Research Announcement

THE MIDPOINT UPWIND SCHEME

Martin Stynes and Hans-Görg Roos

A modified upwind scheme is considered for a singularly perturbed two-point boundary value problem whose solution has a single boundary layer. The scheme is analysed on an arbitrary mesh. It is then analysed on a Shishkin mesh and precise convergence bounds are obtained, which show that the scheme is superior to the standard upwind scheme. A variant of the scheme on the same Shishkin mesh is proved to achieve even better convergence behaviour. Full details appear in [1].

Reference

[1] M. Stynes and H.-G. Roos, *The midpoint upwind scheme* (1996), Appl. Numer. Math. (to appear).

Martin Stynes
Department of Mathematics
University College
Cork

Hans-Görg Roos
Institut für Numerische Mathematik
Technische Universität Dresden
D-01062 Dresden
Germany

Book Review

Introduction to Coding Theory (second edition)

J. H. van Lint
Graduate Texts in Mathematics, Vol. 86
Springer-Verlag 1992, 183 pp.
ISBN 0-387-54894-7

Reviewed by Pat Fitzpatrick

Coding Theory will soon be 50 years old: it dates precisely back to Claude Shannon's fundamental 1948 paper, [14]. For such a young subject it has achieved a great deal, particularly in establishing connections with fundamental mathematics in a wide variety of areas encompassing group theory, finite geometries, combinatorics, number theory, algebraic geometry, algebraic function fields, computational algebra, and complexity theory. These relationships are mainly in the sense that mathematics from other areas is applied to inform the coding theory, for instance in the development of the theory of geometric Goppa codes from curves over $\mathbf{F}_q$, but there have also been some notable applications in the opposite direction, such as in the proof of the non-existence of a projective plane of order 10, [8], and in classical sphere-packing problems, [2]. Coding theory is, in essence, an area of applied mathematics, although it makes use of mathematics which has, until recently, appeared only on the "pure" syllabus. Many researchers in coding theory are engineers and many of the fundamental concerns are with specifically engineering questions such as the implementation of finite field arithmetic in logic or the complexity of decoding algorithms.

Not so the present volume! This is a book about mathematics, written for mathematicians. The presentation is condensed almost to the point of terseness, but the writing is superb, reminiscent in style of what one finds in the poet's quintessential "slim volume." The book began life as a set of lecture notes, with the

first published edition written in 1981 and one is immediately impressed by the obvious qualities of elegance and precision that must have imbued those lectures (given not only by the author but also by A. E. Brouwer, H. W. Lenstra, and H. C. A. van Tilborg, among others). It is also apparent that the audience required what the author refers to as a "fairly thorough mathematical background," in abstract algebra certainly, as well as in certain topics from number theory, probability, and combinatorics. Van Lint provides a whirlwind tour through the necessary background in the first chapter, setting up notation and quoting results without proof on algebraic structures, finite fields, combinatorics and probability, but giving a little more detail on the rather less well known theory of Krawtchouk polynomials (of which more later). He then sets out a basic five chapter course in coding theory followed by five further chapters on what he regards as important topics (and we have every reason to be convinced of the soundness of his judgement).

An $[n, M, d]$ block code C over the finite field $\mathbf{F}_q$ is a subset of size M of the n-dimensional vector space $\mathbf{F}_q^n$. In general the code is not required to have any structure, but if it forms a subspace of dimension k (so that $|M| = q^k$) then it is called an $[n, k, d]$ linear codes. The ambient space is equipped with the Hamming distance

$$d_H(u, v) = |\{i : u_i \neq v_i\}|$$

and the parameter d denotes the minimum distance between codewords in C. The value of $\log_q M/n$ (or k/n) is known as the rate of the code as it represents the rate at which information can be transmitted via an embedding $C \to \mathbf{F}_q^n$. It is easy to see that a codeword $c \in C$ sent over a noisy channel (in which errors are introduced independently of position) and received as $\hat{c}$ can be decoded uniquely to c with maximum likelihood provided that the number of errors $d_H(\hat{c}, c) \leq \lfloor \frac{d-1}{2} \rfloor$. This begs the question of whether any decoding algorithm can be carried out effectively—direct comparison of $\hat{c}$ with every codeword is of exponential complexity and therefore useless in practice. Consequently, two of the major themes in coding theory research are to define and analyse

classes of good codes having relatively large minimum distance, and to find codes for which one can construct polynomial time decoding algorithms.

