ASPECTS OF CODING THEORY

SUDHIR R. GHORPADE

Contents

Coding Theory has its origins in the problem of information transmission across what is called a noisy channel. A solution is found by encoding the message by suitably adding redundancies in such a way that errors can be detected and corrected. In effect, the message may consist of (up to) k symbols from a finite alphabet set Q (for example, $Q = \{0, 1\}$) and encoding is simply an injective map $Q^k \to Q^n$, where $n \geq k$. The image of this map thus consists of the codewords, and these constitute an error correcting code or simply, a code of length n. When Q is a finite field and the encoding is given by a linear map, the corresponding code is said to be linear. In these lectures, we shall focus mainly on certain mathematical aspects of the theory of error correcting codes, especially linear codes which are the most widely studied classes of codes. For more on information theoretic aspects and the origins of coding theory, we refer to the first section of [14, Ch. 1] and the references therein. It may also suffice to glance at the table of contents and over 2000 pages of this *Handbook* [14] or the 750 page treatise of MacWilliams and Sloane [13] which is of an older vintage, to have some idea of the expanse the subject and deduce an obvious corollary that these notes are only a selective, and not comprehensive, introduction to the subject.

These notes are meant for limited distribution to participants of the AIS/IST on Gröbner Bases and Their Applications at IIIT Delhi during December 11–23, 2017 as an accompaniment to the lectures of the author on Gröbner Bases and Coding Theory. Much of the material is borrowed from the informal notes of a series of six lectures given at the Advanced Instructional School on Algebraic Combinatorics held at the Indian Statistical Institute, Bangalore during June 24–July 13, 2013. Thanks are due to Mrinmoy Datta, Kannappan Sampath and Prasant Singh for taking and TeXing a preliminary version of the Bangalore notes. These notes will only cover some basic aspects of coding theory. For connections to Gröbner bases, we refer to the notes of Carvalho [5].

1. Basic Notions

We begin with the general definition of (possibly nonlinear) codes. Throughout this section Q denotes a finite set and n a positive integer.

1.1. Definition. A *code* of length n over Q is a subset of Q^n. If C is a code of length n, then the number of elements in C (denoted $|C|$) is called the *size* of C. Further, if $q = |Q|$, then C is called a *q-ary code* and $k = \log_q |C|$ is called the *dimension* of C. The ratio k/n is called the *rate* of transmission of C. Usually, the elements of Q are called *alphabets* and the elements of C are called *codewords*. We often write "C is a q-ary (n, M)-code over Q" to mean that C is a code of length n and size M over an alphabet set of size q.

There is a simple, but useful, notion of distance in the ambient space Q^n.

1.2. Definition. For $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in Q^n, the *Hamming distance* between x and y is the number of positions where they differ, i.e.,

$$d(x, y) = |\{i : x_i \neq y_i\}| \, .$$

It is easy to see that d defines a metric on Q^n.

1.3. Definition. Let C be a code of length n over Q. The *minimum distance* of C is denoted by $d(C)$ and defined by

$$d(C) = \min\{d(x, y) : x, y \in C, x \neq y\}.$$

If $d = d(C)$, then the ratio d/n is called the *relative distance* of C.

Evidently, the rate as well as the relative distance of a code are positive real numbers ≤ 1. It is usually of interest to construct codes for which the rate as well as the relative distance are as large as possible. These are often conflicting requirements, and moreover there are several limitations. The simplest of these is the following.

1.4. Proposition (Singleton bound). *Let $A_q(n, d)$ denote the maximum possible size of a q-ary code of length n and minimum distance d. Then $A_q(n, d) \leq q^{n-d+1}$. In other words, for any q-ary code of length n, dimension k and minimum distance d, we must have $k \leq n - d + 1$ or equivalently, $d \leq n - k + 1$.*

Proof. Let C be a code of length n over a set Q with q elements. If $d = d(C)$, then the projection map $C \longrightarrow Q^{n-d+1}$ on the last $n - d + 1$ coordinates, given by $(x_1, \ldots, x_n) \mapsto (x_d, x_{d+1}, \ldots, x_n)$, must be injective. Hence $|C| \leq q^{n-d+1}$. $\square$

2. Linear Codes

By $\mathbb{F}_q$ we shall denote "the" finite field with q elements. As before, n denotes a positive integer.

2.1. Definition. A q-ary *linear code* of length n is a subspace of $\mathbb{F}_q^n$.

Note that if C is a q-ary linear code of length n, then the dimension of C is $\dim_{\mathbb{F}_q} C$. Henceforth, we will usually write "C is a $[n,k]_q$- code" to mean that C is a q-ary linear code of length n and dimension k.

2.2. Definition. For $x = (x_1, \dots, x_n) \in \mathbb{F}_q^n$, the *support* of x is the set

$$\operatorname{supp}(x) = \{i : x_i \neq 0\}$$

and the *Hamming weight* of x is the nonnegative integer

$$\mathrm{w_H}(x) = |\operatorname{supp}(x)| = |\{i : x_i \neq 0\}|.$$

Note that $d(x,y) = \mathrm{w_H}(x-y)$ for any $x, y \in \mathbb{F}_q^n$. Thus for a linear code C,

$$d(C) = \min\{\mathrm{w_H}(x) : x \in C \text{ and } x \neq 0\}.$$

In other words, the minimum distance of a linear code is the minimum weight of its nonzero codewords. For a linear code C, elements $x \in C$ with $\mathrm{w_H}(x) = d$ are called *minimum weight codewords* of C.

2.3. Definition. A $[n,k]_q$-code C is said to be a *MDS code* or a *maximum distance separable code* if the Singleton bound is met, i.e., if $d(C) = n - k + 1$.

2.4. Exercise. Given positive integers k, n with $k \leq n$ determine (i.e., find a formula for) the number of $[n,k]_q$-codes.

2.5. Problem.[1] Given n, k, q as above, determine the number of $[n,k]_q$-MDS codes.

2.6. Exercise. Solve the above problem for $k = 1$ and $k = 2$.

In coding theory, it is customary to regard the elements of $\mathbb{F}_q^n$ as row vectors of length n, and we shall do so in the sequel. For $x \in \mathbb{F}_q^n$, we will write x^T to denote the transpose of x, i.e., the corrsponding column vector.

2.7. Definition. Let C be a $[n,k]_q$-code. A $k \times n$ matrix G with entries in $\mathbb{F}_q$ is called a *generator matrix* C if the rows of G forms a basis of C.

Note that if G is a generator matrix of a $[n,k]_q$-code C, then G has rank k and $C = \{uG : u \in \mathbb{F}_q^k\}$. Conversely, if G is any $k \times n$ matrix over $\mathbb{F}_q$ of rank k, then $\{uG : u \in \mathbb{F}_q^k\}$ is a $[n,k]_q$-code. Further, if $G, \tilde{G}$ are $k \times n$ matrices over $\mathbb{F}_q$ of rank k, then G and $\tilde{G}$ are generator matrices of the same code C if and only if $\tilde{G} = EG$ for some $E \in \mathrm{GL}_k(\mathbb{F}_q)$. In particular, a generator matrix G of a $[n,k]_q$-code C can be chosen in such a way that it is in reduced row-echelon form; in this case G is uniquely determined by C. In addition, if the pivotal 1's are in the first k columns, then $G = [I_k|A]$ for some $k \times (n-k)$ matrix A over $\mathbb{F}_q$ and we then say that G is in *standard form.* Here, and hereafter, I_m denote the $m \times m$ identity matrix, where m is any positive integer. Also for any matrix A, we denote by A^T its transpose.

In the remainder of this section, C will denote a $[n,k]_q$-code.

[1] A Problem is an Exercise whose solution is not known. In the case of Problem 2.5, one may consult [8] for more information. A solution to Exercise 2.6 can also be found there.

2.8. Definition. A $(n-k)\times n$ matrix H with entries in $\mathbb{F}_q$ is called a *parity check matrix* of C if C is its nullspace, i.e., $C = \{x \in \mathbb{F}_q^n : Hx^T = 0\}$.

It is clear that the rank of a parity check matrix of a $[n,k]_q$-code is $n-k$.

