Wiley Series in Discrete Mathematics and Optimization

Fourth Edition

THE PROBABILISTIC METHOD

NOGA ALON • JOEL H. SPENCER

WILEY

THE PROBABILISTIC METHOD

WILEY SERIES IN
DISCRETE MATHEMATICS AND OPTIMIZATION

A complete list of titles in this series appears at the end of this volume.

THE PROBABILISTIC METHOD

Fourth edition, July 2015, Tel Aviv and New York

NOGA ALON
School of Mathematics,
Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel Aviv University,
Tel Aviv, Israel.

JOEL H. SPENCER
Courant Institute of Mathematical Sciences,
New York University,
New York, USA

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Alon, Noga.
The probabilistic method / Noga Alon, Joel H. Spencer. – Fourth edition.
pages cm
Includes bibliographical references and index.
ISBN 978-1-119-06195-3 (cloth)
1. Combinatorial analysis. 2. Probabilities. I. Spencer, Joel H. II. Title.
QA164.A46 2016
511′.6–dc23

2015021599

Typeset in 10/12pt TimesLtStd by SPi Global, Chennai, India.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

4 2016

To Nurit and Mary Ann

Contents

PREFACE **xiii**

ACKNOWLEDGMENTS **xv**

PART I METHODS **1**

1 The Basic Method **3**

1.1 The Probabilistic Method, 3
1.2 Graph Theory, 5
1.3 Combinatorics, 9
1.4 Combinatorial Number Theory, 11
1.5 Disjoint Pairs, 12
1.6 Independent Sets and List Coloring, 13
1.7 Exercises, 16

The Erdős–Ko–Rado Theorem, 18

2 Linearity of Expectation **19**

2.1 Basics, 19
2.2 Splitting Graphs, 20
2.3 Two Quickies, 22
2.4 Balancing Vectors, 23

2.5 Unbalancing Lights, 25
2.6 Without Coin Flips, 26
2.7 Exercises, 27

Brégman's Theorem, 29

3 Alterations **31**

3.1 Ramsey Numbers, 31
3.2 Independent Sets, 33
3.3 Combinatorial Geometry, 34
3.4 Packing, 35
3.5 Greedy Coloring, 36
3.6 Continuous Time, 38
3.7 Exercises, 41

High Girth and High Chromatic Number, 43

4 The Second Moment **45**

4.1 Basics, 45
4.2 Number Theory, 46
4.3 More Basics, 49
4.4 Random Graphs, 51
4.5 Clique Number, 55
4.6 Distinct Sums, 57
4.7 The Rödl nibble, 58
4.8 Exercises, 64

Hamiltonian Paths, 65

5 The Local Lemma **69**

5.1 The Lemma, 69
5.2 Property B and Multicolored Sets of Real Numbers, 72
5.3 Lower Bounds for Ramsey Numbers, 73
5.4 A Geometric Result, 75
5.5 The Linear Arboricity of Graphs, 76
5.6 Latin Transversals, 80
5.7 Moser's Fix-It Algorithm, 81
5.8 Exercises, 87

Directed Cycles, 88

6 Correlation Inequalities **89**

6.1 The Four Functions Theorem of Ahlswede and Daykin, 90
6.2 The FKG Inequality, 93
6.3 Monotone Properties, 94

6.4 Linear Extensions of Partially Ordered Sets, 97
6.5 Exercises, 99

Turán's Theorem, 100

7 Martingales and Tight Concentration 103

7.1 Definitions, 103
7.2 Large Deviations, 105
7.3 Chromatic Number, 107
7.4 Two General Settings, 109
7.5 Four Illustrations, 113
7.6 Talagrand's Inequality, 116
7.7 Applications of Talagrand's Inequality, 119
7.8 Kim–Vu Polynomial Concentration, 121
7.9 Exercises, 123

Weierstrass Approximation Theorem, 124

8 The Poisson Paradigm 127

8.1 The Janson Inequalities, 127
8.2 The Proofs, 129
8.3 Brun's Sieve, 132
8.4 Large Deviations, 135
8.5 Counting Extensions, 137
8.6 Counting Representations, 139
8.7 Further Inequalities, 142
8.8 Exercises, 143

Local Coloring, 144

9 Quasirandomness 147

9.1 The Quadratic Residue Tournaments, 148
9.2 Eigenvalues and Expanders, 151
9.3 Quasirandom Graphs, 157
9.4 Szemerédi's Regularity Lemma, 165
9.5 Graphons, 170
9.6 Exercises, 172

Random Walks, 174

PART II TOPICS 177

10 Random Graphs 179

10.1 Subgraphs, 180

10.2 Clique Number, 183
10.3 Chromatic Number, 184
10.4 Zero–One Laws, 186
10.5 Exercises, 193

Counting Subgraphs, 195

11 The Erdős–Rényi Phase Transition 197

11.1 An Overview, 197
11.2 Three Processes, 199
11.3 The Galton–Watson Branching Process, 201
11.4 Analysis of the Poisson Branching Process, 202
11.5 The Graph Branching Model, 204
11.6 The Graph and Poisson Processes Compared, 205
11.7 The Parametrization Explained, 207
11.8 The Subcritical Regions, 208
11.9 The Supercritical Regimes, 209
11.10 The Critical Window, 212
11.11 Analogies to Classical Percolation Theory, 214
11.12 Exercises, 219