Shannon's main theorem establishes the existence of good random codes that for sufficiently large values of n can be used in principle to make decoding error probability arbitrarily small (at appropriate rates). This is clearly the cornerstone of the theory and van Lint makes sure that it has a prominent place in his treatment. However, since Shannon's codes are completely unstructured, considerations of practicality form a competing requirement and it has proved difficult to find codes with practical decoding algorithms that achieve anything like the error probability promised by Shannon's theorem.

After the introduction of some analytical tools such as weight enumerators (essentially generating functions for the numbers of codewords of given weights in a code), the dual code, and the fundamental MacWilliams identities relating a weight enumerator of a code with that of its dual, some specific classes of codes are described. The ubiquitous Golay codes are included, of course. Next come Reed–Muller codes which are not as good as some, but whose advantage is that they are easy to decode. More importantly from van Lint's mathematical perspective they link coding theory with finite geometries and Boolean functions and we are introduced to the automorphism group of a code (the coordinate permutations that preserve it). A short section (added in the second edition) on Kerdock codes, which are subcodes of certain Reed–Muller codes, confirms our belief in the author's instincts, since one of the major developments of the 1990's is the discovery by Hammons *et al*, [6], that the (nonlinear, binary) Kerdock codes can be represented as images of linear codes over $\mathbf{Z}_4$.

A restricted version of the decoding problem for a code with minimum distance d is to decode up to $t = \lfloor \frac{\delta-1}{2} \rfloor$ errors for some $\delta < d$ (and record a decoding error if any received word does not lie within t errors of a codeword). One type of code for which such a bounded distance, incomplete decoding algorithm is the class of BCH codes, a subset of which is formed by the Reed-Solomon codes that are widely used in terrestrial and satellite communi-

cations, compact disks and computer disk drives. A BCH code C over $\mathbf{F}_q$ may be conveniently defined as an ideal in the polynomial algebra $A = \mathbf{F}_q[x]/(x^n - 1)$, where n and q are relatively prime to avoid repeated factors in the decomposition of $x^n - 1$ into irreducibles. Thus, C is generated by a polynomial g dividing $x^n - 1$ and this means that the code is cyclic in the sense that every cyclic shift of a codeword is again a codeword. The theory of general cyclic codes is ultimately derived from the decomposition of the semisimple algebra A as a sum of minimal ideals based on a system of orthogonal primitive idempotents. Van Lint covers this—as do most coding theory books— from first principles, without appealing to general results. The BCH theorem says that if γ is a primitive n-th root of unity in an extension of $\mathbf{F}_q$ and if g contains the consecutive set $\gamma, \gamma^2, \ldots, \gamma^{\delta-1}$ among its roots then the code C has minimum distance at least δ. Extensions of this result, proved by Hartmann and Tzeng, [5], and Roos, [13], are based on the existence of *several* consecutive sets of powers of γ being among the roots of g. The best known bound of this type was proved by van Lint and Wilson, [10]; it is included in the section on BCH codes and the earlier results are derived as special cases.

In general the problem of finding the minimum distance of a given code or class of codes (and hence their exact error correcting capability) is difficult and the determination of upper and lower bounds is another principal theme of the theory. Linearity is not assumed and the function $A(n, d)$ is defined as the maximum value of M for which an $[n, M, d]$ code exists. A code C with $|C| = A(n, d)$ is said to be optimal. The study of this function is the central problem of combinatorial coding theory and van Lint provides an overview of the known bounds. Of particular interest is the construction of classes of asymptotically good codes with parameters $[n_i, M_i, d_i]$ such that the rate k_i/n_i and the relative distance d_i/n_i are both bounded away from zero as $i \to \infty$. None of the classes used in practice (such as the BCH codes) have this property, but an outstanding example discovered by Justesen, [7], and representing a major achievement of the 1970's, has its place in van Lint's treatment. The introduction of Justesen

codes requires the author to develop the notion of concatenation of codes which is valuable in itself, since this technique (in which the codewords of an inner code are used as information vectors to a second outer code) is widely used in practice (in compact disk and deep space telemetry, for example). Also, the recent generalization to what are known as turbo–codes, [1], [4], has produced some of the potentially best performing practical coding schemes known today.

