CODING THEORY

Thomas J. Laffey

Coding Theory deals with the accurate transmission of information tthrough an imperfect (or "noisy") communication channel. The idea is to encode the message (thus enabling one to introduce several redundant (or "check" symbols), transmit the coded message and decode the received message. If the code is constructed sufficiently cleverly, then one is able to deduce from the received message, even if a "small" number of errors have occurred in transmission, what the transmitted message was.

An <u>ALPHABET</u> F is a set of symbols in which our messages are written. Usually F is taken to be a finite field and more particularly, the field Z_2 of two elements 0,1 (binary case) or Z_3 of three elements 0,1,2, (ternary case). [We recall that if F is a finite field, then F has q elements for some prime-power q and conversely for such q, there is essentially one field with q elements.] We denote by F^k the set of all k-tuples $u_1 \dots u_k$ where $u_1,\dots,u_k$ belong to F.

An (n,k) code is a (one,one) function $\mathcal{C}$ from a subset $\mathcal{M}$ of F^k to F^n. The image of $\mathcal{C}$ (that is $\{\mathcal{C}(m) | m \in \mathcal{M}\}$) is called the set of code-words and is also denoted by $\mathcal{C}$, and also referred to as the code. [$\mathcal{M}$ is the set of messages, given a message $m = u_1 \dots u_k$ $\mathcal{C}(m) = x_1 \dots x_n$ is the code-word corresponding to m.] Here n is called the length of the code.

This is the text of a lecture delivered at the Workshop on "Current Developments in Operations Research" in the College of Commerce, Dublin Institute of Technology, June 1981.

Example 1

$$F = Z_2, \mathcal{M} = F^2 \quad \text{and}$$

$$\mathcal{C}\Big\downarrow \begin{array}{cccc} 0\,0 & 0\,1 & 1\,0 & 1\,1 \\ 0\,0\,0 & 0\,1\,1 & 1\,0\,1 & 1\,1\,0 \end{array}$$

$\mathcal{C}(u_1u_2) = u_1\,u_2\,u_3$ where $u_3 \in F$ is such that $u_1+u_2+u_3 = 0$ (in F). $\mathcal{C}$ introduces a parity-check.

Given two n-tuples c_1, c_2, the (Hamming) distance $d(c_1,c_2)$ is the number of places where c_1, c_2 differ.. For example, if $c_1 = 10101$, $c_2 = 11011$, $d(c_1,c_2) = 3$. The weight $w(c)$ of an n-tuple c is the number of non-zero entries in c. Above $w(c_1) = 3$, $w(c_2) = 4$. We call an (n,k) code $\mathcal{C}$ an (n,k,d) code if d is the least distance between distinct elements of $\mathcal{C}$.

The most studied types of codes are linear codes. A (n,k) code is linear if F is a field and the set of code-words is a vector space over F (i.e. if c_1, c_2 are code-words, so is c_1+c_2 and ac_1 for $a \in F$ where code-words are added by adding corresponding entries and ac is obtained from c by multiplying all the entries of c by a). The dimension of this vector-space is called the dimension of the code. Usually in considering linear codes, we assume $\mathcal{M} = F^k$ in which case k is the dimension of $\mathcal{C}$.

Suppose $\mathcal{C}$ is a linear (n,k) code (of dimension k). Then $\mathcal{C}$ can be described by matrices in two ways. There exists a kxn matrix G such that

$$\mathcal{C}(w) = wG$$

for all $w = u_1 \ldots u_k \in F^k$. The dimension condition means that G has rank k. G is called the generating matrix for $\mathcal{C}$. We can also describe $\mathcal{C}$, the set of code-words, as $\{x = x_1 \ldots x_n \in F^n \mid Hx^T = 0\}$ (where T denotes transpose) and H is an $(n-k) \times n$ matrix of rank $n-k$. H is called the parity-check matrix of the code. The code $\mathcal{C}$ is called systematic if $\mathcal{C}(u_1 \ldots u_k) = u_1 \ldots u_k\, u_{k+1} \ldots u_n$, $(u_i \in F)$, that is the first k entries in the code-word corresponding to the message $m = u_1 \ldots u_k$ from m itself. If $\mathcal{C}$ is linear and systematic, then H is of the form

$$H = (X \mid I_{n-k})$$

where X is an $(n-k) \times k$ matrix. In this case,

$$G = (I_k \mid - X^T).$$

Example 1 is an example of a linear systematic $(3,2)$ code, $H = (1\ 1\ 1)$, $G = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ (since we are working over Z_2 in this example $-1 = 1$).