Long paths in the supercritical regime, 220

12 Circuit Complexity 223

12.1 Preliminaries, 223
12.2 Random Restrictions and Bounded-Depth Circuits, 225
12.3 More on Bounded-Depth Circuits, 229
12.4 Monotone Circuits, 232
12.5 Formulae, 235
12.6 Exercises, 236

Maximal Antichains, 237

13 Discrepancy 239

13.1 Basics, 239
13.2 Six Standard Deviations Suffice, 241
13.3 Linear and Hereditary Discrepancy, 245
13.4 Lower Bounds, 248
13.5 The Beck–Fiala Theorem, 250
13.6 Exercises, 251

Unbalancing Lights, 253

14 Geometry 255

14.1 The Greatest Angle Among Points in Euclidean Spaces, 256

14.2 Empty Triangles Determined by Points in the Plane, 257
14.3 Geometrical Realizations of Sign Matrices, 259
14.4 ϵ-Nets and VC-Dimensions of Range Spaces, 261
14.5 Dual Shatter Functions and Discrepancy, 266
14.6 Exercises, 269

Efficient Packing, 270

15 Codes, Games, and Entropy 273

15.1 Codes, 273
15.2 Liar Game, 276
15.3 Tenure Game, 278
15.4 Balancing Vector Game, 279
15.5 Nonadaptive Algorithms, 281
15.6 Half Liar Game, 282
15.7 Entropy, 284
15.8 Exercises, 289

An Extremal Graph, 291

16 Derandomization 293

16.1 The Method of Conditional Probabilities, 293
16.2 d-Wise Independent Random Variables in Small Sample Spaces, 297
16.3 Exercises, 302

Crossing Numbers, Incidences, Sums and Products, 303

17 Graph Property Testing 307

17.1 Property Testing, 307
17.2 Testing Colorability, 308
17.3 Testing Triangle-Freeness, 312
17.4 Characterizing the Testable Graph Properties, 314
17.5 Exercises, 316

Turán Numbers and Dependent Random Choice, 317

Appendix A Bounding of Large Deviations 321

A.1 Chernoff Bounds, 321
A.2 Lower Bounds, 330
A.3 Exercises, 334

Triangle-Free Graphs Have Large Independence Numbers, 336

Appendix B Paul Erdős **339**

B.1 Papers, 339
B.2 Conjectures, 341
B.3 On Erdős, 342
B.4 Uncle Paul, 343
The Rich Get Richer, 346

Appendix C Hints to Selected Exercises **349**

REFERENCES **355**

AUTHOR INDEX **367**
SUBJECT INDEX **371**

Preface

The probabilistic method is one of the most powerful and widely used tools applied in combinatorics. One of the major reasons for its rapid development is the important role of randomness in theoretical computer science and in statistical physics.

The interplay between discrete mathematics and computer science suggests an algorithmic point of view in the study of the probabilistic method in combinatorics, and this is the approach we tried to adopt in this book. The book thus includes a discussion of algorithmic techniques together with a study of the classical method as well as the modern tools applied in it. The first part of the book contains a description of the tools applied in probabilistic arguments, including the basic techniques that use expectation and variance, as well as the more recent applications of martingales and correlation inequalities. The second part includes a study of various topics in which probabilistic techniques have been successful. This part contains chapters on discrepancy and random graphs, as well as on several areas in theoretical computer science: Circuit Complexity, Computational Geometry, Graph Property Testing and, Derandomization of randomized algorithms. Scattered between the chapters are gems described under the heading "The Probabilistic Lens." These are elegant proofs that are not necessarily related to the chapters after which they appear and can usually be read separately.

The basic probabilistic method can be described as follows: In order to prove the existence of a combinatorial structure with certain properties, we construct an appropriate probability space and show that a randomly chosen element in this space has the desired properties with positive probability. This method was initiated by Paul Erdős, who contributed so much to its development over a 50-year period that it seems appropriate to call it "The Erdős Method." His contribution can be measured not only by his numerous deep results in the subject but also by the many intriguing problems and conjectures posed by him that stimulated a big portion of the research in the area.

It seems impossible to write an encyclopedic book on the probabilistic method; too many recent interesting results apply probabilistic arguments, and we do not even try

to mention all of them. Our emphasis is on methodology, and we thus try to describe the ideas, and not always to give the best possible results if these are too technical to allow a clear presentation. Many of the results are asymptotic, and we use the standard asymptotic notation: for two functions f and g, we write $f = O(g)$ if $f \leq cg$ for all sufficiently large values of the variables of the two functions, where c is an absolute positive constant. We write $f = \Omega(g)$ if $g = O(f)$ and $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$. If the limit of the ratio f/g tends to zero as the variables of the functions tend to infinity we write $f = o(g)$. Finally, $f \sim g$ denotes that $f = (1 + o(1))g$; that is, f/g tends to 1 when the variables tend to infinity. Each chapter ends with a list of exercises. The more difficult ones are marked by $(*)$. The exercises enable readers to check their understanding of the material and also provide the possibility of using the book as a textbook.

