there are also three partitions of 5 into *odd* integers: 5, $3+1+1$ and $1+1+1+1+1$. Check!

Suppose a hypothetical physical situation depends critically on the partition of n elementary particles or quanta over a number of bins or eigenstates. Then the purely arithmetic relation (22.13) would tell us that we could never distinguish between two different interpretations or "statistics", namely "distinct" *versus* "odd". Distinctly odd!

Could the reader have discovered the "parity" law (22.13) without generating functions? He may wish to prove his prowess with generating functions by proving the following simple generalization of (22.13): The number of partitions of n with no part repeated more than r times is equal to the number of partitions of n with no part divisible by $r+1$.

For $r=1$, we obtain the above result (2.13). But for $1<r<n-3$, interesting new results emerge. For example, for $n=6$ and $r=2$, there are 7 (out of a total of 11) partitions that do not use the integers 3 and 6: $5+1=4+2=4+1+1=2+2+2=2+2+1+1=2+1+1+1+1=1+1+1+1+1+1$. And there are also 7 partitions not using any integer more than twice $6=5+1=4+2=3+3=4+1+1=3+2+1=2+2+1+1$.

Let us turn the argument around and write down certain polynomial identities and ask what they mean in terms of partitions. Consider the wellknown (?) identity:

$$\frac{1}{1-z} = (1+z)(1+z^2)(1+z^4)(1+z^8)\dots. \tag{22.14}$$

Its correctness can be established by multiplying it with $(1-z)$, which yields

$$1 = (1-z^2)(1+z^2)(1+z^4)(1+z^8)\dots.$$

And now a mathematical chain reaction sets in: Combining the first two factors yields

$$(1-z^4)(1+z^4)(1+z^8)\dots,$$

which in turn yields

$$(1-z^8)(1+z^8)\dots, \text{ etc.,}$$

which converges to 1 for $|z|<1$.

But what does (22.14) imply in terms of partitions? The left side is the generating function of the integers themselves; i. e., it generates each integer exactly once. The right side of (22.14) is the partition of the integers into powers of 2 without repetition. Thus (22.14) tells us that each integer can be represented uniquely as a sum of powers of 2 (including $2^0 = 1$). How comforting to know!

Manipulations of generating functions, as illustrated by the identity (22.14), lead to several fascinating interpretations as partitions. For example, Euler proved that

$$\begin{aligned}(1-z)(1-z^2)(1-z^3)\cdots &= \sum_{n=-\infty}^{\infty} (-1)^n z^{n(3n+1)/2} \\ &= 1 - z - z^2 + z^5 + z^7 - z^{12} - z^{15} + \ldots, \quad (22.15)\end{aligned}$$

which, by multiplying with $F_{\text{dist}}(z)$ from (22.12), can be converted into a more easily recognizable form:

$$\frac{(1-z^2)(1-z^4)(1-z^6)\cdots}{(1-z)(1-z^3)(1-z^5)\cdots} = 1 + z + z^3 + z^6 + z^{10} + \ldots, \qquad (22.16)$$

where the exponents are the familiar triangular numbers $n(n+1)/2$ (see Sect. 7.4). (In other words, the left side of (22.16) is the generating function of the triangular numbers.) The exponents in (22.15) are the pentagonal numbers $g_5(n)$ and $g_5(-n)$, see Sect. 7.4.

The identity (22.15) has a striking combinatorial interpretation. The coefficient of z^n on the left side of (22.15) is

$$\sum (-1)^k,$$

where the summation is over all partitions into distinct integers and k is the number of integers in such a partition. Thus, the left side of (22.15) represents the difference between partitions (into distinct integers) with an even and an odd number of integers, respectively.

Calling the number of partitions of n into an *even* number of *distinct* integers $p_{\text{even}}(n)$, and the number of partitions into an odd number of distinct integers $p_{\text{odd}}(n)$, we get

$$p_{\text{even}}(n) - p_{\text{odd}}(n) = \begin{cases} (-1)^k & \text{for } n = k(3k \pm 1)/2 \\ 0 & \text{otherwise,} \end{cases} \qquad (22.17)$$

where $n = k(3k \pm 1)/2$ corresponds to the terms occurring on the right side of (22.15).

Example: $n = 8$, which does not appear as an exponent on the right side of (22.15). Thus, $p_{\text{even}}(8)$ should equal $p_{\text{odd}}(8)$. And indeed, there are three partitions into distinct integers of each kind:

$$\begin{aligned}&7+1 = 6+2 = 5+3 && \text{(even number of distinct integers)},\\ &8 = 5+2+1 = 4+3+1 && \text{(odd number of distinct integers)}.\end{aligned}$$

The identity (22.15) also leads to a recurrence formula [22.5] for $p(n)$:

$$p(n) = \frac{1}{n} \sum_{k=1}^{n} \sigma(k) p(n-k), \quad p(0) = 1, \tag{22.18}$$

where $\sigma(k)$ is the sum of the divisors of k (see Chap. 11).

With the asymptotic behaviour of $\sigma(k)$ (Sect. 11.4), one obtains an asymptotic law [22.5] for $p(n)$ for large n:

$$p(n) \approx \exp[\pi(2n/3)^{\frac{1}{2}}]/4n3^{\frac{1}{2}} \tag{22.19}$$

Example: for $n = 243$ (21.19) gives $p(n) \approx 1.38 \cdot 10^{14}$. The exact result,

$$p(243) = 133978259344888,$$

contradicts a famous conjecture by S. Ramanujan [22.6, 7] (see also [22.1], p. 289) that *if*

$$d = 5^a 7^b 11^c$$

and

$$24n \equiv 1 \pmod{d},$$

then

$$p(n) \equiv 0 (\operatorname{mod} d).$$

But $p(243)$ is *not* divisible by $d = 7^3 = 343$, although $24 \cdot 243 \equiv 1 \pmod{343}$.

However, Ramanujan found a number of striking congruence relations for $p(n)$ that stood the test of time (and proof). One of them involves the modulus 5:

$$p(5m+4) \equiv 0 (\operatorname{mod} 5). \tag{22.20}$$

In the meantime, many more congruence relations for partitions (and other combinatorial identities) have been discovered and proved, stimulating a lot of good mathematics [22.8].

We conclude this chapter with another striking discovery by Ramanujan involving a (generalized) continued fraction (Chap. 5):

$$1 + \frac{\mathrm{e}^{-2\pi}}{1+} \frac{\mathrm{e}^{-4\pi}}{1+} \frac{\mathrm{e}^{-6\pi}}{1+} = \mathrm{e}^{-2\pi/5} (5^{1/4} g^{1/2} - g)^{-1}, \tag{22.21}$$

where g is the Golden ratio:

$$g = \frac{1}{2}(5^{1/2} + 1) = 1.618\ldots .$$

Chapter 23
From Error Correcting Codes to Covering Sets

23.1 Covering Sets in Coding Theory

Covering sets are well known from coding theory, including the digitizing of analog data. For example, for 2-dimensional data distributed over R_2, to be quantized into m digital words, regions around m representative points must *cover* the entire support of the analog data. To minimize the mean-square error for a given value of m, the representation points for each region must be at the centers of gravity of the different regions and the border lines between two adjacent regions must be half-way between the representation points. The resulting tessellation of the plane is called a Voronoi diagram.

For one-dimensional data, the corresponding quantizing scheme is called a Max-Lloyd quantizer; in which the support of the data is divided into m contiguous segments, each segment being represented by a representative (quantized) value equal to the mean of the data for that segment and border lines half-way between them.

While Claude Shannon established upper limits for the error-free information rate of a given communication channel, his proofs were *existence* proofs. It fell to *coding theory* to develop digital codes which, for a given information rate, coding delay (block length) and channel noise, gave an actual error rate. To minimize the error rate, a wide range of *error-correcting codes* were invented. For some situations the problem was akin to densest sphere packing as epitomized, in 3 dimensions, by the stacking of oranges in a grocery store (see J.H. Conwey and N.J.A. Sloan: *Sphere Packings, Lattices and Groups*).

One of the oldest (and simplest) codes is three-fold repetition in which the information at the sending station, e.g. a single bit, is repeated 3 times and the receiving station takes a "majority vote". Of course this is guaranteed to work only for single errors.

An early, more sophisticated error-correcting code was designed by Richard Hamming in 1949 at Bell Laboratories: the 7-4 Hamming code in which 4 information bits are supplemented by 3 control bits yielding a total of 7 bits for one code word. In general, allowing m check bits, the Hamming code has length $2^{\mathrm{m}} - 1$ of which $2^{\mathrm{m}} - 1 - \mathrm{m}$ bits are information bits. For the original Hamming code discussed above m equals 3. If $\mathrm{m} = 2$, we obtain the triple-repetition with length

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_23, © Springer-Verlag Berlin Heidelberg 2009

$2^2-1=3$ and $2^2-1-2=1$ information bit. Thus, the triple-repetition code is a Hamming code. Its transmission efficiency (number of information bits divided by code length) equals $\varepsilon = 1/3$. For the $\mathrm{m}=4$ Hamming code, $\varepsilon = 11/15 = 0.73\ldots$ In general, $\varepsilon = 1-\mathrm{m}/(2^{\mathrm{m}}-1)$. The efficiency is therefore seen to increase monotonically to 1 as m goes to infinity.

However, no matter how large m, the Hamming codes can at most correct a single error. Larger error rates can be accommodated by the *Simplex codes*, see Sect. 1.5. The Simplex codes are the duals of the Hamming codes, in which the role of information bits and check bits are interchanged. Thus for $\mathrm{m}=5$ and code length $2^5-1=31$, the Simplex code has 5 information bits and 26 check bits. Its transmission efficiency ε equals only $5/31=0.16$ but it can correct $2^{\mathrm{m}-2}-1=7$ errors. As m goes to infinity, an astonishing 25% of the errors can be corrected, see Sect. 28.2.

We now turn to a different kind of covering set:

23.2 Discrete Covering Sets

The even numbers, i.e. the numbers n for which $n \equiv 0 \pmod 2$ cover all integers except the odd integers: $\ldots -1, 1, 3, 5 \ldots$ By adding a second covering set, defined by $n \equiv 0 \pmod 3$, we leave uncovered all integers not divisible by either 2 or 3: i.e. $\ldots -1, 1, 5, 7, 11, 13, \ldots$ Now adding a third covering set $n \equiv 1 \pmod 4$ covers $\ldots -3$ (which is already covered by $n \equiv 0 \pmod 3$), 1, 5, 9 (which is also already covered by $n \equiv 0 \pmod 3$), 13, 17, ...). The numbers still not covered by these three sets are $\ldots -1, 7, 11, 19, 23, \ldots$ which suggests the additional covering sets $n \equiv 1 \pmod 6$ and $n \equiv 11 \pmod{12}$. It is relatively easy to show that these five sets cover all integers *ad infinitum*. (One only has to check a range equal to the least common multiple of 2, 3, 4, 6, 12, i.e. 12, beyond which patterns repeat.)

At the Joint Mathematics Meeting (of the American Mathematical Society and the Mathematical Association of America) in January, 2008 in San Diego, Carl Pomerantz of Dartmouth College described recent progress in the *theory of coverings*, i.e. collections of integer sequences that include every integer from minus infinity to plus infinity. My friend and former Bell Labs colleague Jeffrey Lagarias, now at the University of Michigan, stressed the continuing importance of discrete covering problems—which are easy to state but difficult to solve, illustrating how little we know about *additive* number theory (think of the infamous Goldbach conjecture and other enduring unsolved additive problems—see Sects. 7.7 and 7.8).

There are many different ways to cover the integers using $n \equiv 0$ mod 2, $n \equiv 0 \pmod 3$ and $n \equiv 1 \pmod 4$. But for smallest steps sizes much larger than 2, the covering problem quickly becomes difficult, so difficult in fact that Paul Erdös called it his "favorite problem", namely to prove that, even for an *arbitrarily large* smallest step size, there are always solutions with a *finite* number of integer sequences. But so far the smallest step size of 36 is the largest one for which a solution is known. According to Pace Nielson of the University of Iowa the number of sequences required with 36 as the smallest step size may exceed 10^{40} (!).—The reader may try to find a solution with the smallest step size equal to 3. To avoid trivial solutions, we require that each step size be used only once, otherwise the two sequences n $\equiv 0$ (mod 2) and $n \equiv 1$ (mod 2) would already cover all integers. One possible approach to the problem is a "greedy algorithm", in which one always chooses as the next additional set the set which contains the largest number of uncovered elements.

Covering sets are part of *additive* number theory, a notoriously difficult field. Here are a few recent results from the strange field of additive number theory: In 1960 W. Sierpinski of Warsaw University showed that, for infinitely many values of k, numbers the form $k2^n + 1$ are composite. But according to J. Selfridge k may have to be as large as 78 557. A smaller candidate value of k, $k = 33\,661$, was recently (2007) shown by S. Sunde to yield a prime for $n = 7\,031\,232$, a number with more than 2 million digits!

Part VIII
Cyclotomy and Polynomials

Chapter 24
Cyclotomic Polynomials

Cyclotomy, the art of dividing a circle into equal parts, was a Greek specialty, and the only tools allowed were a straightedge and a compass. The subject is deeply related to number theory, as we saw in our discussion of Fermat primes in Sect. 3.8. In addition, *cyclotomic polynomials* play an important role in modern digital processes and fast computation (Sect. 26.3).

We begin this chapter with an informal discussion on dividing a circle into equal parts and constructing regular polygons. After that we explore some of the important properties of cyclotomic polynomials and their applications in physics, engineering and, in Chap. 30, artistic design.

24.1 How to Divide a Circle into Equal Parts

A circle in the complex number plane $z = x + \mathrm{i}y$, centered at the origin $z = 0$, is described by the equation

$$z = r\mathrm{e}^{\mathrm{i}\alpha}, \tag{24.1}$$

where r is the radius of the circle and α the angle between the direction of z and the real axis. As α goes from 0 to 2π, the circle is traversed exactly once.

If we only care about angles α, and not about the radius r of the circle, we might as well set $r = 1$ and consider the unit circle,

$$z = \mathrm{e}^{\mathrm{i}\alpha}. \tag{24.2}$$

Dividing the circle into, say, n equal parts means finding an angle $\alpha_{1/n}$ such that

$$\alpha_{1/n} = \frac{2\pi}{n}. \tag{24.3}$$

Then the incongruent $(\bmod 2\pi)$ multiples of $\alpha_{1/n}$ divide the circle into n equal parts.

Of course, $\alpha_{1/n}$ may not be unique. In fact, for $(k,n) = 1$,

$$\alpha_{k/n} = k\frac{2\pi}{n} \tag{24.4}$$

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_24, © Springer-Verlag Berlin Heidelberg 2009

will also give the desired "Cyclotomy" – the learned expression for circle division. Since there are $\phi(n)$ incongruent coprime values of k, there are $\phi(n)$ primitive, or generating, angles $\alpha_{k/n}$, but of course, knowledge of one of these suffices to do the job.

The n equidistant points on the unit circle

$$z_m = e^{i2\pi m/n}, \quad m = 0, 1, \ldots, n-1, \tag{24.5}$$

not only divide the circle into equal parts but are also the corners of a regular n-gon inscribed in the circle. Thus, we see that the tasks of equal division of the circle and construction of a regular n-gon are related.

Mathematically, finding the z_m of (24.5) is the same as finding the roots of the polynomial

$$z^n - 1 = 0, \tag{24.6}$$

i.e., we have converted our geometrical problem into an arithmetic one. Upon setting z in (24.6) equal to z_n from (24.5), we obtain

$$(e^{i2\pi m/n})^n - 1 = e^{i2\pi m} - 1 = 1 - 1 = 0,$$

for *all* integer values of z_m, including the n incongruent ones:

$$m = 0, 1, \ldots, n-1. \tag{24.7}$$

But how do we find the z_m, or at least *one* primitive root?

The Case $n = 2$.

Dividing the circle into two parts ($n = 2$) is simple: the equation

$$z^2 - 1 = 0 \tag{24.8}$$

has the well-known solutions $z = 1$ and $z = -1$, corresponding to the angles $\alpha = 0$ and $\alpha = \pi$ (180°), which indeed divides the circle into two equal parts. In other words, a straight line through the (known!) centre of the circle divides it into two equal parts.

The Case $n = 3$.

To divide the circle into three equal parts, the equation we have to solve is

$$z^3 - 1 = 0, \tag{24.9}$$

with the solutions

$$z_{m/3} = e^{i2\pi m/3}, \quad m = 0, 1, 2, \tag{24.10}$$

corresponding to the angles 0, 120° and 240°.

Now comes a crucial question. The reader may recall that we started out with a geometrical problem: the equal division of a circle or the construction of a regular polygon. It is therefore legitimate – and quite natural – to ask whether we can, in fact, solve $z^n - 1 = 0$ by "geometric means", i.e., with a *straightedge* and a *compass* but no other tools.

Let us look more closely at the problem for $n = 3$, the regular trigon – otherwise known as an equilateral triangle. First we notice that $z^3 - 1$ can be factored

$$z^3 - 1 = (z-1)(z^2 + z + 1). \tag{24.11}$$

Here the first factor gives the universal solution (valid for all n) $z = 1$. To find the two remaining solutions, we have to solve a *quadratic* equation, namely

$$z^2 + z + 1 = 0, \tag{24.12}$$

which is easy:

$$z = -\frac{1}{2} \pm \frac{1}{2}\sqrt{-3} = \mathrm{e}^{\pm \mathrm{i}2\pi/3}, \tag{24.13}$$

as expected.

But the important point now is that we only had to solve a *quadratic* equation – and this is something that can be accomplished purely by straightedge and compass, as desired by the purist Greeks. In fact, (24.13) tells us that we should bisect the negative real axis between $z = 0$ and $z = -1$ (the second intersection of the straight line through $z = 0$ and $z = 1$, both of which can be selected arbitrarily). Erecting the normal at $z = -1/2$ (also possible with the means at hand) we find that it intersects the circle at $x = -1/2$ and $y = \pm\sqrt{3/2}$.

Unwittingly, by this construction we have solved the quadratic equation

$$y = \sqrt{1^2 - \left(\frac{1}{2}\right)^2} = \pm\frac{1}{2}\sqrt{3}, \tag{24.14}$$

because by erecting the normal at $z = -1/2$ and locating its intersection with the circle $|z| = 1$, we have constructed a right triangle with a hypotenuse of length 1 and one side of length $1/2$. The length of the other side is then of course given by (24.14).

The Case $n = 4$.

How about the regular tetragon, widely known (in mathematics and other disciplines) as a square? Since we have already solved the case $n = 2$, a bisection of the angle between the two solutions $z = 1$ and $z = -1$ will give the two additional solutions for $n = 4$: $z = \mathrm{i}$ ($\alpha = 90°$) and $z = -\mathrm{i}$ ($\alpha = 270°$).

In arithmetic terms, we can express our ability to solve the case $n = 4$ geometrically as follows. The equation whose roots we need is $z^4 - 1$, which factors as follows:

$$z^4 - 1 = (z^2 - 1)(z^2 + 1). \tag{24.15}$$

Thus, instead of an equation of the fourth degree, we only have to deal with second-degree equations which, as we already know, pose no problems to geometrical purists employing only straightedge and compass.

This process of bisection can be repeated arbitrarily often. Thus, if we have a solution for the case n, we "automatically" have solutions for $n \cdot 2^k$ with $k = 1, 2, 3, \ldots$. So far, we have therefore shown how to construct regular polygons with $n = 2, 3, 4, 6, 8, 12, 16, 24, 32$, etc. edges. In some of these cases there are particularly simple solutions, as for $n = 6$, the regular hexagon, whose edge length equals the radius of the circle.

The Case $n = 5$.

The lowest-degree case so far unsolved is that of the regular pentagon. The equation $z^5 - 1$ factors into

$$z^5 - 1 = (z - 1)(z^4 + z^3 + z^2 + z + 1). \tag{24.16}$$

At first blush, this case looks impossible because of the fourth-degree polynomial on the right of (24.16). But this polynomial must be factorizable into polynomials of no higher degree than quadratic, or the Greeks couldn't have constructed the regular pentagon as they did.

Now instead of giving the Greek solution, we will use the case $n = 5$ to present the gist of Gauss's sensational solution of the case $n = 17$ [24.1]. This makes the case $n = 5$ look slightly more difficult than necessary, but we can then see how Gauss went about solving $n = 17$ and the other basic cases still possible, namely when n equals a Fermat prime, i.e., a prime of the form $n = 2^{2^m} + 1$. For $m = 0$ we get the triangle, for $m = 1$ the pentagon and for $m = 2$ the heptadecagon ($n = 17$). The only other *known* basic cases are $m = 3$ ($n = 257$) and $m = 4$ ($n = 65537$).

24.2 Gauss's Great Insight

Gauss noticed that if n is a *Fermat prime*, then $n - 1$ is a power of 2 (in fact, it must be a *power* of a power of 2). Furthermore, Fermat primes, like all primes, have *primitive roots*. And Gauss proceeded to order the roots of $z^n - 1$ according to powers of a primitive root of n, say w. We shall illustrate this, as we said, for $n = 5$ because it is easier to write down than the case $n = 17$ – but the principle is the same.

First, as always, we factor out the universal factor $z - 1$ (giving the solution $z = 1$):

$$z^5 - 1 = (z - 1)(z^4 + z^3 + z^2 + z + 1).$$

Now we are looking for 4 values of z such that

$$z + z^2 + z^3 + z^4 = -1. \tag{24.17}$$

A primitive root of 5 is $w = 2$. (Check:$w^k \equiv 2,4,3,1 \pmod 5$ for $k = 1,2,3,4$. Check!) We now combine exponents in (24.17) according to alternating terms in the series $w^k = 2,4,3,1$, i.e., we write

$$z + z^2 + z^3 + z^4 = r_1 + r_2, \quad \text{where} \tag{24.18}$$

$$r_1 = z^2 + z^3 \quad \text{and} \quad r_2 = z^4 + z^1. \tag{24.19}$$

Can we determine r_1 and r_2? Yes, we can! With (24.17) we have

$$r_1 + r_2 = -1 \tag{24.20}$$

and with (24.19),

$$r_1 \cdot r_2 = z^6 + z^3 + z^7 + z^4, \tag{24.21}$$

which, because of $z^5 = z^0$, can be written

$$r_1 \cdot r_2 = z^1 + z^3 + z^2 + z^4.$$

Here the sum on the right is the sum of all different fifth roots of 1, except 1 itself. Since the sum of *all* roots equals zero, we have

$$r_1 \cdot r_2 = -1. \tag{24.22}$$

Now (24.20) and (24.22) are of the form of the two solutions r_1 and r_2 of a *quadratic* equation in r:

$$r^2 + ar + b = 0, \tag{24.23}$$

where, according to Vieta's rule [24.2],

$$a = -(r_1 + r_2) \quad \text{and} \tag{24.24}$$

$$b = r_1 \cdot r_2 = -1. \tag{24.25}$$

Hence our quadratic equation (24.23) for r is

$$r^2 + r - 1 = 0 \tag{24.26}$$

with the solution

$$r_1 = -\frac{\sqrt{5}+1}{2} \quad \text{and} \quad r_2 = \frac{\sqrt{5}-1}{2}, \tag{24.27}$$

where, incidentally, r_2 equals the reciprocal of the Golden ratio, $g = 1.618...$, and r_1 equals the negative of the Golden ratio (Sect. 5.5 and Fig. 5.4).

With these intermediate solutions it is easy to obtain the final ones. Setting $z^4 = z^{-1}$ in (24.19), we may write

$$z^{-1} + z = r_2 = \frac{1}{g}, \tag{24.28}$$

another *quadratic* equation! Its solutions are

$$z_{1,2} = \frac{1}{2g} \pm \mathrm{i}\sqrt{1 - \frac{1}{4g^2}}, \tag{24.29}$$

where, of course, $|z_{1,2}| = 1$ and the real part is

$$x_{1,2} = \frac{1}{2g} = \frac{\sqrt{5}-1}{4}. \tag{24.30}$$

Thus, two of the roots are obtained by constructions of $\sqrt{5}$ (as the hypotenuse of a right triangle with sides of lengths 1 and 2) and by proceeding according to (24.30). Since all roots of the polynomial $z^4 + z^3 + z^2 + z + 1$ are *primitive* roots, the remaining roots can be found geometrically by taking integral multiples of the angle associated with, say, $z_{1,2}$:

$$\alpha_{1,2} = \arccos \frac{1}{2g} = \pm \frac{2\pi}{5} = \pm 72^\circ. \tag{24.31}$$

This equation tells us, incidentally, that 72° is an angle whose cosine is not transcendental but is a quadratic irrational.

Thus, we have succeeded in factoring $z^4 + z^3 + z^2 + z + 1$ using only quadratic irrationals:

$$z^4 + z^3 + z^2 + z + 1 = (z - z_1)(z - z_2)(z - z_3)(z - z_4) \tag{24.32}$$

with $z_{1,2}$ as given by (24.29) and $z_{3,4}$ given as the solutions of

$$z + z^{-1} = r_1 = -g,$$

namely

$$z_{3,4} = \frac{g}{2} \pm \mathrm{i}\sqrt{1 - \frac{g^2}{4}}, \tag{24.33}$$

with associated angles

$$\alpha_{3,4} = \arcos\left(\frac{-g}{2}\right) = \pm \frac{4\pi}{5} = \pm 144^\circ. \tag{24.34}$$

If our solutions are correct, they must lie on the unit circle, which they do, and their sum must equal -1. And indeed,

$$z_1 + z_2 + z_3 + z_4 = \frac{1}{g} - g = -1. \quad \text{Check!}$$

The case $n = 17$ proceeds quite analogously but requires one extra step of the kind used in (24.19), namely, assigning powers of z to two different sets. *Rademacher* [24.3] wrote a lucid exposition of the case $n = 17$.

The Case n = 6.

The polynomial $z^6 - 1$ can be written as follows:

$$z^6 - 1 = (z^2)^3 - 1, \tag{24.35}$$

which resembles the case for $n = 3$, except that z in $z^3 - 1$ has been replaced by z^2 . Thus, $n = 6$ can be solved geometrically by obtaining the angle for $n = 3$, as already described, and halving it. Further halvings crack the cases $n = 12$, 24, 48, etc.

The Case $n = 7$.

The polynomial $z^7 - 1$ factors over the rationals into

$$z^7 - 1 = (z-1)(z^6 + z^5 + z^4 + z^3 + z^2 + z + 1). \tag{24.36}$$

Since 7 is a prime, the second factor can be shown to be irreducible over the rationals. But 7 is not a *Fermat prime*, and 6 is not a power of 2. Thus, the grouping exemplified by (24.19), which works for $n = 5, 17, 257$ and 65537, leading to nested *quadratic* equations, does not work for $n = 7$, and the regular heptagon can therefore not be constructed by straightedge and compass. This insight by Gauss should have ended a vain search that had lasted over 2000 years, but some people are still trying!

The Case $n = 8$.

The case $n = 8$ follows from the case $n = 4$ by angle halving, as do the cases $n = 16, 32, 64$, etc.

The Case $n = 9$.

While $9 - 1 = 8$ *is* a power of 2, the number 9 unfortunately is composite. The polynomial $z^9 - 1$ can be written as

$$z^9 - 1 = (z^3)^3 - 1, \tag{24.37}$$

which requires solving the case $n = 3$ and then trisecting the resulting angle (120°), which unfortunately is still not possible. Alternatively, we could say that solving $n = 9$ requires taking a third root, which is impossible geometrically. Thus, $n = 9$ also fails, but for a different reason than $n = 7$: $n = 7$ *has* a primitive root but $7 - 1$

is not a power of 2, while $n = 9$ exceeds a power of 2 by 1, as required, but 9 is composite. *Both* conditions must be fulfilled, and they are fulfilled only by the Fermat primes.

The Case $n = 10$.

The case $n = 10$ reduces to the case $n = 5$ and angle halving, which also yields geometrical solutions for $n = 20, 40, 80$, etc.
The Cases $n = 11$ and $n = 13$ are like $n = 7$: impossible.

The Case $n = 15$.

The case $n = 15$ is interesting. The composite $15 = 3 \cdot 5$ is the product of precisely two Fermat primes. As we know, we can construct geometrically the angle for the equilateral triangle (120°) and the regular pentagon (72°), and half the difference angle, $(120° - 72°)/2 = 24°$, is the angle for the regular 15-gon. In like manner, we can solve any n that is the product of *different* Fermat primes, such as 51, 85, 255, etc.

The Case $n = 17$.

The case $n = 17$ is the famous case first solved by *Gauss* [24.1, 3] using a primitive root of 17 and "noting" that $17 - 1 = 16$ is a power of 2. (See the discussion under $n = 5$.)

The Case $n = 257$.

The cases $n = 257$ and 65537 are similar to $n = 5$ and $n = 17$ but require one and two more decomposing steps, respectively, than the case $n = 17$.

The next *new and interesting* case would be given by a prime $n = 2^m + 1$ which requires m to be a power of 2. The next *candidate* is $m = 2^{32}$, i.e., $n = 4294967297$, but unfortunately this is composite. This is the end of the road for geometrical cyclotomy until another Fermat prime is found. (But see Sect. 3.8 for the odds of this being successful.)

After this brief excursion into the realm of cyclotomy, we shall study, in a somewhat more formal manner, some general properties of cyclotomic polynomials that give rise to a host of enticing applications (fast algorithms, periodic sequences with unique and very useful spectral properties, etc.).

24.3 Factoring in Different Fields

In numerous and diverse applications it will be important to factor the polynomial

$$P_N(z) := z^N - 1. \tag{24.38}$$

In fact, the applications we shall discuss here are so diverse that we seek factors whose coefficients are from three different number fields:

1) The complex number field $\mathbb{C}$
2) The rational number field $\mathbb{Q}$
3) Finite (Galois) number fields GF(p^m)

In the following we will continue our discussion of cyclotomy and related factoring on a somewhat more formal level, and in view of the importance of the subject, we will even permit ourselves a bit of redundancy.

24.4 Cyclotomy in the Complex Plane

Factorization of $z^N - 1$ is easy if we admit complex coefficients. Setting

$$P_N(z) = z^N - 1 = 0, \tag{24.39}$$

we see immediately that the zeros of $P_N(z)$ are the N distinct Nth roots of unity

$$W_N^r, \quad r = 0, 1, \ldots, N-1, \tag{24.40}$$

where W_N is a *primitive* Nth root of unity, for example:

$$W_n = \mathrm{e}^{2\pi\mathrm{i}/N}. \tag{24.41}$$

Insisting on a primitive root here insures that W_N^r will indeed run through *all* N distinct Nth roots as r goes from 0 to $N-1$. For example, for $N = 4$,

$$W_4 = \mathrm{e}^{2\pi\mathrm{i}/4} = \mathrm{i} \tag{24.42}$$

is a primitive 4th root, and so is

$$W_4^3 = \mathrm{e}^{6\pi\mathrm{i}/4} = -\mathrm{i}, \tag{24.43}$$

because each will generate the remaining two roots of $z^4 - 1 = 0$, namely $z = 1$ and $z = -1$.

All the primitive Nth roots are given by

$$W_N^r, \quad \text{where } (r, N) = 1. \tag{24.44}$$

Thus, we see that there are $\phi(N)$ primitive Nth roots.

Since if $(r,N) = 1$, so is $(N-r,N)$, and because r and $N-r$ are distinct for $N > 2$, the primitive roots come in "conjugate" pairs, which are seen to be conjugate *complex*:

$$W_N^r \quad \text{and} \quad W_N^{N-r} = W_N^{-r} = W_N^{r*}, \tag{24.45}$$

where $*$ stands for conjugate complex in this book. (Physicists and engineers need the horizontal bar for averages.) It is interesting to note that the term "conjugate complex" comes from the more general algebraic concept of "conjugate" and not vice versa, as one might be led to think if one is familiar only with *complex* conjugates.

Thus, our factorization of $P_N(z)$ is

$$z^N - 1 = \prod_{n=0}^{N-1} (z - W_N^n), \tag{24.46}$$

which is as far as one can go, because the individual factors are all linear in z.

Replacing the primitive root W_N above by another primitive root, W_N^r, with $(r,N) = 1$, changes only the order of the factors. Thus, the above factorization is essentially unique.

24.5 How to Divide a Circle with Compass and Straightedge

How to divide a circle into N equal parts by compass and straightedge, i.e., by purely geometric means, was one of the classical problems of geometry first brought into prominence by the ancient Greeks. They discovered that a circle could be so divided if $N = 3$ or 5. Since angles can also be halved geometrically, additional possibilities are $N = 6, 12, 24, \ldots$ and $N = 5, 10, 20, \ldots$. Finally, it is not difficult to see that by dividing a circle both into 3 *and* 5 equal parts, the angle $2\pi/3 \cdot 5$ can also be obtained, thereby permitting division of the circle into 15 equal parts.

In general, according to Euclid, a circle can be divided into N equal parts by compass and straightedge if

$$N = 2^k \cdot 3^m \cdot 5^n, \tag{24.47}$$

where k is any nonnegative integer and n and m are either 0 or 1.

Since dividing a circle into N equal parts is equivalent to constructing a regular N-gon, i.e., a polygon with N vertices (or edges), the lowest-order regular N-gon that could not be constructed was the 7-gon or *heptagon*.

Since the time of Euclid, for almost 2000 years, mathematicians and amateurs alike had been trying to smash the boundary at $N = 7$, but in vain. Then, on March 30, 1796, an 18-year-old Brunswick (Germany) youth scribbled in his brand-new notebook, at the top of page 1, (see Fig. 5.1) freely translated from the Latin that he was using: "How to divide the circle by geometric means into 17 equal parts." In other words Gauss (the name of the young man) had just discovered that the numbers 3 and 5 of the ancient Greeks had to be supplemented by 17 and, in general, by primes of the form

$$2^{2^n} + 1, \tag{24.48}$$

i.e., the Fermat primes F_n of which then, and to this day, only 5 are known: $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65537$.

Of equal significance, Gauss proved that the *only* regular N-gons that can be constructed by geometric means are of the form

$$N = 2^k \prod_n F_n, \tag{24.49}$$

where the product is over *distinct* Fermat primes. Thus, apart from the factor 2^k, there are at present $2^5 - 1 = 31$ different regular odd N-gons that can be constructed geometrically: from the triangle and the pentagon to the 17-gon and the $3 \cdot 5 \cdot 17 \cdot 257 \cdot 65537 = 4\,294\,967\,295$-gon. (The author strongly advises against attempting the latter case. Gauss said it is possible and that should suffice. On the other hand, there is a *suitcase* at the Mathematics Institute of the University of Göttingen which is jam-packed with the details of constructing the regular 65537-gon.)

How is Gauss's great discovery related to our factorization? One way to state it is to say that the primitive Nth root W_N involves only *square* roots, and square roots can be constructed with a compass and a straightedge (following old Pythagoras and his right triangle).

Example: $N = 5$:

$$W_5 = \mathrm{e}^{2\pi \mathrm{i}/5} = \cos(2\pi/5) + \mathrm{i}\sin(2\pi/5), \quad \text{where} \tag{24.50}$$

$$\cos(2\pi/5) = (\sqrt{5} - 1)/4,$$

$$\sin^2(2\pi/5) = (\sqrt{5} + 5)/8. \tag{24.51}$$

24.5.1 Rational Factors of $z^N - 1$

One of our main aims is to factor

$$P_N(z) = z^N - 1 \tag{24.52}$$

into polynomials with rational coefficients. This is trivial for

$$P_1(z) = z - 1 \tag{24.53}$$

(which leaves nothing to factor) and easy for $P_2(Z)$:

$$P_2(z) = (z - 1)(z + 1). \tag{24.54}$$

Furthermore, remembering geometric series, we write

$$P_3(z) = (z - 1)(z^2 + z + 1). \tag{24.55}$$

But now we have reached the end of Easy Street – unless, that is, we recall Möbius and the second inversion formula involving his function (Chap. 21).

To be able to apply Möbius, we factor $P_N(z)$ as follows (perhaps as a long shot, not knowing what else to do):

$$z^N - 1 = \prod_{n|N} C_n(z), \tag{24.56}$$

where according to the inversion formula (20.19), the factors $C_n(z)$ are given by

$$C_n(z) = \prod_{d|n} (z^d - 1)^{\mu(n/d)}. \tag{24.57}$$

Since $\mu = 0, \pm 1$, these factors of $z^N - 1$ are indeed rational. Also, they are not further reducible over the rationals. The degree of $C_n(z)$ is

$$\deg[C_n(z)] \sum_{d|n} d\,\mu\left(\frac{n}{d}\right) = \phi(n). \tag{24.58}$$

Hence the degree of the product is

$$\deg\left[\prod_{n|N} C_n(z)\right] = \sum_{n|N} \phi(n) = N. \quad \text{Check!} \tag{24.59}$$

Example: $N = 6$. With (24.56):

$$P_6(z) = z^6 - 1 = C_1(z) \cdot C_2(z) \cdot C_3(z) \cdot C_6(z)$$

and with (24.57):

$$\begin{aligned} C_1(z) &= z - 1 \\ C_2(z) &= \frac{z^2 - 1}{z - 1} = z + 1 \\ C_3(z) &= \frac{z^3 - 1}{z - 1} = z^2 + z + 1 \\ C_6(z) &= \frac{(z-1)(z^6-1)}{(z^2-1)(z^3-1)} = z^2 - z + 1. \end{aligned}$$

Check: $z^6 - 1 = (z-1)(z+1)(z^2+z+1)(z^2-z+1)$. Check! Here the first factor, $z - 1$, "captures" the common root of all $P_n(z) : z_1 = 1$. The second factor, $C_2(z) = z + 1$, captures the remaining root of $z^2 - 1 = 0$, namely the primitive root $z_2 = \mathrm{e}^{2\pi\mathrm{i}/2} = -1$. The third factor, $z^2 + z + 1$, captures the two primitive roots of $z^3 - 1$, namely $z_3 = \mathrm{e}^{2\pi\mathrm{i}/3}$ and $z_4 = \mathrm{e}^{4\pi\mathrm{i}/3}$. Finally, the fourth factor, $z^2 - z + 1$, captures the two primitive roots of $z^6 - 1$, namely $z_5 = \mathrm{e}^{2\pi\mathrm{i}/6}$ and $z_6 = \mathrm{e}^{10\pi\mathrm{i}/6}$. The angles in the complex plane subtended between the real axis and these roots are, in the above order, 0°, 180°, 120°, 240°, 60° and 300°, i.e., all distinct multiples of 60° – as expected.

24.6 An Alternative Rational Factorization

Another way of representing the factors of $C_n(z)$ is as follows. We first replace d by n/d, which only changes the order of the factors:

$$C_n(z) = \prod_{d|n} (z^{n/d} - 1)^{\mu(d)}. \tag{24.60}$$

Now it is apparent that we need only those divisors of n which are squarefree, because otherwise $\mu = 0$. These squarefree divisors can be put into classes depending on the number of distinct prime factors. We introduce the product

$$\prod\nolimits_k := \prod_{d_k} (z^{n/d_k} - 1). \tag{24.61}$$

where the d_k are all the squarefree divisors that have exactly k distinct prime factors. Thus, for example,

$$\begin{aligned} \prod\nolimits_0 &= z^n - 1, \\ \prod\nolimits_1 &= \prod_{p_i} (z^{n/p_i} - 1), \\ \prod\nolimits_2 &= \prod_{p_i \neq p_j} (z^{n/p_i p_j} - 1), \quad \text{etc.,} \end{aligned} \tag{24.62}$$

where the p_i are the prime factors of n.

With this notation, and because by definition $\mu(d_k) = (-1)^k$, we get the factorization

$$C_n(z) = \frac{\prod_0 \prod_2 \prod_4 \cdots}{\prod_1 \prod_3 \prod_5 \cdots}, \tag{24.63}$$

which has several interesting theoretical applications. Specifically, it can be shown that $C_n(z)$ has only *integral* coefficients. Furthermore, for $n < 105$, the only coefficients that appear in $C_n(z)$ other than 0 are ± 1.[1]

24.7 Relation Between Rational Factors and Complex Roots

We have found two seemingly independent factorizations of $z^N - 1$ so far, one involving the complex Nth roots of 1 but with factors linear in z, and the other involving higher powers of z but with rational (in fact, integer) coefficients. How are

[1] $C_{105}(z)$ is the lowest-order cyclotomic polynomial in which other coefficients (-2 in two places) appear. The reason is that 105 is the smallest integer that is the product of three distinct odd primes [24.4].

these two factorizations related? Can we express the cyclotomic polynomials $C_n(z)$ in terms of the nth roots of 1, W_n? Yes, we can:

$$C_n(z) = \prod_{\substack{0<r<n \\ (r,n)=1}} (z - W_n^r). \tag{24.64}$$

Here we recall that W_n, is the “first” nth root of 1:

$$W_n = \mathrm{e}^{2\pi \mathrm{i}/n}, \tag{24.65}$$

which is a *primitive* root, and the roots appearing in the above factorization are *all* the primitive nth roots of 1. There are precisely $\phi(n)$ such roots, so that the degree of $C_n(z)$ becomes $\phi(n)$. Check!

