ENCyclopedia of Mathematics and its Applications

4  W. Miller, Jr.  Symmetry and separation of variables
6  H. Minc.  Permanents
11  W. B. Jones and W. J. Thron.  Continued fractions
18  H. O. Fattorini.  The Cauchy problem
21  W. T. Tutte.  Graph theory
22  J. R. Bionda.  Field extensions and Galois theory
23  J. R. Cannon.  The one-dimensional heat equation
35  A. Salama.  Computer and autonomous systems
36  N. White (ed.).  Theory of matroids
37  N. H. Bingham, C. M. Goldie and J. L. Teugels.  Regular variation
38  P. P. Pethe.  and V. A. Popov.  Rational approximation of real functions
39  N. White (ed.).  Combinatorial geometry
40  M. Prat and H. Zassenhaus.  Algorithmic algebraic number theory
41  J. A. Aczel and J. D. Henshaw.  Functional equations containing several variables
43  R. V. Ambartzumian.  Factorization calculus and geometric probability
44  G. Ciesielski, S. V. Londen and O. Staffans.  Volterra integral and functional equations
45  G. Glauber and M. Rahman.  Basic hypergeometric series
46  E. Zerger and D. Spies.  Comparison of statistical experiments
47  A. Neumaier.  Interval methods for systems of equations
48  N. Korneichuk.  Exact constants in approximation theory
49  R. A. Brualdi and H. J. Ryser.  Combinatorial matrix theory
50  N. White (ed.).  Material applications
51  S. Sakai.  Operator algebras in dynamical systems
52  W. Hodge.  Ideal theory
53  H. Stahl and V. Totik.  General orthogonal polynomials
54  R. Schneider.  Convex bodies
55  G. Da Prato and J. Zabczyk.  Stochastic equations in infinite dimensions
56  A. Broman.  Mas.  Las Vergnas, B. Sturmfels, N. White and G. Ziegler.  Oriented matroids
57  E. A. Edgar and L. Sucheston.  Shopping times and directed processes
58  C. Sims.  Computation with finitely presented groups
59  T. Palmer.  Banach algebras and the general theory of *-algebras
60  F. Benarowski.  Handbook of categorical algebras I
61  F. Benarowski.  Handbook of categorical algebras II
62  F. Benarowski.  Handbook of categorical algebras III
63  A. Katok and B. Hasselblatt.  Introduction to the modern theory of dynamical systems
64  V. A. Sakhnov.  Combinatorial methods in discrete mathematics
65  V. N. Sakhnov.  Probabilistic methods in discrete mathematics
66  P. M. Cohn.  Skew fields
67  Richard J. Gardner.  Geometric tomography
68  George A. Baker, Jr. and Peter Graves-Morris.  Padé approximants
69  Jan Krajcik.  Bounded arithmetic, propositional logic, and complexity theory
70  Hans G. Opdam.  Finite dimensional optimization and control theory
71  A. C. Thompson.  Minkowski geometry
72  R. B. Bapat and T. E. S. Raghavan.  Nonnegative matrices and applications
73  K. Engel.  Spectral theory
74  D. Cvetkovic, P. Rowlinson and S. Simic.  Eigenspaces of graphs
75  F. Bergeron, G. Labelle and P. Leroux.  Combinatorial species and tree-like structures
76  R. Goodman and N. Wallach.  Representations of the classical groups
78  A. Pethö and J. Wunz.  Orevalued systems and Banach space geometry
79  George E. Andrews, Richard Askey and Ranjan Roy.  Special Functions
80  R. Tucci.  Quantum field theory for mathematicians
81  A. A. Ivanov.  Geometry of superspace I 2 ed.
83  O. Steenmark.  La’s Structural Approach to FDE Systems
84  C. F. Dunkl and Y. Xu.  Orthogonal polynomials of several variables
85  J. Mayberry.  The foundations of mathematics in the theory of sets
87  B. Bollobás and G. Erdős.  Geometric and Diophantine integrals
88  D. Kaminseki and R. B. Paris.  Asymptotics and Mellin–Barnes integrals

© Cambridge University Press
Contents

Editor's statement  page viii
Section editor's foreword  ix
Preface to the first edition  x
Preface to the second edition  xii
Introduction
Problems  12
Notes  13

Part one: Information theory

1 Entropy and mutual information  17
  1.1 Discrete random variables  17
  1.2 Discrete random vectors  33
  1.3 Nondiscrete random variables and vectors  37
      Problems  44
      Notes  49