This is the fourth edition of the book; it contains several improved results and covers various additional topics that developed extensively during the last few years. A breakthrough approach to the Local Lemma is described in Chapter 5. A new algorithmic approach to the "six standard deviations" result in discrepancy theory is presented in Chapter 12. A novel proof for the study of Property B, based on a random greedy coloring, appears in Chapter 3. In all the above cases, the algorithmic proofs provide essentially new arguments for the existence of the desired objects. A new, short section on graph limits has been added to Chapter 9. A technique for counting independent sets in graphs and its application in a graph coloring problem is described in Chapter 1. Further additions include a new Probabilistic Lens, several additional exercises, and a new appendix with hints to selected exercises.

As in the previous editions, it is a special pleasure to thank our wives, Nurit and Mary Ann. Their patience, understanding, and encouragement have been key ingredients in the success of this enterprise.

Noga Alon
Joel H. Spencer
Tel Aviv and New York, 2015

Acknowledgments

We are very grateful to all our students and colleagues who contributed to the creation of this fourth edition through joint research, helpful discussions, and useful comments. These include Simon Blackburn, Miklós Bóna, Steve Cook, Ehud Friedgut, Oded Goldreich, Omri Ben-Eliezer, Krzysztof Choromanski, Oliver Cooley, Ohad Feldheim, Naomi Feldheim, Asaf Ferber, Laura Florescu, Lior Gishbboliner, Matan Harel, Danny Hefetz, Timo Hirscher, Rani Hod, Mihyun Kang, Joel Lewis, Nati Linial, Guy Moshkovitz, Dhruv Mubayi, Tahl Nowik, Roberto Oliveira, Ron Peled, Will Perkins, Oliver Riordan, Guy Rutenberg, Jeffrey Shallit, Asaf Shapira, Clara Shikhelman, Philipp Sprüssel, Aravind Srinivasan, John Steinberger, Elmar Teufl, Shai Vardi, Amit Weinstein, Jed Yang, Mariano Zelke and Yufei Zhao, who pointed out various inaccuracies and misprints, and suggested improvements in the presentation as well as in the results. Needless to say, the responsibility for the remaining mistakes, as well as the responsibility for the (hopefully not many) new ones, is solely ours.

PART I

METHODS

1

The Basic Method

What you need is that your brain is open.
–Paul Erdős

1.1 THE PROBABILISTIC METHOD

The probabilistic method is a powerful tool for tackling many problems in discrete mathematics. Roughly speaking, the method works as follows: trying to prove that a structure with certain desired properties exists, one defines an appropriate probability space of structures and then shows that the desired properties hold in these structures with positive probability. The method is best illustrated by examples. Here is a simple one. The *Ramsey number* $R(k, \ell)$ is the smallest integer n such that in any two-coloring of the edges of a complete graph on n vertices K_n by red and blue, either there is a red K_k (i.e., a complete subgraph on k vertices all of whose edges are colored red) or there is a blue K_ℓ. Ramsey (1929) showed that $R(k, \ell)$ is finite for any two integers k and ℓ. Let us obtain a lower bound for the diagonal Ramsey numbers $R(k, k)$.

Proposition 1.1.1 *If* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, *then* $R(k, k) > n$. *Thus* $R(k, k) > \lfloor 2^{k/2} \rfloor$ *for all* $k \geq 3$.

Proof. Consider a random two-coloring of the edges of K_n obtained by coloring each edge independently either red or blue, where each color is equally likely. For any

The Probabilistic Method, Fourth Edition. Noga Alon and Joel H. Spencer.

fixed set R of k vertices, let A_R be the event that the induced subgraph of K_n on R is *monochromatic* (i.e., that either all its edges are red or they are all blue). Clearly, $\Pr[A_R] = 2^{1-\binom{k}{2}}$. Since there are $\binom{n}{k}$ possible choices for R, the probability that at least one of the events A_R occurs is at most $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$. Thus, with positive probability, no event A_R occurs and there is a two-coloring of K_n without a monochromatic K_k; that is, $R(k,k) > n$. Note that if $k \geq 3$ and we take $n = \lfloor 2^{k/2} \rfloor$, then

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1$$

and hence $R(k,k) > \lfloor 2^{k/2} \rfloor$ for all $k \geq 3$. ■

This simple example demonstrates the essence of the probabilistic method. To prove the existence of a good coloring, we do not present one explicitly, but rather show, in a nonconstructive way, that it exists. This example appeared in a paper of P. Erdős from 1947. Although Szele had applied the probabilistic method to another combinatorial problem, mentioned in Chapter 2, already in 1943, Erdős was certainly the first to understand the full power of this method and apply it successfully over the years to numerous problems. One can, of course, claim that the probability is not essential in the proof given above. An equally simple proof can be described by counting; we just check that the total number of two-colorings of K_n is larger than the number of those containing a monochromatic K_k.

