TC10 / **2. Linear codes**

S. Xambó

- Preliminaries on finite fields
- Basic notions
- Generating matrix
- Coding with a generating matrix
  – Example: RS codes
- Dual code
- Parity-check matrix
  – Example: a check matrix for RS codes
  – Hamming codes
- Syndrome decoding
- The Gilbert—Varshamov existence condition
- The MacWilliams identities

Preliminaries on finite fields

Let F be a finite field, and let q be the cardinal of F.

- q is necessarily the power of a prime number, $q = p^r$ (we say that p is the *characteristic* of F).
- If q is a power of a primer number, there is a unique field (up to isomorphism) with cardinal q. This field is denoted $\mathbb{F}_q$ or $GF(q)$.
- If p is a prime number, then $\mathbb{F}_p = \mathbb{Z}_p$
  (the field of residues mod p).
- If we select a monic *irreducible* polynomial
  $$f = X^r + a_1 X^{r-1} + \cdots + a_r \in \mathbb{Z}_p[X]$$
  then $F = \mathbb{Z}_p[X]/(f)$ is a field of cardinal p^r. In fact, setting $x = [X], 1, x, \dots, x^{r-1}$ is a basis of F as a $\mathbb{Z}_p$-vector space:
  $$F = \{\alpha = \alpha_0 + \alpha_1 x + \cdots + \alpha_{r-1} x^{r-1} | \alpha_0, \alpha_1, \dots, \alpha_{r-1} \in \mathbb{Z}_p\}.$$

- The product in F is obtained by performing the product as polynomials in x and reducing the monomials x^j $(j \geq r)$ using the relation

$$x^r = -(a_1 x^{r-1} + \cdots + a_r)$$

obtained from the fact that $f(x) = 0$ (by construction).

Examples

- $X^r + X + 1 \in \mathbb{Z}_2[X]$ is irreducible for $r = 2,3,4,6,7$. This allows us to construct $\mathbb{F}_4$, $\mathbb{F}_8$, $\mathbb{F}_{16}$, $\mathbb{F}_{64}$, $\mathbb{F}_{128}$.
- $\mathbb{F}_{32}$ can be constructed with the polynomial $X^5 + X^2 + 1$ and $\mathbb{F}_{256}$ with the polynomial $X^8 + X^7 + X^3 + X + 1$.

Basic notions

Consider codes $C \subseteq F^n$ that are vector subspaces. Such codes will be said to be *linear*.

Note that if $k = \dim_F(C)$, then k is also the dimension of C, as $|C| = q^k$ (count linear combinations of a basis of C over F).

The *weight* of an element $x \in F^n$, denoted $|x|$, is the number of non-zero entries of x, or also $hd(x, 0)$. It is a norm, as

$$|0| = 0,\ |x| > 0 \text{ if } x \neq 0, \text{ and } |x + y| \leq |x| + |y|.$$

The *minimum weight* of C, w_C, is the minimum of the weights $|x|$ for $x \in C, x \neq 0$.

Lemma. $w_C = d_C$.

Proof. If $x \in C,\ x \neq 0,$ then $|x| = hd(x, 0) \geq d_C,$ and so $w_C \geq d_C$. On the other hand, if $x, x' \in C, x \neq x'$, then

$$hd(x, x') = |x - x'| \geq w_C,$$

and hence $d_C \geq w_C$.

Remark. For a general code C of cardinal M, the determination of d_C involves the computation of the $M(M-1)/2$ Hamming distances between its pairs of distinct elements. The lemma above tells us that if C is linear then the determination of d_C involves only the computation of $M-1$ weights.

Generating matrix

Given a code C of type $[n, k]$, we will say that a matrix $G \in M_n^k(F)$ is a *generating matrix* of C if the rows of G form a linear basis of C.

Conversely, given a matrix $G \in M_n^k(F)$, the subspace $\langle G \rangle \subseteq F^n$ generated by the rows of G is a code of type $[n, k']$, where k' is the rank of G. We will say that $\langle G \rangle$ is *the code generated by* G.

