Book Reviews

Error-Correcting Codes (2nd ed.)—W. Wesley Peterson and E. J. Weldon, Jr. (Cambridge, Mass., and London: M.I.T. Press, 1972, xi + 560 pp., $18.50).

JAMES L. MASSEY

Meeting a treasured and much-admired friend after a long separation can be a bittersweet experience. With the joy of rediscovering his old virtues is mingled the sadness of finding that they are not of the heroic proportions one had remembered. Such has been my encounter with the second edition of W. W. Peterson's coding classic. Like most coding theorists', my own education in the field came largely from the 1961 first edition, which I consider one of the most extraordinary technical books ever written. That book gave a lucid and masterful account that went to the very frontiers of, and distilled to its essence, an exciting new field. My hopes for the new edition rose through two years of premature birth notices, and were further raised by the fact that Peterson now enjoyed the collaboration of E. J. Weldon, Jr., another distinguished coding theorist (and Hawaiian colleague). Thus was I disappointed to find that this new edition was just another good book.

The book has doubled in size and contains a wealth of new results from the past decade. To name but a few of the more welcome additions: graph-theoretic codes, quadratic-residue codes, Burton burst-correcting cyclic codes, Berlekamp–Preparata–Massey and Iwadare burst-correcting convolutional codes, negacyclic codes, comma-free codes, Mandelbaum–Barrows arithmetic codes, the Elias bound on distance, the MacWilliams weight identities, threshold decoding, and Berlekamp's decoding algorithm for the BCH codes. The treatment of convolutional codes has been much expanded (and the label "recurrent codes" has been abandoned altogether). The authors generally deal with this new material deftly and with clarity even in the most difficult parts of the theory. The hand of the old master is still evident.

The novice in coding (but not the merely curious) will find this book his single best educational resource, and the old hand will surely also find much that is new and interesting. I was particularly impressed by the excellent account of threshold decoding (including the clever use of De Morgan's theorem on page 314). And after reading Chapter 8 I now understand the fuss that algebraists have made over the doubly transitive affine group of permutations.

Having said all this and attested to the general excellence of the book, I hope that I may be excused if I devote the remainder of this review to pointing out its defects.

The increased bulk of the book is not fully justified. Much fat could have been trimmed. The coverage tends to be encyclopedic rather than selective. Too much of the original has been lifted intact into the revision. Avenues which since 1961 have come to a dead end, such as the modular representation of codes, should have been excised but are found in the same place of prominence as before. In places one finds long and tedious arguments where a short and crisp one is available. As one instance, the authors present Berlekamp's BCH decoding algorithm via this reviewer's shift-register synthesis approach. Unfortunately, they rely on an early manuscript whose approach was several times longer and more awkward than that in the final paper [1]. As another instance, the "Mattson–Solomon polynomials" of Section 8.3 can be obtained simply and directly by the ordinary technique for solving difference equations applied to the recursion satisfied by codewords of a cyclic code.

Certain weaknesses of the original edition have been perpetuated. On page 216, one finds the same misleading statement on repeated roots of polynomials as in the original edition, which ignores the fact that formal derivatives of order p (the field characteristic) and greater all vanish. The Varshamov–Gilbert bound is again imprecisely formulated. Theorem 4.7 seems to be saying "given n and k, there is a $d\cdots$," to which $n = 4$ and $k = 2$ provides a counterexample. The theorem should have been stated as "given $n - k$ and d there is an $n\cdots$," which would have had the additional desirable effect of making it clear that the Varshamov bound is not the same sort of animal as the Gilbert bound. Chapter 4 on distance bounds has the same defect as its 1961 predecessor, namely that upper bounds applying to all codes are proved only for linear codes. To say, as the authors do, that a given upper bound "can also be proved for nonlinear codes" does not justify the omission and suggests misleadingly that the proof of the general bound is more difficult. In fact the authors' very succinct proof of the Elias upper bound makes no use whatsoever of linearity, and the other general upper bounds are equally accessible.

No book of this size could be entirely free of errors, and I catalog here those small errors that caught my attention. The (23,18) code used on page 110 as "an elementary example of a cyclic code" is in fact not cyclic. The usual delay operator and z transform operator are related as $D = z^{-1}$, not $D = z$ as alleged on page 189. Further on in the same paragraph it is said that "multiplication by X^{-1} delays them (the coefficients) one position. Thus in a sense $D = X^{-1}$" which is difficult to reconcile with the statement on page 174 that with premultiplication by X^r "in a sense, the output is delayed r units of time." In Theorem 8.14 "divisible evenly" means only "divisible." In the formal definition of a BCH code at the bottom of page 271, one requires the further proviso "but $\alpha^{m_0+d_0+1}$ is not a root of $g(X)$" to exclude for instance a primitive binary BCH code of length $n = 31$ with $d_0 = 9$. There seems no reason to justify the restriction "$m = m_0 = 1$" for Reed–Solomon codes that is made on page 277. The $GF(q)$ subfield subcode of a $GF(q^s)$ code with $s > 1$ is *not* a subspace of the latter and is always a proper subcode of the latter, two facts contrary to assertions appearing on page 350. The subfield subcode is of course a vector space in its own right over a different field, namely $GF(q)$.