Moreover, since the vast majority of the probability spaces considered in the study of combinatorial problems are finite, this claim applies to most of the applications of the probabilistic method in discrete mathematics. Theoretically, this is indeed the case. However, in practice the probability is essential. It would be hopeless to replace the applications of many of the tools appearing in this book, including, for example, the second moment method, the Lovász Local Lemma and the concentration via martingales by counting arguments, even when these are applied to finite probability spaces.

The probabilistic method has an interesting algorithmic aspect. Consider, for example, the proof of Proposition 1.1.1, which shows that there is an edge two-coloring of K_n without a monochromatic $K_{2\log_2 n}$. Can we actually find such a coloring? This question, as asked, may sound ridiculous; the total number of possible colorings is finite, so we can try them all until we find the desired one. However, such a procedure may require $2^{\binom{n}{2}}$ steps; an amount of time that is exponential in the size $\left[= \binom{n}{2}\right]$ of the problem. Algorithms whose running time is more than polynomial in the size of the problem are usually considered impractical. The class of problems that can be solved in polynomial time, usually denoted by **P** (see, e.g., Aho, Hopcroft and Ullman (1974)), is, in a sense, the class of all solvable problems. In this sense, the exhaustive search approach suggested above for finding a good coloring of K_n is not acceptable, and this is the reason for our remark that the proof of Proposition 1.1.1 is nonconstructive; it does not supply a constructive, efficient,

and deterministic way of producing a coloring with the desired properties. However, a closer look at the proof shows that, in fact, it can be used to produce, effectively, a coloring that is very likely to be good. This is because, for large k, if $n = \lfloor 2^{k/2} \rfloor$, then

$$\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!}\left(\frac{n}{2^{k/2}}\right)^k \leq \frac{2^{1+\frac{k}{2}}}{k!} \ll 1.$$

Hence, a random coloring of K_n is very likely not to contain a monochromatic $K_{2\log n}$. This means that if, for some reason, we *must* present a two-coloring of the edges of K_{1024} without a monochromatic K_{20}, we can simply produce a random two-coloring by flipping a fair coin $\binom{1024}{2}$ times. We can then deliver the resulting coloring safely; the probability that it contains a monochromatic K_{20} is less than $2^{11}/20!$, probably much smaller than our chances of making a mistake in any rigorous proof that a certain coloring is good! Therefore, in some cases the probabilistic, nonconstructive method does supply effective probabilistic algorithms. Moreover, these algorithms can sometimes be converted into deterministic ones. This topic is discussed in some detail in Chapter 16.

The probabilistic method is a powerful tool in combinatorics and graph theory. It is also extremely useful in number theory and in combinatorial geometry. More recently, it has been applied in the development of efficient algorithmic techniques and in the study of various computational problems. In the rest of this chapter, we present several simple examples that demonstrate some of the broad spectrum of topics in which this method is helpful. More complicated examples, involving various more delicate probabilistic arguments, appear in the rest of the book.

1.2 GRAPH THEORY

A *tournament* on a set V of n players is an orientation $T = (V, E)$ of the edges of the complete graph on the set of vertices V. Thus for every two distinct elements x and y of V, either (x, y) or (y, x) is in E, but not both. The name "tournament" is natural, since one can think of the set V as a set of players in which each pair participates in a single match, where (x, y) is in the tournament iff x beats y. We say that T has the property S_k if, for every set of k Players, there is one that beats them all. For example, a directed triangle $T_3 = (V, E)$, where $V = \{1, 2, 3\}$ and $E = \{(1, 2), (2, 3), (3, 1)\}$, has S_1. Is it true that for every finite k there is a tournament T (on more than k vertices) with the property S_k? As shown by Erdős (1963b), this problem, raised by Schütte, can be solved almost trivially by applying probabilistic arguments. Moreover, these arguments even supply a rather sharp estimate for the minimum possible number of vertices in such a tournament. The basic (and natural) idea is that, if n is sufficiently large as a function of k, then a *random* tournament on the set $V = \{1, \ldots, n\}$ of n players is very likely to have the property S_k. By a random tournament we mean here a tournament T on V obtained by choosing, for each $1 \leq i < j \leq n$, independently, either the edge (i, j) or the edge (j, i), where each of these two choices is equally

likely. Observe that in this manner, all the $2^{\binom{n}{2}}$ possible tournaments on V are equally likely; that is, the probability space considered is symmetric. It is worth noting that we often use in applications symmetric probability spaces. In these cases, we shall sometimes refer to an element of the space as a *random element*, without describing explicitly the probability distribution . Thus, for example, in the proof of Proposition 1.1.1 random two-colorings of K_n were considered; that is, all possible colorings were equally likely. Similarly, in the proof of the next simple result we study random tournaments on V.

Theorem 1.2.1 *If* $\binom{n}{k}(1-2^{-k})^{n-k} < 1$*, then there is a tournament on n vertices that has the property* S_k.