Perhaps the most significant of the distance bounds, especially in terms of motivating new research, is the Gilbert (or Gilbert-Varshamov) bound. This concerns the asymptotic rate

$$\alpha(\tfrac{d}{n}) = \lim_{n\to\infty} \sup n^{-1} \log_q A(n, \tfrac{d}{n})$$

of an optimal code and establishes the existence for $1 \leq \frac{d}{n} \leq \frac{q-1}{q}$ of codes with

$$\alpha(\tfrac{d}{n}) \geq 1 - H_q(\tfrac{d}{n}),$$

where H is the entropy function

$$H_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x).$$

For many years this bound on $\alpha(\frac{d}{n})$ was thought to be best possible, until 1982, when in a remarkable development Tsfasman, Vlădut and Zink, [17], discovered a class of codes improving the bound for $q \geq 49$. Their codes, constructed from algebraic curves over $\mathbf{F}_q$, are based on the pioneering work of Goppa in the early 1980's (see [3]) and as a consequence of their discovery there has been an enormous amount of research over the past ten years in the development of new algebraic geometry (or AG) codes and the search for efficient decoding algorithms. An alternative function field approach to these codes is the subject of a book by Stichtenoth, [15], reviewed recently in these pages by Gary McGuire, [11]. Van Lint manages to give a good flavour of the geometric ideas in just a few pages appended to the original first edition section on Reed–Solomon codes, of which the algebraic geometry codes are a natural generalization.

A notable feature of van Lint's overall treatment of coding theory is the prominent position given to the Krawtchouk polynomials. For fixed values of n and q, this class of orthogonal polynomials is defined as

$$K_k(x;n,q) = \sum_{j=0}^{k} \binom{x}{j}\binom{n-x}{k-j}(q-1)^{k-1}$$

where

$$\binom{x}{j} = \frac{x(x-1)\cdots(x-j+1)}{j!}, \quad x \in \mathbf{R}.$$

Properties of these polynomials are used in several places, such as in the analysis of weight enumerators, referred to earlier, and in the classification of perfect codes. Defining a code over a general alphabet Q rather than just over $\mathbb{F}_q$ a perfect t-error correcting code C of length n has the the property that the Hamming spheres $S_t(x) = \{c \in C | d(c,x) \leq t\}$ are disjoint and completely fill the space Q^n. It was shown by Tietäväinen, [16], and van Lint, [9], that the only nontrivial t-error correcting perfect codes with $t > 1$ and $|Q|$ a prime power are the Golay codes. In the book van Lint proves the binary case using a remarkable sufficient condition, known as Lloyd's Theorem (see [9]), that if a binary perfect t-error correcting code of length n exists then the polynomial $\Psi_t(x) = K_t(x-1; n-1, 2)$ has t distinct zeros among the integers 1, 2, ..., n. This chapter also contains a study of binary uniformly packed codes (which generalize perfect codes) using certain sequences of numbers defined from linear functionals on the group algebra $\mathbf{CF}_2^n$, as well as further properties of the Krawtchouk polynomials.

In the last two chapters of the book van Lint departs from the prevailing theme of block codes to introduce the reader to topics with radically different flavours. First there is a brief look at arithmetic codes which are used in the detection and correction of errors in ordinary arithmetic computations. Much more important from a practical point of view, convolutional codes are considered. In these codes the information sequence is potentially

infinite and the encoded stream is formed by interleaving the convolutions of the input stream with two or more finite sequences (in practice of length no more than about 10). As van Lint notes in his introduction to this chapter "the mathematical theory of convolutional codes is not well developed ...[and this is] one of the reasons that mathematicians find it difficult to become interested [in them]." But convolutional codes are widely used in practice, often concatenated with an outer Reed–Solomon code, and moreover, the well known Viterbi decoding algorithm that is used for convolutional codes also plays a significant role in getting rid of intersymbol interference in the read–write channel for computer disk drives. (Permitting such intersymbol interference in a controlled manner is essentially what has led to the enormous increases over recent years in the density of data storage.) So these codes are not only very open to mathematical analysis but also very important in view of their applications. A particularly interesting and potentially fruitful avenue is in the investigation of the automorphism groups of convolutional codes pioneered by Piret, [12], and true to form van Lint hits the right note by dealing with that aspect in a short final section of this last chapter.

Van Lint's book might almost be regarded as a collection of "edited highlights" of coding theory, in many of which he has been personally involved. One wants to read and re-read in order to fully digest and savour their excellence. There is no doubt that the reader will have to work at this book, but the rewards are handsome.