But that does not suffice to establish the correctness of the factorization. Instead of proving the above factorization of $C_n(z)$ in terms of the $(z - W_n^r)$, we will first illuminate it for a special case, $n = 6$ (see also the above example). The Möbius inversion gives us

$$C_6(z) = \frac{(z-1)(z^6-1)}{(z^2-1)(z^3-1)}. \tag{24.66}$$

Here the factor $z^3 - 1$, for example, can be written

$$z^3 - 1 = (z-1)(z-W_3^1)(z-W_3^2). \tag{24.67}$$

However, we shall express W_3 by W_6:

$$W_3 = W_6^2. \tag{24.68}$$

Thus,

$$z^3 - 1 = (z-1)(z-W_6^2)(z-W_6^4), \tag{24.69}$$

or, more generally, we write

$$z^d - 1 = (z-1) \prod_{(k,n)=\frac{n}{d}} (z - W_n^k). \tag{24.70}$$

In the special case of $C_6(z)$, this factorization is as follows:[2]

$$C_6(z) = \frac{(z-W_6^1)(z-W_6^2)\dots(z-W_6^5)}{(z-W_6^3)(z-W_6^2)(z-W_6^4)}. \tag{24.71}$$

[2] Leaving out the four factors $(z-1)$ which always cancel, except for $C_1(z) = z - 1$, because, as we derived earlier,

$$\sum_{d|n} \mu(d) = 0 \quad \text{for } n > 1.$$

Here, in the numerator the roots W_6^k, $k = 1,2,3,4,5$ appear, i.e., *all* the 6th roots of 1 except 1 itself. In the denominator all those W_6^k appear, where $(k,6) > 1$. Thus, $C_6(z)$ is the product of precisely all factors W_6^r with $0 < r < 6$ and $(r,6) = 1$:

$$C_6(z) = \prod_{\substack{0<r<n \\ (r,6)=1}} (z - W_6^r). \tag{24.72}$$

i.e., $C_6(z)$ is the product of linear factors whose roots are precisely all the *primitive* roots of $z^6 - 1$.

The demonstration would have been essentially the same for other squarefree orders. For non-squarefree orders, some of the factors are missing. We leave it to the reader to generalize this result to arbitrary orders n.

24.8 How to Calculate with Cyclotomic Polynomials

For prime order $n = p$, one has

$$C_p(z) = \prod_{r=1}^{p-1} (z - W_p^r) = \frac{z^p - 1}{z - 1}, \tag{24.73}$$

or,

$$\text{RULE I}: \qquad Cp(z) = z^{p-1} + z^{p-2} + \cdots + z + 1. \tag{24.74}$$

Without proof,

$$\text{RULE II}: \qquad C_{mp^k}(z) = C_{mp}\left(z^{p^{k-1}}\right). \tag{24.75}$$

For $(m,p) = 1$,

$$\text{RULE III}: \qquad C_{mp} = \frac{C_m(z^p)}{C_m(z)}. \tag{24.76}$$

For odd $n > 2$,

$$\text{RULE IV}: \qquad C_{2n}(z) = C_n(-z), \quad \text{for odd } n \geq 3. \tag{24.77}$$

For $z = 1$, one has

$$\text{RULE V}: \qquad C_n(1) = \begin{cases} 0 & \text{for } n = 1 \\ p & \text{for } n = p^k \\ 1 & \text{otherwise.} \end{cases} \tag{24.78}$$

These rules expedite the calculation of cyclotomic polynomials:

$$\begin{aligned}
&C_1(z) = z-1 \\
&C_2(z) = z+1 && \text{(by RULE I)} \\
&C_3(z) = z^2+z+1 && \text{(by RULE I)} \\
&C_4(z) = C_2(z^2) = z^2+1 && \text{(by RULE II)} \\
&C_6(z) = C_2(z^3)/C_2(z) = z^2-z+1 && \text{(by RULE III)}
\end{aligned}$$

$$\begin{aligned}
&\text{Check :} \quad C_6(z) = C_3(-z). \quad \text{Check!} && \text{(by RULE IV)} \\
&\text{Check :} \quad C_6(1) = 1. \quad \text{Check!} && \text{(by RULE V)} \\
&\text{Check :} \quad C_4(1) = 2. \quad \text{Check!} && \text{(by RULE V)}
\end{aligned}$$

And, incidentally,

$$C_{17}(z) = z^{16} + z^{15} + \cdots + z + 1, \tag{24.79}$$

so that

$$z^{17} - 1 = (z-1)C_{17}(z). \tag{24.80}$$

Now, while $C_{17}(z)$ cannot be further factored into polynomials with rational coefficients, Gauss observed that $C_{17}(z)$ can be decomposed into a set of nested *quadratic* equations, each solvable by “geometric means”, leading to the constructibility of the regular 17-gon with compass and straightedge (Sect. 24.2).

Chapter 25
Linear Systems and Polynomials

One of the main applications of polynomial theory occurs in the analysis of linear electrical circuits and the many other physical situations that are customarily represented by linear-circuit analogs. With the advent of computers and digital signal processing, *time-discrete* systems have taken on a special significance, and these are effectively represented by polynomials called z transforms that are akin to *generating functions* in other branches of mathematics. The application of cyclotomic polynomials, in particular, leads to fast computational algorithms, excellent error-correcting codes, and special signals for precision measurement (Chap. 28).

25.1 Impulse Responses

If the capacitor C in Fig. 25.1 is charged up to voltage y_0 at time $t = 0$ and thereafter allowed to discharge via the resistor R, its voltage as a function of time will decay exponentially:

$$y(t) = y_0 \mathrm{e}^{-t/\tau} \quad \text{for } t > 0. \tag{25.1}$$

Here the "time constant" (in seconds) is $\tau = RC$, where R is the resistance (in ohms) and C is the capacitance (in farads). (The symbols R and C do double duty here.)

If the charging impulse was of very short duration and the capacitor was "empty" for negative times, then

$$y(t) = 0 \quad \text{for } t < 0. \tag{25.2}$$

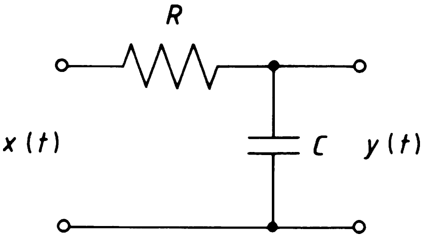


Fig. 25.1 A simple linear passive system

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_25, © Springer-Verlag Berlin Heidelberg 2009

If the charging voltage is a Dirac delta function $\delta(t)$, an idealized pulse which is 0 everywhere except for $t=0$ and represents a unit charge, then the output of the linear system is called its *impulse response* $h(t)$. For the circuit in Fig. 25.1,

$$h(t) = \begin{cases} 0 & \text{for } t<0 \\ \frac{1}{C}\mathrm{e}^{-t/\tau} & \text{for } t>0. \end{cases} \tag{25.3}$$

The term "linear" in linear system means that its parameters (here the values R and C) do not change with the relevant variables (here electrical current, voltage and charge). Roughly: the resistor does not heat up and the capacitor does not break down.

If an arbitrary voltage as a function of time, $x(t)$, is applied to the input of the linear system of Fig. 25.1, the output voltage is a *convolution integral* [25.1]:

$$y(t) = \int x(t')h(t-t')\mathrm{d}t'. \tag{25.4}$$

This expression follows directly from the assumed linearity of the system (and its time invariance and passivity; i. e., its parameters do not change with time and do not *generate* energy).

25.2 Time-Discrete Systems and the z Transform

If we want to simulate linear systems on digital computers, we have to "discretize", i. e., represent by rational numbers, all constants and variables, including time. The impulse response of a time-discrete system is a sequence h_n, which is its output if at "time" $n=0$ a unit impulse was applied to the quiescent system.

The impulse response h_n is often represented by its z transform (Chap. 22):

$$H(z) := \sum_{n=-\infty}^{\infty} h_n z^{-n}, \tag{25.5}$$

which (except for the sign in the exponent of z) is nothing but the familiar *generating function* for the sequence h_n. If we set $z=\mathrm{e}^{\mathrm{i}\omega T}$, then $H(z)$ becomes the Fourier transform, where ω is the radian frequency and T is the "sampling interval", i. e., the time interval corresponding to a difference of 1 in the "time" index n.

25.3 Discrete Convolution

If we apply an input sequence x_n to a linear system with impulse response h_n, the output sequence (because of linear superposition) becomes

$$y_n = \sum_k x_k h_{n-k}. \tag{25.6}$$

Such a *discrete convolution* is often abbreviated by the convolution star:

$$y_n =: x_n * h_n, \tag{25.7}$$

a convenient notation.

But if we express the above sequences y_n, x_n and h_n by their z transforms, then the (somewhat messy) discrete convolution becomes a simple product:

$$Y(z) = X(z) \cdot H(z), \tag{25.8}$$

which is verified by comparing equal powers of z on the two sides of the equation.

25.4 Cyclotomic Polynomials and z Transform

For many applications it is either appropriate or convenient to consider the data and the impulse responses to be periodic or *circular* [25.2]:

$$x_n = x_{n+N}, \tag{25.9}$$

where N is the (fundamental) period. More generally, we may write

$$\left.\begin{aligned} x_n &= x_m \\ h_n &= h_m \end{aligned}\right\} \quad \text{for} \quad n \equiv m \pmod{N}, \tag{25.10}$$

and consequently also

$$y_n = y_m \quad \text{for } n \equiv m \pmod{N}.$$

Circular convolution looks like this:

$$y_n = \sum_{k=0}^{N-1} x_k h_{\langle n-k \rangle_N}, \tag{25.11}$$

where we have used the acute brackets $\langle\rangle_N$ to indicate a least nonnegative residue modulo N.

For circular data, the z transform contains the factor $z^N - 1$, and the possibility of factorizing $z^N - 1$ leads to a reduction in the number of multiplications. We will not pursue the details here, but simply cite a theorem by Winograd which states that the number of multiplications can be reduced by as much as m, where m is the number of irreducible factors of $z^N - 1$, i. e., $m = N$ if we factor over the complex number field, and $m = d(N)$ if we factor over the rationals [25.3].

Chapter 26
Polynomial Theory

In this chapter we further develop our acquaintance with polynomials, especially those with integer coefficients and discover that in many ways, their arithmetic is akin to that of integers: there are coprime polynomials, divisor polynomials, and Diophantine equations in polynomials, and there is even a version of Euclid's algorithm and a Chinese Remainder Theorem for polynomials.

The most important application of integer polynomials is in the construction of finite number fields, also called Galois fields (Chap. 27), which play a dominant role in today's digital world.

26.1 Some Basic Facts of Polynomial Life

Many of the things we have learned about numbers (residue systems, primitive elements, etc.) can be generalized to polynomials – with virtually unbounded applications. Since we will subsequently need at least the elements of *polynomial theory*, we will familiarize ourselves with this fascinating extension now. A polynomial $p(z)$ of *degree n over a field F* looks like this [26.1]:

$$p(z) = \sum_{k=0}^{n} a_k z^k, \quad n \geq 0, \quad a_n \neq 0, \tag{26.1}$$

where the coefficients a_k are elements of the number field F, for example the complex number field $\mathbb{C}$, the rational number field $\mathbb{Q}$, or some *finite* number field such as GF(2), the Galois field of order 2 consisting typically of the elements 0 and 1.

If the leading term $a_n = 1$, the polynomial is said to be *monic*. The degree of $p(z)$ is denoted by $\deg[p(z)]$. With $a_n \neq 0$, $\deg[p(z)] = n$.

A polynomial $d(z)$ *divides* another polynomial $p(z)$ if there exists a polynomial (over the same field!) such that $p(z) = q(z)d(z)$ holds. The polynomial $d(z)$ is then called a divisor of $p(z)$, and we write $d(z)|p(z)$. A polynomial whose only divisors are of degree 0 or $\deg[p(z)]$ is called *irreducible* (over the chosen field) [26.2].

For example, $z^2 + 1$ *is* reducible over $\mathbb{C}$: $z^2 + 1 = (z + i)(z - i)$ but irreducible over $\mathbb{Q}$. Interestingly, $z^2 + 1$ *is* reducible over the finite field GF(2): $z^2 + 1 = (z + 1)^2$,

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_26, © Springer-Verlag Berlin Heidelberg 2009

because $1+1=0$ in GF(2). Thus, possible factorizations depend on the field chosen, and moreover they do so in a seemingly haphazard manner.

Just as integers can be uniquely factored into primes, so polynomials $p(z)$ can be uniquely factored into irreducible polynomials $p_i(z)$:

$$p(z) = K \prod_{i=1}^{s} [p_i(z)]^{r_i}, \tag{26.2}$$

where K is a constant that allows us to restrict the $p_i(z)$ to *monic* polynomials. The degree of $p(z)$ is given by

$$\deg[p(z)] = \sum_{i=1}^{s} r_i \deg[p_i(z)]. \tag{26.3}$$

The "ultimate" factorization is, of course, into polynomials $p_i(z)$ of the first degree – if that is possible in the chosen field. As a result of the Fundamental Theorem of Algebra[1] the number of such first-degree factors, including multiplicity, is precisely equal to $\deg[p(z)]$. Polynomials that have no common factors are called mutually prime or *coprime*.

26.2 Polynomial Residues

It is always possible to write, for two polynomials $p(z)$ and $d(z)$:

$$p(z) = q(z)d(z) + r(z), \tag{26.4}$$

and this representation is *unique*, if

$$\deg[r(z)] < \deg[d(z)]. \tag{26.5}$$

Here, for obvious reasons, $r(z)$ is called the *remainder polynomial*.

In a well-wearing fashion of notation we will also express the above relationship by [26.3]:

$$r(z) = \langle p(z) \rangle_{d(z)}\,. \tag{26.6}$$

These latter notations, of course, recommend themselves if we care not about the *quotient polynomial* $q(z)$ but only about the remainder. (It is peculiar that in this field of human endeavor – as in few others – *remainders* should be the main pickings.)

The operation of obtaining $r(z)$ from $p(z)$ and $d(z)$ is called *polynomial residue reduction*. Two polynomials $p_1(z)$ and $p_2(z)$ are said to be congruent modulo $d(z)$ if they leave the same remainder $r(z)$ upon residue reduction. Equivalently, they are

[1] First complete proof by Gauss in his Ph.D. thesis at Helmstedt, Germany.

congruent modulo $d(z)$ if their difference is divisible by $d(z)$: We write

$$p_1(z) \equiv p_2(z) \pmod{d(z)}, \tag{26.7}$$

or equivalently,

$$d(z)|(p_1(z) - p_2(z)). \tag{26.8}$$

Example:

$$\left\langle z^2 \right\rangle_{z+1} = 1, \tag{26.9}$$

or

$$z^2 \equiv 1 \pmod{(z+1)}, \tag{26.10}$$

because

$$(z+1)|(z^2-1), \tag{26.11}$$

namely:

$$\frac{z^2-1}{z+1} = z-1. \tag{26.12}$$

More generally,

$$\langle p(z) \rangle_{z-a} = p(a), \tag{26.13}$$

and even more generally, for monic divisor polynomials

$$d(z) = z^n + \sum_{k=0}^{n-1} d_k z^k, \tag{26.14}$$

we get:

$$\langle z^n \rangle_{d(z)} = -\sum_{k=0}^{n-1} d_k z^k, \tag{26.15}$$

i. e., the highest power can always be reduced to the negative of the “tail” of the divisor polynomial.

Example:

$$\left\langle x^4 + x^6 \right\rangle_{x^4+x+1} = \langle x^4 \rangle_{x^4+x+1} (1+x^2) = -(x+1)(1+x^2).$$

Of course, if $d(z)$ is not monic, $d(z)/d_n$ will be, and we can proceed as above.

26.3 Chinese Remainders for Polynomials

Whenever it comes to *simultaneous* congruences (and a few other delicacies) we turn to the Chinese, and sure enough, there is a Chinese Remainder Theorem for polynomials also. To wit:

There exists a unique polynomial satisfying

$$y(z) \equiv y_i(z) \pmod{P_i(z)}, \quad i = 1, 2, \ldots, m, \tag{26.16}$$

such that

$$0 \le \deg[y(z)] < \sum_{i=1}^{m} \deg[P_i(z)], \tag{26.17}$$

provided the monic polynomials $P_i(z)$ are pairwise coprime.

For a constructive proof, let us first consider a simpler problem: find m "inverse" polynomials $R_i(z)$ such that each satisfies the following m congruences:

$$R_i(z) \equiv \delta_{ij} \pmod{P_j(z)}, \quad j = 1, 2, \ldots, m. \tag{26.18}$$

With these, the solution is obviously

$$y(z) = \sum_{i=1}^{m} R_i(z) y_i(z) \pmod{P(z)}, \tag{26.19}$$

where

$$P(z) = \prod_{i=1}^{m} P_i(z). \tag{26.20}$$

Now, the inverse polynomials must be of the form

$$R_i(z) = S_i(z) P(z)/P_i(z), \tag{26.21}$$

where $S_i(z)$ is obtained from the *single* congruence

$$S_i(z) P(z)/P_i(z) \equiv 1 \pmod{P_i(z)}. \tag{26.22}$$

This will automatically satisfy the m congruences for $R_i(z)$.

The above congruence can be written as

$$S_i(z) P(z)/P_i(z) + T_i(z) P_i(z) = 1, \tag{26.23}$$

which is the generalization to polynomials of the Diophantine equation that we studied earlier and solved by Euclid's algorithm.

26.4 Euclid's Algorithm for Polynomials

To solve the above Diophantine equation, following Euclid, we expand $P_i^2(z)/P(z)$ into a continued fraction:

$$\frac{P_i^2(z)}{P(z)} = C_0(z) + \cfrac{1}{C_1(z) + \cfrac{1}{C_2(z) + \cdots + \cfrac{1}{C_k(z)}}}, \tag{26.24}$$

where the $C_r(z)$ are determined successively by long division. Truncating the continued fraction yields the polynomial convergents $A_r(z)$ and $B_r(z)$, which obey the recursions

$$A_r(z) = C_r(z)A_{r-1}(z) + A_{r-2}(z) \tag{26.25}$$

and

$$B_r(z) = C_r(z)B_{r-1}(z) + B_{r-2}(z). \tag{26.26}$$

The solution is then given by

$$S_i(z) = (-1)^k K A_{k-1}(z) \tag{26.27}$$

and

$$T_i(z) = (-1)^k K B_{k-1}(z), \tag{26.28}$$

where K is the leading coefficient of $A_k(z)$ and $B_k(z)$.

In addition, the Euclidean algorithm gives the greatest common divisor of the two polynomials: $B_{k-1}(z)$.

Example:

$$R(z) \equiv \begin{cases} 1 & (\mathrm{mod}(z^2+z+1)) \\ 0 & (\mathrm{mod}(z-1)) \end{cases} \tag{26.29}$$

i. e., with (26.23) (and dropping the index $i = 1$),

$$S(z)(z-1) + T(z)(z^2+z+1) = 1. \tag{26.30}$$

We thus need the continued-fraction expansion

$$\frac{z^2+z+1}{z-1} = z + 2 + \cfrac{1}{\frac{1}{3}(z-1)}, \tag{26.31}$$

yielding the approximants

$$\frac{A_0(z)}{B_0(z)} = z+2 \quad \text{and} \tag{26.32}$$

$$\frac{A_1(z)}{B_1(z)} = \frac{\frac{1}{3}(z^2+z+1)}{\frac{1}{3}(z-1)}, \tag{26.33}$$

and consequently, since $k = 1$,

$$S(z) = -\frac{1}{3}(z+2). \tag{26.34}$$

Thus, the solution is

$$R(z) = S(z)(z-1) = -\frac{1}{3}(z^2+z+2). \tag{26.35}$$

Check:

$$\langle R(z)\rangle_{z-1} = -\frac{1}{3}(1^2+1-2) = 0. \quad \text{Check!}$$

$$\langle R(z)\rangle_{z^2+z+1} = -\frac{1}{3}(-1-2) = 1. \quad \text{Check!}$$

The greatest common divisor of z^2+z+1 and $z-1$ equals $B_{k-1}(z) = B_0(z) = 1$. Check!

These rules and algorithms are about all we need to operate successfully with polynomials.

Part IX
Galois Fields and More Applications

Chapter 27
Galois Fields

In Galois fields, full of flowers, Primitive elements dance for hours ...[1]

A residue system modulo a prime p forms a finite number field of order p. For many applications, we need number fields of order p^m. Here, with the knowledge acquired in Chaps. 25 and 26, we learn how to construct and represent them, and how to calculate in them.

27.1 Prime Order

We have already encountered finite number fields of prime order p. We now designate them by $\mathrm{GF}(p)$, where GF stands for *Galois field* [27.1]. They consist, for example, of the elements $0,1,2,\ldots,p-1$, for which addition, subtraction, multiplication and division (except by 0) are defined, obeying the usual commutative, distributive and associative laws. Thus, in $\mathrm{GF}(3)$, for example, consisting of the elements 0, 1 and 2, we have, by way of illustration, $1+2=0$, $1-2=2$, $2\cdot 2=1$, $1/2=2$, etc.

27.2 Prime Power Order

We shall now construct finite number fields of order equal to a prime power p^m, designated $\mathrm{GF}(p^m)$. These Galois fields have virtually unlimited application in such diverse fields as physics (diffraction, precision measurements), communications (error-correcting codes, cryptography) and artistic design (necklaces, etc.).

All realizations of $\mathrm{GF}(p^m)$ are isomorphic. We will choose as field elements either m-tuples ("vectors"), $m\times m$ matrices, or polynomials of degree $m-1$, all with components and coefficients, respectively, from $\mathrm{GF}(p)$.

[1] S. B. Weinstein, found in [27.1].

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_27, © Springer-Verlag Berlin Heidelberg 2009

Since $\mathrm{GF}^*(p^m)$, i. e., $\mathrm{GF}(p^m)$ without the 0 element, forms a cyclic group with multiplication as the group operation, we can also represent $\mathrm{GF}^*(p^m)$, which has order $p^m - 1$, by a primitive element α and its $p^m - 1$ distinct powers:

$$\alpha, \alpha^2, \alpha^3, \ldots, \alpha^{p^m-1} = 1. \tag{27.1}$$

We shall illustrate this for $p = 2$, i. e., $\mathrm{GF}(2^m)$, which has order 2^m. We begin with the representation by m-tuples with components from $\mathrm{GF}(2)$, i. e., 0 or 1. There are exactly 2^m such m-tuples in agreement with the order 2^m. For $m = 4$, $\mathrm{GF}(2^4)$ then has the following 16 elements:

$$\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ & & \vdots & \\ 1 & 1 & 1 & 1 \end{array} \tag{27.2}$$

The 4-tuples are written down here in binary sequence from 0 to 15 (with the least significant digit on the left, for a change).

Addition and subtraction are defined column-wise modulo 2. Thus, for example, $1100 + 1111 = 0011$ (which does *not* correspond to binary addition, because there are no “carries”).

Multiplication is defined as follows. First, the m-tuples are represented by polynomials of degree up to $m-1$. There are exactly p^m such polynomials, i. e., as many as the order of $\mathrm{GF}(p^m)$. Each polynomial corresponds to one element of the field. For example, the element 1100 corresponds to

$$x^0 + x^1 + 0 + 0 = 1 + x, \tag{27.3}$$

and the element 0101 corresponds to

$$0 + x^1 + 0 + x^3 = x + x^3. \tag{27.4}$$

Now, multiplication in $\mathrm{GF}(p^m)$ is defined as multiplication of polynomials *modulo a given irreducible polynomial* $\pi(x)$ over $\mathrm{GF}(p)$ of degree m. For $p = 2$ and $m = 4$, there are precisely three such polynomials: $1 + x + x^4$, (its “reciprocal”) $1 + x^3 + x^4$ and $1 + x + x^2 + x^3 + x^4$. We choose

$$\pi(x) = 1 + x + x^4, \tag{27.5}$$

as our modulus and define the product of two elements of $\mathrm{GF}(2^4)$, represented by two polynomials $g(x)$ and $h(x)$, as

$$g(x) \cdot h(x) = \langle g(x) \cdot h(x) \rangle_{\pi(x)}, \tag{27.6}$$

where the acute brackets indicate a least remainder, i. e. a polynomial of degree smaller than the degree of $\pi(x)$.

Thus, for example,

$$1101\cdot 1001 \overset{\wedge}{=} (1+x+x^3)(1+x^3) = 1+x+x^4+x^6. \tag{27.7}$$

As we saw before, calculating the remainder modulo $\pi(x)$ is the same as setting $\pi(x)=0$, or replacing x^4 by $1+x$. Thus, we may write (remember the coefficients are from GF(2), i. e., $1+1=0$, etc.):

$$1+x+x^4+x^6 = 1+x+1+x+x^2(1+x) = x^2+x^3. \tag{27.8}$$

Thus, we get

$$1101\cdot 1001 = 0011. \tag{27.9}$$

For the multiplication to have an inverse in $\mathrm{GF}^*(p^m)$, $\pi(x)$ must be irreducible – just as we require that for the multiplication of the integers $1,2,\ldots,n-1$ to have an inverse modulo n, n must be “irreducible”, i. e., prime.

To show that $1+x+x^4$ is, in fact, irreducible, we have to show that it is not divisible by any polynomial over GF(2) up to degree $4/2=2$.

Obviously, $1+x+x^4$ is not divisible by x or $1+x$. It is also not divisible by x^2 or $1+x^2$ [because $(1+x^2)^2 = 1+x^4$], which leaves only $1+x+x^2$ as a potential divisor. By long division we get

$$\begin{array}{l}
\;x^4+x\;+\;1 : x^2+x+1 = x^2+x\\
x^4+x^3\;+\;x^2\\
\qquad x^3\;+\;x^2+x+1\\
\qquad x^3\;+\;x^2+x\\
\qquad\qquad\quad \text{remainder}:\;\; 1\neq 0.
\end{array} \tag{27.10}$$

Thus, $1+x+x^4$ is indeed irreducible.

Another polynomial of degree 4 over GF(2) is the *reciprocal* defined by

$$\widehat{\pi}(x) := x^{\deg[\pi(x)]}\pi(x^{-1}), \tag{27.11}$$

or, for $\pi(x) = 1+x+x^4$:

$$\widehat{\pi}(x) = 1+x^3+x^4. \tag{27.12}$$

In general, in a reciprocal polynomial, the exponents are “flipped”, i. e., the exponent k becomes the exponent $\deg[\pi(x)]-k$.

27.3 Generation of GF(2^4)

We shall now try to generate the finite number field GF(2^4) of order 16 by making use of the irreducible polynomial $\pi(x) = 1+x+x^4$ and the primitive element

$$\alpha = 0100 \overset{\wedge}{=} x. \tag{27.13}$$

We start with the 0 element $(\alpha^{-\infty})$ and the 1 element (α^0) and proceed by multiplying with α, i. e., multiplying by the polynomial x (corresponding to a right shift in the m-tuple) and residue reduction modulo $1+x+x^4$, which corresponds to adding 1's (modulo 2) to the two left-most places of the m-tuple when a 1 "disappears" on the right:

α power	polynomial	$4-$tuple
$-\infty$	0	0000
0	1	1000
1	x	0100
2	x^2	0010
3	x^3	0001
4	$1+x$	1100
5	$x+x^2$	0110
6	x^2+x^3	0011
7	$1+x+x^3$	1101
8	$1+x^2$	1010
9	$x+x^3$	0101
10	$1+x+x^2$	1110
11	$x+x^2+x^3$	0111
12	$1+x+x^2+x^3$	1111
13	$1+x^2+x^3$	1011
14	$1+x^3$	1001
15	1	1000

(27.14)

Thus, we see that $\alpha^{15} = \alpha^0 = 1$ and $\alpha^n \neq 1$ for $0 < n < 15$, as is expected of a primitive element. And we also see that the process has generated all $15+1 = 16$ binary-valued 4-tuples and all polynomials over GF(2) of maximal degree 3.

It is interesting to note that the above listing also contains the representation of GF(2^4) by 4×4 matrices with elements 0 or 1. Consider the matrix consisting of the third to sixth 4-tuples:

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}. \tag{27.15}$$

M is a primitive element of this representation. The matrix representation is redundant, but its advantage is that the other nonzero elements are generated by ordinary matrix multiplication! For example,

$$M^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \tag{27.16}$$

which corresponds to the fourth to seventh 4-tuple in the listing above. In general, M^{n+1} is obtained from M^n by deleting the upper row of M^n and appending the next row from the (cyclical) list (27.14) to the bottom of M^n. In fact, M^k is the 4×4 matrix whose top row is α^k.

Note that we have a set of matrices here whose multiplication is commutative.

27.4 How Many Primitive Elements?

In generating GF(2^4), we used α as the primitive element, generating all 15 nonzero elements by α^k, $k = 1,2,\ldots,15$. How many such distinct primitive elements are there? Obviously $\phi(15) = 8$, because with α, α^n is also a primitive element *iff* $(n,15) = 1$. Thus, α^2, α^4, α^7, α^8, α^{11}, α^{13} and α^{14} are the other primitive elements.

In general, the order of an element α^n of GF(p^m) equals $(p^m - 1)/(n, p^m - 1)$. Thus, for GF(2^4), α^3, α^6, α^9 and α^{12} have order 5, while α^5 and α^{10} have order 3. The number of elements having order T, where T must be a divisor of $p^m - 1$, is $\phi(T)$, as we saw in Chap. 13. As always, there is one element of order $T = 1$, namely $\alpha^0 = 1$.

27.5 Recursive Relations

Still another way of looking at GF(p^m) is to focus on the first column of the m-tuple representation. Skipping the 0 element, we obtain a binary-valued periodic sequence $\{a_k\}$, one period of which for GF(2^4) looks like this:

$$\{a_k\} = 1\,0\,0\,0\,1\,0\,0\,1\,1\,0\,1\,0\,1\,1\,1. \tag{27.17}$$

It is easy to see that this sequence is generated by the recursion

$$a_{k+4} = a_{k+1} + a_k \tag{27.18}$$

with the initial condition $a_1 = 1$, $a_2 = a_3 = a_4 = 0$.

The recursion is a direct consequence of having generated the m-tuples by means of polynomial residue reduction. For our example of GF(2^4) we had chosen for the residue reduction the polynomial modulus $\pi(x) = x^4 + x + 1$ which, as we saw, corresponds to setting

$$x^4 = x + 1 = x^1 + x^0. \tag{27.19}$$

Now, multiplication by x engenders a right shift in the m-tuple. This polynomial equation therefore corresponds directly to the recursion

$$a_{k+4} = a_{k+1} + a_{k+0},$$

as already stated.

The above recursion has a "memory" of four binary digits. Once four consecutive a_k are specified, the sequence is uniquely determined by the recursion. One way to implement a physical generator for the above sequence is by a *finite-state* machine, a so-called shift register as shown in Fig. 27.1.

At the time of each clock pulse, the content of a register (0 or 1) is transferred to its next neighbor to the right. The recursion is implemented by the linear feedback connection to the input of the first register, also shown in Fig. 27.1.

Since this finite-state machine has $m = 4$ registers which can hold either a 0 or a 1, there are $2^4 = 16$ possible states. However, the state 0 0 0 0 never generates anything but more 0's and is therefore of no interest to us here. Thus, there remain $p^m - 1 = 2^4 - 1 = 15$ distinct nonzero states and the above sequence *must* repeat after at *most* 15 steps – which it does. In this connection it is instructive to check that the consecutive 4-tuples of $\{a_k\}$ cover all 15 possibilities except the all-zero 4-tuple 0 0 0 0.

It can be shown that such *maximum-length* sequences exist for all periods of length $p^m - 1$ [27.1]. What is needed is an irreducible polynomial of degree m, of which there are, in fact, see also [27.2]:

$$\frac{1}{m}\sum_{d|m}\mu\left(\frac{m}{d}\right)p^d.$$

For $p = 2$ and $m = 4$, there are three irreducible polynomials, as already noted (Sect. 27.2).

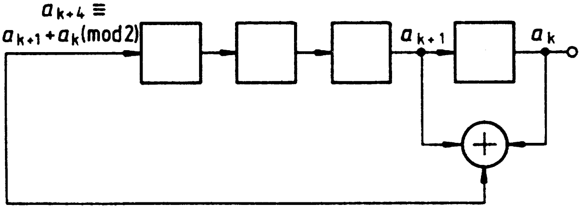


Fig. 27.1 Four-stage linear shift register generating maximum-length Galois sequence corresponding to the finite field GF(2^4). The contents of each of the four registers are transferred to the next neighbor to the right upon application of a clock pulse (not shown). The sequence of pulses repeats after $2^4 - 1 = 15$ clock pulses

If the sequence is to be generated by a shift register, as illustrated in Fig. 27.1, then, in addition, the root x of the irreducible polynomial must be a primitive element. Polynomials with this property are called *primitive polynomials*. There are exactly

$$\frac{\phi(p^m - 1)}{m} \tag{27.20}$$

such polynomials, each leading to a distinct maximum-length sequence. For $p^m > 4$, these sequences come in pairs, generated by pairs of reciprocal polynomials, that are mirror images of each other (reflected in "time").

For $p = 2$ and $m = 4$, (27.20) tell us there are two primitive polynomials. Hence one of the three irreducible polynomials of degree 4 cannot be primitive. Indeed, the polynomial $1+x+x^2+x^3+x^4$, while irreducible, is not primitive because it has no roots that are primitive elements. For example, the root x generates only 5 different elements, i. e. it only has order 5. To wit: $x^5 = x\,x^4 = x+x^2+x^3+x^4 = 1$. The corresponding shift register connection or recursion

$$a_{n+4} = a_{n+3} + a_{n+2} + a_{n+1} + a_n \tag{27.21}$$

produces sequences with period 5. In fact, starting with 1 0 0 0, we get

1 0 0 0 1 1 0 0 0 1 1 0 0 0 1 etc.

Irreducible polynomials that are not primitive are also "distinguished" by the fact that they are factors of $1+x^n$ with $n < p^m - 1$. Thus, for $p = 2$ and $m = 4$, $1+x+x^2+x^3+x^4$ is indeed a factor not only of $1+x^{15}$ but also of $1+x^5$:

$$\left(1+x+x^2+x^3+x^4\right)(1+x) = 1+x^5.$$

This fact can be used to show, in a different way, why the corresponding recursion has only period 5. First one shows (easily) that the recursion produces a sequence that corresponds to the *reciprocal* of the corresponding polynomial. This, in turn, can be written as

$$\left(1+x+x^2+x^3+x^4\right)^{-1} = \frac{(1+x)}{1+x^5} = (1+x)\left(1+x^5+x^{10}+\dots\right),$$

which has obviously period 5. As a binary sequence, the right side corresponds to 1 1 0 0 0 1 1 0 0 0, etc. which agrees, within a shift, with the above sequence found directly by the recursion (27.21).

By the same token, a *primitive* polynomial is a factor of $1+x^n$ with $n = 2^m - 1$ *and no smaller n*. Thus, its reciprocal repeats after $2^m - 1$ terms and not before!

Interestingly, sequences generated by linear-recursion that do not have maximum length depend on the initial condition (even when lateral shifts are discounted). For example, the initial condition 1 1 1 1, with the above recursion (27.21), produces a different sequence:

1 1 1 1 0 1 1 1 1 0 1 1 1 0 etc.

By contrast, different initial conditions (excepting 0 0 … 0), for a recursion based on a *primitive* polynomial, always give the same sequence (starting of course with the initial condition). This must be so because a *maximum-length* sequence runs through all m-tuples (except 0 0 … 0), and each possible initial condition therefore occurs in the sequence; there is no variety to be obtained by different choices of initial conditions other than a shift.

27.6 How to Calculate in GF($\mathrm{p^m}$)

For adding and subtracting elements of GF(p^m), the m-tuple representation is recommended. The m-tuples are added column-wise modulo p. For multiplying, dividing, and exponentiating, the α-power representation is preferable.

Examples from GF(2^4):

$$\begin{aligned}&0111 \cdot 1111 \mathrel{\hat{=}} \alpha^{11}\alpha^{12} = \alpha^{23} = \alpha^{8} \mathrel{\hat{=}} 1010.\\ &(1010)^{-1} \mathrel{\hat{=}} (\alpha^{8})^{-1} = \alpha^{7} \mathrel{\hat{=}} 1101.\\ &(0110)^{1/2} \mathrel{\hat{=}} (\alpha^{5})^{1/2} = (\alpha^{20})^{1/2} = \alpha^{10} \mathrel{\hat{=}} 1110.\end{aligned}$$

The Zech logarithm $Z(n)$, defined by

$$\alpha^{Z(n)} := 1 + \alpha^{n}, \tag{27.22}$$

permits *both* adding and multiplying using the α-power notation without antilogging!

Example:

$$(\alpha^9 + \alpha^{10})^2 = \alpha^3 + \alpha^5 = \alpha^3(1 + \alpha^2),$$

and, from a table of $Z(n)$, $Z(2) = 8$. Thus, $1 + \alpha^2 = \alpha^8$ and

$$(\alpha^9 + \alpha^{10})^2 = \alpha^{11}.$$

Check: $(\alpha^9 + \alpha^{10})^2 \mathrel{\hat{=}} (0101 + 1110)^2 = (1011)^2 \mathrel{\hat{=}} (\alpha^{13})^2 = \alpha^{11}$. Check! But the latter (checking) operation takes three table-lookups (instead of one) and requires going from α powers to m-tuples and back again to α powers. So *Zech* does something quite sobering for us.

27.7 Zech Logarithm, Doppler Radar and Optimum Ambiguity Functions

The Zech logarithm has a noteworthy property that leads directly to several interesting applications in some rather unlikely fields of human endeavor. Imagine, if you can, the year 1943 – not a prime year. Radar had been developed to a fine tool for locating U-boats and air planes. But then a new trick was first tried that foiled conventional radars: metal foil strips, called chaff, tuned to the searching radar's wavelength. When they were released by the target, they so overwhelmed the radar scope with phoney reflections that the real target was all but invisible and often escaped unharmed.

Of course, for every weapon there is a counter weapon, and for the chaff (code-named *Düppel-Streifen* in German) it was the Doppler effect, named after the Austrian physicist Johann Christian Doppler (1803–1853) who discovered it. The Doppler effect says that a wave emitted or reflected from a moving target suffers a change in its wavelength or frequency. (Interestingly, but perhaps not surprisingly, considering that Doppler was born in Salzburg, the city of Mozart, his effect was first verified in 1845 by using *trumpeters* on a moving train.)

To separate dangerous targets from harmless chaff, engineers (electrical) soon hit upon the idea to measure the frequency shift of the radar echo; only the frequency-shifted echoes were harvested for further action. But early Doppler radars, as they were called, using frequency-modulated ("chirp") pulses had a disturbing trait: they confused range (target distance) with range rate (speed of approach). The same echo could signal either stationary chaff at 10 km, say, or a fast approaching plane at 9 km – a potentially fatal ambiguity. To lasso this electronic double entendre into a neat mathematical form, theoretically inclined engineers defined an *ambiguity function*. After that, everyone in the Doppler arena began hunting for the *ideal* ambiguity function, dubbed thumb-tack function: a function of range and range rate that vanishes everywhere except at one sharp point, thereby defining both distance and speed unambiguously.

What does an ideally unambiguous radar pulse look like? First, radar pulses are periodic, with period T, so that target range and speed can be periodically updated, say every $T = 10$ milliseconds, corresponding to a maximal range of 15 km. Radars must also have a certain range accuracy $\Delta r = 2c \cdot \Delta t$, where c is the speed of light and Δt the corresponding timing accuracy (typically a fraction of a microsecond). Without imposing any severe constraints on the design of the radar, we shall assume an integer relationship between T and Δt:

$$T = (p^m - 1)\,\Delta t, \tag{27.23}$$

where p is a prime and m is an integer.