2.9. Lemma. *If C has a generator matrix $G = [I_k|A]$ in standard form, then $H = [-A^T|I_{n-k}]$ is a parity check matrix of C.*

Proof. Suppose $G = [I_k|A]$ is a generator matrix of C and $H = [-A^T|I_{n-k}]$. Then $HG^T = -A^T + A^T = 0$. This implies that $C \subseteq \{x \in \mathbb{F}_q^n : Hx^T = 0\}$. Since $\operatorname{rank}(H) = n-k$, the linear map $H : \mathbb{F}_q^n \longrightarrow \mathbb{F}_q^{n-k}$ defined by $x \mapsto Hx^T$ is surjective. Consequently, its kernel, i.e., the nullspace of H, has dimension k. It follows that $C = \{x \in \mathbb{F}_q^n : Hx^T = 0\}$. $\square$

2.10. Lemma. *Let H be a parity check matrix of C and t a positive integer. If $x \in C$ with $\mathrm{w_H}(x) = t$ and $supp(x) = \{j_1, j_2, \ldots, j_t\}$, then the columns of H indexed by $j_1, j_2, \ldots, j_t$ are linearly dependent. Conversely if some t columns of H are linearly dependent, but no proper subset of these columns is linearly dependent, then C contains a codeword of weight t.*

Proof. Let H_j denote the j^{th} column of H $(1 \le j \le n)$. For $x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$,

$$x \in C \Longleftrightarrow Hx^T = 0 \Longleftrightarrow \sum_{j=1}^{n} x_j H_j = 0.$$

This implies the desired result. $\square$

The above lemma leads to a useful characterization of the minimum distance.

2.11. Corollary. *Let H be a parity check matrix of C and d be a nonnegative integer. Then $d = d(C)$ if and only if every set of $d-1$ columns of H is linearly independent and there exist d columns of H that are linearly dependent.*

Proof. Clearly, $d = d(C)$ if and only if $\mathrm{w_H}(y) \ge d$ for all $y \in C$ and there exists $x \in C$ with $\mathrm{w_H}(x) = d$. Thus Lemma 2.10 yields the desired result. $\square$

Note that the Singleton bound for linear codes can also be deduced from Corollary 2.11. Indeed, if H is a parity check matrix of C, then $\operatorname{rank}(H) = n-k$ and so every set of $n-k+1$ columns of H is linearly dependent. Hence $d(C) \le n-k+1$, thanks to Corollary 2.11.

3. Duality

On $\mathbb{F}_q^n$, we have a nondegenerate symmetric bilinear form $\langle\ ,\ \rangle$ given by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i \quad \text{for } x = (x_1, \ldots, x_n) \text{ and } y = (y_1, \ldots, y_n) \text{ in } \mathbb{F}_q^n.$$

The *dual* of a $[n,k]_q$-code C is defined to be the linear code $C^\perp$ of length n given by

$$C^\perp = \{y \in \mathbb{F}_q^n : \langle x, y\rangle = 0 \text{ for all } x \in C\}.$$

3.1. Exercise. Show that if C is a $[n,k]_q$-code, then
(i) $\dim C^\perp = n-k$, and (ii) $(C^\perp)^\perp = C$.

3.2. Remark. The notion of positive definiteness does not make sense for the bilinear form $\langle\ ,\ \rangle$ defined above. Thus it may happen that a a $[n,k]_q$-code C satisfies $C \subseteq C^\perp$ or even $C = C^\perp$. In the first case, the code C is said to be *self-orthogonal* and in the second case it is said to be *self-dual.* For an interesting connection of self-dual codes with invariant theory, we refer to the *Monthly* article of Sloane [15].

3.3. Exercise. Show that a $(n-k)\times n$ matrix H with entries in $\mathbb{F}_q$ is parity check matrix of $[n,k]_q$-code C if and only if the rows of H form a basis of $C^\perp$.

3.4. Exercise. Show that a $[n,k]_q$-code C is self-dual if and only if it is self-orthogonal and $n = 2k$.

3.5. Exercise. Show that a linear code C is MDS if and only if $C^\perp$ is MDS.

3.6. Definition. Let C be a $[n,k]_q$-code. The *weight distribution* or the *spectrum* of C is given by the sequence $(A_0(C), A_1(C), \ldots, A_n(C))$, where

$$A_i(C) = |\{x \in C : \mathrm{w_H}(x) = i\}| \qquad \text{for } i = 0, 1, \ldots, n.$$

A useful way to describe the weight distribution of C is by means of the homogeneous polynomial $W_C(X,Y)$ in two variables X and Y defined by

$$W_C(x,y) = \sum_{i=0}^{n} A_i(C) X^{n-i} Y^i = \sum_{c\in C} X^{n-\mathrm{w_H}(c)} Y^{\mathrm{w_H}(c)}.$$

This is called the (two variable) *weight enumerator polynomial* of C.

There is a beautiful relationship between the weight distribution of a code and of its dual. This is given by the *MacWilliams identities*, which will be stated and proved a little later. As a warm-up, consider a $[n,k]_q$-code C, and let $A_i = A_i(C)$ and $B_i = A_i(C^\perp)$ for $i = 0, 1, \ldots, n$. Note that $A_0 = 1 = B_0$ and also that

$$\sum_{i=0}^{n} A_i = \sum_{i=0}^{n} \sum_{\substack{c\in C \\ \mathrm{w_H}(x)=i}} 1 = \sum_{c\in C} 1 = |C| = q^k = q^k B_0.$$

Next, consider the $q^k \times n$ matrix $\mathcal{M}$ whose rows are the (coordinates of the) codewords of C. The row corresponding to a codeword of weight j has $n-j$ zeros and there are A_j such codewords. Thus

$$\text{the number of zeros in } \mathcal{M} = \sum_{j=0}^{n} (n-j) A_j.$$

On the other hand the j^{th} column of M is the zero vector if and only if $x_j = 0$ for all $x \in C$ or equivalently, $e_j \in C^\perp$, where e_j denotes the j^{th} standard basis vector of $\mathbb{F}_q^n$. Now e_j's and its nonzero scalar multiples are precisely the elements in $\mathbb{F}_q^n$ of weight 1. It follows that the number of $j = 1, \ldots, n$ for which the j^{th} column consists only of zeros is $B_1/(q-1)$. In each of the remaining columns, every scalar appears exactly q^{k-1} times. To see this, note that projection map $C \longrightarrow \mathbb{F}_q$ given by $x \mapsto x_j$ is a nonzero linear map having q^{k-1} elements in its kernel. Thus we conclude that

$$\sum_{j=0}^{n}(n-j)A_j = \frac{B_1}{q-1}q^k + \left(n - \frac{B_1}{q-1}\right)q^{k-1} = B_1 q^{k-1} + nq^{k-1} = q^{k-1}\sum_{j=0}^{1}\binom{n-j}{n-1}B_j.$$

We will prove a string of such identities in Section 5. But first let us see some examples.

4. Examples

Let r be a positive integer and let $n = \frac{q^r-1}{q-1} = |\mathbb{P}^{r-1}(\mathbb{F}_q)|$ Let $H_r(q)$ be a $r \times n$ matrix with entries in $\mathbb{F}_q$ such that any two columns are linearly independent. In effect, the columns of $H_r(q)$ are obtained by representatives in $\mathbb{F}_q^r$ of distinct points of $\mathbb{P}^{r-1}(\mathbb{F}_q)$. Observe that the columns of $H_r(q)$ include some nonzero scalar multiples of the standard basis vectors of $\mathbb{F}_q^r$, and hence the rank of $H_r(q)$ is r. This leads to the following example(s).

4.1. Example. Define $\mathscr{H}_r(q)$ to be $[n, n-r]_q$-code with $H_r(q)$ as its parity check matrix and $\mathscr{S}_r(q)$ to be $[n, r]_q$-code with $H_r(q)$ as its generator matrix. These are called *Hamming code* and *simplex code*, respectively.

4.2. Exercise. Find $d(\mathscr{H}_r(q))$ and $d(\mathscr{S}_r(q))$. Also show that $\mathscr{H}_r(q)^\perp = \mathscr{S}_r(q)$.