Examples

a) The repetition code of length n is generated by $\mathbf{1}_n$.

b) A generating matrix for the Hamming code $[7,4,3]$ is $G = I_4|R^T$, where the columns of R are the binary vectors of length 3 of weight at least 2 (in some order).

c) If $C \subseteq F^n$ is a code of dimension k, let $C \subseteq F^{n+1}$ be the image of C by the linear map

$F^n \to F^n \times F = F^{n+1}$ such that $x \mapsto x| - s(x)$,

where $s(x) = \sum_i x_i$. Then $\bar{C}$ is a code of type $[n+1, k]$ which is called the *parity extension* of C (the symbol $-s(x)$ is called the *parity check symbol* of the vector x). If G is a generating matrix of C, then the matrix $\bar{G}$ obtained by appending to G the column consisting of the parity check symbols of its rows is a generating matrix of $\bar{C}$. The matrix $\bar{G}$ will be called the *parity completion* (or parity extension) of G.

d) The elements $x \in F^{n+1}$ such that $s(x) = 0$ form a code C of type $[n+1, n]$ (it is called the *zero-parity code* of length $n+1$, or of dimension n). Then the matrix $I_n|\mathbf{1}_n^T$ is a generating matrix of C.

Coding with a generating matrix

It is clear that if G is a generating matrix of C, then the map

$$f: F^k \to F^n,\ u \mapsto uG,$$

induces an isomorphism of F^k onto C and hence we can use f as a coding map for C.

If $A \in M_k(F)$ is an invertible matrix (in other words, $A \in GL_k(F)$), then AG is also a generating matrix of C. From this it follows that for each code C there exists an equivalent code [N1] which is generated by a matrix that has the form $G = (I_k|P)$, where I_k is the identity matrix of order k and $P \in M_{n-k}^k(F)$. Since in this case $f(u) = uG = (u|uP)$, for all $u \in F^k$, we see that the coding of u amounts to appending the vector $uP \in F^{n-k}$ to the vector u (we may think of uP as a "redundancy" vector appended to the "information vector" u).

The codes C (this time not necessarily linear) of dimension k for which there are k positions in which appear, when we let x run in C, all sequences of k symbols, are said to be *systematic* (with respect to those k positions). According to the preceding paragraph, each linear code of dimension k is equivalent to a systematic code with respect to the first k positions.

Example (Reed—Solomon codes). Let $\boldsymbol{\alpha} = \alpha_1, \ldots, \alpha_n \in F$ a sequence of n distinct elements of F $(1 \leq n \leq q)$.

For every integer $k > 0$, let $F[X]_k$ be the F-vector space whose elements are polynomials of degree $< k$ with coefficients in F. We have

$$F[X]_k = \langle 1, X, \ldots, X^{k-1} \rangle_F, \ \dim(F[X]_k) = k.$$

If $k \leq n$, the map

$$\varepsilon : F[X]_k \to F^n, \ f \mapsto (f(\alpha_1), \ldots, f(\alpha_n))$$

is injective, since the existence of a non-zero polynomial of degree $< k$ vanishing on all the α_i implies $n \leq k - 1$. The image of ε is therefore a linear code C of type $[n, k]$.

Proposition. The minimum distance of C is $n - k + 1$.

Proof. Indeed, a non-zero polynomial f of degree $< k$ can vanish on at most $k - 1$ of the elements α_i and hence the weight of $(f(\alpha_1), \ldots, f(\alpha_n))$ is not less than $n - (k - 1) = n - k + 1$. Thus

$d_C = w_C \geq n - k + 1$. On the other hand, $d_C \leq n - k + 1$ by the Singleton bound. Notice that it is also easy to produce an f such that $\varepsilon(f)$ has weight $n - k + 1$, as for instance $f = (X - \alpha_1) \cdots (X - \alpha_{k-1})$.

We will say that C is a Reed—Solomon (RS) code of length n and dimension k, and will be denoted $RS_{\boldsymbol{\alpha}}(k)$.

When the α_i can be understood from the context, we will simply write $RS(k)$, or $RS(n, k)$ if we want to display the length n.

It is clear that the RS codes satisfy the equality in the Singleton bound, and so they are examples of MDS codes. On the other hand we have $n \leq q$, and so we will have to take high values of q to obtain interesting codes.

Since $1, X, \ldots, X^{k-1}$ is a basis of $F[X]_k$, the *Vandermonde matrix*

$$V_k(\alpha_1, \ldots, \alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_n^{k-1} \end{pmatrix}$$

is a generating matrix for $RS_{\boldsymbol{\alpha}}(k)$. Notice that

$$V_k(\alpha_1, \ldots, \alpha_n) = \left(\alpha_i^j\right),\ 1 \leq i \leq n,\ 0 \leq j < k$$

Dual code

The linear subspace of F^n orthogonal to a subset $Z \subseteq F^n$ will be denoted $Z^\perp$.

Let us recall that $Z^\perp = \{x \in F^n | \langle x|z\rangle = 0 \text{ for all } z \in Z\}$, where

$$\langle x|z\rangle = x_1 z_1 + \cdots + x_n z_n \, .$$

If C is a code, the code $C^\perp$ is called the *dual code* of C.