To achieve a timing accuracy Δt, the frequency bandwidth of the radar, according to the uncertainty principle, must exceed $1/\Delta t$. If all the frequencies contained in a sufficiently wide frequency band were transmitted simultaneously in one time slot Δt, the peak power limitation of the radar transmitter would be ill used. Instead we assume that during each time slot, the nth one beginning at

$$t_n = n\Delta t, \quad n = 1, \ldots, p^m - 2, \tag{27.24}$$

only *one* discrete frequency is transmitted; every Δt seconds the radar frequency hops to a new value and then remains constant at that value for a time interval Δt. Calling these discrete frequencies

$$f(n) = f_0 + z(n)\,\Delta f, \tag{27.25}$$

the remaining problem in our radar design is to determine in *which order* they are to be radiated. We want to make sure that for each delay difference k any frequency difference occurs at most once. Mathematically speaking, we are looking for a permutation $z(n)$ of the integers from $n = 1$ to $n = p^m - 2$ such that

$$z(n+k) - z(n) \equiv z(n'+k) - z(n') \tag{27.26}$$

for $k \not\equiv 0$ and *all* $n \not\equiv n' \bmod p^m - 1$.

Such "perfect" permutations are not so easy to generate. For example, the "random" permutation $z(n) = 3,1,4,5,2,6$ of the integers from 1 to 6 does not have property (27.26): for $k = 2$, $n = 1$ and $n' = 4$, we have $4 - 3 = 6 - 5$. In fact, perfect permutations, which obey (27.26), form a small minority among all possible permutations.

This is where our Zech logarithm comes in: more than 30 years after the original Doppler radar excitement it was suddenly discovered that primitive roots and Zech logarithms can generate the required permutations with property (27.26) and thus provide a new answer to the ambiguity problem [27.3].

Let us define the Zech logarithm $z(n)$ by

$$\alpha^{z(n)} := 1 - \alpha^n, \tag{27.27}$$

where α is a primitive element of the Galois field $\mathrm{GF}(p^m)$. (Note that, for $p = 2$, $z(n) = Z(n)$ as defined in (27.22). The minus sign in (27.27) was chosen for subsequent notational convenience.)

To illustrate the use of the Zech logarithm, let us construct a permutation of the integers from 1 to 6 using the Galois field $\mathrm{GF}(2^3)$ with the primitive element $\alpha = x \overset{\wedge}{=} 0\;1\;0$ and the modulus polynomial $x^3 = 1 + x$:

$$\begin{array}{ll}
\alpha^{-\infty} & 0\;0\;0 \\
\alpha^{0} & 1\;0\;0 \\
\alpha^{1} & 0\;1\;0 \\
\alpha^{2} & 0\;0\;1 \\
\alpha^{3} & 1\;1\;0 \\
\alpha^{4} & 0\;1\;1 \\
\alpha^{5} & 1\;1\;1 \\
\alpha^{6} & 1\;0\;1
\end{array} \tag{27.28}$$

Now with (27.27) we have, for $n = 1,2,\ldots 6$, $z(n) = 3,6,1,5,4,2$, which is indeed a perfect permutation having property (27.26). This is easily checked, noting that both $z(n)$ and n are taken modulo $p^m - 1 = 7$.

A general (indirect) proof that this is so proceeds as follows. Suppose that (27.26) is violated for some $k \not\equiv 0$ and some pair $n \not\equiv n' \bmod p^m - 1$. Then, with (27.27),

$$\frac{1-\alpha^{n+k}}{1-\alpha^{n}} \equiv \frac{1-\alpha^{n'+k}}{1-\alpha^{n'}} \mod p^m - 1, \quad \text{or}$$
$$\alpha^n(1-\alpha^k) \equiv \alpha^{n'}(1-\alpha^k) \mod p^m - 1.$$

Since $k \not\equiv 0$, and therefore $1-\alpha^k \not\equiv 0$, it follows that

$$\alpha^n \equiv \alpha^{n'} \mod p^m - 1,$$

which, for primitive α, is true if and only if $n \equiv n' \mod p^m - 1$, contradicting our assumption.

To construct a radar pulse of period T and time resolution Δt according to this principle, a prime power p^m near $1 + T/\Delta t$ is first selected. A primitive element α in GF(p^m) is then used to generate the $p^m - 2$ successive frequencies $f(n)$ in the radar pulse according to (27.25) where $z(n)$ is the Zech logarithm of n defined by (27.27). To give the pulse the proper period $T = (p^m - 1)\Delta t$, one empty interval, during which no energy is radiated, is inserted into each period at $n \equiv 0 \mod p^m - 1$.

In addition to the permutation $z(n)$, generated directly by the Zech logarithm, the product $rz(n)$, where r is coprime to $p^m - 1$, yields more permutations with property (27.26). There are thus a total of $\phi(p^m - 1)$ "Zech-like" permutations, anyone of which leads to the desired ambiguity function and can therefore be used in the Doppler radar frequency hopping scheme (27.25). However, while the original permutation $z(n)$ is symmetric, in that $z(z(n)) = n$, this is no longer true for $r \not\equiv 1 \mod p^m - 1$. This absence of symmetry may be an advantage in certain applications where a potential adversary might exploit symmetry for his advantage. The additional degree of freedom offered by different choices of the parameter r could also be helpful in foiling attempts at jamming.

For $m = 1$, another method of generating perfect permutations of the integers $1, 2, \ldots, p-1$ is by means of successive powers of a primitive root g of the prime p: $g^1, g^2, \ldots, g^{p-1}$. This method is, however, limited to primes p, while the Zech logarithm is applicable to all prime *powers* p^m, a denser set. (How dense?)

Quadratic residues modulo an odd prime can also generate integers with the property (27.26). However, for the periodic sequence $n^2 \mod p$, half the integers in $1, 2, \ldots, p-1$ do not occur and the other half occurs twice, violating our requirement of fully utilizing the available frequency space.

27.8 A Unique Phase-Array Based on the Zech Logarithm

Consider the periodic sequence with period $p^m - 1$:

$$a_0 = 0 \tag{27.29}$$
$$a_n = \exp\left[\frac{2\pi \mathrm{i} z(n)}{p^m - 1}\right], \quad \text{for} \quad n = 1, 2, \ldots, p^m - 2,$$

where $z(n)$ is a Zech logarithm in $\mathrm{GF}(p^m)$ defined according to (27.27). All a_n, $n \not\equiv 0 \bmod p^m - 1$, have magnitude 1. The periodic correlation sequence c_n (cf. Sect. 14.8) is given by

$$\begin{aligned} c_0 &= p^m - 2 \\ c_n &= -1 - \exp\left[\frac{-2\pi \mathrm{i} n}{p^m - 1}\right]. \end{aligned} \tag{27.30}$$

This follows from the definition of c_n (14.40), and the (easy to prove) fact that, for $n \not\equiv 0$, the difference $z(k+n) - z(k)$ attains all values from 1 to $p^m - 2$ exactly once, except the value n. The corresponding power spectrum is

$$\begin{aligned} |A_0|^2 &= |A_1|^2 = 1, \\ |A_k|^2 &= p^m, \quad k = 2, 3, \ldots, p^m - 2. \end{aligned} \tag{27.31}$$

Applied to antenna arrays, operated at a wavelength λ, this result means that a periodic array with element spacing $\lambda/2$ in which all individual elements are driven with equal amplitudes and with phase angles according to (27.29), will radiate equal energies into $p^m - 3$ distinct directions (corresponding to the index values $k = 2$ to $k = p^m - 2$ in (27.31)). The "broadside" direction ($k = 0$) and the first "side lobe" ($k = 1$) will receive only $1/p^m$ of the energy going into the other directions.

If the phase angles $2\pi z(n)/(p^m - 1)$ in (27.29) are multiplied by an integer factor r the "undernourished" lobe (other than the broadside) will be the rth lobe. This follows from (27.34), see Sect. 27.9. Figure 27.2 shows the radiation pattern of such an array with element spacing of $9/16$ wavelengths, based on $p = 3$, $m = 2$ and $r = 2$. (Note that for this application r does not have to be coprime to $p^m - 1$, so that $rz(n)$ is no longer a perfect permutation obeying (27.26).)

27.9 Spread-Spectrum Communication and Zech Logarithms

Another interesting application of the Zech logarithms is to spread-spectrum communication (see Sect. 16.9). Consider the periodic sequences $a_n^{(r)}$, $r = 1, 2, \ldots p^m - 2$, with period $p^m - 1$,

$$\begin{aligned} a_0^{(r)} &= 0 \\ a_n^{(r)} &= \exp\left[\frac{2\pi \mathrm{i} r z(n)}{p^m - 1}\right] \quad \text{for } n = 1, 2, \ldots, p^m - 2. \end{aligned} \tag{27.32}$$

The autocorrelation sequences are similar to (27.30):

$$\begin{aligned} c_0 &= p^m - 2 \\ c_n &= -1 - \exp\left[\frac{-2\pi \mathrm{i} r n}{p^m - 1}\right]. \end{aligned} \tag{27.33}$$

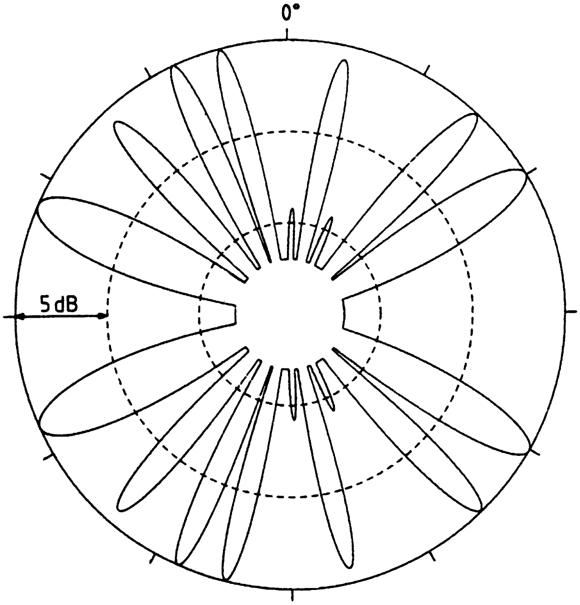


Fig. 27.2 Radiation pattern of constant-amplitude phase array with phase angles based on Zech logarithm for GF(3^2). Array has two periods of period length 8 (total number of nonzero array elements: 14). Spacing between individual antenna elements: $9/16$ wavelengths. Note small lobes in the broadside (0°) and second order (26°, corresponding to $r = 2$) directions. The attenuation of the small lobes corresponds closely to the theoretical value $1/p^m \mathrel{\widehat{=}} -9.5$ dB for infinite arrays. By choosing other values for the parameter r any of the other nonbroadside lobes can be made small. (Prepared by M. Rollwage.)

The corresponding power spectra are flat except for two components:

$$\begin{aligned}&|A_0|^2 = |A_r|^2 = 1\\ &|A_l|^2 = p^m \quad \text{for} \quad k \not\equiv 0 \quad \text{and} \quad k \not\equiv r.\end{aligned} \tag{27.34}$$

Thus, these sequences too, like the original one ($r = 1$), have constant power in time (except for one "silent" time interval per period) and in frequency (except for two spectral components). In addition, and most importantly, the cross-correlation coefficient c_{rs} between two such sequences

$$c_{rs} := \sum_{n=0}^{p^m-2} a_n^{(r)} a_n^{(s)*} \tag{27.35}$$

equals -1 for $r \not\equiv s$. For $p^m \gg 1$, the cross-correlation is thus small compared to the power of these sequences.

These three properties (nearly constant power in time and frequency and small cross-correlation) make these sequences ideal as "carrier waves" or signature sequences in the design of spread-spectrum communication systems where all channels occupy the same frequency band and yet have small mutual interference.

Chapter 28
Spectral Properties of Galois Sequences

Certain periodic sequences with elements from the Galois field GF(p), formed with the help of primitive polynomials over GF(p^m), have unique and much sought-after correlation and Fourier transform properties. These *Galois sequences*, as I shall call them, have found ingenious applications in error-correcting codes, interplanetary satellite picture transmission, precision measurements in physiology and general relativity, and even concert hall acoustics. Other applications are in radar and sonar camouflage, and in noise abatement, because Galois sequences permit the design of surfaces that scatter incoming waves very broadly, thereby making reflected energy "invisible" or "inaudible". A similar application occurs in work with coherent light, where a "roughening" of wavefronts (phase randomization) is often desired (for example, to avoid "speckles" in holograms). Excellent structures for this purpose are light diffusers whose design is based on Galois arrays – in a sense the ultimate in frosted (milk) glass. Finally, Galois sequences allow the design of loudspeaker and antenna arrays with very broad radiation characteristics.

28.1 Circular Correlation

Some of the most interesting properties of Galois sequences $\{a_k\}$ for $p = 2$ emerge when we consider their (circular) correlation. But first we shall convert the a_k into b_k by replacing 0's by 1's and 1's by -1's according to the rule

$$b_k = (-1)^{a_k}, \tag{28.1}$$

or, equivalently,

$$b_k = 1 - 2a_k. \tag{28.2}$$

The circular correlation of the sequence b_k is defined, as before, by

$$c_n = \sum_{k=0}^{2^m-2} b_k b_{k+n}, \tag{28.3}$$

where the indices are reduced modulo the period of the b_k, i. e., $2^m - 1$.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_28, © Springer-Verlag Berlin Heidelberg 2009

For $n \not\equiv 0 \pmod{2^m - 1}$, we get

$$c_0 = 2^m - 1. \tag{28.4}$$

For $n \not\equiv 0$, we shall consider the special case $\mathrm{GF}(2^4)$. The recursion for the b_k is multiplicative:

$$b_{k+4} = b_{k+1} b_k,$$

or

$$b_k = b_{k-3} b_{k-4}, \tag{28.5}$$

which is isomorphic with the corresponding recursion for the a_k (27.18). Now consider the terms in the above correlation sum

$$d_k := b_k b_{k+n}. \tag{28.6}$$

It is easy to see that the product sequence d_k has the same recursion as the b_k:

$$d_k = b_k b_{k+n} = b_{k-3} b_{k-4} b_{k+n-3} b_{k+n-4}. \tag{28.7}$$

By definition, the product of the first and third factors on the right equals d_{k-3}, and the product of the remaining two factors equals d_{k-4}. Thus,

$$d_k = d_{k-3} d_{k-4}, \tag{28.8}$$

as claimed. Since the recursion together with the initial condition determines the sequence uniquely, except for a shift in the index, we have (for $n \not\equiv 0$)

$$d_k = b_{k+s} \tag{28.9}$$

for some shift s.

Using this result in the above correlation, we find

$$c_n = \sum_{k=0}^{2^m-2} b_{k+s} = \sum_{k=0}^{2^m-2} b_k = -1, \tag{28.10}$$

since exactly $2^m/2$ of b_k equal -1 and the remaining $(2^m - 2)/2$ equal $+1$. This is so because in the m-tuple representation of $\mathrm{GF}(2^m)$ all possible 2^m m-tuples appear exactly once. However, the Galois sequence $\{a_k\}$ was constructed without the 0000-tuple, so that one period of $\{a_k\}$ has one extra 1, i.e., in the sequence $\{b_k\}$ there is one extra -1. Hence, the sum over one period of the b_k equals -1.

Most importantly, this result is true for any $n \not\equiv 0$. Thus, our correlation function is two-valued:

$$c_n = \begin{cases} 2^m - 1 & \text{for } n \equiv 0 \bmod (2^m - 1) \\ -1 & \text{otherwise} \end{cases}, \tag{28.11}$$

meaning that the spectrum of the b_k is flat [28.1]:

$$|B_k|^2 = \begin{cases} 2^m & \text{for } k \not\equiv 0 \bmod (2^m - 1) \\ 1 & \text{otherwise} \end{cases}. \tag{28.12}$$

Thus, we have discovered another class of binary-valued periodic sequences with a flat spectrum. But in contrast to the Legendre sequences (Chap. 16), which are not easily generated, our Galois sequences derived from GF(2^m) are generated by a simple *linear* recursion as illustrated in Fig. 27.1. The linearity of the recursion has other important consequences, for example in applications to error-correcting codes (Sect. 28.2).

The periods that are achievable for binary-valued sequences have length $2^m - 1$ and require for implementation an m-stage shift register and a feedback connection based on a primitive polynomial $\pi(x)$ over GF(2) of degree m.

Examples:

$$\begin{aligned} m &= 2 : \pi(x) = 1 + x + x^2 \\ m &= 3 : \pi(x) = 1 + x + x^3 \\ m &= 4 : \pi(x) = 1 + x + x^4 \\ m &= 5 : \pi(x) = 1 + x^2 + x^5 \\ m &= 6 : \pi(x) = 1 + x^2 + x^6 \end{aligned}$$

Note: $1 + x + x^5$ is not irreducible over GF(2); it factors into $(1 + x^2 + x^3)(1 + x + x^2)$.

Primitive polynomials over GF(2) up to high degrees m have been published [28.2]. The period length of the Galois sequence with $m = 168$ equals $2^{168} - 1$, a 51-digit number. With one clock pulse per picosecond, the sequence will repeat after 10^{22} times the age of the universe – long enough for most purposes.

28.2 Application to Error-Correcting Codes and Speech Recognition

Galois sequences are useful for correcting multiple errors in digital representations, for example in transmission and storage [28.3]. In the simplest case, the m-tuples corresponding to the initial condition are determined by the information to be transmitted, while the recursion generates $2^m - 1 - m$ *check bits*. Thus, putting our example of GF(2^4) to work, with the initial condition 1000, the recursion appends 11 check bits for a 15-bit codeword:

$$1\,0\,0\,0,\ \ 1\,0\,0\,1\,1\,0\,1\,0\,1\,1\,1.$$

The efficiency of this code is, of course, only $4/15$, or, to put it differently: of the 2^{15} possible binary-valued codewords of length 15, the code uses only 2^4. But as a reward for this frugality, the code can correct up to three errors.

To see this, we note that the *Hamming distance* between any two distinct codewords is $d = 8$. (The Hamming distance is the number of places in which two codewords differ. Since for the code at hand, the difference (or sum) modulo 2 of two codewords is again a codeword, the above result for d follows immediately.) Now, for an even Hamming distance d, the number of errors that can be corrected is $d/2 - 1$ (by choosing the nearest possible codeword). The occurrence of one additional error can be *detected* but it can no longer be corrected, because there are then two possible codewords equidistant from the erroneous one. This so-called Simplex Code based on $\mathrm{GF}(2^4)$ can correct three errors and detect a fourth one [28.4].

The general result is that a code based on $\mathrm{GF}(2^m)$ in the manner described has efficiency $m/(2^m - 1)$ and corrects $2^{m-2} - 1$ errors per codeword of length $2^m - 1$. Thus, these codes can cope with error rates around 25%! Not bad at all. The fact that the efficiency is low is of little concern in some applications, such as image transmission from interplanetary space vehicles, which take months or years to reach their destinations and have long periods of time available to transmit their pictures of, say, Jupiter's moon Io. But these space communication links need all the error-correcting capabilities they can get, because the typical space probe's transmitter has low power and has to be "heard" over vast distances. In fact, the first good pictures of the "canals" of Mars were obtained with an error-correcting code very similar to the one sketched here. And who wants to confuse a digital error with a new moon of Neptune?

Another advantage of the code described here is that the error correction is quite simple: the received code, converted to the ± 1 representation, is cross-correlated with the above codeword $(1\ 0\ 0 \ldots 0\ 1\ 1\ 1)$, which is also converted to ± 1. If there was no error in transmission, all values of the (circular) cross-correlation will equal -1, except one, which has value 15 and which by its position will indicate which of the 15 allowed codewords was sent.

If there was one error in transmission, the peak value of the correlation will be reduced from 15 to 13 and the other correlation values will be increased from -1 to $+1$. A second error will give correlation values 11 and up to 3, respectively. Three errors will result in correlations 9 and up to 5, respectively. Thus, the peak value (9) is still distinct from the other values (≤ 5), and proper detection is assured.

Four errors will produce correlation values of 7 in the proper location and also 7 in some off-locations, so that correction is no longer possible. But the absence of correlation values larger than 7 is an indication that there are four or more errors.

Error-correcting codes are not only useful for cleansing digital data; they are applicable also to fault finding in other fields. Let us look at (rather than listen to) speech synthesis from text ("talking computers") and *automatic speech recognition* [28.5]. These are the acoustic "bridges" between man and machine.

Speech recognition is a particularly difficult problem for the computer, even for small vocabularies enunciated clearly by "his master's voice". But the *talkwriter* (voice-operated typewriter) is waiting in the wings to replace the irreplaceable human secretary. And one of the advances that will speed the entry of the bloodless transcriber will be the proper exploitation of syntax and semantics.

To illustrate, consider a vocabulary of only 2 words and sentences with a length of 15 words. Without constraints, this strange "language" allows the combination of 2^{15} different sentences. Now suppose that good grammar and meaningful statements limit the number of *useful* sentences to $2048 = 2^{11}$. Then 4 words in each sentence are redundant "check words", allowing $2^4 = 16$ different cases to be distinguished, indicating either no error or which of the 15 words in the sentence was recognized incorrectly.

This is a single-error-correcting Hamming Code [28.4] of length 15, using four check bits. It is the *dual* of the Simplex Code discussed above, which has also length 15 but $15 - 4 = 11$ check bits. The challenge in the application of finite-field theory to speech recognition is to translate linguistic constraints into algebraic structures that permit fast codebook searches. For artificial languages, such as company telephone numbers or stock item labels, one can often start the language design from scratch, matching it to known good codes.

28.3 Application to Precision Measurements

As we saw in the chapter on error-correcting codes, the Galois sequence $\{b_k\}$ with period $2^m - 1$, derived from an irreducible polynomial $\pi(x)$ of degree m over GF(2), has the following amazing properties:

1) Its energy, b_k^2, is distributed perfectly uniformly in "time" (time corresponding to the index k): each "time slot" $k = 0, 1, \ldots, 2^m - 2$ has exactly the same energy, namely 1.
2) The enery of the sequence is also uniformly distributed over all $2^m - 2$ nonzero frequencies.

Property (1) means that a device limited in peak power, such as a transmitter, can output a maximum amount of energy in a given time interval. This makes it possible to measure at extremely low signal-to-noise ratios.

Of course, a sinewave also has its energy stretched out rather uniformly in time, but it contains only a *single* frequency and therefore has poor time resolution. By contrast, because of Property (2), a Galois sequence contains many frequency components, all having the same power. This means that measurements made with Galois sequences as a test signal – such as measurements of interplanetary distances to check out space-time curvature – will have very high time resolution. In fact, the time resolution will be the same as that obtained with a short pulse having the same broad frequency spectrum. The high temporal precision inherent in Galois sequences is brought out by correlation receivers that cross-correlate the received signal (a radar echo from the planet Mercury, for example, or a faint echo in a concert hall) with the outgoing sequence to form sharp peaks at the delays corresponding to pulse travel times.

Thus, these two properties – the full, optimal use of both Fourier domains (time and frequency) – make the Galois sequences, also known as maximum-length sequences or pseudorandom shift-register sequences, ideal candidates for precision measurements with extremely low energies in the presence of strong interfering noises.

28.4 Concert Hall Measurements

Most measurements of concert halls are done in *empty* halls: no live audience wants to listen to the pops, bangs, hisses and howls produced by acousticians to measure the acoustic characteristics of a hall under scrutiny. But measurements made in an empty hall are often irrelevant, because a sizable audience has a noticeable effect on the acoustics of an enclosure. Can one make measurements during an ongoing concert without interfering with the musical enjoyment of the audience?

The answer to this question is, most surprisingly, yes. A Galois sequence can be radiated (from a loudspeaker) into a concert hall (or any other "noisy" environment) at such extremely low sound levels as to be *inaudible* by the audience (owing to an auditory phenomenon called *masking*). The required detection process is cross-correlation – as in the case of the error-correcting code described above – between the received acoustic signal (at some audience location) and the radiated Galois signal (from the stage, for example). By integrating the correlation over not just one fundamental period, but over many periods, namely the duration of the concert (one hour, say), the required signal-to-noise ratios can be realized [28.6].

For concert hall measurements, the fundamental period of the Galois sequence is chosen to equal approximately the reverberation time, say 2 seconds. The clock frequency chosen must be higher than twice the upper audio cutoff frequency, or about 32 kHz. The sequence must therefore have a period length of at least 2 seconds times $32\,\text{kHz} = 64000$. The smallest $2^m - 1$ that is large enough is obtained for $m = 16$, leading to a period of $2^{16} - 1 = 65535$, which is, in fact, a preferred choice.

The author demonstrates this method of measuring acoustic responses in one of his lecture courses at the University of Göttingen by measuring the impulse response between a point near the lectern and some point in the back of the auditorium *while continuing his lecture*. As long as what he says is sufficiently incoherent (with the radiated Galois sequence), his speech signal averages out in the correlation process, and a "noise-free" response builds up on a TV screen watched by the students. (For didactic reasons, the Galois sequence is played at a soft but *audible* level.)

Galois sequences have also been used in physiological measurement of neuronal systems [28.7].

28.5 The Fourth Effect of General Relativity

General Relativity Theory, the theory of gravitation propounded by Einstein in 1915, passed three important experiment tests during Einstein's lifetime. The following effects predicted by him were confirmed:

1) The motion of the perihelion (the point closest to the sun) of the orbit of the planet Mercury, already known from observations in the last century.
2) The bending of light waves near the sun, first observed during a total eclipse in 1919.
3) The *gravitational* red shift, first seen in the light from massive stars, but now measured even on the earth itself using the ultrasensitive Mössbauer effect.

A fourth effect predicted by Einstein's theory (although not considered by Albert himself) was not confirmed until fairly recently [28.8]: the *slowing* of electromagnetic radiation in a gravitational field – as opposed to the *acceleration* of matter as it approaches a heavenly body.[1] The fourth effect was observed by means of radar echoes from the planets Venus and Mercury as they disappeared behind the sun as seen from the earth ("superior conjunction"). In that position, both the outgoing and returning radar waves have to travel very near (indeed around) the sun. Even after taking plasma effects near the sun's surface and other factors into account, physicists found an extra delay of 200 μs – very close to the prediction of general relativity (see Fig. 28.1).

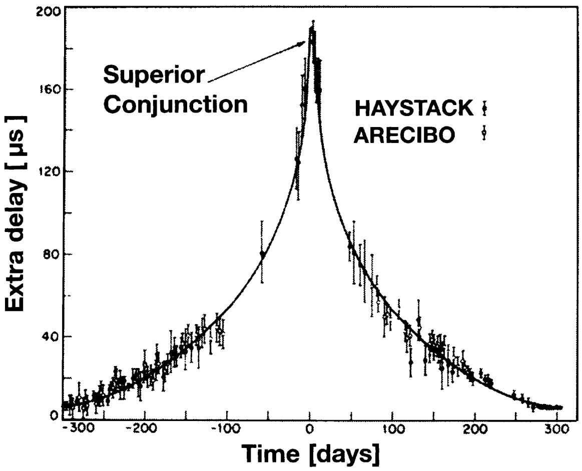


Fig. 28.1 Additional delay of radar echoes from the planet Mercury as predicted by Einstein's general theory of relativity (*solid curve*) and radar measurements from two earth- bound locations

[1] In the long struggle to put his principle of general equivalence of different reference frames into proper mathematical clothing (Riemannian geometry), Einstein discovered – as early as 1909 – that the speed of light could not be constant (as in special relativity) but must depend on the gravitational potential ϕ. Although he had no general theory then, Einstein found that, to first order, $c(\phi) = c + \phi/c$, where c is the usual vacuum velocity of light in field-free space. (Note: $\phi \leq 0$.)

Ironically, the slowing of radiation in gravitational fields, although appreciated very early, was not considered a *testable* proposition until the perfection of radar technology, using Galois sequences, in the second half of this century. The reason for this delay in testing the extra delay was, of course, that no one could picture himself (or anyone else, for that matter) floating next to the sun, stopwatch in hand, clocking the passing photons.

Why was this measurement not done long ago? The reason is that the echo energy from Mercury – exceedingly weak even when visible – drops to 10^{-27} (!) of the outgoing energy as the planet slips behind the sun; in other words, a radar echo would not even suffice to raise the potential of a *single* electron by a thousandth of a volt.

The astounding fact that reliable results have been obtained in spite of these miniscule energies is due mainly to a proper choice of the transmitted sequence of radar pulses, based on irreducible polynomials over finite number fields (called "Galois fields" after the 18th-century French mathematician who died at the age of 20 after a duel over a woman, but not without jotting down some highly ingenious ideas the night before the fatal encounter [28.9]).

28.6 Toward Better Concert Hall Acoustics

Consider the following problem in concert hall acoustics. Recent research, based on a subjective evaluation of the acoustics of 20 major European concert halls [28.10], has shown that many modern halls have poor acoustics because their ceilings are low relative to their widths. Such halls do not provide the listener with enough *laterally* travelling sound waves – as opposed to sound travelling in front/back direction and arriving at the listener's head in his "median" plane (the symmetry plane through his head). Such median-plane sound, of course, gives rise to two very similar acoustic signals at the listener's two ears, and it is thought that the resulting excessive "binaural similarity" is responsible for the poor acoustical quality [28.11].

How can one correct this shortcoming? One component of median-plane sound, the direct sound from the stage, cannot be suppressed because it is needed to localize properly the different instruments on the stage. In principle, the ceiling could be raised so that lateral sound would dominate again, as in the preferred (but unprofitable) old-style high and narrow halls. But modern air-conditioning has made the low ceiling possible – and that is where mounting building costs will keep it.

Still another escape would be to *absorb* the sound at the ceiling and thereby curtail its deleterious effect. But although room acoustics is not usually thought of as an "energy problem", we cannot afford to waste precious musical sound energy, especially in a large modern enclosure, where every "phonon" is needed.

There remains only one alternative, to *disperse* the sound from the ceiling into a lateral pattern. In other words, the ceiling should be what the physicist would call a *reflection phase-grating* – a grating that scatters sound waves comprising many different frequency components into a broad lateral pattern, without absorbing them and with a minimum of specular reflection. What would such a ceiling look like?

The far-field or *Fraunhofer diffraction* of a grating is approximated by the Fourier transform of the complex amplitude of the wave as a function of position as it leaves the grating. Thus, if we want a broad diffraction pattern, we have to look for distributions that have a broad Fourier spectrum. Furthermore, if we want the waves leaving the grating to have uniform magnitude, we need a function of constant magnitude which has a broad Fourier spectrum.

A function which has these properties is given by the Galois sequence $\{b_k\}$ discussed in this chapter; it has a magnitude of 1:

$$|b_k|^2 = 1 \quad \text{for all } k,$$

and the nonzero frequencies of its Fourier spectrum all have the same magnitude:

$$|B_n|^2 = 2^m \quad n \not\equiv 0 \quad \text{mod}(2^m - 1).$$

Here, in the spatial diffraction-pattern application, each frequency component corresponds to a given diffraction order or *spatial frequency* (to use the modern expression from the field of Fourier optics [28.12]). The diffraction angle a_n is given by

$$\sin \alpha_n = n\lambda/L, \qquad |n| \leq \left\lfloor \frac{L}{\lambda} \right\rfloor, \tag{28.13}$$

where λ is the wavelength, L is the period of the grating:

$$L = (2^m - 1)w, \tag{28.14}$$

and w is the "grating constant", i.e., the step size from one element of the grating to the next [28.13].

The next question is: how do we impart amplitudes distributed like Galois sequence b_k to a wave? The answer is by a hard corrugated surface, such as that shown in Fig. 28.2. A normally incident wave is reflected from a hard surface with a *reflection factor* of $+1$. However, in places where the surface is set back by a quarter wavelength, its phase at the reference plane, after having traveled an extra distance of half a wavelength, is shifted by π, i.e., its complex amplitude is $e^{i\pi} = -1$.

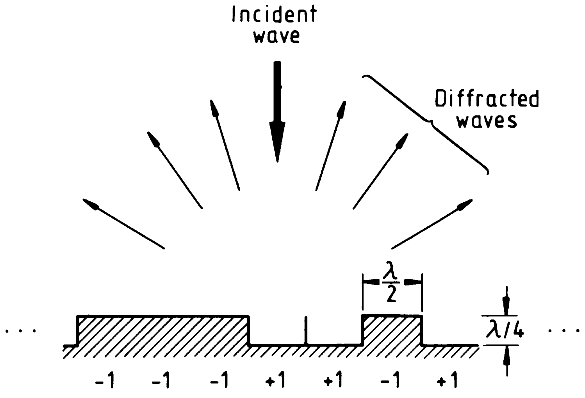


Fig. 28.2 A corrugated surface acting as a reflection phase-grating. If the design is based on a Galois sequence, then the reflected energy is broadly scattered as shown in Fig. 28.3

Of course, the wave reflected from the corrugated structure shown in Fig. 28.2 will not have discontinuous jumps in its amplitude from $+1$ to -1; that would violate the wave equation. We have taken here a kind of Kirchhoff view of diffraction: we consider the diffracted wave to equal the incident wave as modified only in the most obvious manner, without regard to the self-consistency of such assumptions [28.13]. Nevertheless, it is well known that diffraction treated according to Kirchhoff, generally gives results in good agreement with reality, and our case, as proved by subsequent measurements, is no exception (except for some low-frequency loss).

Figure 28.3 shows the diffraction pattern measured from the corrugated surface shown in Fig. 28.2. We see that, indeed, about equal energies are scattered into the different diffraction orders.

The only disadvantage of a ceiling based on a Galois sequence $\{b_k\}$ from $\mathrm{GF}(2^m)$ is that it is binary-valued, $b_k = \pm 1$, so that the corrugated structure only has indentations of one size, namely a quarter wavelength. If we consider an incident wave one octave higher in frequency, i.e., having half the original wavelength, the phase shift it suffers upon reflection by one of the indentations is 2π (instead of π). In other words, there is no phase shift at all, and measurements confirm that there is an almost specular reflection (Fig. 28.4). In fact, Fig. 28.5 shows the reflection from a plane (uncorrugated) surface, and there is hardly any difference between it and the diffraction pattern from the corrugated surface at the octave frequency.

How do we extend the broad diffraction pattern to a broad frequency band comprising several musical octaves? One way would be to make the indentation depths correspond to a quadratic-residue sequence as already discussed (Sect. 16.6). But then we lose the advantage of low specular reflection. Has Galois nothing to offer that works at more than one frequency?

How about considering a finite number field $\mathrm{GF}(p^m)$ with $p > 2$, and perhaps just $m = 1$? For $p = 11$, a primitive element is 2. Using it to construct $\mathrm{GF}(11)$, we get the recursion [see the analogous construction of $\mathrm{GF}(2^4)$, Sect. 26.5]:

$$a_{n+1} = 2a_n \pmod{11}. \tag{28.15}$$

Thus, the Galois sequence for $\mathrm{GF}(11)$, beginning with 1, is

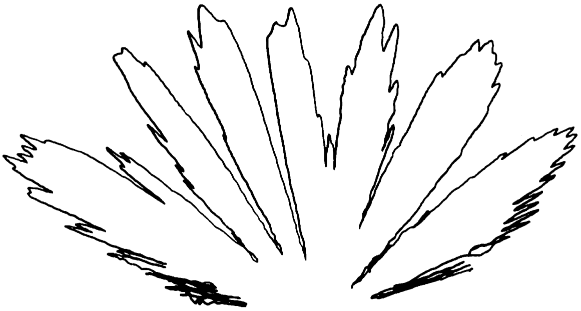


Fig. 28.3 Diffraction pattern from Galois reflection phase-grating shown in Fig. 28.2. Note nearly equal energies being scattered into the (seven) diffraction orders

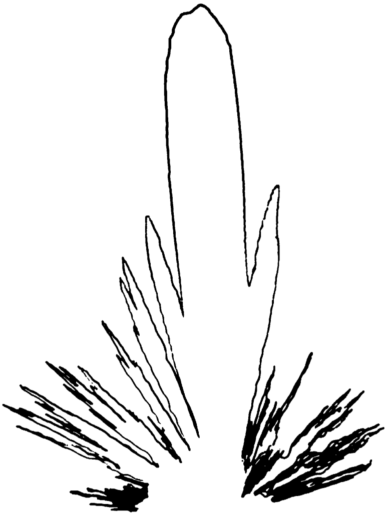


Fig. 28.4 Same as Fig. 28.3 but at half the original wavelength. Note strong specular reflection despite corruption of reflecting surface

$$\{a_n\} = 1,2,4,8,5,10,9,7,3,6. \tag{28.16}$$

However, this sequence is precisely the same as the primitive-root sequence for $p = 11$ and the primitive root $g = 2$ that we studied in Sect. 14.8.

A scale model of a sound-diffusing structure base on GF(11) is shown in Fig. 28.6. It consists of adjacent "wells" of different depths, the depth being proportional to the following sequence of integers:

$$s_n = 2,\ 4,\ 8,\ 5,\ 10,\ 9,\ 7,\ 3,\ 6,\ 1, \tag{28.17}$$

and so on (repeated periodically). It is obvious that the sequence of 10 integers above is a permutation of the integers 1 to 10, i.e., each integer appears exactly

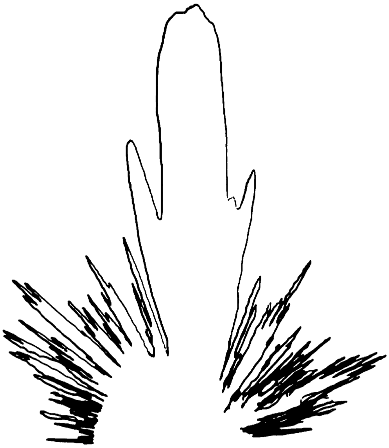


Fig. 28.5 Reflection from plane surface. Note similarity with reflection from corrugated surface shown in Fig. 28.4

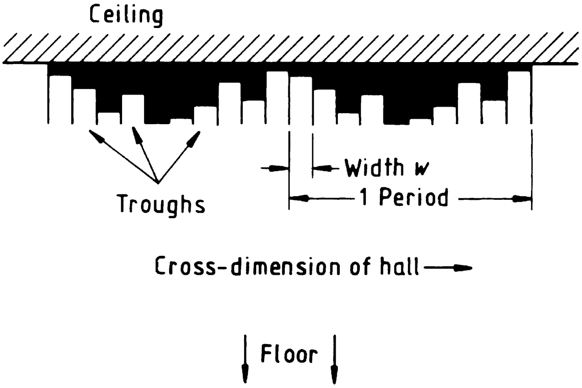


Fig. 28.6 Design of reflection phase-grating based on primitive roots

once. But what is so special about this particular arrangement among the 3,628,800 possible permutations? A little further inspection will reveal that each number in the sequence is twice its predecessor – except when the doubling exceeds 10, in which case 11 is subtracted from it.

In mathematical terms, the sequence s_n is described by the following formula

$$s_n = g^n \pmod{p}, \tag{28.18}$$

where p is a prime number ($p = 11$ in our case) and g is one of its "primitive roots" (here $g = 2$). A primitive root of a prime p is distinguished by the property that the least positive remainder of g^n runs through all the integers from 1 to $p-1$ (in some order) as n goes from 1 to $p-1$.

The prime number 11 has 4 primitive roots ($6, 7, 8$ are the three others besides 2). Thus, *only one* in nearly a *hundred thousand* permutation (not counting cyclic permutations) is generated by primitive roots.

What distinguishes the permutation generated by a primitive root from the thousands of others? The remarkable fact is that the periodic *Fourier transform* of the sequence

$$r_n = \mathrm{e}^{2\pi \mathrm{i} s_n/p} \tag{28.19}$$

has components of *equal* magnitude – except the zeroth, which is smaller. It is this fact that results in the desired wave scattering when these numbers are used in the design of a surface structure as illustrated in Fig. 28.6. The depths of the individual wells are

$$d_n = \frac{s_n \lambda_1}{2p}, \tag{28.20}$$

where λ_1 is the fundamental wavelength.

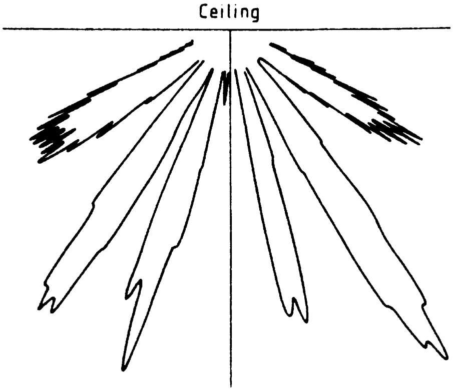


Fig. 28.7 Diffraction pattern from primitive-root design for $p = 7, g = 3$

Figure 28.7 shows the energy scattering from a ceiling structured according to a primitive-root sequence.