Let u be a message, $\mathcal{C}$ a code, $x = \mathcal{C}(u)$ the code word corresponding to x. Suppose that x is transmitted and that y is received. Note that $d(x,y)$ counts the number of errors. Let $e = x-y$. Then e is called the error-vector. Given y we use the code to recover x. Choose $x_0 \in \mathcal{C}$ such that $d(y,x_0)$ is least possible. We decode y as x_0. If $w(e) = t$ (say), we necessarily get $x_0 = x$ provided that $d(y,x') > t$ for all $x' \in \mathcal{C}$, $x' \neq x$. A sufficient condition for this to happen is that $d(c_1,c_2) \geq 2t+1$ for all $c_1,c_2 \in \mathcal{C}$, $(c_1 \neq c_2)$. So $\mathcal{C}$ can "correct" t errors if its minimum distance is at least $2t+1$.

Suppose that $\mathcal{C}$ is a linear code of length n and dimension k and that F has q elements. Suppose that $\mathcal{C}$ can correct t errors. For each $c \in \mathcal{C}$, the "sphere" $S(c,t) = \{x \in F^n | d(c,x) \leq t\}$ contains

$$\binom{n}{0} + \binom{n}{1}(q-1) + \ldots + \binom{n}{t}(q-1)^t$$

elements and these spheres must be disjoint. Hence

$$(*) \qquad q^n \geq q^k[\binom{n}{0} + \binom{n}{1}(q-1) + \ldots + \binom{n}{t}(q-1)^t]$$

[This is known as the Hamming or sphere-backing bound).

If equality holds, the code is called <u>perfect</u>. We now give some examples of perfect codes.

<u>Example 2</u> Let F be the field of q elements and let $H(n,q)$ be the linear code of length $\frac{q^n-1}{q-1}$, dimension $k = \frac{q^n-1}{q-1} - n$ whose parity check matrix is the $n \times (\frac{q^n-1}{q-1})$ matrix whose columns are the distinct non-zero n-tuples whose first non-zero entry is 1. It is easy to check that $H(n,q)$ has minimum distance 3 and that $(*)$ holds with $t=1$. So $H(n,q)$ is a <u>perfect</u> single error-correcting code. $H(n,q)$ is called a Hamming code.

<u>Example 3</u> Let F be a finite field and let $F[x]$ be the set of all polynomials $a_0+a_1x + \ldots + a_{n-1}x^{n-1}$. Let $g(x)$ be a divisor of x^n-1 and let $V(n,g)$ be the set of all polynomials $g(x)f(x)$ $(f(x) \in F[x])$ where we set $x^n = 1$, $x^{n+1} = x$, etc. We think of the elements $v_0+v_1x + \ldots + v_{n-1}x^{n-1} \in V(n,g)$ as n-tuples $v_0v_1 \ldots v_{n-1}$. Then the set $V(n,g)$ becomes a linear code of length n over F. This code is called a cyclic code. By judiciously choosing F,n,g, we get many

interesting codes.

For example, if $F=Z_3$, $n=11, x^{11}-1 = (x-1)g_1(x)g_2(x)$ where $g_1(x)$, $g_2(x)$ are irreducible of degree 5 over F, and taking $g(x) = g_1(x)$, we get a linear code $\mathcal{G}_{11}$ of length 11, dimension 6 and minimum distance 5. An easy calculation shows that $\mathcal{G}_{11}$ is a perfect 2-error correcting code.

If $F=Z_2$, $n=23$, $x^{23}-1 = (x-1)f_1(x)f_2(x)$ where $f_1(x), f_2(x)$ are irreducible of degree 11, and taking $g(x) = f_1(x)$ we get a linear code $\mathcal{G}_{23}$ of length 23, dimension 12 and minimum distance 7. Again, it is easy to check that $\mathcal{G}_{23}$ is a perfect 3-error correcting code. The codes $\mathcal{G}_{11}, \mathcal{G}_{23}$ by introducing a parity check are the famous Golay codes. These are the most remarkable of all codes. They arise in many combinatorial investiagations, and have intimate connections with the Conway simple groups and the recent construction of the Fischer-Griess Monster, etc. Efficient coding and de-coding procedures have been constructed for them and they are used in many communication networks.

Despite the ad-hoc nature of the construction of the perfect codes above (done initially in the 1940's) the following amazing result holds:

Theorem (van Lint, Tietavainen) Let $\mathcal{C}$ be a perfect t-error correcting code over an alphabet F of a prime-power q number of elements. Then $\mathcal{C}$ has the same parameters as $H(n,q)$, $\mathcal{G}_{11}$ or $\mathcal{G}_{23}$.

[Note: Thus $t=1,2$, or 3. If $t = 2$ or 3, then $\mathcal{C}$ is in fact a linear code isomorphic to $\mathcal{G}_{11}$ or $\mathcal{G}_{23}$. However, if $t=1$, $\mathcal{C}$ need not be linear, though it must have length $\frac{q^n-1}{q-1}$, minimum distance 3 and have the same number of elements as $H(n,q)$.]