Since the scalar product $\langle \, | \, \rangle$ is non-degenerate, by linear algebra we know that $C^\perp$ has dimension $n - k$ if C has dimension k. In other words, $C^\perp$ is of type $[n, n-k]$ if C is of type $[n, k]$.

As $C \subseteq C^{\perp\perp}$, tautologically, and both sides of this inclusion have dimension k, we infer that $C^{\perp\perp} = C$.

It often happens that C and $C^{\perp}$ have a non-zero intersection. Even more, it can happen that $C \subseteq C^{\perp}$, including $C = C^{\perp}$. In the latter case we say that C is *self-dual*.

Note that in order to be self-dual it is necessary that $n = 2k$.

Example. If C is the length n repetition code over C, $C^{\perp}$ is the zero-parity code of length n.

Example. If G is a generating matrix of C, then $C \subseteq C^{\perp}$ is equivalent to the relation $GG^T = 0$. If in addition we have $n = 2k$, the relation $GG^T = 0$ is equivalent to $C = C^{\perp}$.

As an application, check that the parity extension of the Hamming code $[7,4,3]$ is a self-dual code.

Parity-check matrix

If H is a generating matrix of $C^{\perp}$ (in which case H is an $(n-k)\times n$ matrix) we have that

$$x \in C \text{ if and only if } xH^T = 0,$$

because if h is a row of H we have

$$xh^T = \langle x|h\rangle.$$

Said in other words, the elements $x \in C$ are exactly those that satisfy the $n-k$ linear equations $\langle x|h\rangle = 0$, where h runs through the rows of H.

Given $h \in C^{\perp}$, the linear equation $\langle x|h\rangle = 0$, which is satisfied for all $x \in C$, is called the *check equation* of C corresponding to h. By the previous paragraph, C is determined by the $n-k$ check equations corresponding to the rows of H, and this is why the matrix H is said to be a *check matrix*, or *control matrix*, of C.

The relation

$$C = \{x \in F^n | xH^T = 0\}$$

can also be interpreted by saying that

C is the set of linear relations satisfied by the rows of H^T,

that is, by the columns of H.

In particular we have the following prescription for finding de minimum distance of a linear code in terms of a check matrix:

Proposition. *If any $r-1$ columns of H are linearly independent, then the minimum distance of C is at least r, and conversely.*

Example (A check matrix for RS codes)

Let $C = RS_{\boldsymbol{\alpha}}(k)$, where $\boldsymbol{\alpha} = \alpha_1, \ldots, \alpha_n$ are distinct nonzero elements of a finite field F. Then we know that

$$G = V_k(\alpha_1, \ldots, \alpha_n)$$

is a generating matrix of C (p. 12). Note that the rows of G have the form

$$\left(\alpha_1^i, \ldots, \alpha_n^i\right), \text{ with } \; i = 0, \ldots, k-1.$$

Now we are going to describe a check matrix H of C, that is, a generating matrix of $C^{\perp}$.

Recall that if we define the *Vandermonde determinant* as the determinant $D(\alpha_1, \ldots, \alpha_n)$ of the Vandermonde matrix $V_n(\alpha_1, \ldots, \alpha_n)$, then $D(\alpha_1, \ldots, \alpha_n) = \prod_{i<j}(\alpha_j - \alpha_i)$.

Define the vector

$$\boldsymbol{h} = (h_1, \ldots, h_n)$$

by the formula

$$h_i = (-1)^{i-1} D(\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n) / D(\alpha_1, \dots, \alpha_n)$$
$$= 1 / \prod_{j \neq i} (\alpha_j - \alpha_i).$$

(Remark that in the last product there are precisely $i - 1$ factors with the indices in reverse order, namely the $\alpha_j - \alpha_i$ with $j < i$).

Theorem. The matrix

$$H = V_{n-k}(\alpha_1, \dots, \alpha_n) \mathrm{diag}(h_1, \dots, h_n)$$

is a check matrix of C.

Proof. To see this, it is enough to show that any row of G is orthogonal to any row of H, because H clearly has rank $n - k$.