What happens if the frequency of the waves is $2, 3, 4, \ldots$ or 10 times higher than the design frequency? Whatever the frequency factor, the structure acts essentially as if the above sequence were multiplied by the frequency factor. Thus, for a frequency 5 times higher, for example, the effective sequence is

$$10, 9, 7, 3, 6, 1, 2, 4, 8, 5 \quad \text{and so on,}$$

which is exactly the old sequence shifted cyclically by 4 places to the left. But, as is well known, the magnitude of the Fourier transform is unaffected by such a shift. Quite a trick for a *primitive* root! Only if we multiply by a multiple of 11 does this method fail. Then all remainders modulo 11 are zero and the structure shown in Fig. 28.6 will act very much like a mirror.

While $2, 6, 7$ and 8 are primitive roots of 11, the number 3, for example, is *not*. To see this, we need only evaluate (28.17) for $g = 3$, which gives

$$3, 9, 5, 4, 1, 3, 9, 5, 4, 1,$$

a sequence that has "forgotten" half the numbers between 1 and 10 ($2, 6, 7, 8$ and 10 are missing!). A ceiling built on the number 3 would lead to poor scattering of sound and lower acoustic quality.

28.7 Higher-Dimensional Diffusors

The wave-scattering principles described in Sect. 28.6 can be applied to two or more dimensions. Suppose the period length of the flat-spectrum sequence has at least two coprime factors m and n, as in $2^4 - 1 = 15 = 3 \cdot 5$ or $31^1 - 1 = 30 = 2 \cdot 15 = 3 \cdot 10 = 5 \cdot 6$. Then the sequence can be converted into a two-dimensional array of length m and width n as illustrated in Fig. 17.1. Such arrays then have the desired spatial correlation and spectrum properties in two dimensions and are useful in noise abatement (to disperse sound waves) or in radar and sonar camouflage and X-ray astronomy (see Sect. 14.10).

If the period length can be factored into three or more coprime factors, correspondingly higher-dimensional designs can be realized.

Other interesting generalizations are described by J. H. van Lint, F. J. MacWilliams, and N. J. A. Sloane [28.14].

For a Galois sequence whose period length $2^m - 1$ is a Mersenne prime, there are of course no coprime factors and higher-dimensional arrays, based on factoring the period, are impossible. Hence, Mersenne primes, which give us perfect numbers (see Sect. 3.7), levy a certain toll when it comes to the construction of higher dimensional arrays – just like the Fermat primes (see Sect. 14.11).

28.8 Active Array Applications

The principles discussed in Sects. 28.6 and 28.7 are also applicable to active arrays (loud-speaker columns, radio and television antennas) in one, two or more dimensions. In fact, since no Kirchhoff approximation is involved, flat-spectrum sequences lead to arrays with highly uniform directional characteristics. Because of the time reversal invariance (T invariance) of mechanics and electromagnetism (also referred to as "reciprocity"), both transmitting *and* receiving arrays with the desired characteristics can be realized [28.15].

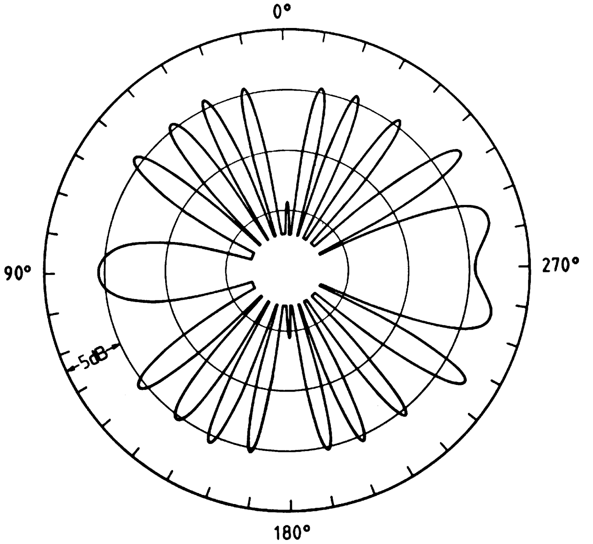


Fig. 28.8 Radiation pattern of 20-element collinear antenna array with half-wavelength spacing. Signals supplied to individual dipoles have constant amplitude across the array but phase angles vary according to primitive-root sequence ($g = 2$, $p = 11$). Note uniformity of radiation in different major lobes, except broadside ($\phi = 0°$ and $180°$) where it is low

One application is to loud-speaker columns that do not focus sound undesirably or whose directional characteristics can be switched from “narrow” to “broad”.

Figure 28.8 shows the directional characteristics of $2 \cdot 10 = 20$ elementary radiators with a spacing of half a wavelength, whose phases have been selected according to the primitive root $g = 2$ of the prime $p = 11$. The uniformity of the radiated energies into the different “side lobes” is impressive. Only the broadside radiation ($\phi = 0^\circ$ and $\phi = 180^\circ$) is weak – as expected for a primitive-root design. Other antenna applications of number theory are discussed in Chap. 30.

Further facts and applications of binary sequences, flat-spectrum or otherwise, can be found in [28.16–18].

Chapter 29
Random Number Generators

"Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin."
– John von Neumann

In contemporary computation there is an almost unquenchable thirst for random numbers. One particularly intemperate class of customers is comprised of the diverse *Monte Carlo* methods.[1] Or one may want to study a problem (or control a process) that depends on several parameters which one doesn't know how to choose. In such cases random choices are often preferred, or simply convenient. Finally, in system analysis (including biological systems) random "noise" is often a preferred test signal. And, of course, random numbers are useful – to say the least – in cryptography.

In using arithmetical methods for generating "random" numbers great care must be exercised to avoid falling into deterministic traps. Such algorithms never produce truly random events (such as the clicks of a Geiger counter near a radioactive source), but give only *pseudo*random effects. This is nicely illustrated by the following (true) story.

An associate of the author, for a study in human vision, wanted to generate visual noise or "snow" (as seen on the screen of a broken TV set). The unwary researcher selected a widely used random number routine available from a renowned software source. The numbers produced by this algorithm had passed a battery of sophisticated tests with random colors flying. Yet when the researcher selected alternative output samples from the "random source" to determine abscissae and ordinates of white dots on a TV screen, the result was not the expected random "snow" but just a few diagonal streaks!

In the following we will sketch some fundamental facts about both random and *pseudo*random number generators. We will be interested in both discrete and (quasi-)continuous distributions of random numbers. (Remember that nothing is *really* continuous in digital computation.) Our main tools will turn out to be congruences and recursions.

[1] The reader may recall that many analytically intractable problems can be solved (albeit with a statistical error) by simulation with a random process and "tabulating" the results of one or more simulation runs.

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_29, © Springer-Verlag Berlin Heidelberg 2009

29.1 Pseudorandom Galois Sequences

In Chap. 27 we became acquainted with a method of generating pseudorandom sequences with elements from $\mathrm{GF}(p)$ and period length $p^m - 1$ [29.1]. For $p = 2$, we obtain binary-valued sequences [29.2] with elements 0 and 1 or, as often preferred, 1 and -1.

These sequences a_n are generated by linear recursions [29.3]:

$$a_n = \langle f_1 a_{n-1} + \ldots + f_m a_{n-m} \rangle_p, \tag{29.1}$$

where addition is modulo p, as indicated by the acute brackets, which signify least positive (or nonnegative) remainders. The coefficients f_k are determined by primitive polynomials in $\mathrm{GF}(p^m)$ (Chap. 28).

These sequences are, of course, not truly random. Each period contains precisely all m-tuples but one (the all-zero tuple). Also, their correlation is not that of a truly random sequence (like, say, the pulses from a Geiger counter near a radioactive source). In fact, the periodic autocorrelation has only two possible values (a highly valued property in numerous applications; see Chap. 28). Nevertheless, by choosing m very large, say $m = 168$, and selecting a relatively small excerpt from a full period, pseudorandomness can approach true randomness under many statistical tests. For $m = 168$ and $p = 2$, for example, the period length exceeds 10^{50} samples, and even as many as 10^{10} samples from such a sequence constitute a very small portion of the entire sequence.

If used in application where higher-order correlations are important, however, trouble can still occur. As an illustration, for $m = 5$ and $p = 2$, and the recursion

$$a_n = \langle a_{n-4} + a_{n-5} \rangle_2, \tag{29.2}$$

or, for $b_n = 1 - 2a_n$,

$$b_n = b_{n-4} \cdot b_{n-5}, \tag{29.3}$$

the third-order correlation coefficient becomes

$$c_{4,5} = \sum b_n \cdot b_{n-4} \cdot b_{n-5} = \sum 1. \tag{29.4}$$

Thus, $c_{4,5}$, instead of being appropriately randomly small, becomess catastrophically large.[2]

The problem with higher-order correlations can be somewhat alleviated by basing the recursion on primitive polynomials with a maximum number of terms. In our example ($m = 5$, $p = 2$), this would mean replacing (29.3) by a recursion involving four terms instead of only two:

[2] Overlooking the importance of higher-order correlation in analyzing nonlinear systems has led to some very misleading results in the animal neurophysiology of hearing when binary pseudorandom sequences of the kind described here have been used as an acoustic input to the animal's ear.

$$b_n = b_{n-1} \cdot b_{n-2} \cdot b_{n-3} \cdot b_{n-5}. \tag{29.5}$$

Now, of course, there is trouble with a fifth-order correlation. But for large m, the recursion can be given so many terms that the nonrandom correlations are usually of an irrelevantly high order.

29.2 Randomness from Congruences

Congruence relations are a preferred device for generating random numbers on a computer [29.4]. A simple congruence

$$x_{n+1} = \langle a x_n + b \rangle_m, \quad x_0 = c, \tag{29.6}$$

has four parameteers: a multiplier a, an additive constant b, a modulus ("base") m, and an initial value ("seed") c.

The integer range of the computer is often chosen as a modulus m, because then if we interpret the acute brackets in (29.6) as meaning "take least absolute residue", the residue reduction is automatic. For example, if the computer is a 32-bit machine, then its integer range is typically from -2^{16} to $2^{16}-1$ (or from $-2^{16}+1$ to 2^{16}), in other words, a least absolute residue system of the modulus $m = 2^{32}$. Since integer "overflow" (an integer exceeding the computer's range) is often automatically handled by taking least absolute residue, the remainder operation in (29.6) therefore does not require additional computation. (Never mind that in computerese residue reduction goes by such, possibly unfamiliar terms as "twos-complement", etc.)

The longest period of the x_n generated according to (29.6) is, of course, m. (Once an earlier value of x_n reappears, the whole sequence of x_n will be repeated; (29.6) represents a first-order (pseudo-)Markov process with no memory beyond its immediate past.) If m is power of 2, $m = 2^k$, then to realize a period length of m, the following conditions must be fulfilled [29.4]:

$$a \equiv 1 \pmod 4 \tag{29.7}$$

and

$$(b, m) = 1, \tag{29.8}$$

i. e., it suffices that b be odd.

With binary computers, a preferred choice for a is $a = 2^5 + 1$, because then multiplication by a is a simple shift opertion and an addition.

If the period length is maximum, then since x_n will eventually assume all m possible different values, the choice of the initial value, x_0, is not crucial.

The random numbers generated by a simple congruence with maximum period length show shorter periods in their least significant digits.[3] If this is troublesome,

[3] Can the reader show why this is so?

and the least significant digits are actually needed (otherwise they could simply be discarded), it is best to adopt another algorithm such as the following.

A more sophisticated random number generator uses r random initial values instead of only one:

$$x_1, x_2, \ldots, x_r, \tag{29.9}$$

which are read out successively. At the same time that an x_n is read out, it is added (modulo m) to the value in the circularly preceding storage location. (In other words, x_r precedes x_1, etc.)

Thus, with the above "seeds", the random sequence will be

$$\begin{aligned}&x_1, x_2, \ldots, x_{r-1}, \langle x_r + x_1\rangle_m,\\ &\langle x_1 + x_2\rangle_m, \langle x_2 + x_3\rangle_m, \quad \ldots\\ &\langle x_1 + 2x_2 + x_3\rangle_m, \quad \text{etc.}\end{aligned} \tag{29.10}$$

29.3 "Continuous" Distributions

If we desire random variables uniformly distributed in the interval $[0,1)$, then simple scaling of the x_n generated by (29.6) or (29.10) will suffice:

$$y_n = \frac{x_m}{m}. \tag{29.11}$$

Of course, the distribution of the y_n is not continuous – nothing is in a computer (or the world, for that matter) – but for large m, the y_n are often taken as a continuous variable.

Distributions that are not uniform can be obtained from a uniform variable as follows. Say $w(z)$ is the desired frequency distribution and $W(z)$ is the corresponding cumulative distribution:

$$0 \leq W(z) \leq 1. \tag{29.12}$$

Then the random variable z is obtained from y by the inverse W^{-1}, of the (nondecreasing) $W(z)$:

$$z = W^{-1}(y). \tag{29.13}$$

For example, for

$$w(z) = \frac{1}{\bar{z}}\mathrm{e}^{-z/\bar{z}}, \quad z \geq 0, \quad \text{and} \tag{29.14}$$

$$W(z) = 1 - \mathrm{e}^{-z/\bar{z}}, \tag{29.15}$$

inversion of $W(z)$ gives

$$z = -\bar{z}\ln(1-y), \tag{29.16}$$

or, more simply (because y is uniform):

$$z = -\bar{z}\ln y. \tag{29.17}$$

Of course, for y near 0, when the logarithm in (29.17) has a singularity, the step size for z is relatively large:

$$\frac{\mathrm{d}z}{\bar{z}} = \frac{|\mathrm{d}y|}{y} = \frac{1}{my}. \tag{29.18}$$

29.4 Four Ways to Generate a Gaussian Variable

One of the most often needed distributions is the *normal* or *Gaussian* distribution:

$$w(z) = \frac{1}{\sqrt{2\pi\sigma}}\,\mathrm{e}^{-(z-\bar{z})^2/2\sigma^2}. \tag{29.19}$$

A Gaussian variable z can be generated according to (29.13), which requires tabulation or computation of the inverse error integral.

In another approach, recourse is had to the Central Limit Theorem of probability theory: an approximation to a Gaussian variable is obtained by summing many uniform variables (at least six, say – but this depends on the application).

A third method uses the fact that a Gaussian variable is obtained by multiplying a Rayleigh variable

$$w(r) = \frac{2r}{\overline{r^2}}\,\mathrm{e}^{-r^2/\overline{r^2}} \tag{29.20}$$

and the cosine of a uniform variable u, uniform in $[0, 2\pi)$. The variable r is obtained with (29.13). Thus, the Gaussian variable g is given by

$$g = \bar{g} + (-2\overline{g^2}\ln y_1)^{\frac{1}{2}} \cdot \cos(2\pi y_2), \tag{29.21}$$

where y_1, y_2 are independent uniform variables in $[0, 1)$.

A second, uncorrelated Gaussian variable can be generated from the same y_1, y_2 by replacing the cosine in (29.21) by sine. More generally, a Gaussian variable with correlation coefficient c can be generated by adding $\arccos(c)$ to the argument of the cosine in (29.21).

In conclusion, we will give a method for generating two independent Gaussian variables that does not require a trigonometric function as in (29.21). Again, we need two uniform variables, y_1, y_2 which will be "centered":

$$u_1 = 2y_1 - 1, \quad u_2 = 2y_2 - 1, \tag{29.22}$$

and *rejected*, unless

$$r = u_1^2 + u_2^2 \leq 1. \tag{29.23}$$

With such r, we then compute the two uncorrelated Gaussian variables with mean 0 and variance 1 as follows:

$$g_1 = \frac{\left(-\frac{2}{r}\ln r\right)^{\frac{1}{2}}}{2} \cdot u_1, \tag{29.24}$$

and

$$g_2 = \frac{\left(-\frac{2}{r}\ln r\right)^{\frac{1}{2}}}{2} \cdot u_2. \tag{29.25}$$

Correlated variables can be obtained by linear combinations of g_1 and g_2.

Of course, none of the methods described here are (or could be!) exact, if only because computation of the logarithm is only approximate. But precision can always be carried as far as needs (and budgets) dictate.

29.5 Pseudorandom Sequences in Cryptography

At first blush, pseudorandom sequences look like a fine tool for srambling or encrypting secret messages. However, great care must be exercised to avoid some terrible pitfalls when relying on *pseudo*randomness.

As early as 1953, *E. N. Gilbert* of Bell Laboratories wrote a memorandum [29.5] in which he pointed out that binary Galois (linear shift-register) sequences (Chap. 27) of (maximum) period lengths $2^m - 1$ can be distinguished from *truly* random sequences on the basis of *much fewer* than 2^m terms; in other words, long before the sequence betrays its pseudorandomness by periodic repetition.

Specifically, Gilbert showed that the rank of a certain matrix constructed from N consecutive terms of the Galois sequence is less than maximal because of the linear dependence between terms. Conversely, the probability that this will happen for a truly random sequence can be made arbitrarily small.

For example, suppose we consider $N = 50$ successive terms from a Galois sequence with $m = 11$. Although the repetition period is 2047, these 50 "bits" will, of course, fail the randomness test based on maximal matrix rank. But the probability that the randomness test will be failed by a truly random sequence with independently distributed 0's and 1's (each with probability $1/2$) is less than 10^{-6} [29.5]. To overcome this inherent weakness of linear shift-register sequences, recent research has concentrated on the problem of increasing the "complexity" of such sequences by taking nonlinear functions of several such sequences [29.6, 7]. This field is presently in considerable flux.

Pronounced biases have also been discovered in pseudorandom (card) shuffles, with interesting number-theoretic implications [29.8].

Good algorithms for pseudorandom permutations [29.9] are of particular interest in cryptography.

New methods for encryption by “random rotations” were described by *Sloane* [29.10] at the Cryptography Workshop held at Burg Feuerstein, Germany, in 1982. Much recent work in cryptography was discussed at that meeting [29.11].

Chapter 30
Waveforms and Radiation Patterns

In numerous technical applications, as well as some theoretical considerations, an enticing problem is encountered. In its most practical (and real) form the problem may be that of designing a radar (or sonar) transmitter *waveform* with a given spectrum such that its peak factor is a minimum. Peak factor is defined here as the *range* of the waveform values divided by their root-mean-square. This is important in radar and sonar, because one often wants to radiate a maximum amount of signal power, of a prescribed spectral shape, with a given peak power limitation on the transmitter.

In many applications, the prescribed spectral shape is a flat (or "white") spectrum. One waveform that has such a flat spectrum is a sharp impulse; but in this case all the power is concentrated at one point in time, leading to a *maximum* peak power. What we want is a minimum peak power. The problem also occurs in talking computers and electronic speech synthesizers in general: synthetic speech made from sharp impulses sounds harsh and reedy; it suffers from a none-too-pleasant "electronic accent". By contrast, computer speech from low peak-factor waveforms sounds smoother.

How can we avoid sharp pulses[1] without modifying the power spectrum? Obviously, we still have the *phase angles* of the Fourier coefficients to play with.

For a given amplitude spectrum (*magnitudes* of Fourier transform coefficients), how does one choose the *phase angles* of the Fourier coefficients in order to achieve the smallest range of magnitudes in the corresponding inverse Fourier transform?

The results on waveforms described in this chapter are also applicable to antenna directivity problems. Specifically, a low peak factor corresponds to an antenna with a wide radiation or receiving directivity pattern.

Another antenna problem solved by number theory is one of "sparse arrays", namely *minimum-redundancy* antennas. Such antennas play an important role in astrophysics and ocean surveillance, where the individual antenna elements (e. g., steerable parabolic "dishes" or submerged hydrophones) are very expensive or costly to control. In such cases, one wishes to construct arrays with the smallest

[1] Also called *Dirac* functions after the famous theoretical physicist who first formulated quantum electrodynamics (QED), one of the most accurate physical theories. Dirac also discovered, on paper so to speak, as consequences of Lorentz invariance, two very fundamental facts of universal reality: elementary particle spin and antimatter. (While spin had been observed before QED, antimatter was another matter and its existence was even doubted by such supreme judges as Wolfgang Pauli.).

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_30, © Springer-Verlag Berlin Heidelberg 2009

number of elements for a given unambiguous target resolution (Sect. 30.5). Minimum redundancy arrays have also become important in real-time diagnostic tomography.

30.1 Special Phases

Early waveforms with low peak factors were found by the author and V. A. Vyssotsky by Monte Carlo computations: several thousand periodic waveforms of a given power spectrum, but different sets of random phase angles, were generated on the computer and sorted according to increasing peak factor. The best waveforms thus found had a peak factor several times smaller than the zero-phase impulse.

Later, the author [30.1] developed a formula for the phase angles α_n of low peak-factor waveforms with a given power spectrum P_n, based on asymptotic spectra of certain frequency-modulated signals:

$$\alpha_n = \alpha_1 - \frac{2\pi}{P}\sum_{k=1}^{n-1}(n-k)P_k, \tag{30.1}$$

where P is the total power:

$$P := \sum_{k=1}^{N} P_k. \tag{30.2}$$

For flat power spectra, $P_k = \text{const}$, (30.1) can be simplified to

$$\alpha_n = \alpha_1 - \frac{\pi}{N}n^2, \tag{30.3}$$

a *quadratic* dependence of phase on harmonic number (frequency) n.

Figures 30.1 and 30.2 illustrate the reduction in peak factor achieved with formula (30.1) for a nonflat power spectrum.

If the low peak-factor waveform is desired to be symmetric in time, then phase angles are restricted to 0 or π. Formula (30.1) is then replaced by [30.1]:

$$\alpha_n = \pi\left\lfloor \sum_{k+1}^{n-1}(n-k)\frac{P_k}{P}\right\rfloor, \tag{30.4}$$

or, for flat spectra,

$$\alpha_n = \pi\left\lfloor \frac{n^2}{2N} + c\right\rfloor, \tag{30.5}$$

where c is a constant that can be adjusted to minimize rounding effects due to the floor function (Gauss bracket) employed in (30.5).

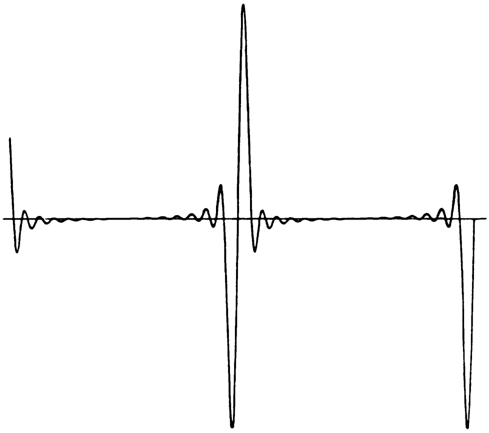


Fig. 30.1 Waveform with "Hamming" power spectrum and all (16) phase angles equal to $\frac{\pi}{2}$

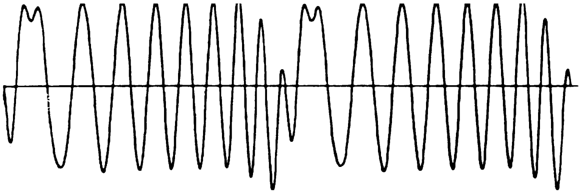


Fig. 30.2 Waveform with same power spectrum as in Fig. 30.1 but phase angles according to (30.1). Note reduction in peak factor (same scale as in Fig. 30.1!)

Figures 30.3 and 30.4 illustrate the 2.64-fold reduction in peak factor achieved with (30.5) for a waveform consisting of 31 harmonics of equal amplitude. The low peak factor is 1.30.

Since this early work on peak factor reduction, several analytical and arithmetic tools have become available which are described in the following sections.

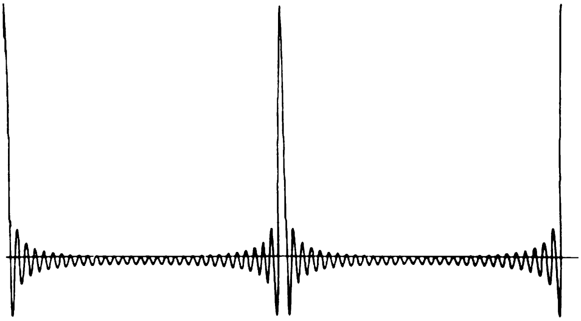


Fig. 30.3 Waveform corresponding to 31 equal-amplitude harmonics in zero phase

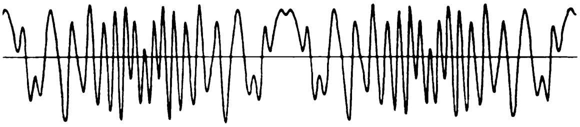


Fig. 30.4 Waveform with same power spectrum as in Fig. 30.3 with phase angles according to (30.5). Reduction in peak factor is by a factor 2.64 (Same scale as in Fig. 30.3)

30.2 The Rudin-Shapiro Polynomials

The so-called Rudin-Shapiro polynomials [30.2] are defined by the following recursion:

$$\begin{aligned} p_0(t) &= q_0(t) = 1, \\ p_{n+1}(t) &= p_n(t) + \mathrm{e}^{\mathrm{i}t2^n} q_n(t), \end{aligned} \tag{30.6}$$

and

$$q_{n+1}(t) = p_n(t) - \mathrm{e}^{\mathrm{i}t2^n} q_n(t). \tag{30.7}$$

It is clear from these definitions that p_n and q_n have period $2\pi/2^n$. In fact, their Fourier transform consists of 2^n frequency components (beginning at zero frequency) of equal magnitude.

The first low-order Rudin-Shapiro polynomials are (abbreviating $\mathrm{e}^{\mathrm{i}t} = z$):

$$\begin{aligned} p_1 &= 1 + z, & q_1 &= 1 - z, \\ p_2 &= 1 + z + z^2 - z^3, & q_2 &= 1 + z - z^2 + z^3, \\ p_3 &= 1 + z + z^2 - z^3 + z^4 + z^5 - z^6 + z^7, \\ q_3 &= 1 + z + z^2 - z^3 - z^4 - z^5 + z^6 - z^7. \end{aligned}$$

It is easy to see from (30.6) and (30.7) that (for $|z| = 1$)

$$|p_n|^2 + |q_n|^2 = 2^{n+1}, \tag{30.8}$$

and therefore

$$\begin{aligned} |p_n|^2 &\leq 2^{n+1}, \\ |q_n|^2 &\leq 2^{n+1}. \end{aligned} \tag{30.9}$$

On the other hand, time averaging over one period yields

$$\overline{|p_n|^2} := \frac{1}{2\pi} \int_0^{2\pi} |p_n(t)|^2 \, \mathrm{d}t = 2^n. \tag{30.10}$$

Thus, the maximum of $|p_n|^2$ exceeds its average value by at most a factor 2 – a property the Rudin-Shapiro polynomials share with a sinusoidal waveform. These polynomials, although perhaps not optimum, are therefore well behaved in terms of their peak factors.

For real waveforms having 2^n frequency components, beginning with frequency m, the real part of $c\,\mathrm{e}^{\mathrm{i}2\pi mt}p_{n-1}(t)$ or $c\,\mathrm{e}^{\mathrm{i}2\pi mt}q_{n-1}(t)$, where c is a complex constant that determines amplitude and initial phase, is a good choice.

The lowest peak factor achieved for $n = 5$ (31 harmonics after elimination of the zero frequency) was 1.35 for the real part of p_5 and 1.36 for the real part of q_5.

Cyclically shifting the sign sequence of the Fourier components and also allowing their phase angles to be shifted by a constant amount improved the peak factors for p and q to 1.29 and 1.31, respectively. These results are comparable to those achieved with the special phases suggested by the author (Sect. 30.1), but at the price of creating waveforms that are asymmetric in time. For symmetric waveforms the best Rudin-Shapiro peak factor (1.35) is noticeably worse than the peak factor (1.30) of the author's phase combination.

30.3 Gauss Sums and Peak Factors

While the Rudin-Shapiro polynomials are limited to waveforms with 2^n frequency components, waveforms $s(t)$ based on Gauss sums can be expected to be well behaved for a number of components equal to $p-1$, where p is a prime:

$$s(t) = \sum_{k=0}^{p-1} \left(\frac{k}{p}\right) \mathrm{e}^{\mathrm{i}kt}, \tag{30.11}$$

where (k/p) is the Legendre symbol, with $(0/p) = 0$.

The reason why $s(t)$ according to (30.11) can be expected to have a low peak factor is the following. The (periodic) "Legendre" sequence (Chap. 16)

$$a_k := \left(\frac{k}{p}\right) \tag{30.12}$$

has a periodic autocorrelation sequence

$$c_n := \sum_{k=0}^{p-1} a_k a_{k+n}, \tag{30.13}$$

which for zero shift is large:

$$c_n = p-1 \quad \text{for} \quad n \equiv 0 \pmod{p}, \tag{30.14}$$

and small otherwise:

$$c_n = -1 \quad \text{for} \quad n \not\equiv 0 \pmod p.$$

This follows from the property of the Legendre symbol

$$\left(\frac{k}{p}\right)\left(\frac{k+n}{p}\right) = \left(\frac{k(k+n)}{p}\right), \tag{30.15}$$

or, after multiplying with

$$\left(\frac{b^2}{p}\right) = 1, \tag{30.16}$$

where for $n \not\equiv 0$, b can be chosen so that

$$bn \equiv 1 \pmod p, \tag{30.17}$$

$$\left(\frac{k}{p}\right)\left(\frac{k+n}{p}\right) = \left(\frac{bk(bk+1)}{p}\right) \quad \text{for} \quad n \not\equiv 0 \pmod p. \tag{30.18}$$

In the sum (30.13) we now replace summation over k by $k' = bk$, which does not affect the result, because k' covers the same values as k. Thus

$$c_n = \sum_{k'=0}^{p-1} \left(\frac{k'(k'+1)}{p}\right) = c_1 \quad \text{for} \quad n \not\equiv 0 \pmod p,$$

and, of course,

$$c_n = p-1 \quad \text{for} \quad n \equiv 0 \pmod p. \tag{30.19}$$

Since

$$\sum_{k=0}^{p-1} \left(\frac{k}{p}\right) = 0, \tag{30.20}$$

we also have

$$\sum_{n=0}^{p-1} c_n = 0, \tag{30.21}$$

and therefore, with (30.19),

$$c_n = -1 \quad \text{for} \quad n \not\equiv 0 \pmod p. \tag{30.22}$$

Now, if the (k/p) are the amplitude spectrum of a periodic complex waveform as defined in (30.11), the squared magnitude of $s(t)$ has the following Fourier representation:

$$|s(t)|^2 = \sum_{n=0}^{p-1} \tilde{c}_n \, \mathrm{e}^{\mathrm{i}2\pi nt}, \tag{30.23}$$

where the $\tilde{c}_n$ are the *aperiodic* correlation coefficients of the (k/p). Nevertheless, with $|c_n|$ small ($|c_n| = 1$ for $n \not\equiv 0$), the $|\tilde{c}_n|$ will also be relatively small. Since c_0 is large ($c_0 = p-1$), $|s(t)|^2$ will be relatively constant with small fluctuations. As a consequence, the real part of the $s(t)$ is expected to stay within relatively narrow bounds – in other words, the peak factor of $s(t)$ constructed according to (30.11) will tend to be small.

Since (30.11) generates only 30 harmonics, a 31st harmonic was added with either sign. The peak factor for either case is 1.66. When applying cyclic shifts, the resulting smallest peak factor over all 30 cyclic shifts of (k/p) is 1.39 – not bad, but not as good as the result obtained with the quadratic ("chirp") phases (Sect. 30.1).

30.4 Galois Sequences and the Smallest Peak Factors

Another type of sequence with two-valued autocorrelation c_n is that of the Galois sequences g_n (Chap. 28) generated by a linear recursion based on a primitive polynomial over GF(2^m). Their period length is $2^m - 1$ and the periodic correlation is given by

$$\begin{aligned} c_n &= 2^m - 1 \quad \text{for} \quad n \equiv 0 \pmod{2^m - 1} \quad \text{and} \\ c_n &= -1 \quad \text{for} \quad n \not\equiv 0 \pmod{2^m - 1}. \end{aligned} \tag{30.24}$$

Thus, a periodic waveform

$$s(t) := \sum_{k=1}^{2^m-1} g_k \, \mathrm{e}^{\mathrm{i}2\pi kt/(2^m-1)} \tag{30.25}$$

constructed with the g_n (or any of their cyclic shifts) will have a relatively constant $|s(t)|^2$, i. e., the peak factor of $s(t)$ will be small (Sect. 30.3).

For $m = 5$ (31 harmonics) there are

$$\frac{\phi(2^5 - 1)}{5} = 6 \tag{30.26}$$

different primitive polynomials over GF(2) of degree 5, namely

$$\begin{array}{lr} 1 + x^2 + x^5 & \text{peak factor} = 1.259 \\ 1 + x^3 + x^5 & = 1.252 \\ 1 + x + x^2 + x^3 + x^5 & = \underline{1.209} \\ 1 + x^2 + x^3 + x^4 + x^5 & = 1.249 \\ 1 + x + x^2 + x^4 + x^5 & = 1.240 \\ 1 + x + x^3 + x^4 + x^5 & = 1.282. \end{array}$$

Here the numbers on the right are the smallest peak factors achieved with each of the six polynomials considering all 31 possible cyclic shifts.[2] As can be seen, all peak factors are smaller than 1.29, and the smallest one, 1.209, is achieved with the third polynomial, giving rise to the recursion

$$g_n = g_{n-1} \cdot g_{n-2} \cdot g_{n-3} \cdot g_{n-5},$$

when used with the initial condition

$$g_1 = g_2 = g_3 = g_4 = -1 \quad \text{and} \quad g_5 = 1.$$

This initial condition happens to give the "idempotent":

$$\begin{gathered} - - - - + - - - + + + - + - + - \\ - + - + + + + - - + + - + + -. \end{gathered}$$

Idempotent means that the polynomial representing this sequence reproduces itself upon squaring. Applied to the sequence, idempotent means that when only the even-numbered terms of the periodically repeated sequence are taken, the sequence is reproduced. Of all the possibilities investigated, the above sequence is the unique choice that gives the smallest peak factor for 31 harmonics with phase angles restricted to 0 or π. The corresponding waveform is shown in Fig. 30.5. This also holds when the same phase angles are applied to a frequency-shifted waveform with harmonics 51 to 81, which has a peak factor of 1.245. In Fig. 30.6, the frequency-shifted

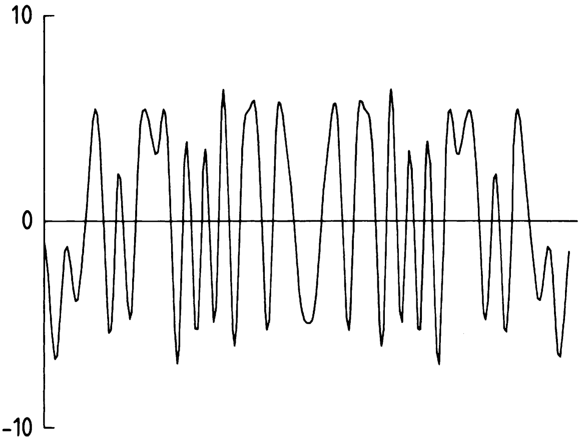


Fig. 30.5 Waveform with 31 constant-amplitude harmonics ($n = 1,2,\ldots,31$) and phase angles according to (idempotent) Galois sequence based on primitive polynomial $1+x+x^2+x^3+x^5$ over GF(2^5). Peak factor is lowest so far found for the given power spectrum

[2] I am grateful to J. L. Hall of Bell Laboratories for calculating these peak factors and for the plots of Figs. 30.5 and 30.6.

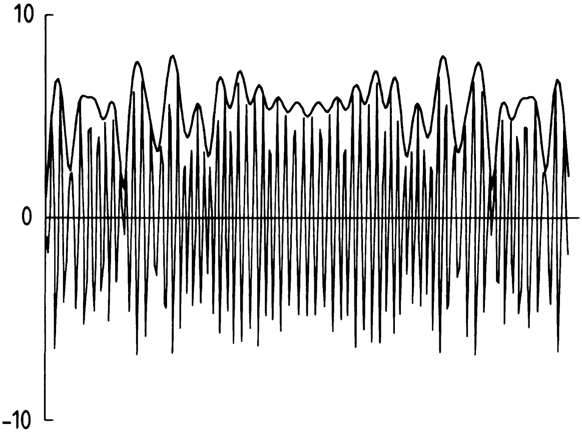


Fig. 30.6 Same as Fig. 30.5 but with harmonic numbers $n = 51, 52, \ldots, 81$. Hilbert envelope of waveform is also shown

waveform is plotted, together with its "Hilbert" envelope. (The Hilbert envelope is defined as the magnitude of the corresponding analytic signal.)

Thus, we see that Galois sequences – already found useful in diffraction problems, precision measurements and error-correction codes – solve yet another practical problem: how to design waveforms with low peak factors for such applications as radar and sonar and computer-synthesized speech. See also [30.3–5].

In the next chapter, still another property of Galois sequences is utilized in the design of necklaces.

30.5 Minimum Redundancy Antennas

The principles of designing waveforms with a low peak factor are also applicable to the design of efficient transmitting or receiving antennas with broad directivity patterns. Given a real or complex sequence having a constant-magnitude Fourier transform, we can construct a linear array with a directivity pattern with constant (or nearly constant) maxima. For the case of a primitive-root sequence (see Sects. 14.8 and 28.6), one resulting directivity pattern is illustrated in Fig. 28.7.

In contrast to such "broadcast" antennas, arrays that favor only *one* direction are often needed. Such antennas have directivity patterns with only one major narrow "lobe". To avoid ambiguity, the spacing between individual elements must be smaller than half the shortest wavelength, λ_{min}, contained in the signal to be received or transmitted. (This requirement follows from the spatial sampling theorem.) For a given directional resolution, the antenna must have a certain minimum size, D. Roughly, for an unambiguous angular resolution $\Delta\alpha$, the antenna size must be

$$D = \frac{\lambda_{\text{min}}}{\Delta\alpha}. \tag{30.27}$$

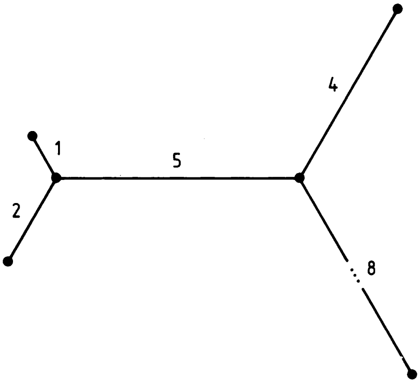


Fig. 30.7 Forked string that can measure all integer distances from 1 to 15

Thus, for a uniform spacing, smaller than $\lambda_{\min}/2$, the number of elements N must be at least

$$N = \frac{2}{\Delta\alpha} + 1. \tag{30.28}$$

For the kinds of resolution required in radio astronomy and ocean surveillance, N can be quite large and given the high cost of individual antenna elements (steerable parabolic reflectors or submerged hydrophones), arrays with uniformly spaced elements can often not be economically realized.

Does this mean we have to give up good radio astronomy or affordable sea security? No. For some applications, especially *detection* tasks, it is sufficent if in the array all required spacings occur at least once. Thus, instead of placing four elements on a straight line with constant spacings of 1 unit, giving an array of size 3 units, the three spacings between the four elements can be chosen to be 1, 3 and 2 units, respectively, giving an array of size 6, which has double the angular resolution for the same number of elements. Unambiguous detection requires that all possible element spacings between 1 and 6 occur at least once, which is the case: the spacings 1, 2, 3 occur between adjacent elements and the remaining three spacings are realized as follows:

$$1+3=4, \quad 3+2=5, \quad 1+3+2=6.$$

Unfortunately, there are no arrays having more than four elements that give every spacing *exactly* once. For $N > 4$, some spacings will be duplicated resulting in an inefficiency or redundancy. Nevertheless, *minimum* redundancy arrays have been designed that give the greatest unambiguous resolution for a given number of elements [30.6]. For example, for $N = 5$, a minimum redundancy configuration has spacings $1,3,3,2$ giving an array size of 9. Except for the spacing 3, all spacings occur only once.

Figure 30.7 shows an interesting string design consisting of two pieces of thread of lengths 3 and 12, connected by a third thread of length 5 as shown. Unsupported by any string theory, this frugal contrivance, communicated to me by H. O. Pollack, can measure all integer distances from 1 to 15. The reader may like to think about generalizations.