Let $\mathcal{C}$ be a linear code of length n and dimension k over a field F with q elements. The dual code $\mathcal{C}^{\perp}$ is the set of all n-tuples $y = y_1 \dots y_n$ such that $x.y = x_1y_1 + \dots + x_ny_n = 0$ for all $x = x_1 \dots x_n \in \mathcal{C}$. Then $\mathcal{C}^{\perp}$ is a linear code of length n and dimension $n-k$.

Example 4 If $\mathcal{C}$ is the Hamming code $H(n,q)$ of Example 2, then $\mathcal{C}^{\perp}$ is a linear code of length $(q^n-1)/(q-1)$ and length n. It is called a first order Reed-Muller code. In particular, the case $q=2$, $n=5$, gives the (31,5) code which was used in transmitting information back to Earth from the Mariner space-craft to Mars.

Let $\mathcal{C}$ be a code of length n and for each $i = 0,1,\dots,n$, let w_i = number of elements $c \in \mathcal{C}$ of weight i. The polynomial

$$W(x,y) = \sum_{i=0}^{n} w_i x^i y^{n-i}$$

is called the weight-enumerator of the code. MacWilliams discovered a beautiful connection between the weight-enumerators of $\mathcal{C}$ and its dual $\mathcal{C}^{\perp}$ for a linear code $\mathcal{C}$. We state the result for the case of a binary code.

MacWilliams' Theorem If $\mathcal{C}$ is a binary linear code, then

$$W_{\mathcal{C}^{\perp}}(x,y) = \frac{1}{|\mathcal{C}|} W_{\mathcal{C}}(x+y,x-y)$$

(where $|\mathcal{C}|$ denotes the number of elements in $\mathcal{C}$).

An elementary proof of this can be obtained using the finite Fourier transform.

$\mathcal{C}$ is called <u>self-dual</u> if $\mathcal{C} = \mathcal{C}^{\perp}$.

<u>Example 5</u> $\mathcal{G}_{12}, \mathcal{G}_{24}$ are self-dual.

MacWilliams' theorem enables one to get information on the weight distribution of a self-dual code. Namely, for $\mathcal{C}$ a self-dual binary code, we have

$$W_{\mathcal{C}}(x,y) = W_{\mathcal{C}}\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right).$$

By considering the group G which leaves $W(x,y)$ invariant and the classical theory of invariants of finite groups, Gleason (and others) have found very detailed information on the structure of self-dual codes, and in particular, were able to prove the non-existence of certain codes.

<u>Example 6</u> The weight enumerator of $\mathcal{G}_{24}$ is

$$x^{24} + 759x^{16}y^{8} + 2576x^{12}y^{12} + 759x^{8}y^{16} + y^{24}.$$

There are many methods by which interesting codes are constructed and we describe here a couple of these.

A nxn matrix H, all of whose entries are ±1 is called a <u>Hadamard matrix</u> if $HH^T = nI$. It is easy to show that if H exists for $n>2$, then n must be a multiple of 4 and it is conjectured that a Hadamard matrix exists for all such n. The smallest n still in doubt is $n=268$. If H_n is a Hadamard matrix of size nxn, multiplying it on left and right by suitable diagonal matrices with entries 1 we get a Hadamard matrix with first row and first column all ones. Such a Hadamard matrix is called normalized.

Suppose H_n is a normalized Hadamard matrix. Deleting the first column of H_n and replacing all its 1's by 0's and its -1's by 1's we get a code A_n of length $n-1$, minimum distance $n/2$ and containing n codewords. Several other codes can be constructed from H_n.

A well-known result on codes is

<u>Plotkin Bound</u> Let m,d be given with $2d>m$ and let $\mathcal{C}$ be a code of length m and minimum distance d. Then $\mathcal{C}$ has at most $2[\frac{d}{2d-m}]$ code-words.

Note that the code A_n achieves this bound for $m=n$, $d=n/2$ and a result of Levensthein shows more strongly that the Plotkin bound is attainable in general if Hadamard matrices H_n exist for all n which are multiples of 4.

We now consider briefly the construction of Hadamard matrices. If H is an $n x n$ Hadamard matrix, then $\left[\begin{matrix} H & H \\ H & -H \end{matrix}\right]$ is a $2n \times 2n$ Hadamard matrix. Starting with (1) we can thus construct Hadamard matrices of size $2^k x 2^k$ $(k \geqslant 1)$.

Suppose n is a multiple of 4 and that $n = q+1$ where q is a prime power. Denote the non-zero elements of the finite field F with q elements by $1,2,\ldots,q-1$.