Since the rows of H have the form

$$(h_1 \alpha_1^j, \dots, h_n \alpha_n^j), \text{ with } j = 0, \dots, n - k - 1,$$

we have to establish that

$$\sum_{l=1}^{n} \alpha_l^{i+j} h_l = 0 \text{ for } 0 \leq i \leq k-1 \text{ and } 0 \leq j \leq n-k-1.$$

Thus it will be enough to make sure that

$$\sum_{l=1}^{n} \alpha_l^{s} h_l = 0 \text{ for } 0 \leq s \leq n-2.$$

To see this, multiply throughout by the (nonzero) determinant

$$D(\alpha_1, \dots, \alpha_n).$$

Taking into account the definition of h_l, we wish to derive that

$$\sum_{l=1}^{n} \alpha_l^{s} (-1)^{l-1} D(\alpha_1, \dots, \alpha_{l-1}, \alpha_{l+1}, \dots, \alpha_n) = 0.$$

But finally this is obvious, because the left hand side coincides with the determinant

$$\begin{vmatrix} \alpha_1^s & \cdots & \alpha_n^s \\ 1 & \cdots & 1 \\ \alpha_1 & \cdots & \alpha_n \\ \vdots & & \vdots \\ \alpha_1^{n-2} & \cdots & \alpha_1^{n-2} \end{vmatrix}$$

(developed along the first row), and this determinant has a repeated row.

Remark. As we will see in a later chapter, the form of the matrix H indicates that $RS_{\boldsymbol{\alpha}}(k)$ is an alternant code, and as a consequence it will be decodable with any of the fast decoders for alternant codes studied in that chapter.

Example (Hamming codes)**.** **A** q-ary code C of type $[n, n-r]$ is said to be a *Hamming code of codimension* r if the columns of a check matrix $H \in M_n^r(F)$ of C form a maximal set among the subsets of F^r with the

property that any two of its elements are linearly independent (of such a matrix we will say that it is a *q-ary Hamming matrix of codimension r*).

There is a straightforward way of constructing such a matrix. For $i = 1, \dots, r$, let H_i be the matrix whose columns are the vectors of F^r that have the form $(0, \dots, 0, 1, \alpha_{i+1}, \dots, \alpha_r)$, with $\alpha_{i+1}, \dots, \alpha_r \in F^r$ arbitrary. Let $H = H_1 | H_2 | \cdots | H_r$. Then H is a q-ary Hamming matrix of codimension r, as any non-zero vector of length r is proportional to exactly one of the columns of H (we will say that such an H is a *normalized* q-ary Hamming matrix of codimension r). Since H_i has q^{r-i} columns, H has $n = (q^r - 1)/(q - 1)$ columns.

It is clear that two Hamming codes of the same codimension are scalar equivalent. We will write $\mathrm{Ham}_q(r)$ to denote any one of them and $\mathrm{Ham}_q^{\perp}(r)$ to denote the corresponding dual code, that is to say, the code generated by the check matrix H used to define $\mathrm{Ham}_q(r)$. The code $\mathrm{Ham}_q(r)$ has dimension

$$k = n - r = (q^r - 1)/(q - 1) - r\ .$$

Its codimension, which is the dimension of $\mathrm{Ham}_q^{\perp}(r)$, is equal to r.

The binary Hamming code of codimension 3, $\mathrm{Ham}_2(3)$, is the code [7,4] defined by (say) the normalized check matrix

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Proposition. *If C is any Hamming code, then $d_C = 3$. In particular, the error-correcting capacity of a Hamming code is* 1.

Proof. If H is a check matrix of C, then we know that the elements of C are the linear relations satisfied by the columns of H. Since any two columns of H are linearly independent, the minimum distance of C is at least 3. On the other hand, C has elements of weight 3, because the sum

of two columns of H is linearly independent of them and hence it must be proportional to another column of H (the columns of H contain all non-zero vectors of F^r up to a scalar factor).

Corollary. *The Hamming codes are perfect* (cf. **P**7).

Proposition. *If $C' = Ham_q^{\perp}(r)$ is the dual of a Hamming code $C = Ham_q(r)$, the weight of any non-zero element of C' is q^{r-1}. In particular, the distance between any pair of distinct elements of C' is q^{r-1}.*

Proof. Let $H = (h_{ij})$ be a check matrix of C. Then the non-zero vectors of C' have the form $z = aH$, $a \in F^r$, $a \neq 0$. So the i-th component of z has the form

$$z_i = a_1 h_{1i} + \cdots + a_r h_{ri} \ .$$

Therefore the condition $z_i = 0$ is equivalent to say that the point P_i of $\mathbb{P}^{r-1} = \mathbb{P}(F^r)$ defined by the i-th column of H belongs to the hyper-plane of $\mathbb{P}^{r-1}$ defined by the equation

$$a_1x_1 + \cdots + a_rx_r = 0\ .$$

Since $\{P_1, \ldots, P_n\} = \mathbb{P}^{r-1}$, it follows that the number of non-zero components of z is the cardinal of the complement of a hyperplane of $\mathbb{P}^{r-1}$. Since this complement is an affine space $\mathbb{A}^{r-1}$, its cardinal is q^{r-1} and so any non-zero element of C' has weight q^{r-1}. □

Codes such that the distance between pairs of distinct elements is a fixed integer d are called *equidistant* of distance d. Thus $\mathrm{Ham}_q^{\perp}(r)$ is equidistant of distance q^{r-1}.

