the spectrum of L to be able to construct admissible Green's functions (which somehow reveals a certain similarity to pseudospectra). In this part of the book, the argumentation gets somewhat technical, but when the unbounded component G_∞ of the spectrum of L is simply connected, everything reduces to a study of a Riemann conformal map, and more familiar results show up. For instance, asymptotically optimal polynomials can then be obtained by means of Faber polynomials. Recall that Faber polynomials can be defined in terms of their generating function. In fact, this is one of Nevanlinna's favorite concepts throughout the entire book, and is stimulated by the resolvent, which is the generating function for the successive approximation polynomials $p_k(L) = L^k$:

$$R(\lambda, L) = (\lambda - L)^{-1} = \sum_{k=0}^{\infty} p_k(L)\lambda^{-k-1}.$$

Chapter 4 deals with the *sublinear case* which is connected to a spectral cluster at $\lambda = 1$. In this case, $I - L$ has no continuous (generalized) inverse, and the fixed point equation is usually called ill-posed (Nevanlinna calls it singular). But why does a sublinear phase occur in the general iteration? In the beginning, the iterative scheme will not "know" whether the problem is ill-posed or not; consequently, there will be no significant difference in its behavior during this phase of the iteration. Nevanlinna argues intuitively: "Imagine looking at the spectrum initially from far away. Typically, most of the spectrum will be in left from 1, maybe within a sectorial set but you don't see whether the problem is singular or not."

Sectorial enclosures for the spectrum pointing at $\lambda = 1$ indeed play a key role in the analysis of the sublinear case. For ill-posed or singular problems, the residual norm will typically converge to zero like $k^{-\alpha}$, and the actual value of α depends on two parameters: the opening angle of the sectorial set, and the "smoothness" of the solution x, i.e., whether x itself belongs to the range of some fractional power of $I - L$. There are a number of interesting results including very tight lower and upper bounds for the convergence rate in this chapter. It also includes the case when L has fixed points, that is, when $I - L$ only admits a generalized inverse (here one might have expected the author to give a reference to the Drazin inverse).

The *superlinear phase* is treated in the final chapter. It occurs when the spectrum of L is separated from $\lambda = 1$ and the optimal asymptotic convergence factor is zero, e.g., when the resolvent of L is meromorphic in $1/\lambda$. This case is examined in a number of modeling examples, e.g., when L is compact and when the eigenvalues of L (in nonincreasing absolute order) form a sequence in ℓ^p. Nevanlinna measures the rate of convergence by *order* ω and *type* τ according to whether bounds like $(\tau e\omega/k)^{k/\omega}$ hold. Note that these terms are different from what is standard in numerical analysis. Instead, they are adopted from entire function theory, namely from order and type of certain interpolating Weierstrass products, with roots at the eigenvalues of L. One of the most fascinating outcomes of this chapter consists in a lower bound for $\|p_k(L)\|$: the result, which uses a lemma by Cartan, essentially states that the optimal order ω of $\|p_k(L)\|$ is intimately connected to the convergence exponent of the spectrum of L, i.e., the infimum of all p for which the eigenvalues of L belong to ℓ^p.

This monograph indeed contains a wealth of material, and it is fascinating to see different areas of mathematics emerge. The book is very well organized, with motivating introductory sections and several examples. However, the reader would have appreciated a thorough check of all cross-references. Also, the bibliography is comparatively meager (only 55 references), and the comments at the end of each section not always clarify the origin of the many results.

The reviewer regrets that Nevanlinna did not include any numerical experiments. For example, on p. 10, the presence of the sublinear phase is supported by writing: "This can be visualized by plotting the evolution of the zeros of the residual polynomials created." Such visualizations would have perfectly fit into a concluding chapter, bringing together the manifold results, and illustrating their impact on numerical linear algebra. Such a chapter would definitely help to present the results to a larger community. Yet it is the reviewer's strong hope that the book will find a considerable audience. It definitely deserves it.

MARTIN HANKE
Universität Karlsruhe

Elements of Information Theory. *By Thomas M. Cover and Joy A. Thomas.* Wiley & Sons,

New York, 1991. xxii+542 pp. $64.95, hardback. ISBN 0-471-06259-6.

Information theory results from the fusion of practice and principles, as encapsulated so aptly in the title of the book [2]. The "practice" part of information theory has to do with the design of codes intended to process a stream of samples from a data source for transmission over a communication channel or for storage in a recording channel. The "principles" part (called Shannon theory in honor of its founder, Claude Shannon [6]) consists of a body of coding theorems detailing the performance of optimal codes for abstract source and channel models. In the information theory book by Thomas Cover and Joy Thomas, the emphasis is upon Shannon theory.

The book consists of 16 chapters. Chapter 1 is a preview of the book, in which the reader is made aware of the impact of information theory in the fields of electrical engineering, computer science, physics, mathematics, and economics. In addition, examples are given illustrating some interpretations of the fundamental concepts of entropy and mutual information in the context of applications.

In Chapter 2, identities and inequalities involving entropy and mutual information are developed for later use.

Chapter 3 is devoted to the proof and interpretation of the result called the Shannon–McMillan theorem by some and the asymptotic equipartition property by others. The Shannon–McMillan theorem is often referred to as the fundamental theorem of information theory because it is used as a tool to prove both source and channel coding theorems. The Shannon–McMillan theorem has also been an important tool for proving coding theorems of ergodic theory, such as D. Ornstein's celebrated result on the isomorphism of Bernoulli shifts.

The concept of the entropy rate of a random process is introduced in Chapter 4, and formulae are presented for the entropy rates of independent and identically distributed (i.i.d.) processes, Markov processes, and finite-state source models.