30.6 Golomb Rulers

Another minimum-redundancy device is the *Golomb ruler*, defined as a shortest straight ruler of integer length L for a given number of marks n so that a maximum number of distinct integer distances can be measured *without repeating any distance*. For example, for $n = 5$, $L = 11$ with marks at positions 1, 4 and 9 and at the endpoints (0 and 11). This ruler can measure all distances between 1 and 11 except 6. Such a 5-mark ruler can measure at most $5 \cdot 4/2 = 10$ distances, as this one does; so we can infer that no distance is duplicated, as required for a Golomb ruler [30.7].

A longer Golomb ruler has length $L = 151$ and 15 marks (elements) at positions 0, 6, 7, 15, 28, 40, 51, 75, 89, 92, 94, 121, 131, 147 and 151. These 15 marks define a maximum of $15 \cdot 14/2 = 105$ distances. Since no distance is duplicated, all 105 possible distances actually occur. But 105 is short by 46 of the 151 distances measurable by a conventional ruler of length 151. Thus 46 distances are "missing", i. e. they cannot be measured with this sparsely marked ruler.

Apart from radio astronomy, such long Golomb rulers are also useful in signal processing tasks where computational effort is at a premium and drastic "undersampling" of waveforms (or spectra) would be very helpful. In fact, in the mentioned 15-element Golomb ruler of length 151 the undersampling factor is about 10, leading to substantial savings in computation. This idea has been successfully applied by D. Püschel and R. Koch (personal communication) to fundamental-frequency extraction ("pitch detection") of running speech. Even speech *compression* seems feasible by means of "Golomb undersampling" and appropriate signal restoration (e. g. by Fourier analysis or linear prediction using its approximate autocorrelation function).

As an example, Fig. 30.8 shows a sine wave sampled by the just described Golomb ruler with 15 sample points. Of course, the sine wave (24.03 periods in the interval of length 151) is not visible to the naked eye, but Fourier analysis, see Fig. 30.9, of the samples reveals its presence at the expected frequency amid, albeit, plenty of "undersampling noise". Signals consisting of several sine waves have been recovered in this manner.

Golomb rulers have also been used to design sparse loudspeaker arrays for active sound absorption in airconditioning ducts [30.8].

The discovery of new Golomb rulers is a task cut out for supercomputers. For example, a Golomb ruler with 16 marks is known; it has length 177. But where to put the 14 interior marks, even if the length was already known? There *are* more than 10^{20} possibilities! [30.9].

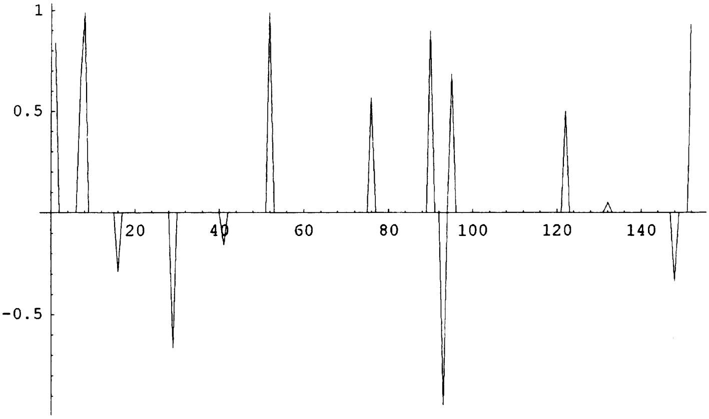


Fig. 30.8 Sine wave sampled by Golomb ruler of length 151

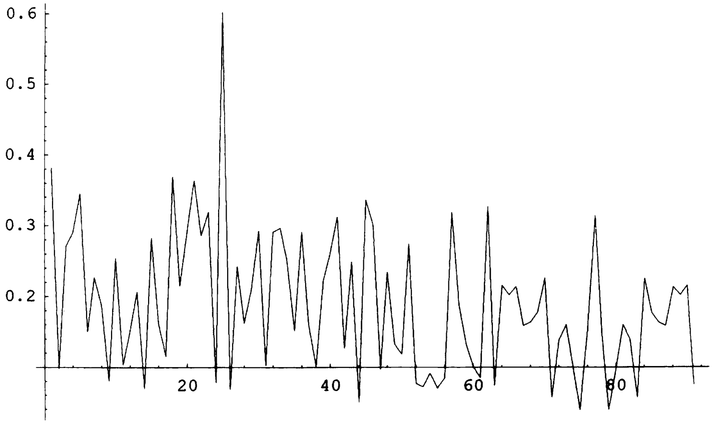


Fig. 30.9 Amplitude spectrum of samples shown in Fig. 30.8. Note the clearly visible peak at the expected frequency 24 Hz above the "undersampling" noise

The number theoretical concept behind such designs is that of *difference sets* discussed in greater detail in [30.10], which also gives solutions to other difficult problems in *additive number theory*.

Chapter 31
Number Theory, Randomness and "Art"

Perfect symmetry is aesthetically boring, and complete randomness is not very appealing either. All works of art show both a degree of regularity and predictability on the one hand, and innovation and "surprise" on the other. In fact, George Birkhoff, of ergodicity fame, once constructed a mathematical formula intended to capture aesthetic value. The formula's two main variables represented regularity and randomness, the aesthetic value function reaching its maximum for ample portions of both ingredients.

While it is probably futile to try to capture artistic achievement in a mathematical formula, there is nevertheless some truth in the basic assumption that both predictability (or a degree of recognizability) and surprise are important in works of art. If we accept this notion, then number theory should be a good store for finding these two ingredients. The distribution of prime numbers, for example, exhibits both some predictability and a certain amount of randomness. Thus, it is not surprising that we can construct, from number-theoretic relationships, some attractive graphic designs and other objects of pleasure. To call them art, however, would be presumptuous.

In this chapter we explore a few of these pleasurable possibilities of number theory.

31.1 Number Theory and Graphic Design

Every third number is divisible by 3, every seventh by 7, etc. In general, the probability that a number picked at random from an infinite range is divisible by the prime number p_i is $1/p_i$, and these probabilities are apparently independent. Thus, the probability that a number is divisible by both 3 and 7 is $1/21$. Similarly, the probability that a given number is *not* divisible by p_i equals $1-1/p_i$.

The probability that *two* numbers picked independently at random are *both* divisible by the *same* prime factor p_i is obviously $1/p_i^2$. And the probability that both are *not* divisible by the same prime factor is consequently $1-1/p_i^2$.

What is the probability P that two randomly chosen numbers have *no* common prime factor, i.e., that they are coprime? If we believe in the independence of prime factors, the answer must be

M. Schroeder, *Number Theory in Science and Communication*, 5th ed.,
DOI 10.1007/978-3-540-85298-8_31, © Springer-Verlag Berlin Heidelberg 2009

$$P = \prod_i \left(1 - \frac{1}{p_i^2}\right), \tag{31.1}$$

where, because we are considering an infinite range, the product is extended over *all* primes. Does P have a finite (nonzero) value? For the first 6 primes $(2,3,5,7,11,13)$ its value is 0.618 (no relation to the Golden ratio!) But that does not teach us much – if we recall, for example, how slowly the distances between successive primes diverge. However, the reciprocal of the infinite product (31.1) can be multiplied out, using the formula for geometric series:

$$\frac{1}{P} = \prod_i \frac{1}{1 - 1/p_i^2} = \prod_i \left(1 + \frac{1}{p_i^2} + \frac{1}{p_i^4} + \frac{1}{p_i^6} + \dots\right). \tag{31.2}$$

The forbidding-looking infinite product of the infinite sum on the right is nothing but *all* the reciprocal squares $1/n^2$, each square appearing exactly *once* (by the Fundamental Theorem). Thus, we have seen in earlier chapters:

$$\frac{1}{P} = \sum_{n=1}^{\infty} \frac{1}{n^2}, \tag{31.3}$$

a well-known sum (the Riemann zetafunction for $z = 2$). Its value, $\pi^2/6$, can be found in most collections of formulae. Hence $P = 6/\pi^2 = 0.6079\dots$.

While formally, the result only holds if we pick pairs of integers from an infinite range, the value $6/\pi^2$ is already very close for the range from 2 to 11: exactly 60 of all 100 possible pairs of integers from 2 to 11 have no common divisor.

Figure 4.8 shows (as white dots) all the integer pairs in the range 2 to 256 that are coprime, i.e., have no common divider. As expected, the plane lattice is fairly uniformly filled with white dots, and their average density is very close to $6/\pi^2$.

Many well-known number-theoretic relations can be observed in such a graphic display, and who knows – close visual inspection might even suggest new theorems [31.1]. The plot is also related to the topic of "visibility": the white dots are precisely those that can be "seen" from the point with coordinates $(1,1)$, while all other points are "hidden" behind one or several points on the straight line connecting them with the point $(1,1)$. To put it more dramatically, if a pellet of paint exploded at $(1,1)$, the paint would only stick on little pegs at lattice points where the white dots are shown in Fig. 4.8.

As we remarked before, the property of divisibility is a *periodic* one. The Fourier transform, of course, is also "sensitive" to periodicities. Thus, the Fourier transform of the "holy" lattice shown in Fig. 4.8, interpreting a white dot as having value $+1$ and the other lattice points as having value -1, should be particularly interesting. Figure 4.9 shows the *magnitude* of the two-dimensional Fourier transform of the number-theoretic "coprimality" function (Fig. 4.8). Not unexpeectedly, we get a cluster of "stars" and reflection symmetries around the two diagonals. But there are also near (broken?) symmetries around the horizontal and vertical axes. They are not so easy to explain (how is *near* symmetry defined?), and we will leave it to the reader to explore this symmetry breaking [31.2].

We may also view Fig. 4.9 as the result of X-ray diffraction from a "moth-eaten" two-dimensional crystal with many atoms missing, but with the remaining ones in perfect lattice positions. But no matter how we look at it, number theory has given us quite a pretty picture here.

31.2 The Primes of Gauss and Eisenstein

Prime numbers can be defined in fields other than the integers. In the complex number field $\mathbb{C}$ we have the *Gauss primes*, a subset of the Gauss integers $n + \mathrm{i}m$, where $\mathrm{i}^2 = -1$. Primes of the form $4k-1$ in $\mathbb{Z}$ are still primes in $\mathbb{C}$, but 2 and primes of the form $4k+1$ can be factored in $\mathbb{C}$!

Examples: $2 = (1+\mathrm{i})(1-\mathrm{i})$, $5 = (2+\mathrm{i})(2-\mathrm{i})$, $13 = (2+3\mathrm{i})(2-3\mathrm{i})$, $17 = (4+\mathrm{i})$ $(4-\mathrm{i})$, etc.

The Gauss primes form a pleasing pattern, shown in Fig. 31.1, which has been used in weaving tablecloths and tiling floors.

Eisenstein, one of Gauss's favorite mathematicians, defined his own primes based on the complex cube root $\omega = (1-\sqrt{-3})/2$. (Note: $1+\omega+\omega^2 = 0$.) They are a subset of the Eisenstein integers $n + \omega m$. The Eisenstein primes form an interesting pattern with hexagonal symmetry (Fig. 31.2). Near the centre of Fig. 31.2, 6 times 6 hexagonal points from 6 hexagons. Are *complete* hexagons formed anywhere else?

The prime 2 and primes of the form $6k-1$ are also Eisenstein primes, but 3 and primes of the form $6k+1$ can be factored.

Examples: $3 = (1-\omega)(1-\omega^2)$, $7 = (2-\omega)(2-\omega^2)$, $13 = (3-\omega)(3-\omega^2)$, $19 = (3-2\omega)(3-2\omega^2)$, etc.

Another pleasing plot is obtained with the set-theoretic *intersection* of the Gauss and Eisenstein *composites*. What are their asymptotic densities?

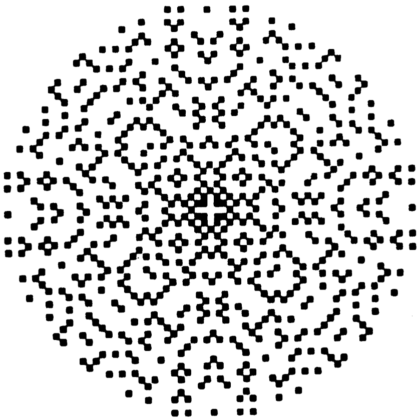


Fig. 31.1 Gauss primes $n + m\mathrm{i}$ with norms $n^2 + m^2$ less than 1000. A favorite design for tablecloths

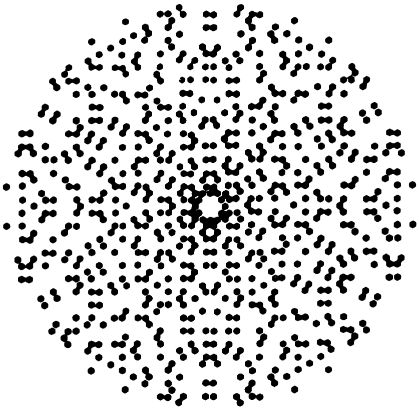


Fig. 31.2 Eisensteins primes $n+m\omega$, where ω is a (primitive) complex root of 3. How would bees react to this pattern?

An even more interesting pattern emerges if the Gauss and Eisenstein primes and their *intersections* are plotted on the same lattice in 3 different colors.

31.3 Galois Fields and Impossible Necklaces

Suppose we inherited a collection of white and black pearls. Could we make a necklace of only 16 pearls (a real choker!) in which all possible variations of 4 adjacent colors appeared? There are $2^4 = 16$ such variations; thus, the necklace would have to have *at least* 16 pearls. But is it really possible?

Figure 31.3 shows the answer, with the 16th pearl, a black one, about to be inserted. With a bit of patience we can find every variation, as promised. For example, the variation of a white pearl followed by a black and two more white ones is found in position 6 to 9.

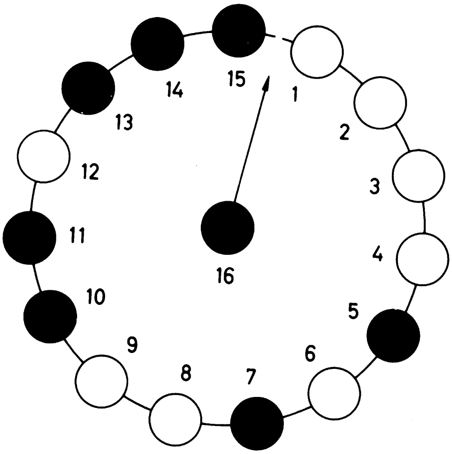


Fig. 31.3 A $15+1$ pearl necklace with all possible quadruplets of black and white pearls

On what principle was the necklace designed? Galois field theory! Or, more specifically, the cyclotomic polynomials over the Galois field GF(2) that we encountered in error-correcting codes and in the confirmation of the fourth general relativistic effect with Mercury or Venus in superior conjunction.

To design the 16-pearl necklace we need a primitive factor of degree 4 of the polynomial $x^{15}+1$. Such a factor is

$$x^4+x^3+1, \tag{31.4}$$

leading to the recursion

$$a_{n+4}=a_{n+3}+a_n. \tag{31.5}$$

With the initial condition $a_1=a_2=a_3=a_4=1$, the recursion gives the following binary sequence of period 15:

$$1\;1\;1\;1\;0\;1\;0\;1\;1\;0\;0\;1\;0\;0\;0;1\;1\;1\;\ldots, \tag{31.6}$$

which contains all variations of length 4 except the quadruplet 0000. This shortcoming can be rectified by adjointing another 0 at the end of the period [(before the semicolon in (31.6)]. If we now identify 1 with a white pearl and 0 with a black pearl and tie one period of the string into a necklace, we obtain the "maximally varied" design shown in Fig. 31.3.

In fact, instead of two colors, we can make such a necklace out of *four* different colors, basing the design on polynomials over GF(4). Of course, now a necklace of 16 pearls can have at most all possible *pairs* of the four colors. That such a necklace is indeed possible is illustrated by Fig. 31.4. The design is based on GF(4) consisting of the four elements $(0,1,\omega,\omega^2)$ with $\omega^3=1$ and $1+\omega+\omega^2=0$ and the primitive polynomial $x^2+x+\omega$. With the initial condition 0, 1 and identifying 0 with red, 1 with blue, ω with yellow and ω^2 with green, we obtain the 16-pearl maximally varied necklace shown in Fig. 31.4 – not, however, before inserting an extra red pearl into the 16th position.

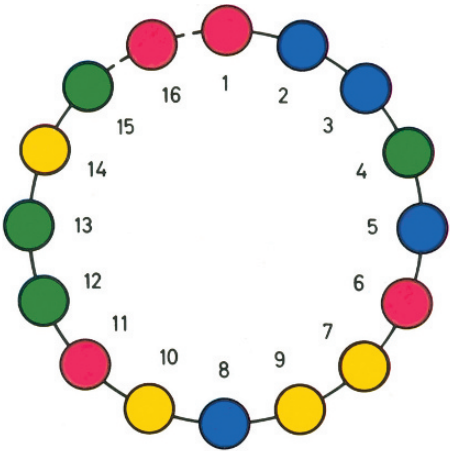


Fig. 31.4 Four-color necklace with 16 pearls exhibiting all possible pairs of colors

Can we go a step further and give each pearl *two* binary properties, such as color and size and have all possible pairs of the four different kinds of pearls (large white, small white, large black, small black) in a necklace of 16 pearls? The answer is yes: we need only identify these four properties with the above four colors to obtain a solution.

What is *not* possible, however, is a necklace with both the stated property for pairs of adjacent pearls and all *quadruplets* of white and black pearls (regardless of size) *and* all quadruplets of large and small size (regardless of color).

If we restrict ourselves to 15 pearls, excusing the occasional absence of a small black pearl, then a solution is possible (see Fig. 31.5). For example, we find the quadruplet black/black/white/black in position 7 to 10, or the quadruplet small/large/small/large in positions 9 to 12. The doublet small white/large black is found in positions 9 to 10. However, the doublets large white/small black and small black/small black are missing unless we sneak in a 16th small black pearl between positions 15 and 1. But even then the doublet small white/small black and the quadruplet black/black/black/black is missing, unless we also allow the 16th pearl to go between positions 5 and 6. In other words, we really need 17 pearls for the *super*-varied necklace. But then some variations occur twice, for example small black/small black – very inelegant! But for that one loose pearl, the arrangement could not be more perfect.

The problem of the loose pearl is deeply related to the fact that 0 has no inverse, not even in a finite field where we expect surprises: for example $3^{-1} = 5$ in GF(7), etc. (Does the reader remember why the inverse of 3 in GF(7) is 5 – and how to calculate it, if it was not so easy to find by trial and error?)

If one pearl is always excused, the super-varied necklace can be extended to, say, $63 = 2^6 - 1$ pearls, having the following six binary properties:[1]

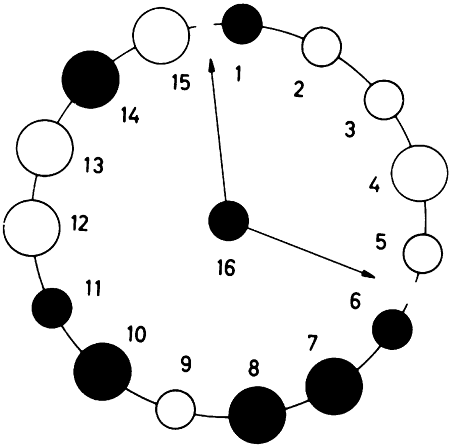


Fig. 31.5 A 16-pearl necklace having pearls of two colors and two sizes. It shows all possible color and size quadruplets and all possible color and size pairs – if pearl number 16 is allowed to move around

[1] Finite mathematicians Jessie MacWilliams and Andrew Odlyzko acted as assistant jewelers to the author in the design of this necklace.

large or small
round or octagonal
solid or pierced
polished or rough
bright or dark
colored or gray .

Except for an occasional missing small-octagonal-pierced-rough-dark-gray-pearl, seen in the centre of Fig. 31.6, the necklace shown contains all possible $2^6 = 64$ sextuplets of *any* of the 6 binary properties listed above, i.e., $6 \cdot 63 = 378$ prescribed sextuplets.

In addition, if two properties are considered at a time, i.e., $\binom{6}{2} = 15$ paired property choices, all $4^3 - 1 = 63$ triplets can be found, or a total of $15 \cdot 63 = 945$ prescribed triplets.

Further, if *three* properties are considered, in any of $\binom{6}{3} = 20$ different ways, the necklace has all $8^2 - 1 = 63$ possible doublets, for a total of $20 \cdot 63 = 1260$ doublets.

Finally, considering six properties at a time, we find a total of 63 different singlets in the necklace.

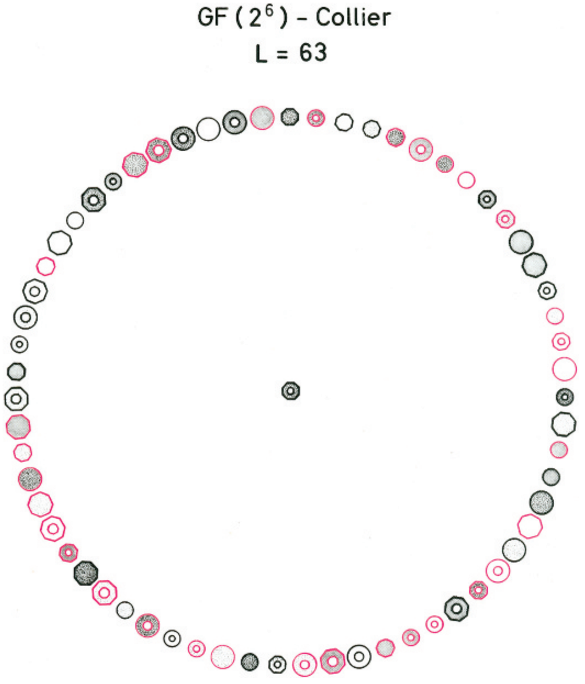


Fig. 31.6 A 64-pearl necklace with 6 binary properties including all 64 different pearls and showing all possible pairs, triplets and sextuplets (a total of 2646 combinations) – if one pearl is allowed to move around

The total number of constraints that the necklace in Fig. 31.6 obeys is a staggering $\left[\binom{6}{1}+\binom{6}{2}+\binom{6}{3}+\binom{6}{6}\right]\cdot 63 = 42\cdot 63 = 2646$. Permitting ourselves occasional cheating with the loose small-octagonal-pierced-rough-dark-gray pearl, all singlets and *any* variation of doublets, triplets and sextuplets of pearls occurs [31.3]. Such a super-necklace would not only embellish the discriminating woman (or man) but its design might even benefit some "experimental designs" (as the psychologists are fond of saying).

31.4 "Baroque" Integers

Nothing, it seems, could be more pedestrian than the sequence of integers 1, 2, 3, 4, 5, 6, 7, ... Yet, as we shall see, the integers can give rise to surprising musical effects.

To begin, write the integers in binary form,

$$0\ 1\ 10\ 11\ 100\ 101\ 110\ 111\ \ldots,$$

then count the number of 1's (in other words, form the digital root),

$$0\ 1\ 1\ 2\ 1\ 2\ 2\ 3\ \ldots. \tag{31.7}$$

This sequence, which I have called the *Dress* sequence, can also be obtained by a simple recursive algorithm [31.4]. Beginning with a single 0, append what you already have augmented by +1:

$$\begin{array}{llll}
0 & & & \\
0\ 1 & & & \\
0\ 1 & 1\ 2 & & \\
0\ 1 & 1\ 2 & 1\ 2\ 2\ 3 & \\
\ldots & & &
\end{array} \tag{31.8}$$

Can the reader see why the recipe (31.8) leads to the same sequence as the one obtained from counting 1's in the binary representation of the integers (31.7)?

The recursive construction of the Dress sequence has two advantages: one, every step *doubles* the length of the sequence instead of adding just a single term. Two, recursive constructions lead to *self-similarity*. Thus, the infinite Dress sequence must be self-similar, which it is indeed: taking every other term, starting with the initial 0 gives

$$0\ 1\ 1\ 2\ \ldots,$$

i. e. the original Dress sequence.

If, instead of taking every second term, we take every *third* term, we obtain a sequence

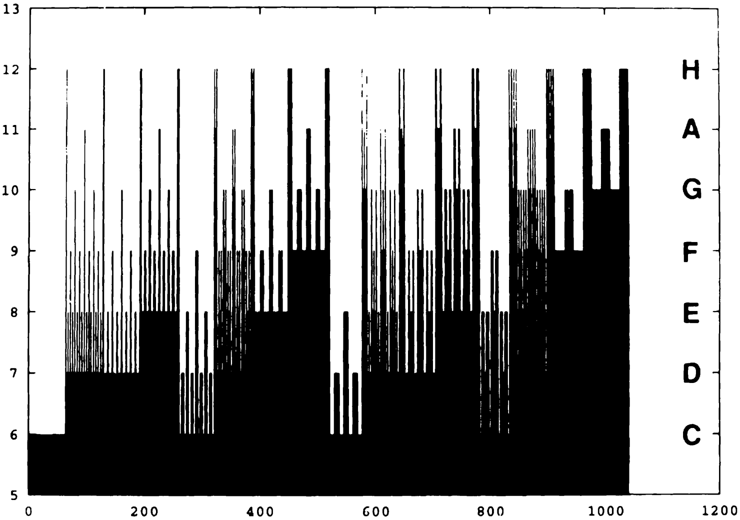


Fig. 31.7 “Baroque” integer sequence obtained from the binary representation of every 63rd integer. When the resulting numbers are identified with musical notes, a Scarlatti-like tune is generated

$$0\,2\,2\,2\,2\,4\,2\,3\,2\,4\,4\,\ldots,$$

whose partial sums exhibit *fractal growth*:

$$0\;2\;4\;6\;8\;12\;14\;17\;19\;23\;27\;\ldots.$$

I leave it to the interested reader to plot this sequence and to discover its fractal growth property, i. e. the exponent with which the partial sums grow.

Similar – in fact *self*-similar – growth properties are also observed for the partial sums of the Dress sequence decimated by other factors that are not powers of 2, especially factors of the form $2^m - 1$. Such decimated sequences have been applied with surprising and pleasing results in the composing of music. My student Lars Kindermann (kindermann@forwiss.uni-erlangen.de) has taken every 63rd term of the Dress sequence and converted it to musical notes in the C major scale ($1 = \text{C}$, $2 = \text{D}$, $3 = \text{E}$, $4 = \text{F}$ etc.). To introduce rhythm into such a number-theoretic “composition” a repeated note is sounded only once and held (see Fig. 31.7). The resulting music has a fascinating and distinct “baroque” quality reminiscent of the Italian composer Domenico Scarlatti (1685–1757). See (and listen to) http://www.forwiss.uni-erlangen.de/~kinderma/musinum.html or visit the Fractal Music Project on the World Wide Web.

Other self-similar sequences that have been used in the composition of “fractal music” are based on sequences of notes with pitches in the ratio $1 : 2^{13/12}$ [31.5–7]. In addition to melody, self-similar sequences, such as the “rabbit” sequence, a binary sequence related to the Fibonacci numbers, have been used to create appealing rhythms [31.5].

Part X
Self-Similarity, Fractals and Art

Chapter 32
Self-Similarity, Fractals, Deterministic Chaos and a New State of Matter

> *"O Julia, mir graut's, wenn man bedenkt,*
> *an dünnen Fädchen nur dein Schicksal hängt.*
> *Dein body ist so thin und doch so schön,*
> *wann werde ich dich sweet heart wiederseh'n."*
> – *Manfred Peschel*

Nature abounds with periodic phenomena: from the motion of a swing to the oscillations of an atom, from the chirping of a grasshopper to the orbits of the heavenly bodies. And our terrestrial bodies, too, participate in this universal minuet – from the heart beat to the circadian rhythm and even longer cycles.

Of course, nothing in nature is *exactly* periodic. All motion has a beginning and an end, so that, in the mathematical sense, strict periodicity does not exist. Nevertheless, periodicity has proved to be a supremely useful concept in elucidating underlying laws and mechanisms.

One reason for the universality of simple harmonic motion is the linearity – or near-linearity – of many physical systems and, more generally, the invariance against displacement in space and time of the laws governing their behaviour.

But there are numerous other phenomena in which these fundamental principles break down and, instead of periodicity, we get aperiodic or even chaotic motion: the smooth waves on a well-behaved lake turn to violent turbulence in the mountain brook and the daily sunrise, the paradigm of predictability, is overshadowed by cloud formations, a haven for *chaos* – albeit *deterministic* chaos.

But no matter how chaotic life gets, with all regularity gone to bits, another fundamental bulwark often remains unshaken, rising above the turbulent rubble: *self-similarity*, an invariance with respect to *scaling*; in other words, invariance not against *additive* translations, but invariance against *multiplicative* changes of scale.

Many laws of nature are independent, or nearly so, of a scaling factor. The fact that scaling usually has a limit (Planck's constant, when things get too small, or the speed of light, when objects fly too fast) does no harm to the usefulness of "thinking self-similar", just as the lack (outside mathematics) of strict periodicity is no great impediment in the real world. In fact, self-similarity is nothing but periodicity on a *logarithmic* scale.

What has number theory got to do with all this? Are there any self-similar principles at work behind the integers? If so, are they useful? Can they shed light into

DOI 10.1007/978-3-540-85298-8_32, © Springer-Verlag Berlin Heidelberg 2009

dark chaos? Surprisingly, the answers to these questions are affirmative; and, in this final chapter, we shall try to show how.

Take the simplest concept from number theory, (“Kronecker’s”) integers. Starting from nothing and counting up, we get

$$0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7\ldots$$

Next, let us represent these numbers in the binary system:

$$0,\ 1,\ 10,\ 100,\ 101,\ 110,\ 111,\ldots$$

Now take the “digital roots” or parity of each binary number (i. e., the sum of its digits modulo 2):

$$0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ \ldots$$

The resulting sequence is called the *Morse-Thue sequence*. It has many remarkable properties. For example, it can be constructed recursively by appending to each subsequence the complemented subsequence, as follows:

$$\begin{array}{l} 0 \\ 0\ 1 \\ 0\ 1\ 1\ 0 \\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1 \end{array}$$

This kind of construction makes it manifestly aperiodic; it never repeats. But it has a much more interesting property: the infinite sequence is *self-similar*. In fact, striking out every second term reproduces the sequence:

$$0\,\cancel{1}\,1\,\cancel{0}\,1\,\cancel{0}\,0\,\cancel{1}\ \ldots = 0\ 1\ 1\ 0\ \ldots \quad . \tag{32.1}$$

The sequence can also be constructed from “nothing” by having each term have a “complementary baby”:

$$\begin{array}{cccccccc} 0 & & & & & & & \\ 0 & & & & 1 & & & \\ 0 & & 1 & & 1 & & 0 & \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}$$

Such sequences are called self-generating [32.1]. Another nonrandom self-generating sequence is obtained from the signs of the Rudin-Shapiro polynomials, see Sect. 30.2. These sequences mimic, in one dimension, the self-organization of matter, such as crystal growth. However, the Morse-Thue sequence is an *aperiodic* self-generating sequence, thus imitating not the growth of a periodic crystal, but a recently discovered new solidification process.

Surprisingly, although the Morse-Thue sequence is aperiodic, its Fourier transform or spectrum is not at all noiselike. The self-similarity of the sequence induces

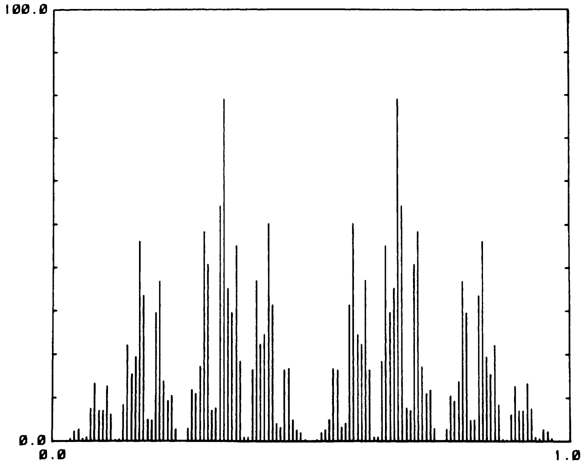


Fig. 32.1 Fourier amplitude spectrum of the self-similar number-theoretic Morse-Thue sequence: 01101001..., defined by the *parity* of the integers $0,1,2,3,\ldots$, when expressed in the binary notation $(0,1,10,11,\ldots)$. The (infinite) sequence reproduces itself when, after each element, the complementary element is inserted. Equivalently, the sequence also reproduces itself upon decimation, namely when every other element is struck out. The scaling factor of the self-similarity is thus 2. Although this sequence is aperiodic (it never repeats itself), the amplitude spectrum shows prominent peaks that would be missing in a truly random sequence, such as an aperiodic sequence without the long-range order of the Morse-Thue sequence. (For the discrete Fourier transform shown here, the sequence was truncated to a length of 256.)

long-range correlations and causes its spectrum to be highly structured with many high peaks, *as if it were periodic*. In other words, the Morse-Thue sequence, being self-similar, is *both* aperiodic *and* has a spectrum resembling that of a *periodic* sequence, see Fig. 32.1.

This fact would perhaps be only mildly interesting if such sequences did not model something that can actually happen in nature, but was discovered only recently: a *new state of matter*!

We are all familiar with *crystals*, in which the individual atoms are arranged in periodic lattices. And we also know *amorphous* substances, such as ordinary glasses or most liquids, in which the atoms are randomly distributed. Until recently, few if any people suspected that there could be another state of matter sharing important aspects with both crystalline *and* amorphous substances. Yet, this is precisely what *D. Shechtman et al.* [32.2] discovered when they recorded electron diffraction patterns, see Fig. 32.2, of a special aluminum-manganese alloy $(\mathrm{Al_6Mn})$. The diffraction pattern, i. e., the two-dimensional Fourier transform, showed *sharp peaks*, just like those for a periodic crystal. But the pattern also showed a five-fold symmetry that periodic crystals simply *cannot* have.

However, as we know from the *number-theoretic* Morse-Thue sequence, sharp spectral peaks and aperiodicity are no contradiction, as long as self-similarity prevails.

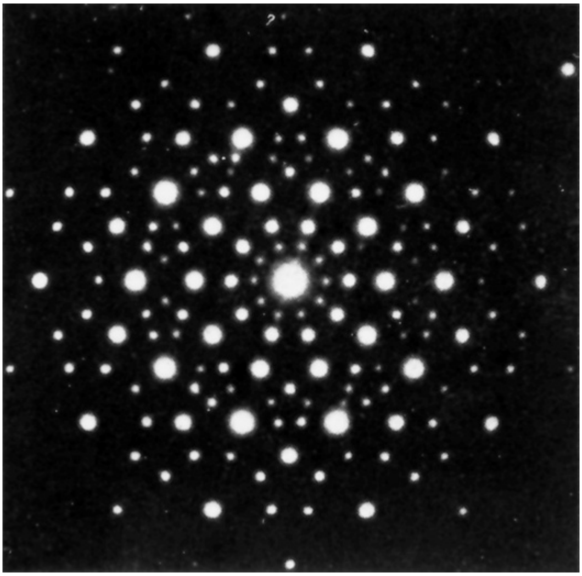


Fig. 32.2 Electron diffraction pattern of an alloy of aluminum and manganese obtained by *Shechtman* et al. [32.2]. The bright spots say that the alloy is not amorphous (like a glass) but "well-ordered". Yet the five-fold symmetry exhibited by the diffraction pattern cannot occur for any periodic crystal. Thus, the alloy must be both aperiodic and well-ordered: a new state of matter that is closely related in its structure to certain self-similar, but aperiodic, sequences known in number theory (see Fig. 32.3)

The number 5, which occurs in the diffraction pattern, points to another self-similar sequence in which perhaps the *Golden ratio* is involved, and this is indeed the case. In fact, it turns out that an aperiodic self-similar sequence derived from the *Fibonacci* numbers is a good model of this new state of matter, showing pronounced peaks in its Fourier spectrum (see Fig. 32.3).

While visiting this newly found Garden of Eden, we will also, in this chapter, make brief calls at an adjoining paradise, that of *Cantor sets* (from which, Hilbert said, we will never be driven away), and at the strangely attractive zoo of *fractal objects*, darlings of a new generation of physical scientists and computer artists.

32.1 Fibonacci, Noble Numbers and a New State of Matter

Let us take another peek at Fibonacci and his rabbits, Fig. 5.9. If we designate a young pair by 0 and a mature pair by 1, Fig. 5.9 (in which this distinction is not shown) becomes this:

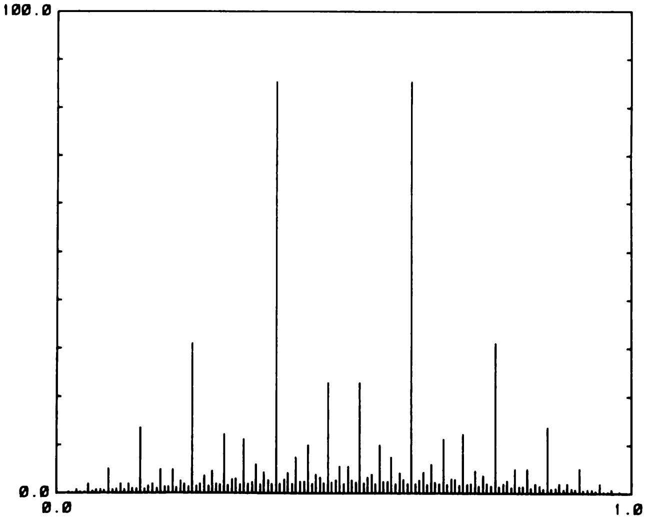


Fig. 32.3 The Fourier amplitude spectrum of the self-similar number-theoretic "rabbit" sequence: 10110101..., which can be generated by replacing each 0 (signifying an immature rabbit pair) by a 1 (mature pair) and inserting after each 1 a 0 (birth of a new pair). This generates the next "generation" of the sequence from the previous one. The lengths of successive sequences equal successive Fibonacci numbers, F_n. (In this illustration, the "breeding" process was terminated at $F_{12} = 144$.) The rabbit sequence can also be generated from a Beatty sequence [32.1] $a_k = \lfloor ng \rfloor = 1,3,4,6,8\ldots$, where $g = 1.618\ldots$ is the Golden ratio, and the a_k specify the indices of the nonzero-terms in the rabbit sequence. (The Beatty sequence $b_k = \lfloor ng^2 \rfloor = 2,5,7,\ldots$ gives the indices of the zero-terms.) Note the pronounced spectral peaks near the two frequencies $1/g$ and $1/g^2$ times the sampling frequency, and the self-similar features of the spectrum involving the (approximate) scaling factors g and g^2. This sequence models, in one dimension, the aperiodic long-range order of the new state of matter recently discovered (see Fig. 32.2)

$$\begin{array}{l} 0 \\ 1 \\ 1\,0 \\ 1\,0\,1 \\ 1\,0\,1\,1\,0 \\ 1\,0\,1\,1\,0\,1\,0\,1 \end{array} \tag{32.2}$$

where each 1 generates a 1 and a 0 in the row below and each 0 gives rise to a 1. Like the Morse-Thue sequence, this is also a self-generating sequence.

Another law of construction for this bit stream (which I shall call the *rabbit sequence*) is quite apparent: after the first two rows, simply append to each row the previous row to form the next one.

Interestingly, each term in the rabbit sequence r_n can be constructed directly without references to prior terms or subsequences. For this purpose, the index n has to be decomposed according to the Fibonacci number system (see Sect. 5.5):

$$n = F_{k_1} + F_{k_2} + \ldots + F_{k_r}. \tag{32.3}$$

The parity of the last index, k_r, determines r_n:

$$r_n = \langle k_r + 1 \rangle_2. \tag{32.4}$$

So, as in the application to the game of Nim (Sect. 5.5), it is the *least* significant Fibonacci number that is important.

By construction, this sequence is aperiodic although arbitrarily long subsequences will appear infinitely often in the sequence. This is of course true even for random sequences. The point is that in the rabbit sequence identical subsequences will be very close by. In fact, for $n > 3$, a given subsequence whose length does not exceed F_n is never farther than F_{n+1} steps away.

But is the rabbit sequence self-similar, and will it have a "peaky" spectrum like that of a periodic sequence? A peaky spectrum is needed in order to model the newly discovered state of matter.

Of course, the self-similarity of the rabbit sequence cannot be as simple as that of the Morse-Thue sequence, where we could strike out every other bit to leave the sequence invariant. The lengths of the rabbit subsequences (32.2) from one row to the next increase in the ratio of successive Fibonacci numbers, tending to the Golden ratio $g = 1.618\ldots$.