2 Discrete memoryless channels and their capacity–cost functions  50
  2.1 The capacity–cost function  50
  2.2 The channel coding theorem  58
      Problems  68
      Notes  73

3 Discrete memoryless sources and their rate-distortion functions  75
  3.1 The rate-distortion function  75
  3.2 The source coding theorem  84
      Problems  91
      Notes  93
## Contents

4 The Gaussian channel and source 95  
4.1 The Gaussian channel 95  
4.2 The Gaussian source 99  
    Problems 105  
    Notes 110  

5 The source–channel coding theorem 112  
    Problems 120  
    Notes 122  

6 Survey of advanced topics for part one 123  
6.1 Introduction 123  
6.2 The channel coding theorem 123  
6.3 The source coding theorem 131  

Part two: Coding theory  

7 Linear codes 139  
7.1 Introduction: The generator and parity-check matrices 139  
7.2 Syndrome decoding on \(q\)-ary symmetric channels 143  
7.3 Hamming geometry and code performance 146  
7.4 Hamming codes 148  
7.5 Syndrome decoding on general \(q\)-ary channels 149  
7.6 Weight enumerators and the MacWilliams identities 153  
    Problems 158  
    Notes 165  

8 Cyclic codes 167  
8.1 Introduction 167  
8.2 Shift-register encoders for cyclic codes 181  
8.3 Cyclic Hamming codes 195  
8.4 Burst-error correction 199  
8.5 Decoding burst-error correcting cyclic codes 215  
    Problems 220  
    Notes 228  

9 BCH, Reed–Solomon, and related codes 230  
9.1 Introduction 230  
9.2 BCH codes as cyclic codes 234  
9.3 Decoding BCH codes, Part one: the key equation 236  
9.4 Euclid’s algorithm for polynomials 244  
9.5 Decoding BCH codes, Part two: the algorithms 249  
9.6 Reed–Solomon codes 253  
9.7 Decoding when erasures are present 266
### Preface

9.8 The (23,12) Golay code  
Problems  
Notes  

10 Convolutional codes

10.1 Introduction  
10.2 State diagrams, trellises, and Viterbi decoding  
10.3 Path enumerators and error bounds  
10.4 Sequential decoding  
Problems  
Notes  

11 Variable-length source coding

11.1 Introduction  
11.2 Uniquely decodable variable-length codes  
11.3 Matching codes to sources  
11.4 The construction of optimal UD codes (Huffman’s algorithm)  
Problems  
Notes  

12 Survey of advanced topics for Part two

12.1 Introduction  
12.2 Block codes  
12.3 Convolutional codes  
12.4 A comparison of block and convolutional codes  
12.5 Source codes  

### Appendices

A Probability theory  
B Convex functions and Jensen’s inequality  
C Finite fields  
D Path enumeration in directed graphs  

### References

1 General reference textbooks  
2 An annotated bibliography of the theory of information and coding  
3 Original papers cited in the text  

Index of Theorems  
Index
Editor’s statement

A large body of mathematics consists of facts that can be presented and described much like any other natural phenomenon. These facts, at times explicitly brought out as theorems, at other times concealed within a proof, make up most of the applications of mathematics, and are the most likely to survive changes of style and of interest.

This ENCYCLOPEDIA will attempt to present the factual body of all mathematics. Clarity of exposition, accessibility to the non-specialist, and a thorough bibliography are required of each author. Volumes will appear in no particular order, but will be organized into sections, each one comprising a recognizable branch of present-day mathematics. Numbers of volumes and sections will be reconsidered as times and needs change.

It is hoped that this enterprise will make mathematics more widely used where it is needed, and more accessible in fields in which it can be applied but where it has not yet penetrated because of insufficient information.

Information theory is a success story in contemporary mathematics. Born out of very real engineering problems, it has left its imprint on such far-flung endeavors as the approximation of functions and the central limit theorem of probability. It is an idea whose time has come.

Most mathematicians cannot afford to ignore the basic results in this field. Yet, because of the enormous outpouring of research, it is difficult for anyone who is not a specialist to single out the basic results and the relevant material. Robert McEliece has succeeded in giving a presentation that achieves this objective, perhaps the first of its kind.

GIAN-CARLO ROTTA
Foreword

Transmission of information is at the heart of what we call communication. As an area of concern it is so vast as to touch upon the preoccupations of philosophers and to give rise to a thriving technology.