Chapter 5 presents noiseless source coding theory, including Kraft's inequality, Huffman codes, arithmetic codes, and the coding theorem equating the entropy rate of a process with the optimal rate at which it can be encoded via prefix codes without information loss.

Chapter 6, entitled Gambling and Data Compression, gives an account of the Cover-King gambling estimate of the entropy of the English language.

Chapter 7 gives a welcome and extensive treatment of the subject of Kolmogorov complexity, a measure of the descriptive complexity of objects that allows an algorithmic approach to coding theorems of information theory, as opposed to the usual model-based approach. To the reviewer's knowledge, the Cover–Thomas book is the first information theory book to devote a chapter to this important subject.

The noisy channel coding theorem is presented in Chapter 8. This theorem equates the optimal rate of reliable channel codes for a discrete memoryless channel with the maximum mutual information for channel input-output pairs (called channel capacity). In Chapters 9 and 10, the noisy channel coding theorem is extended to the case of a Gaussian channel.

Chapter 11, "Maximum Entropy and Spectral Estimation," is centered around Burg's maximum entropy theorem, which is the basis of a popular method for the estimation of spectral densities.

Chapter 12 is a tour-de-force on information theory and statistics, including the method of types, universal noiseless source coding and Lempel–Ziv codes, Sanov's theorem in large deviations, exponential bounds in hypothesis testing, Stein's lemma, the Chernoff bound, Fisher information, and the Cramér–Rao lower bound.

An introduction to rate-distortion theory is presented in Chapter 13. In the rate-distortion problem, one requires that i.i.d. data source samples be encoded block by block so that sufficiently faithful reproductions of the data blocks can be obtained from the encoded blocks. Cover and Thomas give a development of the lossy source coding theorem for this problem, in which a characterization of the optimal encoding rate is obtained in terms of the rate-distortion function of the data source. The Blahut algorithm for computation of the rate-distortion function and the Arimoto–Blahut algorithm for computation of channel capacity are elegantly obtained as special cases of the Csiszár–Tusnády alternating minimization algorithm.

The longest chapter of the book is Chapter 14 on network information theory. Although a complete theory is lacking on the rate-distortion tradeoffs that are achievable in a general network of sources and channels, the authors present a generous portion of the known results on Gaussian multiple user channels, the multiple access channel, the broadcast channel, the relay channel, Slepian–Wolf coding for correlated sources, and source coding with side information.

The book concludes with a chapter on information theory and the stock market and a chapter on inequalities in information theory.

The Cover–Thomas book has many strengths. It gives a simple, yet mathematically precise, approach to nearly all Shannon theory results that one would want to include in a first course on information theory. At the same time, the book is a pleasure to read, because of the authors' entertaining style. Good pedagogical features include end-of-chapter summaries, the naming of each problem to bring out the point that it conveys, and the avoidance of wordy proofs (many proofs consist of carefully chosen strings of inequalities for which the omitted connecting links will be self-evident to the reader).

Although *Elements of Information Theory* does a good and sometimes admirable job of treating each topic that it takes up, the book could have been improved through the addition of a small amount of material. In order to better motivate the reader for what comes later, examples of codes of all types (viz. data compaction codes, data compression codes, data transmission codes, and data translation codes to borrow the convenient classifications of Blahut [2]) should be added to Chapter 1. An extended treatment of the subject of data translation codes is needed, including Adler–Hassner codes, which are important in magnetic and optical recording system design. (The treatment of this important subject consists of just one end-of-chapter problem!) In Chapter 8, the classic result on the efficacy of linear block codes for the binary symmetric channel is not mentioned. In the historical notes section of Chapter 10, the important work of Ungerboeck on trellis-coded modulation should perhaps have been pointed out. Finally, there are no problem solutions at the end of the book.

The reviewer has successfully used the Cover–Thomas book as principal text for a one-quarter course on information theory for first-year graduate students at the University of Minnesota, based on Chapters 1–5, 7, 8, the universal coding parts of Chapters 12, and Chapter 13. A one-semester course could be based on the preceding chapters, plus Chapter 10 and parts of Chapter 14.

Cover and Thomas state that their aim is to provide "a simple and accessible book on information theory" for students in communication theory, computer science, and statistics. They have largely succeeded in their goal. For the reader seeking an introduction to modern information theory, their text is a "must read." Because of its overall excellence, *Elements of Information Theory* can be added to the honor roll of classic information theory books by Ash [1], Gallager [4], McEleice [5], Csiszár–Körner [3], and Blahut [2].

REFERENCES

[1] R. B. ASH, *Information Theory*, Interscience, New York, 1965; Dover, New York, 1990.

[2] R. E. BLAHUT, *Principles and Practice of Information Theory*, Addison-Wesley, Reading, MA, 1987.

[3] I. CSISZÁR AND J. KÖRNER, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[4] R. G. GALLAGER, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[5] R. J. MCELEICE, *The Theory of Information and Coding: A Mathematical Framework for Communication*, Addison-Wesley, Reading, MA, 1977.

[6] C. E. SHANNON, *A mathematical theory of communication*, Bell Sys. Tech. J., 27 (1948), pp. 379–423, 623–656.

JOHN KIEFFER
University of Minnesota

Some Aspects of Brownian Motion, Part 1. *By Marc Yor*, Birkhauser-Verlag, Boston, MA, 1992. 136 pp. $24.50, paperback. Lectures in Math. Series. ISBN 0-8176-2807.

The bibliography for this book lists 31 works by Yor and his coauthors out of a total of 82. Since there are 130 pages of text, if we make the reasonable assertion that the text is primarily a unified exposition of the author's papers, that leaves about four pages per paper. A rapid inspection shows that these papers average over 20 pages in length, so the present work has a compression ratio of at least five to one. Consequently, the fact alone that it is reasonably readable (at least, if one keeps a scratch pad handy) seems in itself quite remarkable.