Syndrome-leader decoding

Let C by an F-linear $[n,k]$ code and H a control matrix of C. Given $y \in F^n$, the element $s = yH^T \in F^{n-k}$ is called *syndrome* of y (with respect to H).

Since $C = \{x \in F^n | xH^T = 0\}$, the elements if C are precisely those whose syndrome is 0. If we set $C_s = \{x \in F^n | xH^T = s\}$ for any $s \in F^{n-k}$, then in particular we have $C = C_0$.

As the linear map $s : F^n \to F^{n-k}$, $y \mapsto yH^T$, is surjective (for H has rank $n-k$) and its kernel is $C_0 = C$, we see that $C_s \neq \emptyset$. Moreover, if we pick any $z_s \in C_s$, then $C_s = z_s + C$. In other words, C_s is a coset modulo C, and we say that C_s is *the coset of the syndrome s*).

Lemma. If $2t < d$, C_s contains at most one vector e such that $|e| \leq t$.

For if $e' \in C_s$ and $|e'| \leq t$, then $x = e - e' \in C_0 = C$, $|x| < d$. So $x = 0$.

Let g be the minimum distance decoder of C, $x \in C$ (sent vector) and $y \in F^n$ (received vector). Then we have, if t is the correcting capacity of C :

Proposition. *The vector y is g-decodable if and only if there exists $e \in F^n$ such that $yH^T = eH^T$ and $|e| \leq t$. If this is the case, then e is unique* (by the lemma) *and $g(y) = y - e$.*

Proof. If y is g-decodable, let $x' = g(y) \in C$ and $e = y - x'$. Then

$$yH^T = (x' + e)H^T = eH^T \text{ and } |e| = hd(x', y) \leq t.$$

Conversely, let $e \in F^n$ be such that $yH^T = eH^T$ and $|e| \leq t$. Let $x' = y - e$. Then $x'H^T = yH^T - eH^T = 0$, so $x' \in C$, and

$$hd(x', y) = |y - x'| = |e| \leq t,$$

which shows that y is g-decodable and that $g(y) = x' = y - e$. □

The proposition above suggests the following decoding scheme. First form a table

$$E = \{s \to e_s\}_{s \in F^{n-k}}, \text{ with } e_s \in C_s \text{ of minimum weight}$$

(The element e_s is said to be a *leader* of the class C_s and E a *leader's table* for C. Note that e_s is unique if $|e_s| \leq t$; otherwise it has to be selected among the vectors of C_s that have minimum weight. Now the *syndrome decoder* can be described as follows:

SYNDROME–LEADER DECODER (SLD)
Input: $y \in F^n$ [received vector]
1. Set $s = yH^T$ [the syndrome of y]
2. Let $e = E(s)$ [the leader of C_s]
3. Return $y - e$.

Proposition. *If* $D = \coprod_{x \in C} B(x, t)$ *is the set of vectors that are decodable by the minimum distance decoder* g*, then* SLD coincides with g for all $y \in D$.

Proof. It is an immediate consequence of the previous proposition.

The Gilbert—Varshamov existence condition

Theorem. Fix positive integers

$$n, k, d \text{ such that } k \leq n \text{ and } 2 \leq d \leq n + 1.$$

If the relation

$$\mathrm{vol}_q(n-1, d-2) < q^{n-k}$$

is satisfied (this is called the Gilbert—Varshamov condition), then there exists a linear code of type $[n, k, d']$ with $d' \geq d$.

Proof. A) It is sufficient to see that the condition allows us to construct a matrix $H \in M_n^{n-k}(F)$ of rank $n-k$ with the property that any $d-1$ of its columns are linearly independent. Indeed, if this is the case, then the code C defined as the orthogonal space of the rows of H has length n, dimension k (from $n-(n-k)=k$), and minimum weight $d' \geq d$ (since there are no linear relations of length less than d among the columns of H, which is a check matrix of C).

B) Before constructing H, we first establish that the Gilbert–Varshamov condition implies that $d-1 \leq n-k$ (note that this is the Singleton bound for the parameters d and k).