Proof. Consider a random tournament on the set $V = \{1, \ldots, n\}$. For every fixed subset K of size k of V, let A_K be the event that there is no vertex that beats all the members of K. Clearly, $\Pr[A_K] = (1-2^{-k})^{n-k}$. This is because, for each fixed vertex $v \in V - K$, the probability that v does not beat all the members of K is $1-2^{-k}$, and all these $n-k$ events corresponding to the various possible choices of v are independent. It follows that

$$\Pr\left[\bigvee_{\substack{K\subset V\\ |K|=k}} A_K\right] \leq \sum_{\substack{K\subset V\\ |K|=k}} \Pr[A_K] = \binom{n}{k}(1-2^{-k})^{n-k} < 1.$$

Therefore, with positive probability, no event A_K occurs; that is, there is a tournament on n vertices that has the property S_k. ■

Let $f(k)$ denote the minimum possible number of vertices of a tournament that has the property S_k. Since $\binom{n}{k} < \left(\frac{en}{k}\right)^k$ and $(1-2^{-k})^{n-k} < e^{-(n-k)/2^k}$, Theorem 1.2.1 implies that $f(k) \leq k^2 \cdot 2^k \cdot (\ln 2)(1+o(1))$. It is not too difficult to check that $f(1) = 3$ and $f(2) = 7$. As proved by Szekeres (cf. Moon (1968)), $f(k) \geq c_1 \cdot k \cdot 2^k$.

Can one find an explicit construction of tournaments with at most c_2^k vertices having property S_k? Such a construction is known but is not trivial; it is described in Chapter 9.

A *dominating set* of an undirected graph $G = (V, E)$ is a set $U \subseteq V$ such that every vertex $v \in V - U$ has at least one neighbor in U.

Theorem 1.2.2 *Let* $G = (V, E)$ *be a graph on n vertices, with minimum degree* $\delta > 1$. *Then G has a dominating set of at most* $n\dfrac{1+\ln(\delta+1)}{\delta+1}$ *vertices.*

Proof. Let $p \in [0, 1]$ be, for the moment, arbitrary. Let us pick, randomly and independently, each vertex of V with probability p. Let X be the (random) set of all vertices picked and let $Y = Y_X$ be the random set of all vertices in $V - X$ that do not have any neighbor in X. The expected value of $|X|$ is clearly np. For each fixed vertex $v \in V$,

$\Pr[v \in Y] = \Pr[v$ and its neighbors are not in $X] \leq (1-p)^{\delta+1}$. Since the expected value of a sum of random variables is the sum of their expectations (even if they are not independent) and since the random variable $|Y|$ can be written as a sum of n indicator random variables χ_v ($v \in V$), where $\chi_v = 1$ if $v \in Y$ and $\chi_v = 0$ otherwise, we conclude that the expected value of $|X| + |Y|$ is at most $np + n(1-p)^{\delta+1}$. Consequently, there is at least one choice of $X \subseteq V$ such that $|X| + |Y_X| \leq np + n(1-p)^{\delta+1}$. The set $U = X \cup Y_X$ is clearly a dominating set of G whose cardinality is at most this size.

The above argument works for any $p \in [0,1]$. To optimize the result we use elementary calculus. For convenience, we bound $1 - p \leq e^{-p}$ (this holds for all nonnegative p and is a fairly close bound when p is small) to give the simpler bound

$$|U| \leq np + ne^{-p(\delta+1)}.$$

Take the derivative of the right-hand side with respect to p and set it equal to zero. The right-hand side is minimized at

$$p = \frac{\ln(\delta+1)}{\delta+1}.$$

Formally, we set p equal to this value in the first line of the proof. We now have $|U| \leq n\dfrac{1+\ln(\delta+1)}{\delta+1}$, as claimed. ∎

Three simple but important ideas are incorporated in the last proof. The first is the linearity of expectation; many applications of this simple, yet powerful principle appear in Chapter 2. The second is perhaps more subtle and is an example of the "alteration" principle that is discussed in Chapter 3. The random choice did not supply the required dominating set U immediately; it only supplied the set X, which has to be altered a little (by adding to it the set Y_X) to provide the required dominating set. The third involves the optimal choice of p. One often wants to make a random choice but is not certain what probability p should be used. The idea is to carry out the proof with p as a parameter giving a result that is a function of p. At the end, that p is selected which gives the optimal result. Here, there is yet a fourth idea that might be called asymptotic calculus. We want the asymptotics of $\min np + n(1-p)^{\delta+1}$, where p ranges over $[0,1]$. The actual minimum $p = 1 - (\delta+1)^{-1/\delta}$ is difficult to deal with, and in many similar cases precise minima are impossible to find in a closed form. Rather, we give away a little bit, bounding $1 - p \leq e^{-p}$, yielding a clean bound. A good part of the *art* of the probabilistic method lies in finding suboptimal but clean bounds. Did we give away too much in this case? The answer depends on the emphasis for the original question. For $\delta = 3$, our rough bound gives $|U| \leq 0.596n$, while the more precise calculation gives $|U| \leq 0.496n$, perhaps a substantial difference. For δ large, both methods give asymptotically $n \ln \delta/\delta$.