4.3. Exercise. Determine the weight distribution of $\mathscr{S}_r(q)$ and write its weight enumerator polynomial.

The next example is the simplest kind of MDS codes, and it is usually meaningful to consider this when q is large.

4.4. Example. Let n, k be positive integers with $n \geq k$ and q be a prime power with $q \geq n$. Fix distinct elements $a_1, a_2, \dots, a_n \in \mathbb{F}_q$ and let

$$C = \{c_f = (f(a_1), f(a_2), \dots, f(a_n)) : f(x) \in \mathbb{F}_q[x] \text{ with } \deg f(x) < k\}.$$

Then one can easily verify that C is a $[n, k]_q$-code with minimum distance $n-k+1$. In other words C is a MDS code. This code C is known as a *Reed Solomon code.*

The next example is one of the most widely studied classes of codes. Classically, it was first studied by Reed and Muller in the binary case (when $q = 2$) by means of the so called boolean functions; see, e.g., [13, Ch. 13]. Here we adopt a more general viewpoint.

4.5. Example. Let m, ν be integers with $m \geq 1$ and $\nu \geq 0$, and let q be a prime power. Denote by $\mathbb{F}_q[X_1, \dots, X_m]_{\leq \nu}$ is the space of all polynomials in $X_1, \dots, X_m$ with coefficients in $\mathbb{F}_q$ of total degree at most ν. Fix a listing $P_1, \dots, P_{q^m}$ of elements of $\mathbb{F}_q^m$ The image of the evaluation map

$$\mathrm{Ev} : \mathbb{F}_q[X_1, \dots, X_n]_{\leq \nu} \to \mathbb{F}_q^{q^m} \quad \text{given by} \quad f \mapsto (f(P_1), \dots, f(P_{q^m}))$$

is a linear code of length $n := q^m$ and it is denoted by $\mathrm{RM}_q(\nu, m)$. It is called the *(generalized) Reed Muller code* of order ν and length q^m.

4.6. Exercise. Show that $\mathbb{F}_q[X_1, \ldots, X_m]_{\le \nu}$ is a finite dimensional vector space over $\mathbb{F}_q$ and find its dimension. Further, show that if $\nu < q$ then the map Ev is injective. Use this to find the dimension of $\mathrm{RM}_q(\nu, m)$ when $\nu < q$.

To determine the dimension of $\mathrm{RM}_q(\nu, m)$ for more general values of ν it is important to understand the distinction between polynomials in m variables over $\mathbb{F}_q$ and their evaluations, i.e., the corresponding functions from $\mathbb{F}_q^m$ to $\mathbb{F}_q$. Indeed two polynomials can give rise to the same function; for example, X_i^q and X_i take the same values on $\mathbb{F}_q^m$. To avoid such situations, one can look at reduced polynomials that are defined as follows. A monomial $X_1^{\alpha_1} \ldots X_m^{\alpha_m}$ in $\mathbb{F}_q[X_1, \ldots, X_m]$ is said to be *reduced* if $\alpha_i \le q-1$ for each $i = 1, \ldots, m$. A polynomial in $\mathbb{F}_q[X_1, \ldots, X_m]$ is said to be *reduced* if it is a $\mathbb{F}_q$-linear combination of reduced monomials. Note that for $f \in \mathbb{F}_q[X_1, \ldots, X_m]$, the condition $\deg f < q$ implies that f is reduced, but the converse is not true. However, if $f \in \mathbb{F}_q[X_1, \ldots, X_m]$ is reduced, then $\deg f \le m(q-1)$.

We have a natural map from the set of all monomials onto the set of all reduced monomials in $\mathbb{F}_q[X_1, \ldots, X_m]$ given by $X_1^{\alpha_1} \ldots X_m^{\alpha_m} \mapsto X_1^{\beta_1} \ldots X_m^{\beta_m}$, where for $i = 1, \ldots, m$, the exponent β_i is obtained from α_i as follows:

$$\beta_i = \begin{cases} \alpha_i & \text{if } 0 \le \alpha_i \le q-1, \\ r_i & \text{if } \alpha_i \ge q \text{ and } \alpha_i = (q-1)s_i + r_i \text{ where } r_i, s_i \in \mathbb{Z} \text{ with } 0 < r_i \le q-1. \end{cases}$$

This map on monomials extends, by linearity, to $\mathbb{F}_q[X_1, \ldots, X_m]$ and the image of $f \in \mathbb{F}_q[X_1, \ldots, X_m]$ under the (extended) map is denoted by $\overline{f}$. Evidently $\overline{f}$ is a reduced polynomial and $\overline{f}(P) = f(P)$ for all $P \in \mathbb{F}_q^m$. Now let

$$\mathfrak{R}_q(m, \nu) = \{f \in \mathbb{F}_q[X_1, \ldots, X_m] : \deg f \le \nu \text{ and } f \text{ is reduced}\}.$$

Observe that $\mathfrak{R}_q(m, \nu)$ is precisely the image of $\mathbb{F}_q[X_1, \ldots, X_m]_{\le \nu}$ under the reduction map $f \mapsto \overline{f}$ and also that $\mathfrak{R}_q(m, \nu) = \mathfrak{R}_q(m, m(q-1))$ if $\nu \ge m(q-1)$. In particular, $\mathrm{RM}_q(\nu, m) = \mathrm{Ev}\,(\mathfrak{R}_q(m, \nu))$ and $\mathrm{RM}_q(\nu, m) = \mathrm{RM}_q\,(\nu, m(q-1))$ if $\nu \ge m(q-1)$. With this in view, one usually restricts to $0 \le \nu \le m(q-1)$ while considering $\mathrm{RM}_q(\nu, m)$.

4.7. Exercise. Show that the restriction of Ev to $\mathfrak{R}_q(m, \nu)$ is injective and find $\dim_{\mathbb{F}_q} \mathfrak{R}_q(m, \nu)$ when $0 \le \nu \le m(q-1)$. Use it to find the dimension of $\mathrm{RM}_q(\nu, m)$ if $0 \le \nu \le m(q-1)$.

A codeword of $\mathrm{RM}_q(\nu, m)$ is an evaluation of a polynomial of degree $\le \nu$, and its Hamming weights are determined by the number of zeros of corresponding polynomial. More precisely, for any $f \in \mathbb{F}_q[X_1, \ldots, X_m]_{\le \nu}$,

$$\mathrm{w_H}(\mathrm{Ev}(f)) = q^m - \#Z(f) \quad \text{where} \quad Z(f) := \{P \in \mathbb{F}_q^m : f(P) = 0\}.$$

4.8. Exercise. If $f \in \mathbb{F}_q[X_1, \ldots, X_m]$ has degree $d < q$, then show that

$$|Z(f)| \le dq^{m-1}.$$

Further show that if $d < q$ and $a_1, \ldots, a_d \in \mathbb{F}_q$ are distinct, then the polynomial $f = (X_1 - a_1) \ldots (X_1 - a_d)$ is an element of $\mathbb{F}_q[X_1, \ldots, X_m]$ with $\deg f = d$ and $|Z(f)| = dq^{m-1}$. Deduce that if $\nu < q$, then $d(\mathrm{RM}_q(\nu, q^m)) = (q - \nu)q^{m-1}$.

Those who like challenges may also attempt to determine $d(\mathrm{RM}_q(\nu, q^m))$ in general, for $0 \le \nu \le m(q-1)$.

Finally in this section, we outline some of the standard constructions used to construct new codes from a given linear code. These together with the above examples furnishes many more examples of linear codes.

Let C be a $[n,k]_q$ code and let $P \subseteq \{1,\dots,n\}$. For any $x = (x_1,\dots,x_n) \in \mathbb{F}_q^n$, let x^P denote the element of $\mathbb{F}_q^{n-|P|}$ obtained from x by removing the x_i's for all $i \in P$. Define

$$C^P = \{x^P : x \in C\} \quad \text{and} \quad C_P = \{x^P : x \in C \text{ with } x_i = 0 \text{ for all } i \in P\}.$$

These are linear codes of length $n-|P|$, called, respectively, the *puncturing* of C at P and the *shortening* of C at P. There is a nice relationship between puncturing, shortening, and taking duals.