Interestingly, the ratio of 1's to 0's (mature rabbits to young rabbits) is also the ratio of successive Fibonacci numbers. Let us therefore try striking out (rab)bits depending on the bit itself: For a 1, let us hop two places before we strike out, and for a 0 let us hop only 1 place. The result is the following:

$$1\;0\;\cancel{1}\;1\;\cancel{0}\;1\;0\;\cancel{1}$$

i. e., the beginning of the row: 10110. For the infinite sequence our striking rule would simply reproduce the sequence itself: *self-similarity*! In fact, this "self-decimating" rule is simply the inverse of the self-generating rabbit creation rule.

Another way to demonstrate the self-similarity of the rabbit sequence is to replace every bit pair 10 by a 1 and every triplet 110 by a 0. This "condensation" also reproduces the rabbits, as does the mapping 101 → 1 and 10 → 0 and infinitely many other mappings with an increasing degree of condensation.

After our experience with the spectrum of the self-similar Morse-Thue sequence, it is therefore small wonder that the rabbit sequence, too, exhibits pronounced spectral peaks, see Fig. 32.3. These peaks are the one-dimensional analogue of the bright spots in the diffraction pattern (Fig. 32.2). The locations of these peaks, moreover, reflect the Golden ratio (many times) as can easily be observed in Fig. 32.3.

For comparison, Fig. 32.4 shows the spectrum of a *random* binary sequence, calculated, like the Morse-Thue and rabbit spectra, by J. L. Hall of AT&T Bell Laboratories. To make the comparison as close as possible, the random sequence was produced by a three-state Markov process modeled on the rabbit sequence. The three states correspond to the three possible bit pairs 11, 10, and 01, see (32.2), with deterministic transitions from 11 to 10 (generating a 0) and from 10 to 01 (generating a 1) and a probabilistic transition from 01 to 11 (with probability $\gamma = (\sqrt{5}-1)/2$,

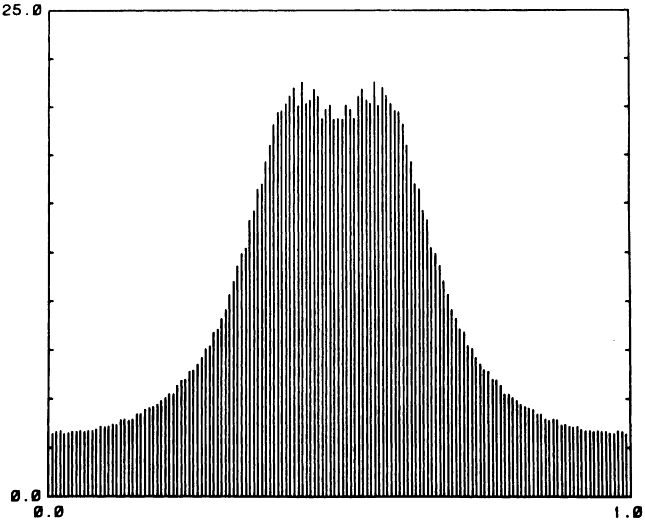


Fig. 32.4 Average Fourier amplitude spectrum of a three-state Markov process having the same transition probabilities as the rabbit sequence (whose spectrum is shown in Fig. 32.3). Note smooth spectral envelope and absence of pronounced single-harmonic spectral peaks, reflecting the absence of long-range order in the Markov process

the reciprocal of the Golden ratio, generating a 1) and a transition from 01 to 10 (with probability γ^2, generating a 0).

The probabilities of the three states are γ^4, γ^3, and γ, respectively, and the probabilities of 1 and 0 are γ and γ^2, respectively. All these probabilities are in conformity with those of the (infinitely long) rabbit sequence. Yet, although the Markov machine mimics the short-range properties of the rabbit sequence, it lacks the *long-range coherence* of the latter, a fact that is reflected in the absence of sharp peaks in its spectrum, see Fig. 32.4.

The existence of pronounced peaks in Fig. 32.3 means that some long-range order or "phase coherence" exists in the rabbit sequence, extending over its entire length. Although aperiodic, the rabbit sequence, when converted to audible sound, should sound tonal! And the diffraction pattern of a physical substance patterned on such self-similar structure would show bright spots, as if it were a periodic crystal! The new form of matter might also exhibit some rather unorthodox and hitherto unsuspected physical properties.

The two highest peaks in the rabbit spectrum (Fig. 32.3) occur at harmonic numbers ("frequencies") F_{n-1} and F_{n-2} for the subsequence of length F_n. Here, the frequency F_{n-1} is simply the mirror image of the frequency F_{n-2}. (Note: $F_{n-1} = F_n - F_{n-2}$). The ratio F_{n-1}/F_n is approximately $\gamma = 1/1.618\ldots$ and the ratio F_{n-2}/F_n tends toward $\gamma^2 = 1/2.618\ldots$. Why do these high peaks, reflecting periodic behaviour, exist and why are they near $1/2.618$ and $1/1.618$ times the "sampling frequency" F_n?

Answer: If we "sample" the rabbit sequence with an average sampling interval of 1.618 samples, we get the same symbol again and again *ad infinitum*! Indeed, the sampled sequences r_k, where k is the largest integer not exceeding n/γ becomes

$$r_k = 1\ 1\ 1\ 1\ 1\ 1\ 1\ \ldots, \qquad k = \left\lfloor \frac{n}{\gamma} \right\rfloor, \tag{32.5}$$

a fact that follows from the above construction rules. Similarly, sampling with an average interval size of 2.618 yields

$$r_k = 0\ 0\ 0\ 0\ \ldots, \qquad k = \left\lfloor \frac{n}{\gamma^2} \right\rfloor. \tag{32.6}$$

The rabbit sequence is also related to the number-theoretic *Beatty sequences* [32.1], defined as follows: if $w < 1$ is a positive irrational, then the two sequences

$$a_m = \left\lfloor \frac{m}{w} \right\rfloor, \qquad m = 1,2,3,\ldots \tag{32.7}$$

$$b_m = \left\lfloor \frac{m}{1-w} \right\rfloor, \qquad m = 1,2,3,\ldots, \tag{32.8}$$

together assume every positive integer value exactly once.

This is easy to see, because for the integer k to be a member of the first sequence, a_m, the following inequalities must obtain for some m (with w irrational, equalities need not be considered):

$$\begin{aligned} &\frac{m}{w} - 1 < k = \left\lfloor \frac{m}{w} \right\rfloor < \frac{m}{w}, \quad \text{or} \\ &m - w < kw < m. \end{aligned} \tag{32.9}$$

On the other hand, for k to be a member of the *second* sequence, b_m, the inequalities are the same as (32.9) with $1-w$ replacing w. Substituting the integer $k-m$ for m, one obtains as the condition for membership in the second sequence that there be some m for which

$$m < kw < m + 1 - w. \tag{32.10}$$

Comparing conditions (32.10) and (32.9), one sees that they are precisely complementary: a given integer k fulfills either (32.9) or (32.10) and is therefore a member of either a_m, or b_m. The "densities" of these two sequences are w and $1-w$, respectively, which add up to 1, as they should if no integer is duplicated or left out.

For a "winding number" (the physicist's term) $w = \gamma = 1/g$, i. e., $1/w = g$ and $1/(1-w) = g^2$, the first sequence becomes

$$\begin{aligned} a_m &= 1,3,4,6,8,9,11,\ldots \quad \text{and the second sequence} \\ b_m &= 2,5,7,10,13,15,18,\ldots. \end{aligned}$$

Now a rabbit sequence term r_k equals 1 if k is a member of the Beatty sequence a_m; otherwise (i. e., if k is a member of b_m) it equals 0. – Who can show the equivalence between this rule and (32.3, 4)? Hint: use (5.30).

With this kind of long-range order it is small wonder that there are spectral peaks near harmonic numbers $\gamma^2 F_n$ and γF_n. The other peaks visible in the spectrum (Fig. 32.3) result from the self-similarity of the rabbit sequence (scaling factor equal to the Golden ratio 1.618...) and the spectral folding ("aliasing") inherent in the Fourier transform of a discrete sequence of finite length.

Beyond the Golden ratio, self-similar sequences and new quasicrystals are generated by the "Silver ratios", $\tau_N^{\pm}$, defined by $\tau_N^{\pm} = N \pm 1/\tau_N^{\pm}$, i. e., all those (quadratic) irrationals that can be expressed by periodic continued fractions with period length one and ± 1 as the numerator. For example, $\tau_2^{+} = 1+\sqrt{2}$ is expected to generate a quasicrystal with the crystallographically forbidden 8-fold symmetry, while $\tau_4^{-} = 2+\sqrt{3}$ would underlie a likewise forbidden 12-fold symmetry axis. In addition to the Golden ratio $g = \tau_1^{+}$, all $\tau_N^{\pm}$, where N is the nth Lucas number (see Sect. 5.8) and the sign superscript equals $(-1)^n$, generate quasicrystals with a 5-fold symmetry!

In another generalization, the *noble numbers*, recently defined as all those numbers whose continued-fraction expansions end in infinitely many 1's, distinguish themselves both in the present case and in the approach to chaos of nonlinear dynamical systems. In this nomenclature the Golden ratio is but the noblest of the noble numbers.

Cassini's divisions in the rings of Saturn are a manifestation of what happens when, instead of noble numbers, simple rationals reign: rocks and ice particles constituting the rings, whose orbital periods are in simple rational relation with the periods of the moons of Saturn, are simply swept out of their paths by the resonance effects between commensurate orbital periods. In fact, the very stability of the solar system depends on the nobility of at least some of the orbital period ratios.

The noble numbers ν, defined by the continued fraction

$$\nu := \left[a_0; a_1, \ldots, a_n, \overline{1}\right], \tag{32.11}$$

can be expressed with the help of the Golden ratio g as follows:

$$\nu = \frac{gA_n + A_{n-1}}{gB_n + B_{n-1}}, \tag{32.12}$$

where A_k and B_k are the numerator and denominator, respectively, of the kth approximating fractions of $[a_0; a_1, \ldots, a_n]$, see (5.3.4). What is the CF of the noble $(81g+334)/(73g+301)$?

A two-dimensional analogue of the new state of matter is the Penrose tiling [32.3] of the plane, see Fig. 32.5. Like the one-dimensional rabbit sequence, it consists of two elements ("arrows" and "kites") that cover the plane without periodic repetition. If we would replace the arrows and kites with 0's and 1's at their centres of gravity, the Fourier transforms of the resulting two-dimensions array would show strong peaks and exhibit the crystallographically forbidden five-fold symmetry apparent in Fig. 32.2. This then is the gist of the paradoxical diffraction pattern shown in Fig. 32.2. For a complete wrap-up of this matter, the three-dimensional case was analyzed by *D. Levine* and *P. J. Steinhardt* [32.4].

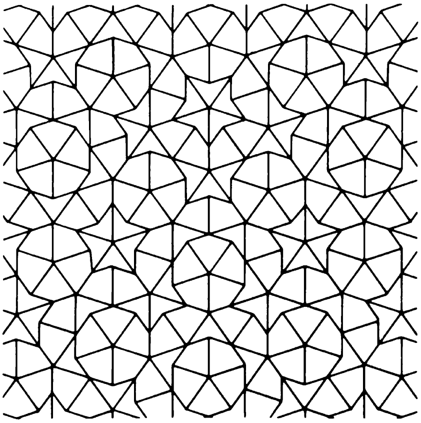


Fig. 32.5 Penrose tiling of the plane with "arrows" and "kites". The five-fold symmetry, forced by the angles of the arrows and kites and their edge-length proportions (equal to the Golden ratio $g = 1.618\ldots$), makes a periodic tiling impossible [32.3]. But there is still a long-range order that is the two-dimensional analog of the recently discovered new state of matter, see [32.2, 4] and Fig. 32.2

32.2 Cantor Sets, Fractals and a Musical Paradox

In the midst of the animated debates on the foundations of mathematics – and the very meaning of the concept *number* – during the last century, Georg Cantor (1845–1918), of set fame, wanted to present his colleagues with a set of numbers between 0 and 1 that has measure zero (i. e., a randomly thrown "dart" would be very unlikely to hit a member) and, at the same time, has *so* many members that it is in fact uncountable (just like *all* the real numbers between 0 and 1).

Many mathematicians, and even Cantor himself for awhile, doubted that such a "crazy" set existed – but it does, and its construction is in fact quite straightforward. Imagine the real line between 0 and 1 (drawn with chalk on a blackboard if you will) and wipe out the middle third. Then wipe out each middle third of each remaining third, and so forth *ad infinitum*. The first seven wipings are illustrated in Fig. 32.6, but there is no way to draw the final result, aptly called *Cantor dust* by *B. Mandelbrot* [32.5].

However, there is a neat number-theoretic way to represent the "dust", namely by *ternary* fractions employing the digits 0, 1 and 2. For example, 0.5 corresponds to $0.111\ldots$ in ternary notation. The ternary representation is unique if we forbid periodic fractions ending in a 1 or $1\overline{2}$ and, for example, write $0.0222\ldots$ instead of 0.1 and 0.2 instead of $0.1\overline{2}$.

Now, the numbers in the interval $[1/3, 2/3]$ are precisely all those numbers that in the ternary system have the digit 1 in the first position to the right of the period. Wiping these numbers away on our route to the Cantor dust, we are left with numbers that begin with 0.0 or 0.2.

Similarly, the second wiping (third line in Fig. 32.6) eliminates all numbers with a 1 in the *second* place to the right of the period. In the end, having arrived at the Cantor dust, we are left with all those numbers that have no 1's in *any* place.

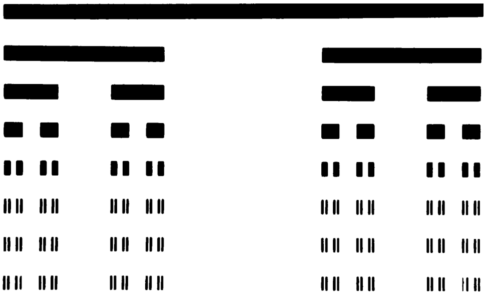


Fig. 32.6 The first steps to the Cantor "dust", the original Cantor set obtained by repeatedly "wiping out" middle thirds of remaining intervals. The resulting set, which can be represented by ternary fractions using only the two digits 0 and 2, has only isolated points and zero measure. Yet its cardinality is as uncountable as that of all the real numbers ("aleph one"). Cantor sets, which are characterized by a (fractal) Hausdorff dimension, play a prominent role in modern nonlinear physics where they describe "strange attractors", i.e., aperiodic orbits in phase space that attract other points from their "basin of attraction". The boundaries of such basins, called Julia sets, are also often fractal (nondifferentiable). The Hausdorff dimension of the original Cantor set equals $d = \log 2/\log 3 = 0.631\ldots$ [32.5]

Incidentally, the members of the Cantor set form a *self-similar* set: take any line in Fig. 32.6, leave out the left or right half and magnify the remainder three-fold. This results in the line immediately above it. More precisely, the Cantor set is invariant, modulo 1, to scaling by a factor 3. In fact, in the ternary notation, this scaling is nothing but a left-shift of all digits and dropping any 2's that protrude to the left of the period. For example, the Cantor number 0.202202 maps into 0.02202, another Cantor number.

Now it is easy to see why the Cantor set has measure zero: the probability that a random digit in $[0,1)$ has not a single 1 in its ternary expansion is of course zero. More precisely, the expression for the probability that there is no 1 in n ternary places equals $(2/3)^n$, which goes to zero when n goes to infinity.

But why are the members of this extremely "thin" set so numerous that they are not even countable (like the integers or rationals)? The reason is that we can bring the members of the Cantor set in a one-to-one correspondence with *all* the real numbers in the interval between 0 and 1. We simply identify with each Cantor number the *binary* number obtained by changing all 2's to 1's. Thus, for example, 0.020222 corresponds to 0.010111 $(= 23/64)$. In this manner, each member of the Cantor set can be mapped into a real number and, conversely, *all* reals between 0 and 1 can be mapped into Cantor numbers, which have thus the same "cardinality" as the reals.

The well-known fact that the reals form an *uncountable* set was proved by Cantor using the "diagonal method" already known to Galileo. The diagonal method is used in an indirect proof that proceeds as follows.

Assume that all the numbers between 0 and 1 form a countable set; they could then be written down, one after another, in the counting sequence. In binary notation, for example, the list might look as follows (outlawing terminating fractions):

$$\begin{array}{l} 0.01011\ldots \\ 0.11110\ldots \\ 0.01011\ldots \\ 0.10001\ldots \\ \text{etc.} \end{array}$$

Now write down a number whose first digit to the right of the period is the complement of the corresponding digit in the first number in the above list, and whose second digit is the complement of the second digit in the second number, and so forth. The resulting number, which begins

$$0.1011$$

cannot be found anywhere in the above list because it differs from each number in at least one place. Thus, the list cannot be complete and our assumption that the reals form a countable set was false.

But how can we quantify by some *finite* number a point set that has measure zero and cardinality "uncountable"? This we often need to do in physical applications, where Cantor sets may characterize the members of "strange attractors" in nonlinear dynamical systems that revel in deterministic chaos (see Sect. 32.6).

Here the concept of the *Hausdorff dimension*, named after the German mathematician F. Hausdorff (1868–1942), comes to our rescue. For a self-similar set, such as the Cantor set, the Hausdorff dimension d is particularly simple to calculate. We simply look at that part of the set that falls within a "length" one rth the original set, where r is the scaling factor of the self-similarity, and ask what fraction f of the set falls into that portion. Then the Hausdorff dimension is defined as

$$d := \frac{\log(1/f)}{\log r}, \tag{32.13}$$

often a *noninteger* value.

For a straight line f equals of course $1/r$ and $d = 1$, as one would expect. For the Cantor set $f \neq 1/r$; in fact, $r = 3$ and $f = 1/2$. Thus, $d = 0.6309\ldots$. The Cantor set having a $d < 1$ reflects its "thinness", its dustlike consistency, compared to a piece of coherent line.

If we calculate the Hausdorff dimension of another self-similar set, the Hilbert curve (Fig. 32.7), then we find for $r = 2$, $f = 1/4$. Thus, $d = \log 4/\log 2 = 2$, which is nice considering the fact that (in the limit) the Hilbert curve goes through every point in the unit square. (But note the monstrous result that a *curve*, something we have always thought of as having just *one* dimension, has now dimension $d = 2$.)

The Hausdorff dimension even allows us to salvage a basic Euclidean theorem that otherwise would go to pieces when applied to fractal objects. Look at Fig. 32.8: it shows a hexagon surrounded by 6 identical hexagons. The total area of the figure is 7 times the area of a hexagon, and its perimeter is 3 times that of a hexagon. Note that the total figure is *not* a hexagon, i. e., the figure is not similar to its parts.

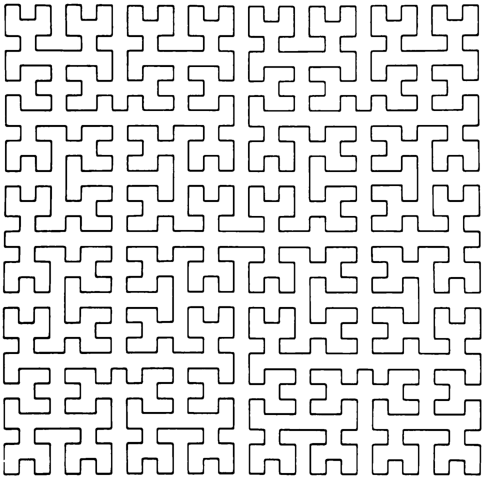


Fig. 32.7 Intermediate stage, on the way to the plane-filling Hilbert curve. The Hausdorff dimension of the Hilbert curve equals $d = \log 4/\log 2 = 2$, as behooves a curve that, in the limit, touches every point in the unit square. The construction of the Hilbert curve accomplishes a *continuous* mapping of the (two-dimensional) unit square onto a (one-dimensional) line, in contrast to the mapping by alternating binary digits, which is discontinuous

But we can make it so! The firt step is to break up each straight line in Fig. 32.8 to be similar to one third of the perimeter of the entire figure. For this purpose we break up each straight line by a zig-zag of three straight lines, each having $1/\sqrt{7}$ length, see Fig. 32.8. By repeating this breaking up process infinitely often, we arrive at a figure, approximated by Fig. 32.9, that is exactly similar to its 7 parts. The area ratio is obviously 7 to 1 and the perimeter is 3 times the perimeter of one of the 7 similar pieces.

Now, following Euclid, areas of similar figures are proportional to their linear dimensions squared. But 3^2 equals 9 and not 7. Where did we go astray? Has Euclid finally been proven wrong??

Well, the old Greeks (with the possible exception of Zeno) can continue to rest in peace. Geometric objects like the one illustrated (imperfectly) in Fig. 32.9 were

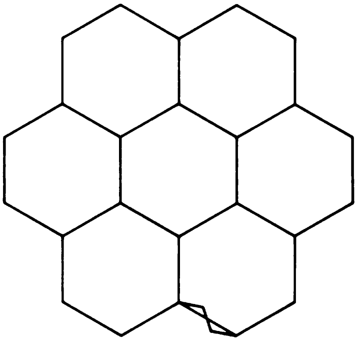


Fig. 32.8 Seven hexagons. Hexagons tile the plane, but they can never form a hexagon again (as opposed to tiling by squares, for example). This tiling defect can be remedied by breaking up each of the 30 straight lines (of unit length) in the figure into 3 straight-line segments, each of length $1/\sqrt{7}$ and repeating the process *ad infinitum*. The result of the first breaking up is illustrated on one edge

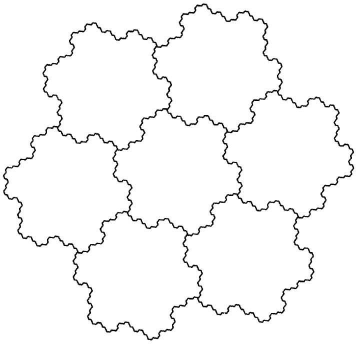


Fig. 32.9 The seven hexagons after the first five edge-mutilating steps. The final figure is similar to its seven parts and has a nondifferentiable fractal boundary (Hausdorff dimension $d = \log 3/\log\sqrt{7} = 1.129\ldots$). The perimeter of the large "hexagon" is 3 times as long as that of a small "hexagon". Yet its area is not $3^2 = 9$ but only 7 times as large: a "non-Euclidean" paradox! (See text.)

never on exhibit in Euclid's school. Its boundaries, while everywhere continuous, are nowhere differentiable. The object is a *fractal*, so named by *Mandelbrot* [32.5], with an easily calculable Hausdorff dimension. Since the scaling factor equals $\sqrt{7}$ by construction and f equals $1/3$, we obtain with (32.13) for the dimension the following noninteger value:

$$d = 1.12915\ldots$$

Although noninteger, the appellation "dimension" for d is quite apt, because with it we can reformulate Euclid's theorem about similar areas and obtain a more generally valid result, applicable to fractals and nonfractals alike:

Theorem: *FOR SIMILAR FIGURES, THE RATIOS OF CORRESPONDING MEASURES ARE EQUAL WHEN REDUCED TO THE SAME DIMENSION.*

Thus, the area ratio in Fig. 32.9 is 7. Since area has dimension 2, reduction to dimension 1 results in $7^{1/2} = 2.64575\ldots$. and the perimeter ratio 3 when reduced to dimension 1 gives the same number. To wit: $3^{1/d} = 3^{1/1.12915} = 2.64575\ldots$. It is because of properties like this that the Hausdorff dimension is such a useful concept. It is just an extension of the concept of dimension to fractal values, which model many phenomena in the real world.

In Fig. 32.9 we have encountered a function, the boundary of the figure, that is everywhere continuous and nowhere differentiable. Such exotic functions were defined by Karl Weierstrass (1815–1897) to show an unbelieving world one of the "impossible" things that could happen with what is innocently called "function". And like Cantor sets, Weierstrass functions are a rich mine of common-sense paradoxes.

For example, a musical chord patterned after a Weierstrass function can have the following weird property. If recorded on magnetic tape and replayed at *twice* the recording speed, the chord will not sound an octave higher in pitch (as every well-behaved recorded sound would) but a semitone *lower*. How is *this* possible?

Let us construct a simple Weierstrass function

$$w(t) = \sum_{k=-\infty}^{\infty} \cos\left(\beta^k t\right). \tag{32.14}$$

If we scale the time dimension t by a factor β, we obtain

$$w(\beta t) = \sum_{k=-\infty}^{\infty} \cos\left(\beta^{k+1} t\right) = \sum_{k=-\infty}^{\infty} \cos\left(\beta^k t\right), \tag{32.15}$$

i. e., $w(\beta t)$ *equals* the unscaled function $w(t)$. Thus, $w(t)$ is self-similar. Obviously, such a function cannot have a nonzero derivative anywhere, because derivatives change with scaling. (It is interesting to note that the concept of self-similarity entered mathematics at two independent points, Cantor sets and Weierstrass functions, at about the same time in history and for similar reasons: to elucidate the foundations of mathematics: numbers and functions.)

Now suppose we select $\beta = 2^{13/12}$ and convert $w(t)$ to audible sound:

$$w(t) = \sum_k \cos\left(2^{k13/12} t\right), \tag{32.16}$$

where k only has to cover the audio range (20 Hz to 20 000 Hz). Playing back $w(t)$ as constructed in (32.16) at twice the speed produces

$$w(2t) = \sum_k \cos\left(2^{k13/12+1} t\right) = \sum_{k'} \cos\left(2^{k'13/12} 2^{-1/12} t\right), \tag{32.17}$$

where $k' = k+1$. Now, if the summations in (32.17) cover the entire audio range, then, as far as the human ear is concerned,

$$w(2t) = w\left(2^{-1/12} t\right). \tag{32.18}$$

Thus, a doubling of the tape speed will produce a sound with a pitch lowered by a factor $2^{-1/12}$, the chord will sound one semitone *lower* rather than an octave higher! This paradox was originally discovered by Jean-Claude Risset, see [29.5].

P. Meyer at the Drittes Physikalisches Institut, University of Göttingen, programmed a desk-top computer to produce a $w(t)$ as in (32.16) with 11 components comprising the frequencies from 10.0 Hz to 18245.6 Hz. By doubling the playback speed the 6th component, for example, will change in frequency from 427.15 Hz to 854.3 Hz. But in comparing the two chords, the human auditory system will identify the doubled 6th component at 854.3 Hz with the *nearest* component of the original chord, namely the 7th at 905.1 Hz. Since 854.3 Hz is a semitone lower than 905.1 Hz, a lowered pitch will be perceived. An astounding paradox, when listened to, but quite easily accounted for.

32.3 The Twin Dragon: A Fractal from a Complex Number System

One of the most widely used number systems in the digital world is the binary system, based on the base 2. In this system, a real number is represented by a sum of powers of 2:

$$x = \sum_k \varepsilon_k 2^k, \tag{32.19}$$

where the binary digits ε_k are either 0 or 1; other values are not needed.

To extend the binary system to *complex* numbers z we could represent both real and imaginary parts by sums like (32.19). But there is a more elegant method: we select $1-\mathrm{i}$ as a base and write

$$z = \sum_k \delta_k (1-\mathrm{i})^k. \tag{32.20}$$

What values for the digits δ_k do we need to cover all z in a given region of the complex plane? Surprisingly, values of δ_k that are either 0 or 1 suffice, just as in the binary system (32.19) for reals. Thus, (32.20) is a *binary* number system for complex numbers!

What area of the complex plane is occupied by the proper fractions, i. e., all those z for which $\delta_k = 0$ for $k \geq 0$? We should not have asked, because the answer is monstrous: the region shown in Fig. 32.10. It has a fractal boundary with surfeit

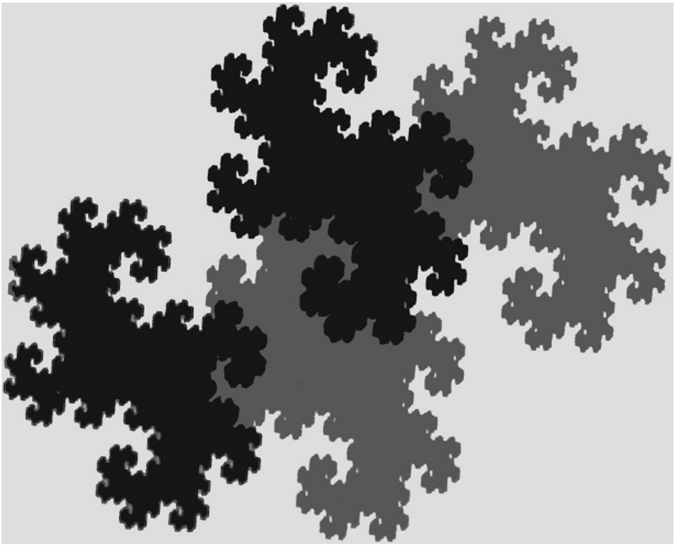


Fig. 32.10 A number-theoretic fractal monster: Donald Knuth's "Twin Dragon" [29.4, 32.5], defined as the loci in the complex plane of all proper fractions in the binary number system with the base $1-\mathrm{i}$. The Twin Dragon tiles the plane, is similar to its parts, and has a fractal perimeter with Hausdorff dimension $d = 1.5236\ldots$, but is otherwise harmless

of self-similarities. It is known as the *Twin Dragon*, and its perimeter has Hausdorff dimensions $d = 1.5236$, a solution of $2^{3d/2} - 2^d = 2$ [32.5].

The Twin Dragon decomposes into similar dragons, indicated by different shadings in Fig. 32.10. Conversely, Twin Dragons tile the plane, but rather perversely, considering their fractal skin.

The physically attractive feature of the Twin Dragon is the fact that *strange attractors* in nonlinear dynamical systems often have "basins of attraction" whose boundary is fractal, much like the dragon skin. Attractors characterize stable orbits. They can have a period of 1, in which case they are called fixpoints, or they can be periodic with periods typically equal to 2, 4, 8, 16, etc., depending on the degree of nonlinearity of the system.

As the nonlinearity is increased, the "bifurcation cascade", i. e., the splitting of one orbit into two, continues until the "period" of the attractor is infinite, in which case it is no longer periodic. Initial conditions within the basin of attraction are attracted to a "thin" set of points, a Cantor set (see Sect. 32.6).

32.4 Statistical Fractals

The fractal objects we have encountered so far are strictly self-similar in a *deterministic* way. But the concept of self-similarity can also apply to statistical laws, resulting in statistical fractals. In fact, Brownian motion, misinterpreted for so long, is a statistical fractal *par excellence* (and so are stock market prices). As is well known, the mean-square displacement $\overline{r^2}$ of a particle suspended in a warm liquid at rest is proportional to time t:

$$\overline{r^2} = 2Dt,$$

where D is the "diffusion constant". If we sample the motion every t seconds and connect the sample values of $r(t)$ by straight lines, we obtain a rather jagged picture. But this picture has almost nothing to do with the actual motion of the particle, which changes direction upon every impact of the liquid's molecules, perhaps every billionth of a billionth of a second. Indeed, if we sample the path of the Brownian particle 100 times as frequently and magnify the result by 10 diameters, we obtain (statistically) the same picture as before.

This process of more frequent sampling and greater magnification can be repeated many times; it will always give a picture that is statistically indistinguishable from the original. And this is precisely what is meant by *statistical self-similarity*. For the Brownian motion, it spans the range from the mean-free path of the particle to the dimensions of the container – a great many orders of magnitude. The Hausdorff dimension equals, by definition (32.13), $d = 2$. In other words, like the Hilbert curve, see Fig. 32.7, the Brownian motion curve has the same dimension as a two-dimensional surface (which is why a chemical reaction on a surface, governed by a two-dimensional random walk, proceeds much faster than the same reaction in three dimensions – something nature discovered during evolution and still employs to advantage in enzymatic reactions).

For those readers who are disappointed that the dimension of the Brownian motion is an integer $(d = 2)$, I will mention the fact that the set of its "zero-crossings" has fractal dimension $d = 1/2$. This "thinness" of zeros is a well-known fact for random walks ("Wiener processes"), but it is hardly ever mentioned in terms of fractal dimensions and Cantor-like sets. (A numerical model for a Cantor set with $d = 1/2$ is given by the quarternary fractions in which only two nonadjacent digits, 1 and 3 say, are allowed.)

It is interesting to recall that correct analysis of Brownian motion established the existence of atoms and molecules of finite size, the diffusion constant D being critically dependent on *atomic* dimensions. And who laid the foundation stone? One Einstein (in 1905).

Figure 32.11 shows a statistical fractal obtained by computer simulation of a diffusion-limited aggregation process [32.6]. Each dot represents an atom or molecule that attached itself to the existing structure following simple probabilistic rules that depend on occupancies of nearby sites. Such statistical fractals, as we saw in the case of Brownian motion, can be characterized by a fractal Hausdorff dimension. This dimension is often determined numerically by counting the number N of atoms inside spheres of different radii r. In two dimensions, for a nonfractal substance, N is proportioned to r^2. But in Fig. 32.11 proportionality of N with $r^{1.7}$ was found. Thus, the fractal has Hausdorff dimension $d = 1.7$.

Electrical discharges on insulators obey laws very similar to diffusion-limited aggregation, and the resulting spark patterns (called Lichtenberg figures, after the 18th-century Göttingen physicist and aphorist who studied them) look indeed very much like Fig. 32.11.

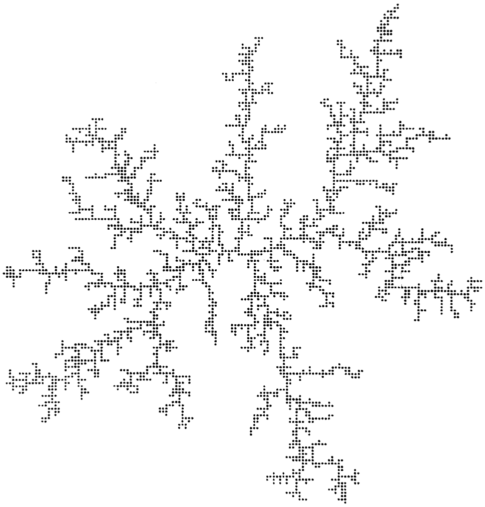


Fig. 32.11 Result of a computer simulation of a two-dimensional diffusion-limited aggregation process, see *Witten* and *Sander* [32.6], a statistical fractal with Hausdorff dimension $d \cong 1.7$, i. e., the number of molecules inside a circle of radius r grows with $r^{1.7}$ (instead of r^2). Two-dimensional electric discharges ("Lichtenberg figures") have similar structures and Hausdorff dimensions. Three-dimensional lightnings (worth looking at) have Hausdorff dimension $d = 2.4$ or greater

Simulations of such scale-invariant processes in three dimensions give fractal Hausdorff dimensions between 2 and 3. The most beautiful (highly branching) lightnings, for example, have Hausdorff dimensions near 2.4. (Have insurance companies become aware and taken note of this?)

Another important application of fractal dimensionality lies in the analysis of time series of a wide spectrum of natural phenomena. Are long-time temperature variations [32.7] on the earth or other aspects of climate evolution describable as fractal phenomena? Can we analyze brain waves (electro-encephalograms, EEGs) and deduce insights into the state of mind of a patient [32.8] from the Hausdorff dimension of the EEG? Perhaps such fractal evaluation will exceed in diagnostic value even linear predictive analysis [32.9].

32.5 Some Crazy Mappings

A long time ago, Isaac Newton suggested finding the zeros of a function $f(z)$ by an iteration based on tangent extrapolations. Given an approximate value of z_n to the solution of $f(z) = 0$, Newton finds the next approximation by calculating

$$z_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}, \tag{32.21}$$

where $f'(z_n) \neq 0$ is the derivative of $f(z)$.

For $f(z) = z^2 - 1$, (32.21) reads

$$z_{n+1} = \frac{1}{2}\left(z_n + \frac{1}{z_n}\right). \tag{32.22}$$

Not surprisingly, for an initial value z_0, with a positive real part, z_n converges toward the positive root of $z^2 - 1 = 0$, namely $z = 1$. Similarly, for a negative real part of z_0, the solution converges to the negative root, $z = -1$.

But what happens for purely imaginary $z_0 = \mathrm{i}r_0$, where $r_0 \neq 0$ is real? Interestingly, it does not converge at all; the iteration (32.22) cannot make up its mind, so to speak, and hops all over the imaginary axis, according to the mapping

$$r_{n+1} = \frac{1}{2}\left(r_n - \frac{1}{r_n}\right). \tag{32.23}$$

For example, $g = 1.618\ldots$ maps into 0.5, -0.5, $0.291\overline{6}$, $-1.56845\ldots$, etc. But some r_0 behave differently, such as $r_0 = 1 + \sqrt{2}$, which maps into 1, 0, ∞, a kind of fixed point, if rather distant.

How can we bring order into this chaotic mapping? A trigonometric substitution will do:

$$r = -\mathrm{cotan}(\alpha\pi), \tag{32.24}$$

which turns (32.23) into the simple

$$\alpha_{n+1} = 2\alpha_n \mod 1. \tag{32.25}$$

Now everything becomes transparent. If we express α_n as a binary fraction, then α_{n+1} will be simply the fractional part of α_n shifted one place to the left. Thus, periodic binary α_0 will lead to periodic orbits. For example, $\alpha_0 = 1/3 = 0.\overline{01}$ will map into $0.\overline{10} = 2/3$, which will map right back into $\alpha_2 = 0.\overline{01} = \alpha_0$. Indeed, $r_0 = -\cot\text{an}(\pi/3) = -1/\sqrt{3}$ maps into $r_1 = \sqrt{3}$, which maps back to $r_2 = -1/\sqrt{3} = r_0$.

We also see by this analysis that any irrational α_0 will lead to an aperiodic "orbit" along the imaginary axis in the z-plane. This leaves us with the nonperiodic rationals. Where will they end up? Continued left-shifting and taking fractional parts will, sooner or later, produce 0 which maps into ∞. For example, $\alpha_0 = 7/8 = 0.111$ will map into $0.11 = 3/4$, then $0.1 = 1/2$, which maps into 0. In fact, the corresponding $r_0 = -\cot\text{an}(7\pi/8) = 1 + \sqrt{2}$ maps into 1, 0, ∞, as we already saw.

Thus, the simple mapping given by Newton's iteration for the function $x^2 - 1$ has rather strange consequences for initial values on the imaginary axis: numbers are classified into

aperiodic binary rationals,
periodic binary rationals,

and

irrationals.

While aperiodic binary rationals "converge" on a fixed point (∞), periodic binary rationals lead to periodic orbits. By contrast, irrational numbers α_0, an uncountable set, give aperiodic orbits: the same value never occurs twice, nor is α_n ever rational.

Thus, we have the curious fact that the tiniest distinction – the difference between rational and irrational numbers – makes a decisive difference in the final fate of a situation. Ordinarily, one should think that physics, and certainly the *tangible* world at large, are untouched by the purely mathematical dichotomy between rationals and irrationals. But in reality this just is not so. While, true enough, everything in the real world can be adequately described by rational numbers, it so happens that a *modeling* that distinguishes between different kinds of numbers is not only extremely useful but catches the true (perhaps hidden) spirit of a physical phenomenon.

More specifically, two different initial conditions that are completely indistinguishable by any finite measurement precision will sooner or later lead to a total divergence as the system evolves in time or space. The essential condition for this to happen is that the corresponding Poincaré map (see Sect. 32.6) be sufficiently *nonmonotonic*, such as (32.25), which has a sawtooth-like nonmonotonicity. The rate of divergence is measured by the so-called *Lyapunov exponent* λ.

In our bare-bones example (32.25), $\lambda = \log(\alpha_{n+1}/\alpha_n)$ and taking natural logarithms gives $\lambda = \log 2 \approx 0.693$.

Newton's method applied to $f(t) = z^3 - 1$ gives

$$z_{n+1} = \frac{2z_n^3 + 1}{3z_n^2}. \tag{32.26}$$

Now the mapping is even crazier. For one thing, the conjecture that all z_0 will converge toward the closest of the three roots is false. For example, $z_0 = -1$ will converge on 1, the root most distant from it.

Figure 32.12 shows the intimately intertwined "basins of attraction" of the three roots – a real, or rather imaginary, crazy quilt. In fact, it can be shown that for the mapping (32.26) two basins (colors) can never meet without the third one being present also. This may sound impossible, and in fact it would be – if it were not for fractal boundaries as intimated in Fig. 32.12 and already encountered in the dragon skin (Fig. 32.10).

One way to begin to understand the basins of attraction in Fig. 32.12 is to consider the neighbourhood of $z = 0$. By (32.26) the point $z = 0$ maps of course to ∞. But points near 0 participate in a devil's dance before approaching one of the three attractors (or being mapped into 0). In fact, for $|z_0| \ll 1$, $z_1 \cong 1/3z_0^2$, a large number. The next iterations make z smaller again: $z_2 \cong 2/9z_0^2$, etc.; it is almost unpredictable what z_n will do as n goes to infinity.