It can easily be deduced from the results in Alon (1990b) that the bound in Theorem 1.2.2 is nearly optimal. A non-probabilistic, algorithmic proof of this theorem can be obtained by choosing the vertices for the dominating set one by

one, when in each step a vertex that covers the maximum number of yet-uncovered vertices is picked. Indeed, for each vertex v, denote by $C(v)$ the set consisting of v together with all its neighbors. Suppose that during the process of picking vertices the number of vertices u that do not lie in the union of the sets $C(v)$ of the vertices chosen so far is r. By the assumption, the sum of the cardinalities of the sets $C(u)$ over all such uncovered vertices u is at least $r(\delta + 1)$, and, hence by averaging, there is a vertex v that belongs to at least $r(\delta + 1)/n$ such sets $C(u)$. Adding this v to the set of chosen vertices, we observe that the number of uncovered vertices is now at most $r(1 - (\delta + 1)/n)$. It follows that in each iteration of the above procedure the number of uncovered vertices decreases by a factor of $1 - (\delta + 1)/n$ and, hence after $n \ln(\delta + 1)/(\delta + 1)$ steps, there will be at most $n/(\delta + 1)$ yet uncovered vertices that can now be added to the set of chosen vertices to form a dominating set of size at most equal to the one in the conclusion of Theorem 1.2.2.

Combining this with some ideas of Podderyugin and Matula, we can obtain a very efficient algorithm to decide whether a given undirected graph on n vertices is, say, $n/3$ edge-connected. A *cut* in a graph $G = (V, E)$ is a partition of the set of vertices V into two nonempty disjoint sets $V = V_1 \cup V_2$. If $v_1 \in V_1$ and $v_2 \in V_2$, we say that the cut *separates* v_1 and v_2. The *size* of the cut is the number of edges of G having one end in V_1 and the other end in V_2. In fact, we sometimes identify the cut with the set of these edges. The *edge connectivity* of G is the minimum size of a cut of G. The following lemma is due to Podderyugin and Matula (independently).

Lemma 1.2.3 *Let $G = (V, E)$ be a graph with minimum degree δ, and let $V = V_1 \cup V_2$ be a cut of size smaller than δ in G. Then every dominating set U of G has vertices in V_1 and in V_2.*

Proof. Suppose this is false and $U \subseteq V_1$. Choose, arbitrarily, a vertex $v \in V_2$, and let $v_1, v_2, \ldots, v_\delta$ be δ of its neighbors. For each i, $1 \leq i \leq \delta$, define an edge e_i of the given cut as follows: if $v_i \in V_1$, then $e_i = \{v, v_i\}$, otherwise $v_i \in V_2$, and since U is dominating, there is at least one vertex $u \in U$ such that $\{u, v_i\}$ is an edge; take such a u and put $e_i = \{u, v_i\}$. The δ edges $e_1, \ldots, e_\delta$ are all distinct and all lie in the given cut, contradicting the assumption that its size is less than δ. This completes the proof. ∎

Let $G = (V, E)$ be a graph on n vertices, and suppose we wish to decide whether G is $n/3$ edge-connected; that is, whether its edge connectivity is at least $n/3$. Matula showed, by applying Lemma 1.2.3, that this can be done in time $O(n^3)$. By the remark following the proof of Theorem 1.2.2, we can slightly improve it and get an $O(n^{8/3} \log n)$ algorithm as follows. We first check if the minimum degree δ of G is at least $n/3$. If not, G is not $n/3$ edge-connected, and the algorithm ends. Otherwise, by Theorem 1.2.2, there is a dominating set $U = \{u_1, \ldots, u_k\}$ of G, where $k = O(\log n)$, and it can in fact be found in time $O(n^2)$. We now find, for each i, $2 \leq i \leq k$, the minimum size s_i of a cut that separates u_1 from u_i. Each of these problems can be solved by solving a standard network flow problem in time $O(n^{8/3})$ (see, e.g., Tarjan (1983)). By Lemma 1.2.3, the edge connectivity of G is simply the minimum between δ and $\min_{2 \leq i \leq k} s_i$. The total time of the algorithm is $O(n^{8/3} \log n)$, as claimed.

1.3 COMBINATORICS

A *hypergraph* is a pair $H = (V, E)$, where V is a finite set whose elements are called *vertices*, and E is a family of subsets of V, called *edges*. It is *n-uniform* if each of its edges contains precisely n vertices. We say that H has *property B*, or that it is two-*colorable*, if there is a two-coloring of V such that no edge is monochromatic. Let $m(n)$ denote the minimum possible number of edges of an n-uniform hypergraph that does not have property B.

Proposition 1.3.1 [Erdős (1963a)] *Every n-uniform hypergraph with less than* 2^{n-1} *edges has property B. Therefore* $m(n) \geq 2^{n-1}$.