4.9. Proposition. *C be a $[n,k]_q$ code and let $P \subseteq \{1,\dots,n\}$. Then*

$$(C^P)^\perp = (C^\perp)_P \quad \textit{and} \quad (C^\perp)^P = (C_P)^\perp.$$

Proof. Let $z \in (C^\perp)_P$. Then $z = y^P$ for some $y \in C^\perp$ with $y_i = 0$ for all $i \in P$. Hence $\langle z, x^P\rangle = \langle y^P, x^P\rangle = \langle y, x\rangle = 0$ for all $x \in C$. Thus $(C^\perp)_P \subseteq (C^P)^\perp$.

To prove the other inclusion, suppose $z \in (C^P)^\perp$. Extend z to $y \in \mathbb{F}_q^n$ by

$$y_i = \begin{cases} z_i & \text{if } i \notin P, \\ 0 & \text{if } i \in P. \end{cases}$$

Then, $\langle y, x\rangle = \langle y^P, x^P\rangle = \langle z, x^P\rangle = 0$ for all $x \in C$. Hence $z \in (C^\perp)_P$. This proves that $(C^P)^\perp \subseteq (C^\perp)_P$.

Thus $(C^\perp)_P = (C^P)^\perp$. Replacing C by $C^\perp$, we obtain $(C^\perp)^P = (C_P)^\perp$. $\square$

5. MacWilliams Identities

As before n, k will always denote positive integers with $k \le n$ and q a prime power. Further, we let $[n]$ denote the set $\{1,\dots,n\}$ of the first n positive integers. The MacWilliams identities alluded to toward the end of Section 3 are as follows.

5.1. Theorem. *Let C be a $[n,k]_q$-code. Let $A_i = A_i(C)$ and $B_i = A_i(C^\perp)$ for $0 \le i \le n$. Then, for $\nu = 0, 1, \dots, n$, we have,*

$$\sum_{j=0}^{n} \binom{n-j}{\nu} A_j = q^{k-\nu} \sum_{j=0}^{n} \binom{n-j}{n-\nu} B_j$$

Proof. Fix $0 \leqslant \nu \leqslant n$. Let N_ν denote the cardinality of the set

$$\{(x, I) : x \in C,\ I \subseteq [n],\ |I| = \nu \text{ and } x_i = 0 \text{ for all } i \in I\}.$$

For $I \subseteq [n]$, let $I^c := [n] \setminus I$ denote the complement of I in $[n]$. Note that for any $x \in \mathbb{F}_q^n$, the condition $x_i = 0$ for all $i \in I$ is equivalent to $\mathrm{supp}(x) \subseteq I^c$. Thus

$$N_\nu = \sum_{x\in C} \sum_{\substack{I\subseteq[n] \\ |I|=\nu \\ \mathrm{supp}(x)\subseteq I^c}} 1 = \sum_{j=0}^{n} \sum_{\substack{x\in C \\ \mathrm{w}_H(x)=j}} \sum_{\substack{I\subseteq[n] \\ |I|=\nu \\ \mathrm{supp}(x)\subseteq I^c}} 1 = \sum_{j=0}^{n} A_j \binom{n-j}{\nu}. \tag{5.1}$$

On the other hand,

$$N_\nu = \sum_{\substack{I\subseteq[n]\\|I|=\nu}} |\{x \in C : x_i = 0 \ \forall\, i \in I\}|\,.$$

Moreover, $x \mapsto x^{I^c}$ defines a linear map $C \to C^{I^c}$ and $\{x \in C : x_i = 0 \ \forall\, i \in I\}$ is the kernel of this map. The cardinality of this kernel is $q^{k-k^{I^c}}$ where $k^{I^c} := \dim C^{I^c}$. Consequently, using Proposition 4.9 and noting that the length of C^{I^c} is ν, we find

$$N_\nu = \sum_{\substack{I\subseteq[n]\\|I|=\nu}} q^{k-k^{I^c}} = q^{k-\nu} \sum_{\substack{I\subseteq[n]\\|I|=\nu}} |(C^{I^c})^\perp| = q^{k-\nu} \sum_{\substack{I\subseteq[n]\\|I|=\nu}} \left|(C^\perp)_{I^c}\right|.$$

Now, $\left|(C^\perp)_{I^c}\right| = \left|\{y \in C^\perp : y_i = 0 \ \forall\, i \in I^c\}\right| = \left|\{y \in C^\perp : \mathrm{supp}(y) \subseteq I\}\right|$. Hence

$$\begin{aligned} q^{\nu-k} N_\nu &= \sum_{\substack{I\subseteq[n]\\|I|=\nu}} \sum_{\substack{y\in C^\perp\\ \mathrm{supp}(y)\subseteq I}} 1 \\ &= \sum_{\substack{I\subseteq[n]\\|I|=\nu}} \sum_{j=0}^{n} \sum_{\substack{y\in C^\perp\\ \mathrm{w}_H(y)=j\\ \mathrm{supp}(y)\subseteq I}} 1 \\ &= \sum_{j=0}^{n} \sum_{\substack{y\in C^\perp\\ \mathrm{w}_H(y)=j}} \sum_{\substack{|I|=\nu\\ I\subseteq[n]\\ \mathrm{supp}(y)\subseteq I}} 1 \\ &= \sum_{j=0}^{n} \sum_{\substack{y\in C^\perp\\ \mathrm{w}_H(y)=j}} \binom{n-j}{\nu-j} \\ &= \sum_{j=0}^{n} \binom{n-j}{\nu-j} B_j. \end{aligned}$$

Thus, we have:

$$N_\nu = q^{k-\nu} \sum_{j=0}^{n} \binom{n-j}{\nu-j} B_j. \tag{5.2}$$

Combining (5.1) and (5.2), we obtain the desired result. □

It may be noted that for $j, \nu \in \{0, 1, \ldots, n\}$, the binomial coefficient $\binom{n-j}{\nu}$ vanishes if $j > n-\nu$, whereas $\binom{n-j}{\nu-j}$ vanishes if $j > \nu$. Thus the MacWilliams identities in Theorem 5.1 can also be written as

$$\sum_{j=0}^{n-\nu} \binom{n-j}{\nu} A_j = q^{k-\nu} \sum_{j=0}^{\nu} \binom{n-j}{n-\nu} B_j \quad \text{for } \nu = 0, 1, \ldots, n.$$

These identities could be expressed in a more compact form using the two variable weight enumerator polynomial as follows.

5.2. Corollary (MacWilliams Identity). *For any $[n,k]_q$-code C,*

$$W_{C^\perp}(X,Y) = \frac{1}{|C|} W_C(X+(q-1)Y, X-Y). \tag{5.3}$$

In particular, in the binary case, $W_{C^\perp}(X,Y) = 2^{-k} W_C(X+Y, X-Y)$.