Other "dangerous" points are the three points $z_0 = (-1/2)^{1/3}$, which land on 0. Since the mapping (32.26) is continuous, the devil's dance around 0 will be mimicked around these three points, as can be seen in Fig. 32.12. In fact, their 9 "preimages", one of which is real ($z_{-1} = -1.1388\ldots$), show again the same devilish behaviour. And the same dance continues for *pre*preimages and so forth *ad infinitum*. Several "generations" of these prepreimages, which form a fractal *Julia set*, are readily visible in Fig. 32.12. The orbits of any points near them become totally irrational (in the common sense of the world). And since all three basins of attraction meet at 0, they also meet at all the preimages – which is everywhere where different basins meet: two basins never meet without the third.

More precisely, the Julia set that forms the boundary between the different basins is the set of all those z that map eventually into 0; it corresponds to the nonperiodic rationals for the map $f(z) = z^2 - 1$ discussed above. And any point z_0 near a member of the Julia set will be mapped in one or several steps into a point near zero, where the *real* dance begins.

The Julia set is self-similar, as is apparent in Fig. 32.12 from the similarity of the neighbourhood of 0 with its preimages. The Julia set is also uncountable and has a fractal Hausdorff dimension between 1 and 2, which in the plane is pretty thin dust – but the dust is there; in fact it holds Fig. 32.12 together.

32.6 The Logistic Parabola and Strange Attractors

Newton's map is not the only map with pathological consequences. Even such a simple mapping law as the "logistic parabola"

$$x_{n+1} = rx_n(1 - x_n) \tag{32.27}$$

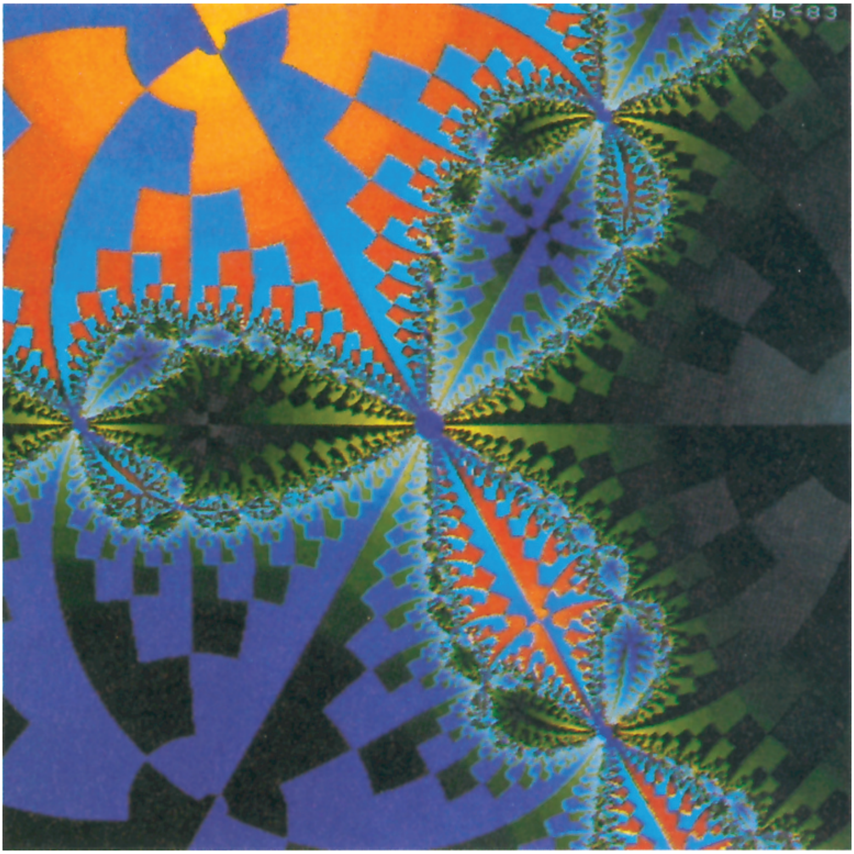


Fig. 32.12 The iteration $z_{n+1} = (2z_n^3 + 1)/3z_n^2$ (Newton's method for finding the roots of $z^3 = 1$). The three different basins of attraction are indicated by 3 different colors. Whenever two colors meet, the third is also present – a seemingly impossible feat. The meeting points form a Cantor-like, fractal Julia set, consisting of the origin $(z = 0)$ and all of its preimages. The different "annular" zones indicate the number of iteration steps necessary to land in one of the (half black) disks surrounding each of the three attractors. (The black portion of each "ring" lands in the black half of the disk.) Each ring has twice as many black parts as its image ring, illustrating the increasingly complex structure of the iteration near the Julia boundary. Because the map is continuous for $z \neq 0$, the figure shows a great deal of near self-similarity around $z = 0$ and its three preimages, $z = (-0.5)^{1/3}$, their nine preimages, and so forth, *ad infinitum*. While periodic repetition was (and is) one of the staples of artistic expression, periodicity on a *logarithmic* scale, i. e., *self-similarity* is now coming into its own – thanks to modern computer graphics which produced this figure entitled *Orchis Tertiarum* by Forschungsgruppe Komplexe Dynamic (H.-O. Peitgen, P. H. Richter), Universität Bremen

show very strange behaviour indeed. Equation (32.27) models the growth (or death) of a population. Without the factor $(1 - x_n)$ and for $r > 1$, the population would grow geometrically. But if resources (think of food) are limited, so is the growth, as modeled by the factor $(1 - x_n)$.

For $1 < r < 3$, (32.27) has a unique nontrivial limit x^*. Any $x_0 \neq 0$ or 1 will be attracted to x^*. For example, for $r = 2$ the attractor x^* equals 0.5. With $x_0 = 0.8$ (32.27) will yield 0.32, 0.435, 0.492, 0.4998 ... which converges on the attractor $x^* = 0.5$.

But for $r > 3$, the attractor is no longer stable. For example, with $r = 3.2$ and $x_0 = 0.8$, (32.27) gives 0.512, 0.7995, 0.51288,...,0.799455, 0.513044. Now there is no longer a single-valued attractor; x^* has *bifurcated* into an orbit of period 2, also called a periodic attractor characterized by two values of x: 0.513... and 0.799.... Other nontrival values of x_0 are attracted to this pair of values.

If we keep increasing the "nonlinearity parameter" r the attractor will bifurcate again and have a period of 4. Further increases of r will lead to periods 8, 16, 32, etc. until, at a critical value r_∞, the period has become infinite. In other words, the orbit is now *aperiodic*, comprising a point set of infinitely many values of x that never repeat. However, other values of x are attracted to this point set, which is in fact a Cantor set with Hausdorff dimension d a little over 0.5. Such attractors have been called *strange*, although once one knows about Cantor sets they are really not so strange after all.

In fact, we even encounter the self-similarity again that we first met (Sect. 32.2) when discussing the original Cantor set with $d \cong 0.631$. The reason is that for the logistic parabola (and many other mapping laws) the interated map $f^{(2)}(x) := f(f(x))$, where $f(x) = rx(1-x)$, has again parabolic regions similar to that of $f(x)$ itself. And the same is true for $f^{(4)}(x)$, $f^{(8)}(x)$, etc. In fact, the important regions of $f^{(2^k)}(x)$ become increasingly similar to each other as k increases. This leads to a self-similar scaling law for the parameter values r_k at which bifurcation takes place: the ratios of successive parameter intervals

$$\frac{r_k - r_{k-1}}{r_{k+1} - r_k}$$

converge to a *universal* constant, the famous (and probably transcendental) *Feigenbaum constant* 4.6692... (originally found by S. Grossmann and S. Thomae). This magic number has earned the epithet "universal" because it applies to many different nonlinear laws, independent of the details of the mapping, as long as the mapping has a parabolic maximum [32.10].

The logistic parabola becomes even more exciting if we allow the parameter r in (32.27) to assume complex values. If we now plot in black all those values in the r-plane for which (32.27) shows a simple periodic behaviour and the rest of the r-plane in different shades (depending on the rapidity of divergence in the iteration), we obtain a fractal image with many self-similarities, dominated by a funny looking, rotund little being called "Apfelmännchen" in German, and Mandelbrot set more formally.

Magnifying ever smaller portions of the r-plane, we will discover our model "Apple Manikin" again and again, hiding in every nook and cranny of the complex plane. One such blowup is shown in our last illustration, Fig. 32.13. We shall have to leave the Apfelmännchen and a detailed discussion of the Julia world it

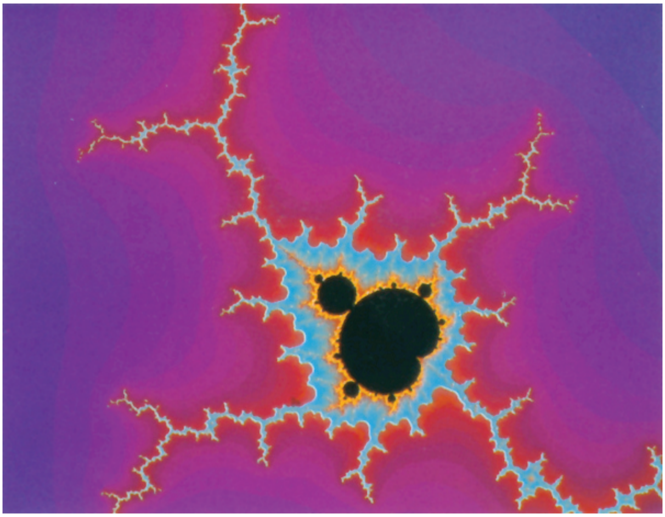


Fig. 32.13 One of the uncountably many *Apfelmännchen* inhabiting the complex plane (and sprouting smaller "Apple Manikins" on its skin). Computer graphic by Forschungsgruppe Komplexe Dynamic (H.-O. Peitgen, P. H. Richter), Universität Bremen

(he?) inhabits for another book, in which the many uses (including artistic) of self-similarity might be the main theme [32.11].

It is interesting to note that the simple logistic parabola models the behaviour of many nonlinear physical systems, including some with rather complex dynamics. The reason is that the often inconceivably complicated orbits in phase space of such systems are adequately described for many purposes by a "cut" through phase space, known as a *Poincaré section* after the famous French mathematician Henri Poincaré (1854–1912). This ingenious trick converts the dynamics into a *mapping* that is much easier to study than the original problem. Especially the chaotic behaviour shown by such complex systems is faithfully modeled by the aperiodic behaviour observable in such simple equations as the logistic parabola.

Just as for Newton's method applied to the solution of $z^2 - 1 = 0$, see Sect. 32.5, the logistic parabola, too, shows aperiodic behaviour that can be understood as left-shifts of binary digits and the taking of fractional parts, as expressed by equation (32.25). In fact, the *deterministic chaos*, which dominates many a physical discussion these days, is ultimately understandable in these simple terms: the evolution of a complex system, in the final analysis, is akin to the left-shifting of digits of a real irrational number. No matter how precisely the initial conditions are known, temporal evolution will sooner or later bring the originally unknowable, far-right, digits (representing the exact initial conditions) into the neighbourhood of the "decimal" point, where they will dominate the system's behaviour, which will then seem chaotic.

This has nothing to do with quantum uncertainty or Planck's constant; unpredictability or "chaos" prevails even for completely deterministic systems, which is why it is called *deterministic* chaos. It reflects the fact that initial conditions, no matter how precisely specified, can never be given with an infinite precision, as represented by a real irrational number with infinitely many nonperiodic digits.

This, in a nutshell, is what modern chaos is all about. (The mythical primeval Chaos has not yet been studied in any great detail and may be quite another matter.)

32.7 Conclusion

We live in a age of increasing overlap between the traditional scientific disciplines. Few are the advances in modern technology – and even basic scientific understanding – that rely on a single field of learning. To give a few examples, our understanding of the physical universe, which itself impinges on fundamental philosophical questions of the beginning and the end of being, depends on progress in astronomy, which in turn results from innovations in computer processing, image enhancement, and a host of other new technologies. Our comprehension of the genetic code is influenced by information theory as well as by physical chemistry, which in turn is based on quantum mechanics, including the exclusion principle. Elementary particles and the even more elementary quarks, not to mention atoms, molecules and crystals, are best described by universal principles of symmetry as embodied in mathematical group theory.

The application of the theory of numbers and finite fields touched upon in this volume – the design of concert halls, precision measurements, error-correcting codes, secret communication, special waveforms and radiation patterns, efficient diffraction gratings, pseudorandom numbers, deterministic chaos and even artistic designs and more playful things – can only provide a flavor of the interdisciplinary endeavors already undertaken, and of those yet to come.

Glossary of Symbols

$a_1 \equiv a_2 (\bmod b)$	a_1 congruent to a_2, modulo b; $a_1 - a_2$ divisible by b
$\mathbb{C}$	the field of complex number
$d(n)$	the number of (positive) divisors of n; $\sigma_o(n)$
$d \mid n$	d divides n; n is a multiple of d; there is an integer q such that $dq = n$
$d \not\mid n$	d does not divide n
e	base of natural logarithms; 2.718281828459045…
exp{ }	exponential function
F_n	Fermat numbers: $2^{2^n} + 1$ Or Fibonacci numbers
$f(x) = 0(g(x))$	$f(x)/g(x) \to 0$ as $x \to \infty$
$f(x) = 0(g(x))$	there is a constant c such that $\lvert f(x) \rvert < cg(x)$
i	square root of -1; $\mathrm{i}^2 = -1$
$\ln x$	natural logarithm of x
(m,n)	GCD (greatest common divisor) of m and n; highest common factor of m and n

$[m,n]$	LCM (least common multiple) of m and n. Also, the block of consecutive integers, $m, m+1, \ldots n$
M_p	Mersenne numbers: $2^p - 1$
$n!$	factorial n; $1 \times 2 \times 3 \times \ldots \times n$
$\binom{n}{k}$	n choose k; the binomial coefficient $n!/k!(n-k)!$
$\left(\frac{p}{q}\right)$ or (p/q)	Legendre symbol, also fraction
$p^a \| n$	p^a divides n, but p^{a+1} does not divide n
p_n	the nth prime, $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$
$\mathbb{Q}$	the field of rational numbers
$r_k(n)$	least number of numbers not exceeding n, which must contain a k-term arithmetic progression
$\lfloor x \rfloor$	Gauss bracket or floor of x; greatest integer not greater than x
$\lceil x \rceil$	ceiling of x; last integer not less than x
$\langle x \rangle_n$	least positive (or nonnegative) remainder of x modulo n
$\mathbb{Z}$	the ring of integers
$\mathbb{Z}_n$	the ring of integers, 0, 1, $2, \ldots, n-1$ (modulo n)
γ	Euler's constant; $0.577215664901532\ldots$
π	ratio of circumference of circle to diameter; $3.141592653589793\ldots$
$\pi(x)$	number of primes not exceeding x, also primitive polynomial

$\pi(x;a,b)$	number of primes not exceeding x and congruent to a, modulo b
$\prod$	product
$\sigma(n)$	sum of divisors of n
$\sigma_k(n)$	sum of kth powers of divisors of n
$\sum$	sum
$\phi(n)$	Euler's totient function; number of positive integers not exceeding n and prime to n
ω	complex cube root of 1, $\omega^3 = 1$, $\omega \neq 1$, $\omega^2 + \omega + 1 = 0$
$\omega(n)$	number of distinct prime factors of n
$\Omega(n)$	number of prime factors of n, counting repetitions

References

Chapter 1

1.1 T. M. Apostol: *Introduction to Analytic Number Theory* (Springer, Berlin, Heidelberg, New York 1976)

1.2 I. Asimov: *Asimov on Numbers* (Doubleday, Garden City, NY, 1977)

1.3 A. O. L. Atkin, B. J. Birch (eds.): *Computers in Number Theory* (Academic, London 1971)

1.4 E. R. Berlekamp, J. H. Conway, R. K. Guy: *Winning Ways* (Academic, London 1981)

1.5 W. Kaufmann-Bühler: *Gauss. A Biographical Study* (Springer, Berlin, Heidelberg, New York 1981)

1.6 P. J. Davis: *The Lore of Large Numbers* (Random House, New York 1961)

1.7 L. E. Dickson: *History of the Theory of Numbers*, Vols. 1–3 (Chelsea, New York 1952)

1.8 U. Dudley: *Elementary Number Theory* (Freeman, San Francisco 1969)

1.9 C. F. Gauss: *Disquisitiones Arithmeticae* [English transl. by A. A. Clarke, Yale University Press, New Haven 1966]

1.10 W. Gellert, H. Küstner, M. Hellwich, H. Kästner (eds.) *The VNR Concise Encyclopedia of Mathematics* (Van Nostrand Reinhold, New York 1977)

1.11 R. K. Guy: *Unsolved Problems in Intuitive Mathematics*, Vol. I, Number Theory (Springer, Berlin, Heidelberg, New York 1981)

1.12 H. Halberstam, C. Hooley (eds.): *Progress in Analytic Number Theory*, Vol. I (Academic, London 1981)

1.13 G. H. Hardy: *A Mathematician's Apology* (Cambridge University Press, Cambridge 1967)

1.14 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 4th ed. (Clarendon, Oxford 1960)

1.15 L. H. Hua: *Introduction to Number Theory* (Springer, Berlin, Heidelberg, New York 1982)

1.16 K.-H. Indlekofer: *Zahlentheorie*, Uni-Taschenbücher 688 (Birkhäuser, Basel 1978)

1.17 K. Ireland, M. Rosen: *A Classical Introduction to Modern Number Theory* (Springer, New York 1990)

1.18 H. Minkowski: *Diophantische Approximationen* (Teubner, Leipzig 1907; reprinted by Physica, Würzburg 1961)

1.19 T. Nagell: *Introduction to Number Theory* (Wiley, New York 1951)

1.20 C. S. Ogilvy: *Tomorrow's Math* (Oxford University Press, Oxford 1962)

1.21 O. Ore: *Number Theory and Its History* (McGraw-Hill, New York 1948)

1.22 H. Rademacher: *Lectures on Elementary Number Theory* (Blaisdell, New York 1964)

1.23 H. Rademacher, O. Toeplitz: *The Equipment of Mathematics* (Princeton University Press, Princeton 1957)

1.24 A. Scholz, B. Schoenberg: *Einführung in die Zahlentheorie*, Sammlung Göschen 5131 (Walter de Gruyter, Berlin 1973)

1.25 C. E. Shannon: Communication theory of secrecy systems. Bell Syst. Tech. J. *28*, 656–715 (1949)

1.26 W. Sierpiński: *250 Problems in Elementary Number Theory* (American Elsevier, New York 1970)

1.27 J. V. Uspensky, M. A. Heaslet: *Elementary Number Theory* (McGraw-Hill, New York 1939)

1.28 D. J. Winter: *The Structure of Fields*, Graduate Texts in Mathematics, Vol. 16 (Springer, Berlin, Heidelberg, New York 1974)

Chapter 2

2.1 L. H. Hua: *Introduction to Number Theory* (Springer, Berlin, Heidelberg, New York 1982)

2.2 R. Plomp, G. F. Smoorenburg: *Frequency Analysis and Periodicity Detection in Hearing* (A. W. Sijthoff, Leiden 1970)

Chapter 3

3.1 C. Pomerance: The search for prime numbers. Sci. Am. *247*, No. 6, 136–147 (1982)

3.2 W. H. Mills: A prime representing function. Bull. Am. Math. Soc. *53*, 604 (1947)

3.3 T. Nagell: *Introduction to Number Theory* (Wiley, New York 1951)

3.4 D. Slowinski: Searching for the 27th Mersenne prime. J. Recreational Math. *11*, 258–261 (1978–79)

3.5 D. B. Gillies: Three new Mersenne primes and a statistical theory. Math. Comp. *18*, 93–97 (1963)

3.6 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 2.5 (Clarendon, Oxford 1984)

3.7 W. Kaufmann-Bühler: *Gauss. A Biographical Study* (Springer, Berlin, Heidelberg, New York 1981)

3.8 C. Chant, J. Fauvel (eds.): *Science and Belief* (Longman, Essex 1981)

Chapter 4

4.1 P. Erdös, M. Kac: The Gaussian law of errors in the theory of additive number theoretic functions. Am. J. Math. *62*, 738–742 (1945)

4.2 P. D. T. A. Elliot: *Probabilistic Number Theory*, Vols. 1–2 (Springer, Berlin, Heidelberg, New York 1980)

4.3 D. Zagier: “Die ersten 50 Millionen Primzahlen” in *Lebendige Zahlen*, ed. by F. Hirzebruch (Birkhäuser, Basel 1981)

4.4 G. Kolata: Does Gödel’s theorem matter to mathematics? Science *218*, 779–780 (1982)

4.5 P. Erdös: On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. Proc. Nat. Acad. Sci. U.S.A. *35*, 374–384 (1949)

4.6 H. M. Edwards: *Riemann’s Zeta Function* (Academic Press, New York 1974)

4.7 Z. Füredi, J. Komlos: The eigenvalues of random symmetric matrices. Combinatorica *1*, 233–241 (1981)

4.8 M.R. Schroeder: A simple function and its Fourier transform. Math. Intelligencer *4*, 158–161 (1982)

4.9 U. Dudley: *Elementary Number Theory* (Freeman, San Francisco 1969)

4.10 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed. Sect. 22.3 (Clarendon, Oxford 1984)

4.11 M. R. Schroeder: Speech Communication *1*, 9 (1982)

4.12 I. M. Vinogradov: *An Introduction to the Theory of Numbers* (Pergamon, New York 1955)

Chapter 5

5.1 C. D. Olds: *Continued Fractions* (Random House, New York 1963)

5.2 H. S. Wall: *Analytic Theory of Continued Fractions* (Van Nostrand, Princeton 1948)

5.3 A. N. Khovanskii: *The Application of Continued Fractions and Their Generalizations to Problems in Approximation Theory* (Noordhoff, Groningen 1963)

5.4 A. Y. Khinchin: *Continued Fractions* (University of Chicago Press, Chicago 1964)

5.5 F. D. M. Haddani: Phys. Rev. Lett. *51*, 605–607 (1983)

5.6 K. Ikeda, M. Mitsumo: Phys. Rev. Lett. *53* 1340–1343 (1984)

5.7 C. J. Bouwkamp, A. J. Duijvestijn, P. Medema: Tables relating to simple squared rectangles (Dept. of Mathematics and Mechanics, Technische Hogeschool, Eindhoven 1960)

5.8 V. E. Hoggatt: *Fibonacci and Lucas Numbers* (Houghton Mifflin, Boston 1969)

5.9 P. H. Richter, R. Schranner: Leaf arrangment. Naturwissenschaften *65*, 319–327 (1978)

5.10 M. Eigen: “Goethe und das Gestaltproblem in der modernen Biologie,” in H. Rössner (ed.): *Rückblick in die Zukunft* (Severin und Siedler, Berlin 1981)

5.11 O. Ore: *Number Theory and Its History* (McGraw-Hill, New York 1948)

5.12 A. Koenig (personal communication)

5.13 W. Gellert, H. Küstner, M. Hellwich, H. Kästner (eds.): *The VNR Concise Encyclopedia of Mathematics* (Van Nostrand Reinhold, New York 1977)

5.14 L. K. Hua, Y. Wang: *Applications of Number Theory to Numerical Analysis IX* (Springer, Berlin, Heidelberg, New York 1981)

5.15 J. C. Lagarias, A. M. Odlyzko: Solving “low-density” subset sum problems. J. Association of Computing Machinery *32*, 229–246 (1985)

5.16 R. L. Graham (personal communication)

5.17 R. K. Guy: *Unsolved Problems in Intuitive Mathematics*, Vol. I, Number Theory (Springer, Berlin, Heidelberg, New York 1981)

5.18 M. Gardner: Mathematical games. Sci. Am. *239*, No. 4, 22–26 (1978)

5.19 R. L. Graham: A theorem on partitions. J. Austral. Math. *4*, 435–441 (1963)

5.20 E. Landau: *Elementary Number Theory* (Chelsea, New York 1958)

5.21 E. H. Neville: *The Farey Series of Order 1025* (Cambridge University Press, Cambridge 1950)

5.22 C. M. Rader: Recovery of undersampled periodic waveforms. IEEE Trans. ASSP-*25*, 242–249 (1977)

5.23 T. L. Mac Donald: Astronomische Nachrichten *241*, 31 (1931)

5.24 M. Gardner: *Wheels, Life and Other Mathematical Amusements* (Freeman, New York 1983)

5.25 R. T. Gregory and E. V. Krishnamurthy: *Methods and Applications of Error-Free Computation* (Springer, New York 1984)

Chapter 6

6.1 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 5.2 (Clarendon, Oxford 1984)

6.2 P. J. Davis: *The Lore of Large Numbers* (Random House, New York 1961)

6.3 L. E. Dickson: *History of the Theory of Numbers*, Vols. 1–3 (Chelsea, New York 1952)

Chapter 7

7.1 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 5.4 (Clarendon, Oxford 1984)

7.2 T. L. Heath: *Diophantus of Alexandria* (Dover, New York 1964)

7.3 C. F. Gauss: *Disquisitiones Arithmeticae* [English transl. by A. A. Clarke, Yale University Press, New Haven 1966]

7.4 R. Tijdeman: On the equation of Catalan. Acta Arith. *29*, 197–209 (1976); "Exponential Diophantine Equations," in Proc. Int. Congr. Math., Helsinki (1978)

7.5 S. Singh: *Fermat's Enigma* (Walker and Company, New York 1987). See also W. Kaufmann-Bühler: *Gauss. A Biographical Study* (Springer, Berlin, Heidelberg, New York 1981)

7.6 M. Abramowitz, I. A. Stegun: *Handbook of Mathematical Function* (Dover, New York 1965)

7.7 M. Schroeder: Math. Intelligencer *16*, No. 4, 19 (1994)

7.8 D. Goldfeld: The Sciences *March/April*, 34 (1996)

7.9 M. D. Hirschhorn: A simple proof of Jacobi's four-square theorem. J. Austral. Math. Soc. *32*, 61–67 (1981)

7.10 M. R. Schroeder: Acustica *75*, 94 (1991)

7.11 H. Minkowski: Peter Gustav Lejeune Dirichlet und seine Bedeutung für die heutige Mathematik. Jahresbericht der Deutschen Mathematiker-Vereinigung *14*, 149–163 (1905)

7.12 M. R. Schroeder: Eigenfrequenzstatistik und Anregungsstatistik in Räumen. Acustica *4*, 45–68 (1954)

Chapter 8

8.1 H. Halberstam, C. Hooley (eds.): *Progress in Analytic Number Theory*, Vol. I (Academic, London 1981)

8.2 R. H. Hudson: A common combinatorial principle underlies Riemann's formula, the Chebyshev phenomenon, and other subtle effects in comparative prime number theory. J. reine angew. Math. *313*, 133–150 (1980)

Chapter 9

9.1 N. J. A. Sloane: "Error-Correcting Codes and Cryptography," in *The Mathematical Gardener*, ed. by D. Klarner (Prindle Weber Schmidt, Boston 1981) pp. 347–382. Republished in Cryptologica *6*, 128–153, 258–278 (1982)

9.2 A. G. Konheim: *Cryptography: A Primer* (Wiley, New York 1981)

9.3 W. Diffie, M. E. Hellman: Privacy and authentication: An introduction to cryptography. Proc. IEEE *67*, 397–427 (1979)

9.4 J. J. Quisquater, C. Couvreur: Fast decipherment algorithm for RSA public-key cryptosystem. Electron. Lett. *18*, 905–907 (1982)

9.5 G. Brassard: A note on the complexity of cryptography. IEEE Trans. IT *25*, 232–233 (1979)

9.6 A. R. Meijer: Groups, Factoring, and Cryptography. Math. Mag. *69*, 103–109 (1996)

9.7 C. S. Ogilvy: *Through the Mathescope* (Oxford University Press, London 1956), cited in M. Dekking, M. Mendés France, A. van der Poorten: Folds! Math. Intelligencer *4*, 130–138 (1982)

9.8 C. H. Meyer, S. Matyas: *Cryptography: A New Dimension in Computer Data Security* (Wiley, New York 1982)

Chapter 10

10.1 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 18.2 (Clarendon, Oxford 1984)

Chapter 11

11.1 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 22.8 (Clarendon, Oxford 1984)

11.2 M. Abramowitz, I. A. Stegun (eds.): *Handbook of Mathematical Functions* (Dover, New York 1970)

11.3 J. Kubilius: *Probabilistic Methods in the Theory of Numbers* Translations of Mathematical Monographs *11* (Auger. Math. Soc., Providence 1964)

11.4 S. W. Graham: The greatest prime factor of the integers in an interval. J. London Math. Soc. (2) *24*, 427–440 (1981)

11.5 C. Couvreur, J. J. Quisquater: An introduction to fast generation of large prime numbers. Philips J. Res. *37*, 231–264 (1982)

Chapter 12

12.1 R. L. Rivest, A. Shamir, L. A. Adleman: A method for obtaining digital signatures and public key cryptosystems. Comm. ACM *21*, 120–126 (1978)

12.2 A. G. Konheim: *Cryptography: A Primer* (Wiley, New York 1981) pp. 331–347

12.3 E. N. Gilbert, F. J. MacWilliams, N. J. A. Sloane: Codes which detect deception. Bell Syst. Tech. J. *53*, 405–424 (1974)

12.4 V. Fak: Repeated use of codes which detect deception. IEEE Trans. IT *25*, 233–234 (1979)

12.5 T. Beth: Sci. Am. *December*, 70 (1995)

12.6 I. Stewart: Sci. Am. *February*, 124 (1996)

12.7 A. Beutelspacher: *Cryptology* (The Math. Assoc. of America, Washington, DC 1994)

Chapter 13

13.1 C. F. Gauss: *Disquisitiones Arithmeticae* [English transl. by A. A. Clarke, Yale University Press, New Haven 1966]

13.2 M. Abramowitz, I. A. Stegun (eds.): *Handbook of Mathematical Functions* (Dover, New York 1970)

13.3 C. S. Ogilvy: *Tomorrow's Math* (Oxford University Press, Oxford 1962)

13.4 L. J. Alex: Solving exponential diophantine equations. Math. Mag. *54*, 259–260 (1981)

13.5 S. W. Golomb: *Shift Register Sequences* (Holden-Day, San Francisco 1967)

13.6 H. E. Rowe: *Signals and Noise in Communication Systems* (Van Nostrand Reinhold, New York 1965)

13.7 M. R. Schroeder, D. Gottlob, K. F. Siebrasse: Comparative study of European concert halls. J. Acoust. Soc. Am. *56*, 1195–1201 (1974)

13.8 M. R. Schroeder: Binaural dissimilarity and optimum ceilings for concert halls: More lateral sound diffusion. J. Acoust. Soc. Am. *65*, 958–963 (1979)

13.9 M. R. Schroeder: Toward better acoustics for concert halls. Phys. Today *33*, No. 10, 24–30, October (1979)

13.10 H. W. Strube: More on the diffraction theory of Schroeder diffusors. J. Acoust. Soc. Am. *70*, 633–635 (1981)

13.11 H. P. Lawther, Jr.: An application of number theory to the splicing of telephone cables. Am. Math. Monthly *42*, 81–91 (1935)

13.12 G. K. Skinner: Sci. Am. *August*, 66 (1988)

Chapter 14

14.1 N. J. A. Sloane: "Error-Correcting Codes and Cryptography," in *The Mathematical Gardener*, ed. by D. Klarner (Prindle Weber Schmidt, Boston 1981) pp. 347–382. Republished in Cryptologia *6*, 128–153, 258–278 (1982)

14.2 R. C. Merkle, M. E. Hellman: Hiding information and signatures in trapdoor knapsacks. IEEE Trans. IT *24*, 525–530 (1978)

14.3 A. Shamir: "A Polynomial Time Algorithm for Breaking Merkle-Hellman Cryptosystems," Internal Report Applied Mathematics, The Weizmann Institute, Rehovot, Israel

14.4 Y. Desmedt, J. Vandewalle, R. Govaerts: “Critical Analysis of the Security of Knapsack Public Key Algorithms,” in *Proceedings of the IEEE International Symposium on Information Theory* (IEEE, New York 1982) pp. 115–116

14.5 J. C. Lagarias (personal communication)

Chapter 15

15.1 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 6.5–7 (Clarendon, Oxford 1984)

15.2 E. Jahnke, R. Emde: *Tables of Functions* (Dover, New York 1945)

15.3 M. Born, E. Wolf: *Principles of Optics* (Pergamon, Oxford 1970)

15.4 M.R. Schroeder, R.E. Gerlach, A. Steingrube, H.W. Strube: Response to “Theory of Optimal Plane Diffusors.” J. Acoust. Soc. Am. *66*, 1647–1652 (1979)

15.5 M.R. Schroeder: Constant-amplitude antenna arrays with beam patterns whose lobes have equal magnitudes. Archiv für Elektronik und Übertragungstechnik (Electronics and Communication) *34*, 165–168 (1980)

15.6 J. E. Mazo: Some theoretical observation on spread-spectrum communications. Bell Syst. Tech. J. *58*, 2013–2023 (1979)

15.7 I. F. Blake, J. W. Mark: A note on complex sequences with low correlations. IEEE Trans. IT *28*, 814–816 (1982)

15.8 R. A. Scholtz: The origins of spread-spectrum communication. IEEE Trans. Communications, *30*, 822–852 (1982); see also other articles in the same issue (May 1982)

15.9 R. M. Lerner: Signals having good Correlation Functions. Western Electronic Show and Convention, San Francisco, August 1961. Paper 9/3

15.10 S. W. Golomb: *Shift Register Sequences* (Holden-Day, San Francisco 1967)

15.11 M. R. Schroeder: Acustica *81*, 364 (1995)

Chapter 16

16.1 L. H. Hua: *Introduction to Number Theory* (Springer, Berlin, Heidelberg, New York 1982)

16.2 C. E. Shannon (personal communication)

16.3 J. H. McClellan, C. M. Rader: *Number Theory in Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ 1979)

16.4 H. J. Nussbaumer: *Fast Fourier Transform and Convolution Algorithms* (Springer, Berlin, Heidelberg, New York 1981)

16.5 B. Gold, C. M. Rader, A. V. Oppenheim, T. G. Stockham: *Digital Processing of Signals* (McGraw-Hill, New York 1969)

16.6 J. W. Goodman: *Introduction to Fourier Optics* (McGraw-Hill, New York 1968)

16.7 C. M. Rader: Discrete Fourier transforms when the number of data samples is prime. Proc. IEEE *56*, 1107–1108 (1976)

Chapter 17

17.1 A. Hedayat, W. D. Wallis: Hadamard matrices and their applications. Ann. Statistics *6*, 1184–1238 (1978)

17.2 M. Harwit, N. J. A. Sloane: *Hadamard Transform Optics* (Academic, New York 1979)

17.3 H. J. Nussbaumer: *Fast Fourier Transform and Convolution Algorithms* (Springer, Berlin, Heidelberg, New York 1981)

Chapter 18

18.1 W. Kaufmann-Bühler: *Gauss. A Biographical Study* (Springer, Berlin, Heidelberg, New York 1981)

18.2 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 8.1 (Clarendon, Oxford 1984)

Chapter 19

19.1 P. Ribenboim: *The New Book of Prime Number Records* (Springer, New York 1996)

19.2 C. Pomerance: Recent developments in primality testing. Math. Intelligencer *3*, 97–105 (1981)

19.3 A. J. van der Poorten, A. Rotkiewicz: On strong pseudoprimes in arithmetic progressions. J. Austral. Math. Soc. *29*, 316–321 (1980)

19.4 A. Rotkiewicz: On Euler–Lehmer pseudoprimes and strong pseudoprimes with parameters *L*, *Q* in arithmetic progressions. Math. Comp. *39*, 239–247 (1982)

19.5 S. Goldwasser, S. Micali: "Probabilistic Encryption and How To Play Mental Poker," in *Proceedings of the 4th ACM Symposium on the Theory of Computing* (Assoc. Comp. Machinery, New York 1982) pp. 365–377

19.6 S. Micali (personal communication)

19.7 A. Shamir, R. L. Rivest, L. M. Adleman: "Mental Poker," in *The Mathematical Gardener*, ed. by D. Klarner (Prindle Weber Schmidt, Boston 1981) pp. 37–43

19.8 S. S. Wagstaff: Large Carmichael numbers. Math. J. Okayama Univ. *22*, 33–41 (1980)

19.9 M. O. Rabin: Probabilistic algorithm for testing primality. J. Number Theory *12*, 128–138 (1980)

19.10 C. F. Gauss: *Disquisitiones Arithmeticae* [English transl. by A. A. Clarke, Yale University Press, New Haven 1966]

19.11 J. D. Dixon: Asymptotically fast factorization of integers. Math. Comp. *36*, 255–260 (1981)

19.12 E. Lucas: *Théorie des Nombres* (Blanchard, Paris 1961)

19.13 J. Brillhart, D. H. Lehmer, J. L. Selfridge: New primality criteria and factorizations of $2^m \pm 1$. Math Comp. *29*, 620–647 (1975)

19.14 L. M. Adleman, C. Pomerance, R. S. Rumely: On distinguishing prime numbers from composite numbers. Ann. Math. (2) *117*, 173–206 (1983). See also: M. J. Coster, B. A. LaMacchia, C. P. Schnorr, J. Stern: Improved low-density subset sum algorithms. J. Computational Complexity *2*, 111–128 (1992)

19.15 I. L. Chuang, R. Laflamme, P. W. Shor, W. H. Zurek: Science *270*, 1633 (1995)

Chapter 20

20.1 L. H. Hua: *Introduction to Number Theory* (Springer, Berlin, Heidelberg, New York 1982)

20.2 T. M. Apostol: *Introduction to Analytic Number Theory* (Springer, Berlin, Heidelberg, New York 1976)

20.3 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 17.7 (Clarendon, Oxford 1984)

Chapter 21

21.1 L. H. Hua: *Introduction to Number Theory* (Springer, Berlin, Heidelberg, New York 1982)

21.2 G. H. Hardy, E. M. Wright: *An Introduction to the Theory of Numbers*, 5th ed., Sect. 17.1–5, 19.3–15 (Clarendon, Oxford 1984)

21.3 L. R. Rabiner, B. Gold: L. R. Rabiner, B. Gold: *Theory and Application of Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ 1975)

21.4 W. Feller: *An Introduction to Probability Theory with Its Applications* (Wiley, New York 1968)

21.5 M. Abramowitz, I. A. Stegun (eds.): *Handbook of Mathematical Functions* (Dover, New York 1970)

21.6 S. Ramanujan: *Notebooks* (Tata Institute of Fundamental Research, Bombay 1956)

21.7 G. H. Hardy: *Ramanujan* (Chelsea, New York 1978)

21.8 D. M. Bressoud: Analytical and combinational generalization of the Rogers-Ramanujan identities. Memoirs Am. Math. Soc. *24*, No. 227, 1–54 (1980)

Chapter 22

22.1 C. F. Gauss: *Disquisitiones Arithmeticae* [English transl. by A. A. Clarke, Yale University Press, New Haven 1966]

22.2 W. Gellert, H. Küstner, M. Hellwich, H. Kästner (eds.): *The VNR Concise Encylopedia of Mathematics* (Van Nostrand Reinhold, New York 1977)

22.3 H. Rademacher: *Lectures on Elementary Number Theory* (Blaisdell, New York 1964)

22.4 J. H. McClellan, C. M. Rader: *Number Theory in Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ 1979)

Chapter 23

23.1 A. Papoulis: *Signal Analysis* (McGraw-Hill, New York 1977)

23.2 J. H. McClellan, C. M. Rader: *Number Theory in Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ 1979)

23.3 S. Winograd: On computing the discrete Fourier transform. Math. Comp. *32*, 175–199 (1978)

Chapter 24

24.1 L. H. Hua: *Introduction to Number Theory* (Springer, Berlin, Heidelberg, New York 1982)

24.2 D. Hilbert: Über die Irreduzibilität ganzer rationaler Funktionen mit ganzzahligen Koeffizienten. J. reine angew. Math. *110*, 104–129 (1892)