Proof. Let $H = (V, E)$ be an n-uniform hypergraph with less than 2^{n-1} edges. Color V randomly by two colors. For each edge $e \in E$, let A_e be the event such that e is monochromatic. Clearly, $\Pr[A_e] = 2^{1-n}$. Therefore,

$$\Pr\left[\bigvee_{e\in E} A_e\right] \leq \sum_{e\in E} \Pr[A_e] < 1$$

and there is a two-coloring without monochromatic edges. ∎

In Section 3.6 we present a more delicate argument, due to Cherkashin and Kozik (2015), which shows that

$$m(n) \geq \Omega\left(\left(\frac{n}{\ln n}\right)^{1/2} 2^n\right).$$

The best known upper bound to $m(n)$ is found by turning the probabilistic argument "on its head." Basically, the sets become random and each coloring defines an event. Fix V with v points, where we shall later optimize v. Let χ be a coloring of V with a points in one color, $b = v - a$ points in the other. Let $S \subset V$ be a uniformly selected n-set. Then

$$\Pr[S \text{ is monochromatic under } \chi] = \frac{\binom{a}{n} + \binom{b}{n}}{\binom{v}{n}}.$$

Let us assume v is even for convenience. As $\binom{y}{n}$ is convex, this expression is minimized when $a = b$. Thus

$$\Pr[S \text{ is monochromatic under} \chi] \geq p,$$

where we set

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}}$$

for notational convenience. Now let $S_1, \ldots, S_m$ be uniformly and independently chosen n-sets, with m to be determined. For each coloring χ, let A_χ be the event in which none of the S_i is monochromatic. By the independence of the S_i

$$\Pr[A_\chi] \le (1-p)^m.$$

There are 2^v colorings, so

$$\Pr\left[\bigvee_\chi A_\chi\right] \le 2^v(1-p)^m.$$

When this quantity is less than 1, there exist $S_1, \ldots, S_m$ so that no A_χ holds; that is, $S_1, \ldots, S_m$ is not two-colorable and hence $m(n) \le m$.

The asymptotics provide a fairly typical example of those encountered when employing the probabilistic method. We first use the inequality $1 - p \le e^{-p}$. This is valid for all positive p, and the terms are quite close when p is small. When

$$m = \left\lceil \frac{v \ln 2}{p} \right\rceil,$$

then $2^v(1-p)^m < 2^v e^{-pm} \le 1$ so $m(n) \le m$. Now we need to find v to minimize v/p. We may interpret p as twice the probability of picking n white balls from an urn with $v/2$ white and $v/2$ black balls, sampling without replacement. It is tempting to estimate p by 2^{-n+1}, the probability for sampling with replacement. This approximation would yield $m \sim v2^{n-1}(\ln 2)$. As v gets smaller, however, the approximation becomes less accurate and, as we wish to minimize m, the tradeoff becomes essential. We use a second-order approximation

$$p = \frac{2\binom{v/2}{n}}{\binom{v}{n}} = 2^{1-n}\prod_{i=0}^{n-1}\frac{v-2i}{v-i} \sim 2^{1-n}e^{-n^2/2v}$$

as long as $v \gg n^{3/2}$, estimating

$$\frac{v-2i}{v-i} = 1 - \frac{i}{v} + O\left(\frac{i^2}{v^2}\right) = e^{-i/v+O(i^2/v^2)}.$$

Elementary calculus gives $v = n^2/2$ for the optimal value. The evenness of v may require a change of at most 2, which turns out to be asymptotically negligible. This yields the following result of Erdős (1964):

Theorem 1.3.2 $m(n) < (1+o(1))\dfrac{e \ln 2}{4} n^2 2^n$.

Let $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ be a family of pairs of subsets of an arbitrary set. We call $\mathcal{F}$ a (k, ℓ)-*system* if $|A_i| = k$ and $|B_i| = \ell$ for all $1 \le i \le h, A_i \cap B_i = \emptyset$ and $A_i \cap B_j \ne \emptyset$ for

all distinct i, j, with $1 \le i, j \le h$. Bollobás (1965) proved the following result, which has many interesting extensions and applications:

Theorem 1.3.3 *If $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^{h}$ is a (k, ℓ)-system then $h \le \binom{k+\ell}{k}$.*

Proof. Put $X = \bigcup_{i=1}^{h}(A_i \cup B_i)$ and consider a random order π of X. For each i, $1 \le i \le h$, let X_i be the event that all the elements of A_i precede all those of B_i in this order. Clearly, $\Pr[X_i] = 1/\binom{k+\ell}{k}$. It is also easy to check that the events X_i are pairwise disjoint. Indeed, assume this is false, and let π be an order in which all the elements of A_i precede those of B_i and all the elements of A_j precede those of B_j. Without loss of generality, we may assume that the last element of A_i does not appear after the last element of A_j. But in this case, all elements of A_i precede all those of B_j, contradicting the fact that $A_i \cap B_j \neq \emptyset$. Therefore, all the events X_i are pairwise disjoint, as claimed. It follows that

$$1 \ge \Pr\left[\bigvee_{i=1}^{h} X_i\right] = \sum_{i=1}^{h} \Pr[X_i] = h/\binom{k+\ell}{k},$$

completing the proof. ∎

Theorem 1.3.3 is sharp, as shown by the family $\mathcal{F} = \{(A, X \setminus A) : A \subset X, |A| = k\}$, where $X = \{1, 2, \ldots, k+\ell\}$.