Proof. Putting $X = Y+Z$, we see that (5.3) is equivalent to:

$$W_{C^\perp}(Y+Z,Y) = \frac{1}{|C|} W_C(Z+qY, Z).$$

Now,

$$\begin{aligned} W_{C^\perp}(Y+Z,Y) &= \sum_{j=0}^{n} B_j (Y+Z)^{n-j} Y^j \\ &= \sum_{j=0}^{n} B_j Y^j \sum_{\nu=0}^{n-j} \binom{n-j}{\nu} Y^{n-j-\nu} Z^\nu \\ &= \sum_{\nu=0}^{n} \left(\sum_{j=0}^{n-\nu} \binom{n-j}{\nu} B_j \right) Y^{n-\nu} Z^\nu \\ &= \sum_{\nu=0}^{n} \left(\sum_{j=0}^{\nu} \binom{n-j}{n-\nu} B_j \right) Y^{\nu} Z^{n-\nu}. \end{aligned}$$

where the penultimate equality is obtained by interchanging summations and the last equality is obtained by changing ν to $n-\nu$. On the other hand,

$$\begin{aligned} \frac{1}{|C|} W_C(Z+qY,Z) &= \frac{1}{q^k} \sum_{j=0}^{n} A_j (Z+qY)^{n-j} Z^j \\ &= \frac{1}{q^k} \sum_{j=0}^{n} \left(\sum_{\nu=0}^{n-j} A_j \binom{n-j}{\nu} Z^{n-j-\nu} q^\nu Y^\nu \right) Z^j \\ &= \sum_{\nu=0}^{n} \frac{1}{q^{k-\nu}} \left(\sum_{j=0}^{n-\nu} \binom{n-j}{\nu} A_j \right) Z^{n-\nu} Y^\nu. \end{aligned}$$

Thus Theorem 5.1 yields the desired result. □

5.3. Remark. Comparing coefficients in (5.3), we obtain, for $j = 0,1,\dots,n$:

$$B_j = \frac{1}{|C|} \sum_{i=0}^{n} K_j(i) A_i$$

where $K_j(X) = K_j^{n,q}(X)$ is the j^{th} *Krawtchouk polynomial* defined by:

$$K_j(X) = \sum_{r=0}^{j} (-1)^r \binom{X}{r} \binom{n-X}{j-r} (q-1)^{j-r}$$

where for any $r \in \mathbf{Z}$, and variable X,

$$\binom{X}{r} := \begin{cases} \frac{X(X-1)\cdots(X-r+1)}{r!} & \text{if } r > 0, \\ 0, & \text{if } r < 0. \end{cases}$$

These Krawtchouk polynomials satisfy the following orthogonality relations.

$$\sum_{i=0}^{n} K_j(i)K_i(k) = q^n\delta_{j,k} \quad \text{and} \quad \sum_{i=0}^{n} \mu(i)K_j(i)K_k(i) = q^n\mu(j)\delta_{j,k},$$

where $\mu(i) := \binom{n}{i}(q-1)^i$ and δ is the Kronecker delta. We refer to [17, Ch. 1] for a quick introduction to Krawtchouk polynomials and their properties.

5.4. Exercise. Let C and A_i, B_i be as in Theorem 5.1. Prove that

$$\sum_{j=\nu}^{n} \binom{j}{\nu} A_j = q^{k-\nu} \sum_{j=0}^{\nu} (-1)^j \binom{n-j}{n-\nu} (q-1)^{\nu-j} B_j \quad \text{for } \nu = 0, 1, \ldots, n.$$

(Hint: Put $Y = X + Z$ in (5.3).)

5.5. Remark. We remark in passing that, (5.3) can be used to obtain Pless power moment formulae, which express the ν^{th} moment $\sum_{j=0}^{n} j^\nu A_j$ in terms of the B_j's and also express $\sum_{j=0}^{n}(n-j)^\nu A_j$ in terms of the B_j's. To this end, it suffices to express the two bases $\{X^j : j \geq 0\}$ and $\{\binom{X}{j} : j \geq 0\}$ in terms of each other (and this can be done using the so called Stirling numbers of the second kind) and using Exercise 5.4. For details, we refer to [14, Ch. 1].

6. Equivalence and Automorphisms of Codes

6.1. Definition. Let C_1 and C_2 be two linear codes of length n. We say that C_1 and C_2 are *permutation equivalent* if there exists $\sigma \in S_n$ such that

$$C_2 = \{(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) : (x_1, \ldots, x_n) \in C_1\}$$

In other words, the map:

$$f_\sigma : \mathbb{F}_q^n \to \mathbb{F}_q^n \quad \text{defined by} \quad (x_1, \ldots, x_n) \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

induces a linear isomorphism of $C_1 \to C_2$.

Equivalently, two codes C_1 and C_2 of length n are permutation equivalent, if there exists a permutation matrix $P \in \mathrm{GL}_n(\mathbb{F}_q)$ such that $f_\sigma(x) = xP$ gives a bijection of C_1 onto C_2.

6.2. Notation. If C_1 and C_2 are permutation equivalent, we write $C_1 \sim C_2$.

For our next definition, recall that an $n \times n$ matrix M is said to be a *monomial matrix* if $M = PD$ where P is a permutation matrix and D a diagonal matrix whose diagonal entries are nonzero.

6.3. Definition. Let C_1 and C_2 be two linear codes of length n. We say that C_1 and C_2 are *(monomially) equivalent* if there exists a monomial matrix $M \in \mathrm{GL}_n(\mathbb{F}_q)$ such that $x \mapsto xM$ gives a bijection of C_1 onto C_2.

6.4. Notation. We write $C_1 \approx C_2$ to denote that C_1 and C_2 are monomially equivalent; we also say that $C_1 \approx C_2$ via M if the monomial matrix M gives the bijection $C_1 \to C_2$ via the map $x \mapsto xM$.

6.5. Definition. A map $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is said to be an *isometry* if it is bijective and

$$d(x, y) = d(f(x), f(y)) \quad \text{for all } x, y \in \mathbb{F}_q^n.$$

Clearly, if $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is linear and bijective, then, f is an isometry if and only if $\mathrm{w}_H(f(x)) = \mathrm{w}_H(x)$ for all $x \in \mathbb{F}_q^n$.

6.6. PROPOSITION. *If $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is a linear isometry, then there exists a monomial matrix $M \in \mathrm{GL}_n(\mathbb{F}_q)$ such that $f(x) = xM$ for all $x \in \mathbb{F}_q^n$.*

Proof. Suppose $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is a linear isometry. Let $\{e_1, \dots, e_n\}$ be the standard basis for $\mathbb{F}_q^n$. Then for each $i = 1, \dots, n$, $f(e_i)$ has Hamming weight 1, and hence

$$f(e_i) = \lambda_i e_{\sigma(i)} \qquad \text{for some } \lambda_i \in \mathbb{F}_q^* \text{ and } \sigma(i) \in [n].$$

Note that $\sigma \in S_n$, since f is a bijection. Let P_σ be the permutation matrix in $\mathrm{GL}_n(\mathbb{F}_q)$ associated to σ, and let $M \in \mathrm{GL}_n(\mathbb{F}_q)$ be the monomial matrix

$$M = P_\sigma D \qquad \text{where} \qquad D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Then $f : \mathbb{F}_q^n \to \mathbb{F}_q^n$ is given by $f(x) = xM$ for $x \in \mathbb{F}_q^n$. This proves the assertion. □

6.7. REMARK.

(1) The notions of permutation equivalence and monomial equivalence coincide for binary code.
(2) If C_1 and C_2 are (permutation or monomial) equivalent, they have the same parameters.
(3) Also, if $C_1 \approx C_2$ via M, and if G_1 is a generator matrix of C_1, then, $G_2 = G_1 M$ is a generator matrix of C_2. In particular, any linear code is (permutation) equivalent to a code whose generator matrix is in standard form.

Equivalences of a code with itself leads to the important notion of automorphism of a code. In the remainder of this section, C denotes a $[n,k]_q$-code.

6.8. DEFINITION. The *permutation automorphism group* of C is:

$$\mathrm{PAut}(C) = \{\sigma \in S_n : (x_{\sigma(1)}, \dots, x_{\sigma(n)}) \in C \text{ for all } (x_1, \dots, x_n) \in C\}$$

Clearly, this group is a subgroup of S_n. Also we have the following isomorphism:

$$\mathrm{PAut}(C) \simeq \{P \in \mathrm{GL}_n(\mathbb{F}_q) : P \text{ a permutation matrix such that } xP \in C \ \forall\, x \in C\},$$

Thus we may think of the permutation automorphism group as a subgroup of $\mathrm{GL}_n(\mathbb{F}_q)$.

6.9. DEFINITION. The *monomial automorphism group* of C is:

$$\mathrm{MAut}(C) = \{M \in \mathrm{GL}_n(\mathbb{F}_q) : M \text{ a monomial matrix and } xM \in C \ \forall\, x \in C\}.$$

Clearly, this is a subgroup of $\mathrm{GL}_n(\mathbb{F}_q)$.

The most general notion of automorphism of a code is given by the following.