Indeed,

$$\begin{aligned} \mathrm{vol}_q(n-1, d-2) &= \sum_{j=0}^{d-2} \binom{n-1}{j} (q-1)^j \\ &\geq \sum_{j=0}^{d-2} \binom{d-2}{j} (q-1)^j \\ &= \big(1+(q-1)\big)^{d-2} = q^{d-2} \end{aligned}$$

and hence $q^{d-2} < q^{n-k}$ if the Gilbert–Varshamov condition is satisfied. Therefore $d-2 < n-k$, or $d-1 \leq n-k$.

C) In order to construct H, we first select a basis $c_1, \ldots, c_{n-k} \in F^{n-k}$. Since $d-1 \leq n-k$, any $d-1$ vectors extracted from this basis are li-

nearly independent. Also, any matrix H of type $(n-k)\times m$ with entries in F which contains the columns c_j^T $(j=1,\dots,n-k)$, has rank $n-k$.

Now assume that we have constructed, for some $i \in [n-k, n]$, vectors $c_1,\dots,c_i \in F^{n-k}$ with the property that any $d-1$ of them

are linearly independent. If $i=n$, it is sufficient to take $H=(c_1^T,\dots,c_n^T)$ and by part A our question is answered.

Otherwise we will have $i<n$. In this case, the number of linear combinations that can be formed with at most $d-2$ vectors from among $c_1,\dots,c_i$ is not greater than $\mathrm{vol}_q(i,d-2)$ (see **P**6). Since $i \le n-1$, $\mathrm{vol}_q(i,d-2) \le \mathrm{vol}_q(n-1,d-2)$. If the Gilbert–Varshamov condition is satisfied, then there is a vector $c_{i+1} \in F^{n-k}$ which is not a linear combination of any subset of $d-2$ vectors extracted from the list $c_1,\dots,c_i$, and our claim follows by induction. □

Remark. Since $A_q(n,d) \geq A_q(n,d')$ if $d' \geq d$, the Gilbert--Varshamov condition shows that $A_q(n,d) \geq q^k$. This lower bound, called the bert–Varshamov bound, often can be used to improve the Gilbert bound.

Example. Gilbert(10,3)=19, as in this case

$$q^n/\mathrm{vol}_q(n,d-1) = 2^{10}/\mathrm{vol}_2(10,2) = 2^{10}/56 \simeq 18.3\,.$$

And GilbertVarshamov(10,3)=64, as $\mathrm{vol}_2(9,1) = 10$ and $10 < 2^{10-6} = 16$ while $2^{10-7} = 8 < 10$.

The MacWilliams identities

The *weight enumerator* of a code C is defined as the polynomial

$$A(t) = \sum_{i=0}^{n} A_i t^i,$$

where A_i is the number of elements of C that have weight i.

It is clear that $A(t)$ can be expressed in the form

$$A(t) = \sum_{x \in C} t^{|x|}$$

for the term t^i appears in this sum as many times as the number of solutions of the equation $|x| = i, x \in C$.

Remark. $A_0 = 1$, since $\mathbf{0}_n$ is the unique element of C with weight 0. On the other hand $A_i = 0$ if $0 < i < d$, where d is the minimum distance of C. The determination of the other A_i is not easy in general and it is one of the basic problems of coding theory.

Theorem (F. J. MacWilliams).[1] *The weight enumerator $B(t)$ of the dual code $C^{\perp}$ of a code C of type $[n, k]$ can be determined from the weight enumerator $A(t)$ of C according to the following identity:*

$$q^k B(t) = (1 + (q-1)t)^n A\left(\frac{1-t}{1+(q-1)t}\right).$$

Remark. In the proof of the theorem we need the notion of *character* of a group G, which by definition is a group homomorphism of G to $U(1) = \{z \in \mathbb{C} | \, |z| = 1\}$, the group of complex numbers of modulus 1.

The constant map $g \mapsto 1$ is a character, called the *unit character*. A character different from the unit character is said to be *non-trivial* and the main fact we will need is that the additive group of finite field $\mathbb{F}_q$ has non-trivial characters. Actually any finite abelian group has non-trivial characters. Let us sketch how this can be established.

It is known that any finite abelian group G is isomorphic to a product of the form

$$\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$$

with k a positive integer and $n_1, \ldots, n_k$ integers greater than 1. For example, in the case of a finite field $\mathbb{F}_q$, we have (if $q = p^r$, p prime), $\mathbb{F}_q \cong \left(\mathbb{Z}_p\right)^r$, since $\mathbb{F}_q$ is a vector space of dimension r over $\mathbb{Z}_p$.