24.3 J. H. McClellan, C. M. Rader: *Number Theory in Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ 1979)

Chapter 25

25.1 F. J. MacWilliams, N. J. A. Sloane: *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam 1978)

25.2 G. Hoffmann de Visme: *Binary Sequences* (The English University Press, London 1971)

25.3 S. W. Golomb and H. Taylor: Constructions and properties of Costas Arrays. Proc. IEEE *72*, 1143–1163 (1984)

Chapter 26

26.1 S. W. Golomb: *Shift Register Sequences* (Holden-Day, San Francisco 1967)

26.2 W. Stahnke: Primitive binary polynomials. Math. Comp. *27*, 977–980 (1973)

26.3 E. R. Berlekamp: *Algebraic Coding Theory* (McGraw-Hill, New York 1968)

26.4 F. J. MacWilliams, N. J. A. Sloane: *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam 1978)

26.5 M. R. Schroeder (ed.): *Speech and Speaker Recognition* (S. Karger, Basel 1985)

26.6 M. R. Schroeder: Integrated-impulse method of measuring sound decay without using impulses. J. Acoust. Soc. Am. *66*, 497–500 (1979)

26.7 A. R. Møller: Use of stochastic signals in evaluation of the dynamic properties of a neuronal system. Scand. J. Rehab. Med., Suppl. *3*, 37–44 (1974)

26.8 I. I. Shapiro, G. H. Pettengill, M. E. Ash, M. L. Stone, W. B. Smith, R. P. Ingalls, R. A. Brockelman: Fourth test of general relativity. Phys. Rev. Lett. *20*, 1265–1269 (1968)

26.9 T. Rothman: The short life of Evariste Galois. Sci. Am. *246*, No. 4, 136–149 (1982)

26.10 M. R. Schroeder, D. Gottlob, K. F. Siebrasse: Comparative study of European concert halls. J. Acoust. Soc. Am. *56*, 1195–1201 (1974)

26.11 M. R. Schroeder: Binaural dissimilarity and optimum ceilings for concert halls: More lateral sound diffusion. J. Acoust. Soc. Am. *65*, 958–963 (1979)

26.12 J. W. Goodman: *Introduction to Fourier Optics* (McGraw-Hill, New York 1968)

26.13 M. Born, E. Wolf: *Principles of Optics* (Pergamon, Oxford 1970)

26.14 S. H. van Lint, F. J. MacWilliams, N. J. A. Sloane: On Pseudo-Random Arrays. SIAM J. Appl. Math. *36*, 62–72 (1979)

26.15 M. R. Schroeder: Constant-amplitude antenna arrays with beam patterns whose lobes have equal magnitudes. Archiv für Elektronik und Übertragungstechnik (Electronics and Communication) *34*, 165–168 (1980)

26.16 G. Hoffmann de Visme: *Binary Sequences* (The English University Press, London 1971)

26.17 F. J. MacWilliams: The structure and properties of binary cyclic alphabets. Bell Syst. Tech. J. *44*, 303–332 (1965)

26.18 F. J. MacWilliams: A table of primitive binary idempotents of odd length. IEEE Trans. IT *25*, 118–121 (1979)

Chapter 27

27.1 F. J. MacWilliams, N. J. A. Sloane: *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam 1978)

27.2 G. Hoffmann de Visme: *Binary Sequence* (The English University Press, London 1971)

27.3 S. W. Golomb: *Shift Register Sequences* (Holden-Day, San Francisco 1967)

27.4 D. E. Knuth: *The Art of Computer Programming*, Vol. 2, Seminumerical Algorithms (Addison-Wesley, Reading, MA 1969)

27.5 E. N. Gilbert: Unpublished notes (1953)

27.6 T. Herlestan: “On the Complexity of Functions of Linear Shift Register Sequences,” in *Proceedings of the International Symposium on Information Theory* (IEEE, New York 1982) p. 166

27.7 H. J. Baker, F. C. Piper: Communications security, a survey of cryptography. IEE Proc. *A 129*, No. 6, 357–376 (1982)

27.8 D. P. Robbins, E. D. Bolker: The bias of three pseudo-random shuffles. Aequationees Math. *22*, 268–292 (1981)

27.9 P. Diaconis, M. Shahshahani: Generating a random permutation with random transpositions. Z. Wahrscheinlichkeitstheorie *57*, 159–179 (1981)

27.10 N. J. A. Sloane: “Encrypting by Random Rotations,” in [Ref. 27.11] pp. 71–128

27.11 T. Beth (ed.): *Cryptography*, Proc. Workshop, Burg Feuerstein, March 29–April 2, 1982, Lecture Notes in Computer Science, Vol. 149 (Springer, Berlin, Heidelberg, New York 1983)

Chapter 28

28.1 M. R. Schroeder: Synthesis of low peak-factor signals and binary sequences with low autocorrelation. IEEE Trans. IT *13*, 85–89 (1970)

28.2 W. Rudin: Some theorems on Fourier coefficients. Proc. Am. Math. Soc. *10*, 855–859 (1959)

28.3 D. A. Shedd, D. V. Sarwate: Construction of sequences with good correlation properties. IEEE Trans. IT-*25*, 94–97 (1979)

28.4 A. Gersho, B. Gopinath, A. M. Odlyzko: Coefficient inaccuracy in transversal filtering. Bell Syst. Tech. J. *58*, 2301–2316 (1979)

28.5 I. F. Blake, J. W. Mark: A note on complex sequences with low correlations. IEEE Trans. IT *28*, 814–816 (1982)

28.6 A. T. Moffet: Minimum redundancy linear arrays. IEEE Trans. Antennas and Propagation *16*, 172–175 (1968)

28.7 A. K. Dewdney: Sci. Am. *December*, 16 (1985)

28.8 D. Guicking, H. Freienstein: Proc. Active ’95, 371 (1995)

28.9 J. Shearer: IEEE Trans. IT *16*, 183 (1990)

28.10 A. Rosa, G. Sabadussi, J. Turgeon (eds.): *Theory and Practice of Combinatories* (North-Holland, Amsterdam 1982)

Chapter 29

29.1 M. R. Schroeder: Images from computers and microfilm plotters. Comm. ACM *12*, 95–101 (1969). See also IEEE Spectrum, March 1969, pp. 66–78

29.2 M. R. Schroeder: A simple function and its Fourier transform. Math. Intelligencer *4*, 158–161 (1982)

29.3 M. R. Schroeder: Number theory in physics, engineering and art. Interdisciplinary Sci. Rev. *6*, No. 3, 239–248 (1980)

29.4 M. R. Schroder: Fractal, Chaos, Power Laws (Freeman, New York 1991)

29.5 J. C. Risset: Proc. 7th Int. Congr. Acoustics *3*, 613 (1971)

29.6 M. R. Schroeder: J. Acoust. Soc. Am. *79*, 186 (1986)

29.7 M. R. Schroeder: Nature *325*, 765 (1987)

Chapter 30

30.1 N. J. A. Sloane: *A Handbook of Integer Sequences* (Academic Press, Orlando, FL 1973)

30.2 D. Shechtman, I. Blech, D. Gratias and J. W. Cahn: Metallic phase with long-range orientational order and no translational symmetry. Phys. Rev. Lett. *53*, 1951–1953 (1984)

30.3 M. Gardner: Extraordinary nonperiodic tiling that enriches the theory of tiles. Sci. Am. *236*, 110–121 (Jan. 1977)

30.4 D. Levine and P. J. Steinhardt: Quasicrystals: A new class of ordered structures. Phys. Rev. Lett. *53*, 2477–2480 (1984)

30.5 B. Mandelbrot: *The Fractal Geometry of Nature* (Freeman, San Francisco 1983)

30.6 T. A. Witten and L. M. Sander: Phys. Rev. Lett. *47*, 1400–1403 (1981); Phys. Rev. B *7*, 5686–5697 (1983)

30.7 C. Nicolis and G. Nicolis: Gibt es einen Klima-Attraktor? Phys. Blätter *41*, 5–9 (1985)

30.8 E. Basar: Toward a physical approach to integrative physiology. I. Brain dynamics and physical causality. Am. J. Physiol. *245* (Regulatory Integrative Comp. Physiol. *14*), R510–R533 (1983); see also A. Abraham, A. Mandel and D. Farmer, in *Proceedings Nonlinear Functions of the Brain* (Santa Barbara 1982)

30.9 M. R. Schroeder: Linear prediction, entropy and signal analysis. IEEE ASSP Magazine *1*, 3–11 (July 1984)

30.10 M. J. Feigenbaum: Universal behavior in nonlinear systems. Los Alamos Science *1*, 4–27 (1981); see also M. J. Feigenbaum: Quantitative universality for a class of nonlinear transformation. J. Statistical Physics *19*, 25–52 (1978)

30.11 M. Schroeder: *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise* (Freeman, New York 1991)

Name Index

Abel, N. H., 17
Abramowitz, M., 182, 186
Adleman, L. M., 264
Albion College, 170
Alexandria, 120
America, 193
AT&T Bell Laboratories, 52, 54, 88, 384

Bach, J. S., 27
Baker, W. O., xix
Behr, J. von, 55, 59
Beijing, 93
Bell Laboratories, 128, 237
Berlin, 13
Bernoulli, N., 83
Bessel, F. W., 100
Beurling, A., 147
Birkhoff, G., 367
Bohr, N., 4
Born, M., viii
Bose, S. N., 188
Brown, R., 3, 395
Brunswick, 14, 42, 77, 298
Burg Feuerstein, 353

Caivano, D., xix
Cantor, G., xvii, 130, 382, 388–390, 392, 393, 395, 396, 401
Carmichael, R. D., 260
Cassini, G. D., 80, 387
Catalan, E. C., 128
Chaos, 403
Chebyshev, P. L., 145
Churchill, W. S., 161
Cohen, H., 264
College of New Jersey, 193
Columbia University, 193
Conan Doyle, Sir Arthur, 45
Conway, J. H., 35
Courant, R., 46
Cryptologia, 170
Culviner, N., xix

David, King, 143
Diophantus of Alexandria, 11, 120
Dirac, P. A. M., 17, 355
Dirichlet, P. G. Lejeune, 56, 134, 144, 272, 273, 275, 276
Doppler, J. C., 325, 326
Dress, A., 374
Drittes Physikalisches Institut, 393
Dudley, H. W., 35

Earth, 23
Egypt, 94
Einstein, A., viii, 3, 42, 188, 336, 337, 396
Eisenstein, G., 369
England, 193
Eratosthenes, 31–33, 46
Erdös, P., 50
Euclid, 28, 31, 104, 121, 298, 313, 390, 391
Euler, L., 5, 13, 14, 36, 39, 42, 75, 130, 139, 141, 161, 163, 176, 220, 230, 256, 268, 270

Faltings, G., 12
Farey, J., 73, 95, 104, 105
Feigenbaum, M. J., 401
Fermat, P. de, 4, 5, 11, 12, 31, 36, 41, 42, 125, 129, 132, 139, 196, 211, 255, 262, 289, 295
Fibonacci, Leonardo (da Pisa), viii, 5, 8, 10, 83, 85–87, 89–91, 102, 382, 383

Fourier, J. B. J., 5, 19, 54, 206, 219, 223, 225, 227, 231, 239–242, 245, 247, 248, 276, 338, 355, 380, 381, 383, 385, 387
Fraunhofer, J. von, 338
Frénicle, B., 125
Fresnel, A. J., 226
Frobenius, F.G., 123

Galois, E., 5, 16, 19, 245, 309, 317, 326, 331–344, 352, 361–363, 370
Gardner, M., xix, 102
Gauss, C. F., vii, 5, 12–14, 21, 31, 34, 42, 48–50, 76, 77, 101, 116, 125, 129, 195, 197, 219, 224–226, 263, 292, 295, 298, 310, 351, 359, 369
Geiger, H., 347
Georg August (King of Hanover), 193
Georg-August University, 45, 193
George II, 193
Germain, S., 129
Gerson, L. B., 128
Gilbert, E. N., 352
Gillies, D. B., 36
Gödel, K., 48
Göhl, G., 51
Goldbach, C., 67, 130
Golomb, S. W., 365
Göttingen, viii, 42, 45, 77, 193
Göttingen Academy of Science, 129
Göttingen Mathematical Society, 134
Graham, R. L., 92, 95
Gregory, R. T., 105
Grossmann, S., 401
Guy, R. K., 35

Hadamard, J. S., 5, 50, 104, 245, 247, 248
Haddani, F. D. M., 75
Hagis, Jr., P., 40
Hall, J. L., xix, 362, 384
Hamburg, 55
Hamming, R. W., 4, 19, 334, 335, 357
Hanauer, S., 54, 128
Hanover, 193
Hardy, G. H., vii, 42, 132
Harwit, M., 248
Hausdorff, F., xvii, 389–392, 394–396, 399, 401
Heidelberg, xix
Heisenberg, W., viii
Helmstedt, 310
Hensel, K., xviii, 105, 106
Hermite, C., 223
Hilbert, D., viii, 51, 52, 363, 382, 390, 391, 395
Hirschhorn, M. D., 132
Holube, I., 101
Hung, E., xix

Ikeda, K., 75
Institute of Electrical and Electronics Engineers, 170
Io, 334

Jacobi, K. G. J., 132
Julia, G., xvii, 389, 399
Jupiter, viii, 334

Khinchin, A. Y., 78
Kindermann, L., 375
King George II, 193
King's College, 193
Kirchhoff, G. R., 208, 340, 344
Kirschmann-Schröder, G., xix
Kneser, M., xix
Knuth, D., 394
Koch, R., 365
Koenig, A., 88
Kohlrausch, A., xix
Königsberg, viii
Krishnamurthy, E. V., 105
Kronecker, L., 3, 5, 245, 247, 249, 270, 380
Kummer, E. E., 129
Kyrene, 32

Lagarias, J. C., xix, 216
Lagrange, J. L., 131
Landau, E., 271
Lavoisier, A. L., 3
Le Blanc ("Monsieur"), 129
Legendre, A. M., 49, 50, 219, 221, 223, 226, 230, 256, 333
Lenstra, Jr., H. W., 37, 216, 264
Leonardo da Pisa (Fibonacci), 8, 87, 89
Levine, D., 387
Lichtenberg, C., xvii, 396
Liebe, L., xix
Lindemann, F. von, viii
Lint, J. H. van, 344
London, 193
Lorentz, H. A., 355
Lotsch, H., xix
Lucas, E. A., xvii, 35, 90, 92, 263
Lyapunov, A. M., 398

Ma Chung-Fan, 93
MacWilliams, F. J., 344, 372
Magnus, W., xix
Mandelbrot, B., 388, 392, 401

Mangoldt, H. von, 270
Manhattan, 193
Markov, A. A., 384, 385
Mars, 334
Mathematical Intelligencer, 265
Mathews, M. V., xix, 28
Maxwell, J. C., 188
Mercury, viii, 23, 335, 337, 338, 371
Mersenne, M., 31, 35–40, 263, 344
Mertens, F., 173, 181, 271, 272
Meyer, P., xix, 393
Michels, R., xix
Mills, W. H., 34, 58
Minkowski, H., vii, viii, 3, 5, 134
Mitsumo, M., 75
Möbius, A. F., 5, 182, 184, 267, 268, 270, 272, 274, 299
Moivre, A. de, 83
Monte Carlo, 347, 356
Monzert, L. A., 102
Mordell, L. J., 12
Morse, H. M., 380, 381
Mössbauer, R., 337
Mozart, W. A., 325
Münchhausen, Baron von, 193

Nagell, T., 34
Neptune, 23
Neumann, J. von, 347
New York Times, 42
Newton, I., 42, 397, 398, 400, 402
Nickel, L., 36
Nobel Prize, 188
Noll, C., 36
North America, 3

Odlyzko, A. M., 52, 216, 272, 372

Paris, 51
Pascal, B., 85
Pauli, W., 355
Peitgen, H.-O., 400, 402
Pell, J., 124, 125
Penrose, R., 103, 387
Penzias, A. A., xix
Pierce, J. R., xix, 28
Planck, M., 188, 379, 402
Pluto, 23
Poincaré, H., 398, 402
Poisson, S. D., 36, 38, 145, 184, 189, 276
Pollak, H. O., xix, 93
Princeton University, 193
Püschel, D., 365
Pushkin, A. S., 15, 42

Pythagoras of Samos, 3, 26, 79, 126, 127, 299

Rabin, M. O., 261
Rademacher, H., 295
Rader, C. M., 242
Ramanujan, S., 282
Rayleigh, Lord (J. W. Strutt), 351
Reeves, A., 28
Rejewski, M., 150
Richter, P. H., 400, 402
Riele, H. J. J. te, 272
Riemann, B., 49, 51, 55, 181, 184, 271, 272, 337, 368
Riesel, H., 51
Risset, J.-C., 393
Roberts, L., 28
Rollwage, M., 329
Rüdenberg, L. (née Minkowski), viii
Rudin, W., 358, 359
Ruhr, 46
Rumely, R. S., 264

Salzburg, 325
Sander, L. M., 396
Saturn, xviii, 80, 387
Scarlatti, D., 375
Schönke, A., xix
Schrödinger, E., viii, 17
Schroeder, A., xix
Selberg, A., 50
Seurat, G., 79
Shamir, A., 216
Shannon, C. E., 237
Shechtman, D., 381, 382
Sherlock Holmes, 45
Skewes, S., 48
Sloane, N. J. A., 163, 214, 248, 344, 353
Slowinski, D., 36
Sorge, R., 14, 157
Spandikow, P. C., xix
St. Petersburg, 13
Stanford University, 28
Star of David, 143
Stegun, I. A., 182, 186
Steinhardt, P. J., 387
Strube, H. W., xix
Sullivan, W., 42
Sun, 23
Sylvester, J. J., 247

Taylor, H., xix
Texas, 102
Thomae, S., 401
Thue, A., 380, 381

Universität Bremen, 400, 402
Universität Göttingen, 299, 336, 393
Urbana, Illinois, 36

Vallée Poussin, C. J. de la, 50
Venus, viii, 337, 371
Vyssotsky, V. A., 356

Wagstaff, Jr., S. S., 37
Waring, E., 136
Weierstrass, K. T., xvii, 392
Weinrich, A., xix
Weinstein, S. B., 317
Wendorff, E. R., 82
Wiener, N., 396
Wigner, E. P., 52
Williams, H. C., 265
Wilson, N. Y., 139, 140, 201, 263
Witten, T. A., 396
Wolfskehl Prize, 129
Wright, E. M., 42, 132

Zürich, viii
Zagier, D., xix, 49
Zech, xviii, 324, 326–328
Zeilinger, A., 156
Zeno of Elea, 391

Subject Index

Abel’s proof, 17
Absolute pseudoprime, 260
Acoustic responses, 336
Acoustical quality, 338
Active array, 344
Additive number theory, 68, 366
Additive problem, 276
Aesthetic value, 367
Alteration (of documents), 194
Aluminum-manganese alloy, 381
Ambiguity function, 324, 325
Amicable number, 40
Amorphous substances, 381
Analytic number theory, 144
Antimatter, 355
Aperiodic, 380
Aperiodic orbit, 398
Aperiodicity, 381
Apfelmännchen, 402
Approximating fraction, 74
Approximating function, 76
Art, 367
Artificial language, 88, 335
Ascents, 152
Astronomy, 403
Atomic dimension, 396
Attractor, 395, 400, 401
Automatic speech recognition, 334

Backscatter, 209
Backward recursion, 102
Bank deposit, 102
Baroque integers, 374–375
Basin of attraction, 395, 399
Beatty sequences, 386
Bee, 88, 370
Bending of light, 337
Bifurcate, 401
Bifurcation cascade, 395
Binaural similarity, 338
Black-body radiation, 188
Bohr’s postulate, 4
Bose-Einstein distribution, 188
Brain waves, 397
Broadside direction, 328
Brownian motion, 3, 395
Brownian particle, 395

Camouflage, 229, 331
Cantor dust, 388
Cantor Set, 388, 389, 401
Cantor’s conjecture, 130
Card shuffle, 353
Cardinality, 389
Cassini’s divisions, 387
Casting out 9’s, 112
Cent (1/1200 of the musical octave), 75
Central limit theorem, 351
Certified signature, 193
Chaff, 325
Chaos, 379
Chaotic behaviour, 402
Chaotic mapping, 397
Chaotic motion, 379
Characteristic function, 276
Cheating, 258
Check-bit, 333
Chemical element, 4
Chemical reaction, 103, 395
Chinese number system, 236
Chinese primality test, 33
Chinese remainder, 312
Chinese remainder system, 105
Chinese remainder theorem, 235, 251, 312

Chirp, 361
Christmas card, 54
Circle of fifths, 27
Circular convolution, 238, 243, 307
Circular correlation, 331
Circulary polarized sound waves, 133
Climate, 397
Codebook, 335
Codeword, 333
Coin tossing by telephone, 258
Columbia University, 193
Combinatorial interpretation, 281
Compass and straightedge, 298
Complete residue system, 112
Completely additive, 185
Completely multiplicative, 272
Complex number system, 394
Complex root, 301
Complexification, 230
Complexity, 352
Composite, 21, 255
Computer graphics, 400
Computer processing, 403
Computer science, 88
Concert hall acoustic, 338
Concert hall acoustics, 207
Concert hall ceiling, 208, 228
Congruence, 115, 119
Congruence notation, 15
Conjugate partition, 277
Continued fraction, 73, 82, 124, 313, 387
Convex region, 134
Convolution, 237, 276
Convolution integral, 306
Conway's prime producing machine, 35
Coprimality, 59, 368
Coprime, 28, 312
Coprime probability, 52, 55
Correlation properties, 219, 331
Countable set, 389
Counter weapon, 325
Counterexample, 272
Coupled oscillators, 99
Covering set, 283
Cross-correlation, 334
Cryptologia (a journal), 170
Crystallographically forbidden, 387
Cumulants, 276
Cyclotomic polynomial, 289, 301, 303, 304, 307
Cyclotomy, 13, 14, 43, 290, 297

Decibel, 28
Decimal periodic fraction, 197
Decrypting exponent, 165, 166
Decrypting without factoring, 155
Derangements, 152
Descents, 152
Deterministic chaos, 9, 379, 390, 402, 403
Deterministic primality testing, 263
Deterministic system, 402
Devil's dance, 399
Diagnostic value, 397
Diagonal method, 389
Difference equation, 84
Difference set, 366
Diffraction, 225
  grating, 219
  pattern, 339, 343, 382, 384
Diffusing sound, 207
Diffusion constant, 395
Diffusion-limited aggregation, 396
Digital encryption, 164
Diophantine equation, 97, 103, 119
Diophantus's Arithmetica, 11
Dirac equation, 17
Dirac-delta function, 306
Directivity pattern, 363
Dirichlet series, 272, 275, 276
Dirichlet's theorem, 145
Discrete convolution, 306
Discrete Fourier Transform (DFT), 206, 239
Disproof, 272
Disquisitiones Arithmeticae, 263
Distances between primes, 63, 66
Distribution of $\omega(n)$, 182
Distributive law, 25
Dividing the circle, 43
Divisibility of Fibonacci numbers, 88
Division of the circle into equal parts, 298
Divisor function, 171
  average, 174
  geometric mean, 174
  summatory function, 175
Doppler effect, 325
Doppler radar, 324, 325
Double-precision, 103

Earth, 99
Easy knapsack, 213
Egyptian fraction, 94
Eigenvalues, 247
Einstein's analysis, 3, 396
Einstein's relativity theory, 336
Eisenstein primes, 369
Electrical discharge, 396
Electrical network, 81
Electro-encephalogram, 397

Electron diffraction, 381
Electronic banking, 194
Elementary method, 45
Elementary particles, 280, 403
Elementary proof, 50
Elliptic curves, 265
Encrypted message, 164, 166
Encrypting modulus, 165
Encryption exponent, 164, 165
Encryption requirements, 169
Entropy, 188
Enzymatic reaction, 396
EPR (Einstein-Podolski-Rosen), 156
Error-correcting code, 305, 333
Error-free computing, 103
Euclid’s algorithm, 28, 121
Euclid’s algorithm for polynomials, 313
Euclid’s proof, 31
Euclid’s school, 392
Euclidean algorithm, 104
Euclidean construction, 211
Euclidean theorem, 390
Euler criterion, 230
Euler identity, 281
Euler pseudoprimes, 256
Euler trap door, 161
Euler’s conjecture, 130
Euler’s constant, 174
Euler’s criterion, 220
Euler’s expansion, 75
Euler’s ϕ function, 13, 141, 268
Euler’s function average, 176
Euler’s theorem, 13, 141, 164
Evolution, 396
Exclusion principle, 403
Experimental design, 374
Exponential congruence, 202
Exponential Diophantine equation, 128
Extended Riemann Hypothesis (ERH), 51

Factoring, 263
Factoring algorithm, 264
Factorization, 299, 302, 310
Faltings’s proof, 12
Farey fraction, 95, 104
Farey series, 96
Farey tree, 97, 100
Fast Fourier Transform (FFT), 128, 238, 248
Fast Hadamard transform, 245
Fast prime-length Fourier transform, 242
Fatal ambiguity, 325
Feigenbaum constant, 401
Fermat *p*, 211
Fermat number, 12
Fermat prime, 12, 41, 211, 292, 298
Fermat’s Last Theorem (FLT), 11, 129
Fermat’s theorem, 12, 139
Ferrer graph, 277
Fibonacci, 382
Fibonacci number, 83, 102, 277
Fibonacci number system, 10, 86
Fibonacci’s rabbits, 8
Field element, 318
Finite field, 16
Finite-state machine, 322
Five-fold symmetry, 381, 387
Fixpoint, 395, 398
Floor function, 34
Flower, 88
Forgery, 194
Formula for primes, 34
Four-color necklace, 371
Fourier optics, 241
Fourier series, 276
Fourier spectrum, 338
Fourier transform, 306
Fourier transformation, 54
Fourth effect of general relativity, 336
Fractal, 379, 400
  boundary, 399
Fractal skin, 395
Fractal value, 392
Fractals, 388
Fraunhofer diffraction, 338
Frequency channel, 232
Frequency hopping, 327
Frequency shift, 325
Frequency-modulated signal, 356
Fundamental theorem of arithmetic, 22

Galois array, 209, 331
Galois field, 17, 309, 317, 331
Galois group, 17
Galois phase-grating, 337
Galois sequence, 318–321, 331, 333, 335–337, 339, 348, 361, 363
  spectral properties, 331
Galois sequences, 211
Gauss bracket, 34
Gauss primes, 369
Gauss sum, 15, 219, 224, 359
Gauss sum incomplete, 225
Gauss’s cryptogram, 101
Gauss’s estimate, 48
Gauss’s Eureka, 125
Gauss’s Insight, 292
Gauss’s notebook, 77
Gaussian variable, 351

Geheimschreiber, 147
General relativity theory, 336
Generalized divisor function, 175
Generalized Euler function, 270
Generalized Fibonacci number, 93
Generalized Fibonacci sequence, 91
Generalized Legendre sequence, 230
Generating function, 272, 275, 276, 306
Genetic code, 403
Geometric interpretation, 120
Geometric mean, 291
"Geometrical" construction, 77
Geometrical interpretation, 132
Geometry of Numbers, 131
GF(p^m), 324
Glass, 382
Glossary, 405
Gödel's theorem, 48
Goldbach's conjecture, 67, 130
Golden ratio, 9, 75, 80, 83, 102, 282, 383
Golden rectangles, 80
Golomb ruler, 365
Göttingen Academy of Science, 129
Göttingen Mathematical Society, 134
Göttingen University, 42
Graham's partition, 95
Graham's sequence, 92
Graphic design, 367
Gravitational coupling, 101
Gravitational field, 337
Gravitational potential, 337
Gravitational red shift, 337
Greatest common divisor, 23
Group theory, 403

Hadamard matrix, 245
Hadamard transform, 245
Hamming code, 335
Hamming distance, 4, 334
Hamming error-correcting code, 19
Hamming power spectrum, 357
Hard knapsack, 214
Harmonic foundation, 28
Harmonic mean of $\Omega(n)$, 188
Harmonic oscillator, 223
Hausdorff dimension, 390, 395, 399, 401
Hearing, 348
Hensel code, 105
Heptagon, 295, 298
Hermite polynomial, 223
Higher-dimensional diffusor, 343
Higher-order correlation, 348
Hilbert curve, 390, 391, 395
Hilbert envelope, 363
Hologram, 331
Home computer, 265
Human ear, 393
Human pitch perception, 25
Hydrophone, 364

Idempotent, 362
Iff, 33
Image enhancement, 403
Impulse response, 305
Impulse response, periodic, 307
Inaudible, 331
Index, 201
Index array, 231
Indirect proof, 389
Infinite precision, 403
Initial condition, 402
Integer, 3, 21
  lattice, 132
  lattice points, 134
Internet, 194
Interplanetary, 334
Intersection, 25
Inverse Möbius transform, 268, 270
Inverse polynomial, 312
Inversion formula, 269–271
Invisible, 331
Io, 334
Irrational frequency, 100
Irrational numbers, 75
Irrationality, 79
Irreducible polynomial, 318, 320
Isotope, 3
Iteration, 397

Jamming, 219
Julia set, 399, 400
Jupiter, 101
Jupiter's moon Io, 334

Key distribution, 157
Kilo, 27
Kilobit, 27
Kirchhoff diffraction, 208, 340
Knapsack encryption, 213
Knapsack ripping, 216
Kronecker, 3
  power, 246
  products, 245
  symbol, 267, 270
Kronecker's constant, 181

Landau's dissertation, 271
Language, 335

design, 335
Lateral sound wave, 207, 338
LCM (The Least Common Multiple), 22
Least remainder, 261
Legendre formula, 49
Legendre sequences, 230, 333, 359
Legendre sequences, Fourier property, 223
Legendre symbol, 221, 226, 360
Li(x), 48, 49
Liber Abaci, 87
Lichtenberg figure, 396
Light diffuser, 331
Lightning, 396
Linear congruence, 111
Linear progression, 56, 144
Linear recursion, 333
Linear shift register, 322
Linear system, 305, 306
Logical "and", 114
Logistic parabola, 399
Long-range coherence, 385
Long-range order, 387
Lorentz invariance, 355
Loud-speaker columns, 344
Low crosscorrelation, 232
Lucas number, 90
Lucas test, 263
Lyapunov exponent, 398

Magdeburg (German cruiser), 162
Major third, 74
Mandelbrot set, 401
Mapping, 397, 402
Markov process, 384
Mars, 334
Martian canals, 334
Masquerading, 194
Mathematical dichotomy, 398
Mathematical logic, 25
Mathematics Institute, University of Göttingen, 299
Matrix representation, 321
Matter wave, 4
Maximal matrix rank, 352
Maximum-length sequence, 225, 322, 336
Maxwell's equation, 188
Mean-free path, 395
Measure, 389
Measure theory, 77
Median, 190
Mediant, 96
Mercury, 23, 335, 337
Mersenne number, 35
Mersenne prime, 31, 35
Merten's theorem, 37, 181
Mertens conjecture, 271
Mill's formula, 34
Minimum redundancy antenna, 363
Minimum redundancy array, 364
Minkowski's theorem, 134
Minkowski's *Urquell*, 3
Möbius function, 182, 184, 267, 272
Möbius inversion, 302
Möbius inversion formula, 299
Möbius transform, 268
Modeling, 398
Modular arithmetic, 111
Monic polynomials, 312
Monsieur Le Blanc, 129
Monster group, 262
Monte Carlo method, 347
Moon, 99
Moore-Penrose inverse, 103
Mordell conjecture, 12
Morse-Thue sequence, 380
Mössbauer effect, 337
Most irrational, 9
Motion, perihelion, 337
Multiplicative function, 141
Multiplicative group, 142
Multiplicative inverses, 17
Multiplicative problem, 276
Music, 28
Musical chord, 392
Musical paradox, 388

Natural number, 21
Necklace, 370, 371
maximally varied, 371
super-varied, 372
Neptune, 23, 334
Neurophysiology, 348
New musical scale, 28
New state of matter, 382
Newton's iteration, 398
Newton's method, 398
Nim, 10
Nim-like games, 87
Noble numbers, 382, 387
Noise abatement, 331
Non-Euclidean paradox, 392
Non-Maxwellian, 188
Nonfractals, 392
Nonlinear Diophantine equation, 124
Nonlinear dynamical system, 390, 395
Nonlinear function, 352
Nonlinear law, 401
Nonlinear physical system, 402

Nonlinearity, 395
  parameter, 401
Nonperiodic rational, 398
Nontotient, 144
Normal-mode degeneracy, 135
Nucleon, 3
Number of divisors, 171
Number of primitive elements, 323
Number of primitive roots, 196
Number-theoretic convolution, 273
Number-theoretic fluke, 28
Number-theoretic logarithm, 201, 230
Numerical computation, 272
Numerical evidence, 272

Oblivious transfer, 194
Ocean surveillance, 364
Octave, 26, 393
One-time-pad, 156
One-way process, 166
Optical spectra, 3
Orbital period, 387
Order, 195

P-adic algebra, 105
P-adic Hensel code, 105
Pallas, 101
Parabolic maximum, 401
Partition, 275, 277
  generating function, 278
  into distinct integer, 279
  into even integer, 280
  into odd integer, 280
  recurrence formula, 282
Peak factor, 355–357, 359, 361
Pell equation, 124
Penrose tiling, 387
Pentagonal number, 281
Percentile, 190
Perfect number, 39
Perfect permutation, 326
Periodic attractor, 401
Periodic continued fraction, 78
Periodic correlation, 206
Periodic crystal, 385
Periodic decimal fraction, 197
Periodic lattice, 381
Periodic orbit, 398
Permutation, 147, 326
Permutation Cycles, 150
Perrin sequence, 35
Phase, 356
  angle, 357
  randomization, 331
Phase coherence, 385
Phase space, 402
Phase-array, 327
Phase-grating, based on primitive roots, 342
Photon, 188
Physiological measurement, 336
Pitch, 25, 393
Pitch detector, 365
Planck's constant, 402
Planck's harmonic oscillators, 188
Pluto, 23
Pocket calculator, 264
Poincaré map, 398
Poincaré section, 402
Poisson distribution, 185, 189
Poisson process, 38, 145
Poisson variables, 276
Poker, 255
Polynomial, 305
  convergent, 313
  reciprocal, 319
  residue, 310
  residue reduction, 310, 321
  theory, 309
Polynomial's Diophantine equation, 312
Polynomial-time algorithm, 167
Positron, 17
Power spectrum, 206
Precision measurement, 335
Predictive analysis, 397
Preimage, 400
Primality horizon, 264
Primality testing, 31, 263
Prime, 31
  counting function, 48
  divisor, 179
  divisor function, 179, 274
  divisor, cumulative distribution, 190
  divisors, number of, 179, 185
  number, 21
  number theorem, 50
  order, 317
  power order, 317
  quadruplets and quintuplets, 62
  residue system, 115
  spacings, 63, 66
  triplet, 61
Primeless expanse, 58
Primitive element, 318
Primitive root, 195, 292, 302
  ceiling, 208, 209
  sequence, 344
Primitive root array, 211
Primitive roots, Fourier property, 206

Princeton University, 193
Principle of symmetry, 403
Probabilistic argument, 45
Probability, 45
  distribution, 276
  theory, 276
Proper fraction, 394
Properties of matter, 134
Protocol, 258
Proton, 264
Pseudoperfect number, 40
Pseudoprime, 116, 255
Pseudorandom, 347
Pseudorandom permutation, 353
Pseudorandom sequence, 336, 348, 352
Pseudorandomness, 352
Public-key encryption, 4, 161, 165, 193
Pythagoras scale, 3
Pythagorean comma, 27
Pythagorean numbers, 11, 127

Quadratic congruence, 219, 251
Quadratic progession, 57
Quadratic reciprocity, 226
Quadratic residue, 219
Quadratic-residue sequence, 227
Quanta, 280
Quantum cryptography, 156
Quantum decrypting, 153
Quantum factoring, 256, 265
Quantum mechanics, 223, 403
Quarks, 403
Quotient polynomial, 310

Rabbit spectrum, 385
Radar, 229, 324, 331, 337
Radar echo, 325, 335
Radiation pattern, 229, 329, 344, 355
Radio astronomy, 364
Radio communication, 229
Railroad switch yard problem, 93
Ramanujan congruence relation, 282
Random number, 349
Random number generator, 347, 350
Random phase angle, 356
Random rotation, 353
Random variables, 276
Random walk, 395
Randomness, 367
Range, 325
  rate, 325
Rate of divergence, 398
Rational factor, 299
Rational factorization, 301
Rayleigh variable, 351
Real-time tomography, 356
Reciprocal polynomial, 319
Recovery of undersampled periodic waveform, 97
Recursion, 321
Recursive relation, 321
Reflection phase-grating, 207, 228, 229, 339
Regular *n*-gon, 298
Regular 17-gon, 76
Regular pentagon, 79
Remainder, 15, 111
  polynomial, 310
Remote coin tossing, 255
Reneging, 194
Repeated Encryption, 168
Repunit, 39
Residue, 4, 111
  reduction, 320
Resonant frequency, 131
Resonator, 131
Restrict Partitions, 279
Reverberation time, 336
Riemann Hypothesis, 51, 272
Riemann zetafunction, 50, 55
Riemann's approximation, 50, 51
Riemannian geometry, 337
Rings of Saturn, 387
Roman number system, 237
Rounding error, 103
Rudin-Shapiro polynomials, 358

Satellite picture transmission, 331
Saturn, 387
Scaling, 379
  factor, 379
  law, 401
Scattered wave, 338
Scattering, 207
Schrödinger wave equation, 17
Search algorithm, 88
Secret, 260
Security, 4
Self-conjugate, 277
Self-decimating, 384
Self-dual, 25
Self-generating sequences, 380
Self-similar sequence, 382, 387
Self-similar set, 389
Self-similarity, 379, 380, 387, 393, 400, 401
Semantics, 335
Semitone, 393
Set theory, 25
Shift register, 322, 333

Shor algorithm, 153
Short-range property, 385
Sieve of Eratosthenes, 32, 46
Signal-to-noise ratio, 336
Signalbuch der Kaiserlichen Marine, 163
Signature code, 232
Signature sequence, 329
Silver ratio, 387
Similar areas, 392
Similar figures, 392
Simplex code, 334
Simultaneous congruence, 235
Sin, 347
Single modulus system, 105
Sino-optical Fourier transformer, 240
Sino-representation, 236, 239
  application, 237
  generalized, 241
"Slaved" oscillator, 99
Slowing of electromagnetic radiation, 337
Smallest prime divisor, 56
Sociable number, 40
Solar systems, 387
Sonar, 229, 331
Sound, 207
Space vehicle, 334
Speckle, 331
Spectroscopy, 248
Specular reflection, 209, 338
Speech recognition, 333
Spin, 17, 355
Spread-spectrum communication, 219, 229, 231, 328
Square number, 126
Squared square, 82
Squarefree, 28, 173
  integer, 267
Squarefree and coprime integers, 59
Stabilities, 387
Stable orbit, 395
Stamp Problem, 122
Star of David, 143
Statistical fractal, 395
Statistical moments, 276
Statistical self-similarity, 395
Statistically indistinguisable, 395
Strahlkörper, 134
Straightedge and a compass, 291
Strange attractor, 389, 395, 399
Strong pseudoprime, 260, 261
Sum of 2 squares, 132
Sum of 3 squares, 131
Sum of 4 squares, 131
Sum of primes, 68
Summatory function, 173, 267
Superior conjunction, 337
Surprise, 367
Symbols, 405
Symmetry, 367
Syntaxes, 335

Talking computer, 334
Talkwriter, 334
Television antenna, 344
Temperament (musical), 26
Ternary fraction, 388, 389
Theorema fundamentale, 226
Third-order correlation, 348
Thumb-tack function, 325
Tiling, 391
Time reversal invariance, 344
Time-discrete, 276
Time-discrete system, 305, 306
Totient function, 141
Trap-door function, 163
Triangular number, 125, 281
Turbulence, 379
Twiddle factor, 248
Twin dragon, 394
Twin prime, 59
Two-dimensional array, 210, 229
Two-squares theorem, 133

U-505, 157
U-boat, 325
Ultimate test for primality, 140
Uncountable, 389, 399
  set, 398
Union, 25
Unit fraction, 94
Universal constant, 401

Venus, 337
Visibility, 368
von Mangoldt function, 270
von Münchhausen's secret message, 193

Waring's problem, 136
Waveform, 355, 357
Weierstrass function, 392
Weyl's theorem, 135
Wiener process, 396
Wiener-Khinchin theorem, 207
Wilson's theorem, 139, 201
Winding number, 386
Wolfskehl Prize, 129

X-ray astronomy, 209

Young tableaux, 277

z transform, 276, 305
z transform, 306, 307
Zahlen-Mystik, 101
Zech logarithm, 324, 327, 328
Zero-crossing, 396
Zetafunction, 50, 181, 184, 186, 187, 272, 368