6.10. Definition. The *semilinear automorphism group* of C is

$$\Gamma\mathrm{Aut}(C) = \left\{ M\mu \;\middle|\; \begin{array}{c} M \in \mathrm{GL}_n(\mathbb{F}_q) \text{ a monomial matrix} \\ \mu \in \mathrm{Aut}(\mathbb{F}_q) \text{ such that } xM\mu \in C \;\forall x \in C \end{array} \right\}$$

Clearly, $\Gamma\mathrm{Aut}(C)$ is a subgroup of the group $\Gamma\mathrm{L}(n, q)$ of semilinear isomorphisms of $\mathbb{F}_q^n$.

In general, if V_1, V_2 are vector spaces over a field $\mathbb{F}$ and $\mu \in \mathrm{Aut}(\mathbb{F})$ is an automorphism of $\mathbb{F}$, then a map $f : V_1 \to V_2$ is said to be *μ-semilinear* if $f(x + y) = f(x) + f(y)$ for all $x, y \in V_1$ and $f(ax) = \mu(a)f(x)$ for all $x \in V_1$ and $a \in \mathbb{F}$. We say that $f : V_1 \to V_2$ is *semilinear* if it is μ-semilinear for some $\mu \in \mathrm{Aut}(\mathbb{F})$. A bijective semilinear map whose inverse is semilinear is called a *semilinear isomorphism.*

In connection with semilinear isomorphisms, we state without proof the following analogue of Proposition 6.6. A proof can be found, for example, in [4] or [9].

6.11. Proposition (MacWilliams). *Let C_1 and C_2 be linear codes of length n and dimension at least 3, and let $f : C_1 \to C_2$ be a map of C_1 into C_2. Then f is a weight preserving bijection that maps r-dimensional subspaces of C_1 onto r-dimensional subspaces of C_2 for each $r \geq 0$ if and only if $f : C_1 \to C_2$ is a semilinear isomorphism.*

In the binary case (i.e., when $q = 2$), we have $\mathrm{PAut}(C) = \mathrm{MAut}(C) = \Gamma\mathrm{Aut}(C)$, and we use $\mathrm{Aut}(C)$ to denote this group. It can be shown, for example, that $\mathrm{Aut}(\mathscr{H}_r(2)) = \mathrm{GL}(r, 2)$. In general, it is difficult to determine the automorphism group of a given class of codes. For the determination of the automorphism group of Reed-Muller codes, we refer to [7, 3, 12].

7. Cyclic Codes

Cyclic codes are an important and well-studied class of linear codes. Here we give a very brief introduction. As before, n denotes a positive integer.

7.1. Definition. A linear code C of length n is said to be *cyclic* if

$$(c_0, c_1, \ldots, c_{n-1}) \in C \Longrightarrow (c_{n-1}, c_0, c_1, \ldots, c_{n-2}) \in C.$$

We have used here $c = (c_0, c_1, \ldots, c_{n-1})$ to denote a typical element of $\mathbb{F}_q^n$. This is in order to identify c with the polynomial $c_0 + c_1X + c_2X^2 + \cdots + c_{n-1}X^{n-1}$ in $\mathbb{F}_q[X]$ or equivalently the image of this polynomial in the ring $R_n := \mathbb{F}_q[X]/\langle X^n - 1\rangle$ obtained from $\mathbb{F}_q[X]$ by "moding out" by the ideal generated by $X^n - 1$. The resulting map from $\mathbb{F}_q^n$ to R_n will be denoted by π. Clearly, $\pi : \mathbb{F}_q^n \to R_n$ is a natural $\mathbb{F}_q$-linear isomorphism.

7.2. Exercise. Let $C \subseteq \mathbb{F}_q^n$. Show that C is a cyclic code of length n if and only if $\pi(C)$ is an ideal of R_n.

Note that $\mathbb{F}_q[X]$ is a PID and the ideals of R_n correspond precisely to the ideals of $\mathbb{F}_q[X]$ containing $\langle X^n - 1\rangle$. In particular, every ideal of R_n is principal. If I is an ideal of R_n, then there is a unique monic polynomial, $g(X) \in \mathbb{F}_q[X]$, such that

$g(X) \mid X^n - 1$ and I is generated by the image of $g(X)$ in R_n. We call $g(X)$ the *generator polynomial* of I or of the corresponding cyclic code $C = \pi^{-1}(I)$.

7.3. Exercise. Let I, $g(X)$ and C be as above. Suppose $\deg g(X) = n - k$. Write $g(X) = g_0 + g_1 X + \cdots + g_{n-k} X^{n-k}$. Show that $\dim(C) = k$ and the $k \times n$ matrix:

$$G = \begin{pmatrix} g_0 & g_1 & \dots & & g_{n-k} & 0 & & \dots & & 0 \\ 0 & g_0 & g_1 & \dots & & g_{n-k} & 0 & & \dots & 0 \\ & & & & & & & & & \\ & \vdots & & & & \vdots & & & & \\ & & & & & & & & & \\ 0 & 0 & \dots & 0 & g_0 & g_1 & & \dots & & g_{n-k} \end{pmatrix}$$

is a generator matrix of C. Further if we let

$$h(X) = \frac{X^n - 1}{g(X)} = h_0 + h_1 X + \cdots + h_k X^k,$$

then show that the $(n - k) \times k$ matrix

$$H = \begin{pmatrix} h_k & h_{k-1} & \dots & & h_0 & 0 & & \dots & & 0 \\ 0 & h_k & h_{k-1} & \dots & & g_0 & 0 & & \dots & 0 \\ & & & & & & & & & \\ & \vdots & & & & \vdots & & & & \\ & & & & & & & & & \\ 0 & 0 & \dots & 0 & h_k & h_{k-1} & & \dots & & h_0 \end{pmatrix}$$

is a parity check matrix of C.

8. Bounds for Codes

By a $[n, k, d]_q$ *code* we shall mean a $[n, k]_q$-code C with $d(C) = d$.

8.1. Theorem (Griesmer). *Let C be a $[n, k, d]_q$ code. Then*

$$n \geqslant \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil$$

.

Before proving, we show that, the singleton bound follows as a corollary:

8.2. Corollary (Singleton Bound). $n \geqslant d + k - 1$.

Proof of Corollary. We simply note that:

$$\left\lceil \frac{d}{q^0} \right\rceil = d \quad \text{and} \quad \left\lceil \frac{d}{q^i} \right\rceil \geqslant 1 \text{ for all } i = 1, \dots, k - 1.$$

Hence Theorem 8.1 yields the Singleton bound. □

The proof of Griesmer's bound involves iterating the construction of a code C' from a linear code C and a minimum weight codeword x of C. The resulting code C' is sometimes called the *residue* of the code C at the word $x \in C$. We outline this construction in the following lemma.

8.3. Lemma. *Let C be a $[n, k, d]_q$ code with $k > 0$ and let $x \in C$ be such that $\mathrm{w_H}(x) = d$. If $P = \mathrm{supp}(x)$, then $C' = C^P$ is a $[n-d, k-1, d']$-code with $d' \geqslant \left\lceil \frac{d}{q} \right\rceil$.*

Proof. By passing to an equivalent code, if necessary, we may assume:

$$x = (\underbrace{1, \ldots, 1}_{d \text{ times}}, 0, \ldots, 0) \in C.$$

Then $P = \mathrm{supp}(x) = \{d, d+1, \ldots, n\}$ and the map given by $y \mapsto y^P$ is a linear map of C onto C' that has x in its kernel. Hence, $\dim(C') \leqslant k - 1$. Further, if $\dim(C') < k - 1$, then, there is y in the kernel such that $y \neq \lambda x$ for all $\lambda \in \mathbb{F}_q$ and

$$y_{d+1} = \cdots = y_n = 0.$$

But, then, $y - \lambda x$ is a nonzero codeword of C whose weight is at most $d - 1$, for some $\lambda \in \mathbb{F}_q$. This is a contradiction. So, $\dim(C') = k - 1$.