We owe to the genius of Claude Shannon* the recognition that a large class of problems related to encoding, transmitting, and decoding information can be approached in a systematic and disciplined way: his classic paper of 1948 marks the birth of a new chapter of Mathematics.

In the past thirty years there has grown a staggering literature in this fledgling field, and some of its terminology even has become part of our daily language.

The present monograph (actually two monographs in one) is an excellent introduction to the two aspects of communication: coding and transmission.

The first (which is the subject of Part two) is an elegant illustration of the power and beauty of Algebra; the second belongs to Probability Theory which the chapter begun by Shannon enriched in novel and unexpected ways.

MARK KAC

General Editor, Section on Probability

---

Preface to the first edition

This book is meant to be a self-contained introduction to the basic results in the theory of information and coding. It was written during 1972–1976, when I taught this subject at Caltech. About half my students were electrical engineering graduate students; the others were majoring in all sorts of other fields (mathematics, physics, biology, even one English major!). As a result the course was aimed at nonspecialists as well as specialists, and so is this book.

The book is in three parts: Introduction, Part one (Information Theory), and Part two (Coding Theory). It is essential to read the introduction first, because it gives an overview of the whole subject. In Part one, Chapter 1 is fundamental, but it is probably a mistake to read it first, since it is really just a collection of technical results about entropy, mutual information, and so forth. It is better regarded as a reference section, and should be consulted as necessary to understand Chapters 2–5. Chapter 6 is a survey of advanced results, and can be read independently. In Part two, Chapter 7 is basic and must be read before Chapters 8 and 9; but Chapter 10 is almost, and Chapter 11 is completely, independent from Chapter 7. Chapter 12 is another survey chapter independent of everything else.

The problems at the end of the chapters are very important. They contain verification of many omitted details, as well as many important results not mentioned in the text. It is a good idea to at least read the problems.

There are four appendices. Appendix A gives a brief survey of probability theory, essential for Part one. Appendix B discusses convex functions and Jensen's inequality. Appeals to Jensen's inequality are frequent in Part one, and the reader unfamiliar with it should read Appendix B at the first opportunity. Appendix C sketches the main results about finite fields needed in Chapter 9. Appendix D describes an algorithm for counting paths in directed graphs which is needed in Chapter 10.
Preface

A word about cross-references is in order: sections, figures, examples, theorems, equations, and problems are numbered consecutively by chapters, using double numeration. Thus “Section 2.3,” “Theorem 3.4,” and “Prob. 4.17” refer to section 3 of Chapter 2, Theorem 4 of Chapter 3, and Problem 17 of Chapter 4, respectively. The appendices are referred to by letter; thus “Equation (B.4)” refers to the fourth numbered equation in Appendix B.

The following special symbols perhaps need explanation: “□” signals the end of a proof or example; “iff” means if and only if; \[<x\] denotes the largest integer \(< x\); and \[\geq x\] denotes the smallest integer \(\geq x\).

Finally, I am happy to acknowledge my debts: To Gus Solomon, for introducing me to the subject in the first place; to John Pierce, for giving me the opportunity to teach at Caltech; to Gian-Carlo Rota, for encouraging me to write this book; to Len Baumert, Stan Butman, Gene Rodemich, and Howard Rumsey, for letting me pick their brains; to Jim Lesh and Jerry Heller, for supplying data for Figures 6.7 and 12.2; to Bob Hall, for drafting the figures; to my typists, Ruth Stratton, Lillian Johnson, and especially Dian Rapchak; and to Ruth Flohn for copy editing.

Robert J. McEliece
Preface to the second edition

The main changes in this edition are in Part two. The old Chapter 8 (“BCH, Goppa, and Related Codes”) has been revised and expanded into two new chapters, numbered 8 and 9. The old chapters 9, 10, and 11 have then been renumbered 10, 11, and 12. The new Chapter 8 (“Cyclic codes”) presents a fairly complete treatment of the mathematical theory of cyclic codes, and their implementation with shift register circuits. It culminates with a discussion of the use of cyclic codes in burst error correction. The new Chapter 9 (“BCH, Reed–Solomon, and Related Codes”) is much like the old Chapter 8, except that increased emphasis has been placed on Reed-Solomon codes, reflecting their importance in practice. Both of the new chapters feature dozens of new problems.