Although I was personally delighted to see the much expanded coverage of convolutional codes in this revision, I must aim my sharpest criticism at what I consider to be a shaky and misleading treatment in many places. The statement on page 51 that "every convolutional code is equivalent to a systematic convolutional code" should have been accompanied by loud shouts that the "equivalence" is for the first constraint length only. Of the several distance measures for convolutional codes, only the "feedback decoding minimum distance" (which is the only distance measure for convolutional codes used in this book), is not impaired by the restriction to systematic codes. Most convolutional coding theorists would I think agree that the single most important parameter of a convolutional code is the minimum distance between two distinct semi-infinite encoded sequences, a distance measure for which this reviewer coined the term "free distance" some five years ago. The superiority of nonsystematic convolutional codes over systematic codes in terms of the criterion of free distance is well-known. This surprising contrast between convolutional codes and linear block codes, where systematicity entails no loss of optimality, should have been stressed. The authors fail to distinguish between the two distinct forms of "error propagation" in convolutional coding, namely "catastrophic" (which is an easily-avoidable *encoder* property equally inimical to feedback decoders, definite decoders, Viterbi decoders, and sequential decoders) and "ordinary" (which is a hard-to-avoid *decoder* property to which definite decoders are not susceptible). When they say "error propagation," the authors ordinarily mean "ordinary error propagation." Adding to this confusion is the erroneous claim on page 418 that the Viterbi decoder is not subject to error propagation since it is a definite decoder—which it most assuredly is not. In fact the Viterbi decoder for the simple (4,2) single-error-correcting code coincides with the usual feedback decoder for that code, not the definite decoder. Perhaps most misleading of all is the statement found on page 6 that "Block codes and convolutional codes

have similar error-correcting capabilities and the same fundamental limitations," which ignores the fact that convolutional codes have been proved to be better than block codes in important ways [2]. The one area in which it can be shown that convolutional codes offer no fundamental advantage (or disadvantage) *vis-a-vis* block codes is in burst-correction. Curiously, this is the one place where the authors claim superiority for convolutional codes (page 115), at least when n is large. They failed to notice that the convolutional codes required much longer guard spaces than the block codes to which they were being compared.

The heart of this book is the theory of cyclic linear codes. Thus it was surprising to find that Chapter 15, Arithmetic Codes, the last chapter in the book, did not pursue its subject from the cyclic code viewpoint. All the arithmetic codes with $AB = 2^n - 1$ are true cyclic codes, including the Brown single-error-correcting codes and the Mandelbaum–Barrows large distance codes. The authors missed a splendid opportunity to tie the material of this chapter into the remainder of the book.

Sadly, this second edition appeared just too soon to miss the most exciting development in coding theory since 1961. "There is no known coding system for which it has been proved that d/n remains nonzero as n approaches infinity with the rate k/n held fixed" (from page 100) is a precise statement of the fundamental impasse that defied the breaching efforts of coding theorists for more than 20 years until Justesen succeeded with his brilliantly simple construction [3]. This Danish discovery has opened the floodgates to a tide of new results that may well require a "third edition" of this book in the not distant future. Also too recent to be included and of lesser importance but equal interest to coding devotees was the recent demonstration that no perfect codes exist beyond those already known; the authors cannot be entirely excused for omitting the earlier contributions by van Lint that proved to be the key to this problem.

This review would be incomplete without some comparison to the two other major books in English on coding theory, namely those of Berlekamp [4] and Lin [5]. The Peterson–Weldon book is more readable but less imaginative than the Berlekamp book. Of the two, Peterson–Weldon would be a better textbook and Berlekamp would be a better companion for a long, lonely evening. Lin is less thorough and less rigorous than Peterson–Weldon, but is the easiest of all three to read and the one most suitable for the nonspecialist in coding.

References

[1] J. L. Massey, "Shift-register synthesis and BCH decoding," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 122–127, Jan. 1969.
[2] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 260–269, Apr. 1967.
[3] J. Justesen, "A class of constructive asymptotically-good algebraic codes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 652–656, Sept. 1972.
[4] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.
[5] S. Lin, *An Introduction to Error-Correcting Codes*. Englewood, N.J.: Prentice-Hall, 1970.