Let $z \in C'$ be a nonzero codeword. Then, $z = y^P$ for some $y = (y_1, \ldots, y_n) \in C$. Look at $y_1, \ldots, y_d$. By the Pigeonhole Principle, there exists $\alpha \in \mathbb{F}_q$ such that at least $\lceil d/q \rceil$ of the y_is are equal to α. Hence

$$\mathrm{w_H}(y - \alpha x) \leq d - \left\lceil \frac{d}{q} \right\rceil + \mathrm{w_H}(y_{d+1}, \ldots, y_n) = d - \left\lceil \frac{d}{q} \right\rceil + \mathrm{w_H}(z).$$

On the other hand, $y - \alpha x \neq 0$, by the choice of y, and so $\mathrm{w_H}(y - \alpha x) \geq d$. Thus it follows that $\mathrm{w_H}(z) \geq \left\lceil \frac{d}{q} \right\rceil$. This proves that $d' = d(C') \geqslant \left\lceil \frac{d}{q} \right\rceil$. □

Proof of Theorem. We induct on k. The cases $k = 0$ and $k = 1$ are trivial. Assume that $k > 1$ and that, the result holds for $k - 1$. Choose $x \in C$ such that $\mathrm{w_H}(x) = d$ and let C' be as in the previous lemma. By the induction hypothesis,

$$n - d \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^i} \right\rceil \geq \sum_{i=0}^{k-2} \left\lceil \frac{d}{q^{i+1}} \right\rceil = \sum_{i=1}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil.$$

Thus $n \geq \sum_{i=0}^{k-1} \lceil d/q^i \rceil$, as desired. □

8.4. Exercise. Show that the simplex codes meet the Griesmer bound.

Before proving more bounds, we make the following definitions:

8.5. Definition. Let Q be a finite set with q elements. For $x \in Q^n$ and $t \in \mathbf{R}$, the (solid) *sphere* of radius t centered at x is:

$$\mathbb{S}_t(x) = \{y \in Q^n : d_H(x, y) \leqslant t\}.$$

Clearly, the number $V_q(n, t)$ of points in the sphere $\mathbb{S}_t(x)$ (this is also the volume of the sphere $\mathbb{S}_t(x)$ with respect to the counting measure in Q^n) is given by:

$$V_q(n, t) := \sum_{i=0}^{t} \binom{n}{i} (q-1)^i.$$

Note that, this number is independent of the center $x \in Q^n$ of the sphere $\mathbb{S}_t(x)$. In most applications, Q will be the finite field $\mathbb{F}_q$ and this may be tacitly assumed when we consider linear codes.

We now look at the question of how large can a code be, given its desired properties (some natural constraints on its parameters)? In this context, the following notation is relevant:

8.6. NOTATION. For a prime power q and integers n, d with $n \geq 1$ and $d \geq 0$, let

$$A_q(n,d) := \max\{|C| : C \text{ } q\text{-ary code of length } n \text{ and minimum distance } d\}, \text{ and}$$
$$B_q(n,d) := \max\{|C| : C \text{ } q\text{-ary linear code of length } n \text{ and minimum distance } d\}.$$

Clearly, $B_q(n,d)$ is defined only when q is a prime power and in this case, $B_q(n,d) \leqslant A_q(n,d)$. The following theorem gives an upper bound on $A_q(n,d)$, which is obtained by "packing" C with spheres of radius t.

8.7. THEOREM (Hamming Bound/Sphere-packing Bound). *Let q, n, d be any positive integers with $d \leq n$. Then*

$$A_q(n,d) \leqslant \frac{q^n}{V_q(n,t)}, \quad \textit{where} \quad t := \left\lfloor \frac{d-1}{2} \right\rfloor.$$

Proof. Let C be a q-ary code of length n and minimum distance d. We observe that, the spheres $\mathbb{S}_t(c)$ as c varies over C are disjoint. Indeed, if, $x \in \mathbb{S}_t(c_1) \cap \mathbb{S}_t(c_2)$, where $c_1, c_2 \in C$ and $c_1 \neq c_2$, then

$$d(c_1,c_2) \leq d(c_1,x) + d(x,c_2) \leq 2t \leq d-1 < d,$$

which is a contradiction. Thus, $\coprod_{c \in C} \mathbb{S}_t(c) \subseteq \mathbb{F}_q^n$. This shows that

$$|C| \cdot V_q(n,t) \leq q^n \quad \text{and hence} \quad |C| \leq \frac{q^n}{V_q(n,t)}$$

which yields the desired bound for $A_q(n,d)$. □

Recall that, the codes that meet singleton bound were called MDS codes. Here is yet another notion of a "good" code.

8.8. DEFINITION. A $[n,k,d]$-code C that meets sphere packing bound is called a *perfect* code.

Evidently, if C is perfect code on an alphabet set Q, then,

$$\coprod_{c \in C} \mathbb{S}_t(c) = Q^n.$$

8.9. EXAMPLE. Trivial examples of perfect linear codes include $\{0\}$ and $\mathbb{F}_q^n$.

8.10. EXAMPLE. The binary Hamming code $\mathscr{H}_r(2)$ is perfect. Here $d = 3$, so $t = 1$. Set $n = 2^r - 1$, the length of $\mathscr{H}_r(2)$; and let $M = |\mathscr{H}_r(2)| = 2^{n-r}$. We compute the volume of a unit sphere in $\mathbf{F}_2^n$:

$$\binom{n}{0}(q-1)^0 + \binom{n}{1}(q-1) = 1 + n(2-1) = 2^r$$

Now, it is easy to see $\mathscr{H}_r(2)$ is perfect:

$$M \cdot 2^r = 2^n$$
$$\coprod_{c \in \mathscr{H}_r(2)} \mathbb{S}_1(c) = \mathbf{F}_2^n$$

8.11. Exercise. Show that $\mathscr{H}_r(q)$ is perfect for any prime power q.

8.12. Remark.

(1) Firstly, if C is perfect, $C^\perp$ is not necessarily perfect. For example, the binary simplex code is not perfect, but the binary Hamming code is.

(2) The Golay Codes

$$\begin{array}{ll} G_{23} & [23,12,7]_2\text{-code} \\ G_{24} & [24,12,8]_2\text{-code} \\ G_{11} & [11,6,5]_3\text{-code} \\ G_{12} & [12,6,6]_3\text{-code} \end{array}$$

have the property that, G_{23} and G_{11} are perfect. It is said, the following curious observation,

$$\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} = 2^{11} = 2^{23-12}$$

led to the construction of these codes. We note in passing that, G_{23} is a cyclic code whose generator polynomial is one of the irreducible factors of the polynomial

$$\frac{X^{23}-1}{X-1}.$$

The Golay codes are usually defined by describing explicitly their generator matrices (that happen to be in standard form). For more on these, we refer to the *Handbook* [14]. See also the book of Conway and Sloane [6] for Golay codes, their automorphism groups, and many other fascinating topics.

8.13. Theorem (Gilbert Bound/Sphere Covering Bound). *Let q, n, d be any positive integers with $d \leq n$. Then*

$$A_q(n,d) \geqslant \frac{q^n}{V_q(n,d-1)}$$

Equivalently, $\log_q(A_q(n,d)) \geqslant n - \log_q(V_q(n,d-1))$.

Proof. Let $C = \{c_1, \ldots, c_M\}$ be a code of length n over a finite alphabet set Q with q elements such that $M = A_q(n,d)$.

Since C is optimal, if $x \in Q^n$, then, $d(x, c_i) < d$ for some i (for otherwise, if $d(x, c_i) \geqslant d$ for all i, then, the code $C \cup \{x\}$ has length n and minimum distance d, contradicting the optimality of C). Consequently,

$$Q^n \subseteq \bigcup_{i=1}^M \mathbb{S}_{d-1}(c_i) \quad \text{and so} \quad q^n \leq \sum_{i=1}^M V_q(n,d-1) = MV_q(n,d-1).$$

Thus, $M \geq q^n / V_q(n,d-1)$, as desired. $\square$

For a linear code, we may do slightly better. To do this, we appeal to linear algebra in the following slightly technical lemma:

8.14. Lemma. *Let n, d be integers with $2 \leqslant d \leqslant n$. If k is a positive integer such that:*

$$V_q(n-1, d-2) < q^{n-k} \tag{8.1}$$

then, there exists a $[n,k]_q$-code C with $d(C) \geqslant d$.