References

[1] C. Berrou, A. Glavieux and P. Thitimajshima, *Near Shannon limit error correcting coding and decoding: TURBO codes*, Proc. ICC '94 (1993), 737-740.

[2] J. H. Conway and N. J. A. Sloane, Sphere Packings, Lattices, and Groups. Springer: New York, 1988.

[3] V. D. Goppa, Geometry and Codes. Mathematics and its Applications, Vol. 24. Kluwer Academic: Dordrecht, 1988.

[4] J. Hagenauer, E. Offer and L. Papke, *Iterative decoding of binary block and convolutional codes*, IEEE Trans. Info. Thy **42** (1996), 429-445.

[5] C. R. P. Hartmann and K. K. Tzeng, *Generalizations of the BCH bound*, IEEE Trans. Info. Thy **20** (1972), 489-498.

[6] A. R. Hammons, P. V. Kumar, A. R. Calderbank, N. J. A. Sloane and P. Solé, *The $\mathbb{Z}_4$ linearity of Kerdock, Preparata, Goethals, and related codes*, IEEE Trans. Info. Thy **40** (1994), 301-319.

[7] J. A. Justesen, *A class of constructive asymptotically good algebraic codes*, IEEE Trans. Info. Thy **18** (1972), 652-656.

[8] C. W. H. Lam, *The search for a finite projective plane of order 10*, Amer. Math. Monthly **98** (1991), 305-318.

[9] J. H. van Lint, *A survey of perfect codes*, Rocky Mountain J. Math. **5** (1975), 199-224.

[10] J. H. van Lint and R. M. Wilson, *On the minimum distance of cyclic codes*, IEEE Trans. Info. Thy **32** (1986), 23-40.

[11] G. McGuire, Review of *Algebraic Function Fields and Codes* by H. Stichtenoth, Irish Math. Soc. Bulletin **30** (1994), 64-73.

[12] Ph. Piret, Convolutional Codes: An Algebraic Approach. MIT Press: Cambridge, Mass., 1988.

[13] C. Roos, *A new lower bound on the minimum distance of a cyclic code*, IEEE Trans. Info. Thy **29** (1983), 330-332.

[14] C. E. Shannon, *A mathematical theory of communication*, Bell Syst. Tec. J. **27** (1948), 379–423, 623–656.

[15] H. Stichtenoth, Algebraic Function Fields and Codes. Springer: Berlin, 1993.

[16] A. Tietäväinen, *On the nonexistence of perfect codes over finite fields*, SIAM J. Appl. Math. **24** (1973), 88-96.

[17] M. A. Tsfasman, S. G. Vlădut and Th. Zink, *On Gopppa codes which are better than the Varshamov-Gilbert bound*, Math. Nachr. **109** (1982), 21-28.

Patrick Fitzpatrick,
Department of Mathematics,
University College Cork,
email: fitzpat@ucc.ie

INSTRUCTIONS TO AUTHORS

The Bulletin is typeset with TEX. Authors should if possible submit articles to the Bulletin as TEX input files; if this is not possible typescripts will be accepted. Manuscripts are not acceptable.

Articles prepared with TEX

Though authors may use other versions of TEX, It is preferred that they write plain TEX files using the standard IMS layout files. These files can be received by sending an e-mail message to `listserv@irlearn.ucd.ie`. The body of the message should contain the three lines:

```
get imsform tex
get mistress tex
get original syn
```

Instructions on the use of these is contained in the article on *Directions in Typesetting* in issue number 27, December 1991.

The TEX file should be accompanied by any non-standard style or input files which have been used. Private macros, reference input files and both METAFONT and TEX source files for diagrams should also accompany submissions.

The input files can be transmitted to the Editor either on an IBM or Macintosh diskette, or by electronic mail to the following Bitnet or EARN address:

RODGOW@OLLAMH.UCD.IE

Two printed copies of the article should also be sent to the Editor.

Other Articles

Authors who prepare their articles with word processors can expedite the typesetting of their articles by submitting an ASCII input file as well as the printed copies of the article.

Typed manuscripts should be double-spaced, with wide margins, on numbered pages. Commencement of paragraphs should be clearly indicated. Hand-written symbols should be clear and unambiguous. Illustrations should be carefully prepared on separate sheets in black ink. Two copies of each illustration should be submitted: one with lettering added, the other without lettering. Two copies of the manuscript should be sent to the Editor.