In any case, it is clear that if we know how to find a non-trivial character of $\mathbb{Z}_{n_1}$, then we also know how to find a non-trivial character of G (the composition of the non-trivial character of $\mathbb{Z}_{n_1}$ with the projection of $\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$ onto $\mathbb{Z}_{n_1}$ gives a non-trivial character of G).

Finally note that if n is an integer greater than 1 and $\xi \neq 1$ is an n-th root of unity then the map $\chi : \mathbb{Z}_n \to U(1)$ such that $\chi(k) = \xi^k$ is well defined and is a non-trivial character of $\mathbb{Z}_n$.

Proof. Let χ be a non-trivial character of the additive group of $F = \mathbb{F}_q$ (see the preceding Remark). Thus we have a map

$$\chi : F \to U(1)$$

such that $\chi(\alpha + \beta) = \chi(\alpha)\chi(\beta)$ for any $\alpha, \beta \in F$ with $\chi(\gamma) \neq 1$ for some $\gamma \in F$. Observe that $\sum_{\alpha \in F} \chi(\alpha) = 0$, since

$$\chi(\gamma) \sum_{\alpha \in F} \chi(\alpha) = \sum_{\alpha \in F} \chi(\alpha + \gamma) = \sum_{\alpha \in F} \chi(\alpha)$$

and $\chi(\gamma) \neq 1$.

Now consider the sum

$$S = \sum_{x \in C} \sum_{y \in F^n} \chi(x, y) t^{|y|}$$

where $\chi(x, y) = \chi(\langle x|y \rangle) = \prod_{i=1}^{n} \chi(x_i y_i)$. After reordering we have

$$S = \sum_{y \in F^n} t^{|y|} \sum_{x \in C} \chi(x, y)$$

If $y \in C^{\perp}$, then $\langle x|y \rangle = 0$ for all $x \in C$ and so $\chi(\langle x|y \rangle) = 1$ for all $x \in C$, and in this case $\sum_{x \in C} \chi(x, y) = |C| = q^k$.

If $y \notin C^{\perp}$, then the linear map $C \to F$ such that $x \mapsto \langle x|y \rangle$ takes each value of F the same number of times, say r, and hence, using that $\sum_{\alpha \in F} \chi(\alpha) = 0$, we have that $\sum_{x \in C} \chi(x, y) = r \sum_{\alpha \in F} \chi(\alpha) = 0$. Putting the two cases together we have that

$$S = |C| \sum_{y \in C^{\perp}} t^{|y|} = q^k B(t).$$

On the other hand, for any given $x \in C$, and making the convention, for all $\alpha \in F$, that $|\alpha| = \begin{cases} 1 \text{ if } \alpha \neq 0 \\ 0 \text{ if } \alpha = 0 \end{cases}$, we have

$$\sum_{y\in F^n}\chi(x,y)t^{|y|}=\sum_{y_1,..,y_n\in F}\chi(x_1y_1+\cdots+x_ny_n)t^{|y_1|+\cdots+|y_n|}$$
$$=\sum_{y_1,..,y_n\in F}\prod_{i=1}^{n}\chi(x_iy_i)\,t^{|y_i|}$$
$$=\prod_{i=1}^{n}\Big(\sum_{\alpha\in F}\chi(x_i\alpha)\,t^{|\alpha|}\Big).$$

But $\sum_{\alpha\in F}\chi(x_i\alpha)\,t^{|\alpha|}=\begin{cases}1+(q-1)t & \text{if } x_i=0\\ 1-t & \text{if } x_i\neq 0\end{cases}$

because $\sum_{\beta\in F^*}\chi(\beta)=-1$. Consequently

$$\prod_{i=1}^{n}\sum_{\alpha\in F}\chi(x_i\alpha)\,t^{|\alpha|}=(1+(q-1)t)^n\left(\frac{1-t}{1+(q-1)t}\right)^{|x|}$$

and summing with respect to $x\in C$ it is clear that

$$S=(1+(q-1)t)^nA\left(\frac{1-t}{1+(q-1)t}\right).$$ □

Example (Weight enumerator of the zero-parity code). The zero-parity code of length n is the dual of the repetition code C of length n. Since the weight enumerator of C is $A(t)=1+(q-1)t^n$, the weight enumerator $B(t)$ of the zero-parity code is given by the relation

$$qB(t) = (1+(q-1)t)^n \left(1+(q-1)\frac{1-t}{1+(q-1)t}\right)^n$$
$$= (1+(q-1)t)^n + (q-1)(1-t)^n$$

In the binary case we have

$$2B(t) = (1+t)^n + (1-t)^n,$$

and therefore

$$B(t) = \sum_{i=0}^{i\leq n/2} \binom{n}{2i} t^{2i}.$$

Note that this could have been written directly, for the binary zero-parity code of length n has only even-weight words and the number of those having weight $2i$ is $\binom{n}{2i}$.