James L. Massey received the B.S.E.E. degree from the University of Notre Dame, Notre Dame, Ind., and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1956, 1960, and 1962, respectively.

His monograph, Threshold Decoding, *published by the M.I.T. Press, received the Paper Award of the Group on Information Theory in 1963. He is currently the Frank Freimann Professor of Electrical Engineering at the University of Notre Dame. He is a Fellow of the IEEE, has served as Chairman of the Group on Information Theory, and is currently Associate Editor for Algebraic Coding of the* IEEE TRANSACTIONS ON INFORMATION THEORY.

Estimation Theory With Applications to Communication and Control—Andrew P. Sage and James L. Melsa (New York: McGraw-Hill, 1971, 529 pp., $17.50).

K. YAO

When one sees a new technical book, it is often not easy to state whether it is a reference, research monograph, or a textbook. The book by Sage and Melsa clearly has no such classification problem. It is an introductory graduate level textbook for a course in modern estimation theory for engineering students. Despite the existence of numerous books in this area, this book represents a contribution in the teaching and popularization of this subject to an audience of the widest scope with minimal mathematical prerequisites.

Chapter 1 consists of a brief introduction and summary of each of the following eight chapters. In Chapters 2 and 3, the most basic formulas of probability theory and stochastic processes, as needed for later chapters, are stated without proofs or motivations. The material in these two chapters, with possibly the exception of 3.5, is usually covered in a standard senior-level introductory probability-stochastic process course. Thus, besides providing common notations and definitions, the true usefulness of these two chapters is not clear. Certainly, if one is not already familiar with these basic concepts, by seeing a summary of the basic equations, one will not be in a position to understand the rest of the book.

Chapter 4 first deals with elementary properties of Markov processes, including the Fokker–Planck equation. Then a heuristic treatment of stochastic differential equations driven by white noise is presented. Various comments on Wiener processes, stochastic integral and differential equations, and Itô calculus [as considered in Kailath–Frost (1968)] are given. These comments should generally be quite helpful to readers who are being exposed to this material for the first time. Finally, it concludes with a discussion of mean and variance propagation in nonlinear systems. This chapter certainly contains material not usually found in an introductory estimation theory book.

Chapter 5 deals with decision theory. The beginning sections consider briefly the usual Bayes criterion and related topics already found in many detection theory books. The last section deals with the detection of Markov signals in Gaussian noise and attempts to form a bridge between detection theory and estimation theory. It seems that the incorporation of the estimator-correlator receiver concept (e.g., Kailath (1969), etc.) could have enhanced the relevance of this section. This concept would show the reader that the material considered in Chapter 4 (e.g., the Itô calculus, etc.) is useful in detection problems.

The next three chapters form the heart of this book and can be used in a one-term elementary estimation theory course. Chapter 6 deals with basic point estimation theory. This chapter starts with standard Bayes estimation, and successively relaxes requirements on prior statistical knowledge about the parameters under estimation. Then the MAP and the ML estimators are considered. A rather detailed error analysis for these estimators when errors exist in the prior means and variances is given. This material is generally not found in standard statistical estimation theory books. This chapter also covers the linear minimum-error-variance and least-square estimators, as well as a rather brief discussion of properties of estimators and the Cramèr–Rao bound. Sixteen informative examples, ranging from the obvious to some practical system-motivated problems, are in this chapter.

Chapters 7 and 8 deal with the Kalman–Bucy filter problem, or more precisely, the linear minimum-error-variance sequential state-estimation problem. Chapter 7 is restricted to basic developments with white observation noise, while Chapter 8 deals with colored observation noise and other extensions. All the results in these two chapters are well known. Most of the basic results in Chapter 7 are derived by more than one method. On page 252, the authors state: "The justification for presenting several different developments, although one would be sufficient, since the basic results are identical, is that they illustrate the many ways in which a given estimation problem may be viewed. By approaching the problem from several avenues, it is hoped that a deeper understanding of the physical and statistical features of the results is achieved." From a pedagogic point of view, this approach adopted by the authors is reasonable. In Chapter 7, the discrete-time filter is derived by using the orthogonal principle and the MAP estimation method. The continuous-time filter is derived by a limiting argument applied to the discrete case, by using the calculus of variations method, and also by using the Wiener–Hopf equation. Finally, the stationary solution of the continuous-time filter is compared to the classical Wiener filter, and the asymptotic stability issue is also discussed. In Chapter 8 previously derived estimation algorithms for filtering, prediction, and smoothing are extended to the case of colored observation noise. The innovations approach is used for the