Let Q be the qxq matrix (indexed by the elements 0,1,...,q-1 of F) whose (i,j) entry

$$q_{ij} = 0 \text{ if } i=j$$
$$= 1 \text{ if } i\neq j \text{ and } i-j \text{ is a square in } F$$
$$= -1 \text{ if } i\neq j \text{ and } i-j \text{ is not a square in } F.$$

Let H be the nxn matrix

$$\begin{bmatrix} 1 & 1 & . & . & . & 1 \\ 1 & & & & & \\ \cdot & & & & & \\ \cdot & & & Q-I_q & & \\ \cdot & & & & & \\ 1 & & & & & \end{bmatrix}$$

Then H is a Hadamard nxn matrix. This is called the Paley construction. If we take n=12, the matrix obtained in this way is related to the generating matrix of the Golay code $\mathcal{G}_{24}$.

Recently, work of Goethals-Seidel has led to a method related to the quaternions and Turyn sequences for constructing Hadamard matrices. It is conjectured that this method will work in all cases. For details (though beware the typographical errors), see M. Hall's paper in the Proceedings of the Santa Cruz Group Theory Conference (AMS 1981).

Given a Hadamard matrix H, we can replace the entries 1, -1, by 0, 1 respectively and then consider the binary code spanned by the rows. These codes have also been studied, the ones corresponding to the Paley-type Hadamard matrices are examples of quadratic residue codes.

We conclude by giving an application of coding theory to combinatorics. Suppose $\mathcal{P}$ is a finite projective plane of order 10. [The problem of whether $\mathcal{P}$ exists is a famous, (still) open question.] Then $\mathcal{P}$ consists of 111 points, 111 lines, each point lies on exactly 11 lines, each line has exactly 11 points. Each pair of points lies on a unique line and two distinct lines meet in exactly one point. Let $A = (a_{ij})$ be the incidence matrix of $\mathcal{P}$. Thus A is an 111×111 matrix and if P_i, L_i are the points, lines, resp. of $\mathcal{P}$,

$$\begin{aligned} a_{ij} &= 1 \text{ if } P_i \text{ lies on } L_j, \\ &= 0 \text{ if } P_i \text{ does not lie on } L_j. \end{aligned}$$

Let $\mathcal{A}$ be the binary code with A as its generating matrix. It has been shown that $\mathcal{A}$ has minimum distance 11. Let $\hat{\mathcal{A}}$ be the code obtained by adding a parity check to $\mathcal{A}$. Then $\hat{\mathcal{A}}$ has minimum distance 12, length 112 and it is easy to show that $\hat{\mathcal{A}} \subseteq \hat{\mathcal{A}}^{\perp}$. Thompson has shown that in fact $\hat{\mathcal{A}}$ has dimension 56 and thus $\hat{\mathcal{A}}$ is a self-dual code. It has been shown that the weight distribution of $\hat{\mathcal{A}}$ would be known if w_{12}, w_{15} and w_{16} were known. MacWilliams, Sloane, Thompson, have shown that $w_{15} = 0$. By considering $\hat{\mathcal{A}}$ more closely, recently, Anstee, Hall, Thompson, have shown that $\mathcal{P}$ has no automorphism of order 5. This, with work of Whitesides shows that the automorphism group of $\mathcal{P}$ has order a power of 3 and is now conjectured to be trivial. So $\mathcal{P}$, if it exists, does not appear to have any of the symmetry we usually associate with a geometry.

References

For a general survey of coding theory, the best sources are the following books:

van Lint. Coding Theory (Lecture Notes in Mathematics No.201). Springer-Verlag, 1971.

McEliece. The Theory of Information and Coding (Encyclopaedia of Mathematics & its Applications, Vol.3). Addison-Wesley, 1977.

MacWilliams and Sloane. The Theory of Error-Correcting Codes. Parts I,II. North-Holland, 1977. *(This book also contains a very comprehensive set of references.)*

The book:

Shu Lin. An Introduction to Error-Correcting Codes. Prentice Hall, 1970, *discusses the problems of constructing efficient programmes to implement the various codes.*

The book:

Blake and Mullin. An Introduction to Algebraic and Combinatorial Coding Theory. Academic Press, 1976, *gives a quite concise account of coding theory with particular reference to its relation to combinatorics. It also contains a nice account of the algebraic machinery (particularly the theory of finite fields and their automorphisms) required for constructing codes.*

The work of Anstee, Hall, Thompson, *referred to above appears in* Journal of Combinatorial Theory, Series A, 29 (1980, 39-58.

The Golay Codes *are discussed at length in the* MacWilliams-Sloane *book. The material on the nonexistence of perfect codes referred to above is also available there, or in* van Lint's *paper in* Combinatorics (Proceedings of the NATO Adv. Study Institute, Breukelen, The Netherlands, 1974), Riedel Publ. Co., 1975. *This book also contains a very nice account of Gleason's and MacWilliams' theorems on weight-enumerators by* Sloane *is well as interesting papers on coding theory by* Delsarte and McEliece.