Example (Weight enumerator of the Hamming codes). We know that the dual Hamming code $\mathrm{Ham}_q^{\perp}(r)$ is *equidistant*, with minimum distance q^{r-1}. This means that its weight enumerator, say $B(t)$, is the polynomial

$$B(t) = 1 + (q^r - 1)t^{q^{r-1}},$$

because $q^r - 1$ is the number of non-zero vectors in $\mathrm{Ham}_q^{\perp}(r)$ and each of these has weight q^{r-1}. Now the MacWilliams identity allows us to determine the weight enumerator $A(t)$ of $\mathrm{Ham}_q(r)$. Setting $\hat{q} = q - 1$, we have:

$$\begin{aligned}
q^r A(t) &= (1 + \hat{q}t)^n B\left(\tfrac{1-t}{1+\hat{q}t}\right) \\
&= (1 + \hat{q}t)^n + (q^r - 1)(1 + \hat{q}t)^n \tfrac{(1+t)^{q^{r-1}}}{(1+\hat{q}t)^{q^{r-1}}} \\
&= (1 + \hat{q}t)^n + (q^r - 1)(1 + \hat{q}t)^{n-q^{r-1}}(1 + t)^{q^{r-1}} \\
&= (1 + \hat{q}t)^{n-q^{r-1}}\left((1 + \hat{q}t)^{q^{r-1}} + (q^r - 1)(1 - t)^{q^{r-1}}\right) \\
&= (1 + \hat{q}t)^{\frac{q^{r-1}-1}{q-1}}\left((1 + \hat{q}t)^{q^{r-1}} + (q^r - 1)(1 - t)^{q^{r-1}}\right)
\end{aligned}$$

since $n = (q^r - 1)/(q - 1)$ and $n - q^{r-1} = (q^{r-1} - 1)/(q - 1)$.

In the binary case the previous formula yields that the weight enumerator $A(t)$ of $\mathrm{Ham}(r)$ satisfies

$$2^r A(t) = (1+t)^{(2^{r-1}-1)}\left((1+t)^{(2^r-1)} + (2^r-1)(1-t)^{2^{r-1}}\right)$$

For $r = 3$, $\mathrm{Ham}(3)$ has type $[7,4]$ and

$$8A(t) = (1+t)^3((1+t)^4 + 7(1-t)^4),$$

$$A(t) = 1 + 7t^3 + 7t^4 + t^7.$$

So $A_0 = A_7 = 1,\ A_3 = A_4 = 7,\ A_1 = A_2 = A_5 = A_6 = 0.$

Actually it is easy to find, using the description of this code, that the weight 3 vectors of $\mathrm{Ham}(3)$ are

[1,1,0,1,0,0,0], [0,1,1,0,1,0,0], [1,0,1,0,0,1,0], [0,0,1,1,0,0,1],

[1,0,0,0,1,0,1], [0,1,0,0,0,1,1], [0,0,0,1,1,1,0]

and the weight 4 vectors are

[1,1,1,0,0,0,1], [1,0,1,1,1,0,0], [0,1,1,1,0,1,0], [1,1,0,0,1,1,0],

[0,1,0,1,1,0,1], [1,0,0,1,0,1,1], [0,0,1,0,1,1,1]

Note that the latter are obtained from the former, in reverse order, by interchanging 0 and 1.

Notes

1. http://www.awm-math.org/noetherbrochure/TOC.html

Notes

N1. Two codes $C, C' \subseteq T^n$ of length n are said to be *equivalent* if there is a permutation σ of $1, \dots, n$ and permutations $\tau_1, \dots, \tau_n$ of T such that the map $\rho : T^n \to T^n$ given by

$$(x_1, \dots, x_n) \mapsto \left(\tau_1\left(x_{\sigma(1)}\right), \dots, \tau_n\left(x_{\sigma(n)}\right)\right)$$

induces a 1-to-1 map between C and C'. Note that equivalent codes have the same parameters, as $hd(x, y) = hd(\rho x, \rho y)$ for any $x, y \in C$. If we can choose $\tau_1 = \cdots = \tau_n = Id$, then C and C' are called *permutationally equivalent*. If T is a finite field and

$$\tau_j(t) = \alpha_j t$$

for some non-zero α_j, then we say that C and C' are *scalarly equivalent*.