1.4 COMBINATORIAL NUMBER THEORY

A subset A of an abelian group G is called *sum-free* if $(A + A) \cap A = \emptyset$, that is, if there are no $a_1, a_2, a_3 \in A$ such that $a_1 + a_2 = a_3$.

Theorem 1.4.1 [Erdős (1965a)] *Every set $B = \{b_1, \ldots, b_n\}$ of n nonzero integers contains a sum-free subset A of size $|A| > \frac{1}{3}n$.*

Proof. Let $p = 3k + 2$ be a prime that satisfies $p > 2\max_{1 \le i \le n}|b_i|$, and put $C = \{k + 1, k + 2, \ldots, 2k + 1\}$. Observe that C is a sum-free subset of the cyclic group Z_p and that

$$\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}.$$

Let us choose at random an integer x, $1 \le x < p$, according to a uniform distribution on $\{1, 2, \ldots, p-1\}$, and define $d_1, \ldots, d_n$ by $d_i \equiv xb_i (\mathrm{mod}\ p)$, $0 \le d_i < p$. Trivially, for every fixed i, $1 \le i \le n$, as x ranges over all numbers $1, 2, \ldots, p-1$, d_i ranges over all nonzero elements of Z_p, and hence $\Pr[d_i \in C] = |C|/(p-1) > \frac{1}{3}$. Therefore, the expected number of elements b_i such that $d_i \in C$ is more than $n/3$. Consequently, there is an x, $1 \le x < p$, and a subsequence A of B of cardinality $|A| > n/3$, such that $xa (\mathrm{mod}\ p) \in C$ for all $a \in A$. This A is clearly sum-free, since, if $a_1 + a_2 = a_3$ for some $a_1, a_2, a_3 \in A$, then $xa_1 + xa_2 \equiv xa_3 (\mathrm{mod}\ p)$, contradicting the fact that C is a sum-free subset of Z_p. This completes the proof. ∎

Remark. The above proof works whenever p is a prime that does not divide any of the numbers b_i. This can be used to design an efficient deterministic algorithm for finding a sum-free subset A of size bigger than $|B|/3$ in a given set B as above. In Alon and Kleitman (1990), it is shown that every set of n nonzero elements of an arbitrary abelian group contains a sum-free subset of more than $2n/7$ elements, and that the constant $2/7$ is the best possible. For quite some time it was not clear whether or not the constant $1/3$ in Theorem 1.4.1 can be replaced by a larger constant, until Eberhard, Green and Manners (2013) proved that the constant $1/3$ is tight. The problem of deciding whether or not every set of n nonzero integers contains a sum-free subset of cardinality at least $n/3 + w(n)$, where $w(n)$ tends to infinity with n, remains open. It will be very surprising if there is no such $w(n)$.

1.5 DISJOINT PAIRS

The probabilistic method is most striking when it is applied to prove theorems whose statement does not seem to suggest at all the need for probability. Most of the examples given in the previous sections are simple instances of such statements. In this section we describe a (slightly) more complicated result, due to Alon and Frankl (1985), which solves a conjecture of Daykin and Erdős.

Let $\mathcal{F}$ be a family of m distinct subsets of $X = \{1, 2, \ldots, n\}$. Let $d(\mathcal{F})$ denote the number of disjoint pairs in $\mathcal{F}$, that is

$$d(\mathcal{F}) = |\{\{F, F'\} : F, F' \in \mathcal{F}, \quad F \cap F' = \emptyset\}|.$$

Daykin and [Erdős] conjectured that, if $m = 2^{(1/2+\delta)n}$, then for every fixed $\delta > 0$, $d(\mathcal{F}) = o(m^2)$, as n tends to infinity. This result follows from the following theorem, which is a special case of a more general result:

Theorem 1.5.1 *Let $\mathcal{F}$ be a family of $m = 2^{(1/2+\delta)n}$ subsets of $X = \{1, 2, \ldots, n\}$, where $\delta > 0$. Then*

$$d(\mathcal{F}) < m^{2-\delta^2/2}. \tag{1.1}$$

Proof. Suppose (1.1) is false; pick independently t members $A_1, A_2, \ldots, A_t$ of $\mathcal{F}$ with repetitions at random, where t is a large positive integer, to be chosen later. We will show that with positive probability $|A_1 \cup A_2 \cup \cdots \cup A_t| > n/2$ and still this union is disjoint to more than $2^{n/2}$ distinct subsets of X. This contradiction will establish (1.1).

In fact,

$$\begin{aligned} &\Pr[|A_1 \cup A_2 \cup \cdots \cup A_t| \le n/2] \\ &\le \sum_{S \subset X, |S| = n/2} \Pr[A_i \subset S, i = 1, \ldots, t] \\ &\le 2^n (2^{n/2}/2^{(1/2+\delta)n})^t = 2^{n(1-\delta t)} . \end{aligned} \tag{1.2}$$