Proof. We will construct columns $h_1, \dots, h_n$ of a $(n-k) \times n$ matrix H such that any $d-1$ of these columns are linearly independent. Then, the code C with H as its parity check matrix will have the desired property.

The trick is to get the greedy algorithm to work: that is, we show that, we can satisfy our greed!

First, take h_1 to be an arbitrary non-zero column in $\mathbb{F}_q^{n-k}$. Having chosen $h_1, \dots, h_j \in \mathbb{F}_q^{n-k}$ such that any $d-1$ of them are linearly independent and $j < n$, we choose h_{j+1} from the complement of the set of all linear combinations of $d-2$ of the j vectors $h_1, \dots, h_j$.

We just need to show that, such a choice can always be made: to show this, we count the number $\mathcal{L}$ of linear combinations of $d-2$ of the j vectors $h_1, \dots, h_j$:

$$\mathcal{L} = \sum_{i=0}^{d-2} \binom{j}{i} (q-1)^i \leq \sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i = V_q(n-1, d-2) < q^{n-k}$$

Thus, we can find $h_{j+1} \in \mathbb{F}_q^{n-k}$ which is different from any of these linear combinations. This finishes the proof. □

8.15. Corollary (Varshamov Bound). *Let q be a prime power and n, d be any positive integers with $d \leq n$. Then*

$$B_q(n,d) \geqslant q^{n - \lceil \log_q (1 + V_q(n-1, d-2)) \rceil}.$$

Further, the above bound implies the following weaker, but simpler bound:

$$B_q(n,d) \geqslant \frac{q^{n-1}}{1 + V_q(n-1, d-2)}.$$

Proof. First, suppose $d = 1$ Then $C = \mathbb{F}_q^n$ is evidently a q-ary linear code of length n with $d(C) = 1$, and hence $B_q(n,d) \geq q^n$ (in fact, $B_q(n,d) = q^n$), which implies both the assertions, since $V_q(n-1, -1) = 0$. Now suppose $d \geq 2$. Consider

$$k := n - \lceil \log_q (1 + V_q(n-1, d-2)) \rceil, \quad \text{i.e.,} \quad n - k = \lceil \log_q (1 + V_q(n-1, d-2)) \rceil.$$

Note that k is a positive integer $\leq n$, since $1 \leq V_q(n-1, d-2) \leq q^{n-1}$, i.e., $2 \leq 1 + V_q(n-1, d-2) \leq 1 + q^{n-1} < q^n$, and so $1 \leq \lceil \log_q (1 + V_q(n-1, d-2)) \rceil \leq n$. Further, from the definition of k and the fact that $\lceil x \rceil \geq x$ for all $x \in \mathbb{R}$, we obtain

$$q^{n-k} \geq q^{\log_q (1 + V_q(n-1, d-2))} = 1 + V_q(n-1, d-2) > V_q(n-1, d-2).$$

Hence by Lemma 8.14, there exists a $[n,k]_q$-code C with $d(C) \geqslant d$. We can see that this code C can be suitably punctured and extended to obtain a $[n,k]_q$-code $\widehat{C}$ with $d(\widehat{C}) = d$. Indeed, if $d_1 := d(C)$ and $d_1 > d$, then we may choose $x \in C$ with $x \neq 0$ and $\mathrm{w_H}(x) = d_1$. Pick a subset P of $\mathrm{supp}(x)$ such that $|P| = d_1 - d$. Consider the map $\phi : C \to \mathbb{F}_q^n$ that associates to $c = (c_1, \dots, c_n) \in C$, the n-tuple $\widehat{c} = (\widehat{c}_1, \dots, \widehat{c}_n)$, where $\widehat{c}_i = c_i$ if $i \notin P$ and $\widehat{c}_i = 0$ if $i \in P$, Evidently, ϕ is a linear map and it is injective because a nonzero element $c \in \ker(\phi)$ iwould satisfy $\mathrm{w_H}(c) \leq d_1 - d < d_1$, which is a contradiction. Hence the image of ϕ is a $[n,k]_q$-code, say $\widehat{C}$. Also for any nonzero $\widehat{c} = \phi(c) \in \widehat{C}$, it is clear that $\mathrm{w_H}(\widehat{c}) \geqslant \mathrm{w_H}(c) - (d_1 - d) \geqslant d_1 = (d_1 - d) = d$, and moreover, $\mathrm{w_H}(\widehat{x}) = d$. It follows that $d(\widehat{C}) = d$. This implies that

$$B_q(n,d) \geqslant q^k = q^{n - \lceil \log_q(1 + V_q(n-1,d-2)) \rceil},$$

and the first assertion is proved. This implies the weaker, but simpler bound, since

$$q^{n - \lceil \log_q(1 + V_q(n-1,d-2)) \rceil} \geqslant \frac{q^{n-1}}{q^{\log_q(1+V_q(n-1,d-2))}} = \frac{q^{n-1}}{1 + V_q(n-1,d-2)},$$

where the first inequality follows by noting that $\lceil \theta \rceil \leq \theta + 1$ for all $\theta \in \mathbb{R}$. □

The bounds obtained in this section lead to important asymptotic bounds for the largest possible rate of a family of q-ary codes having lengths going to ∞ and relative distances approaching δ, that is, for the function

$$\alpha(\delta) = \limsup_{n \to \infty} \frac{\log_q A_q(n, \lfloor \delta n \rfloor)}{n}.$$

These are not difficult, but rather technical. We thus skip them here, but refer the interested reader to the *Handbook* [14] or to [16].

9. Generalized Hamming Weights

The notion of generalized Hamming weight, also known as higher weight, is a natural generalization of the notion of minimum distance of a code. It is closely connected to questions about intersections of hypersurfaces of a given degree on a projective algebraic variety over a finite field and also has applications to cryptography. We provide some basics of the theory here.

Recall that for any codeword x of a code C, we defined the support of x. We extend this notion to subcodes, or more generally, subsets of C.

9.1. Definition. Let C be a q-ary linear code of length n. For any $D \subseteq C$, the *support* of D is defined by

$$\mathrm{supp}(D) := \{i \in \{1, \dots, n\} : \text{there is } x \in D \text{ with } x_i \neq 0\},$$

and the *Hamming weight* of D is defined by $\mathrm{w_H}(D) := |\,\mathrm{supp}(D)|$.

It is clear that if D is one-dimensional subspace of C spanned by x, then $\mathrm{supp}(D) = \mathrm{supp}(x)$ and $\mathrm{w_H}(D) = \mathrm{w_H}(x)$.

9.2. Definition. Let C be a $[n,k]_q$-code and r be a positive integer $\leq k$. The rth *generalized Hamming weight* or the rth *higher weight* of C is defined by

$$d_r(C) := \min\{\mathrm{w_H}(D) : D \text{ a subspace of } C \text{ with } \dim D = r\}.$$

The k-tuple $(d_1(C), \dots, d_k(C))$ is called the *weight hierarchy* of C.

Observe that $d_1(C) = d(C)$.

9.3. Exercise. Compute the weight hierarchy of the simplex codes.

If C is any $[n,k]_q$-code, then is clear that $d_1(C) \le d_2(C) \le \ldots \le d_k(C) \le n$. Actually, more is true.

9.4. Proposition (Monotonicity). *Let C be a $[n,k]_q$-code. Then*

$$d_1(C) < d_2(C) < \ldots < d_k(C) \le n.$$

Proof. Write $d_j = d_j(C)$ for $1 \le j \le k$. Fix an integer r with $1 \le r < k$. Clearly $d_r \le d_{r+1} \le n$. Suppose D is a subcode of C of dimension $r+1$ such that $\mathrm{w_H}(D) = d_{r+1}$, and suppose $i \in \mathrm{supp}(D)$. Consider $E = \{x \in D\colon x_i = 0\}$. Then E is the kernel of the ith projection map $\pi_i : D \longrightarrow \mathbb{F}_q$. Since $i \in \mathrm{supp}(D)$, the map π_i is nonzero, and hence $\dim E = r$. Moreover, $\mathrm{supp}(E) \subseteq \mathrm{supp}(D) \setminus \{i\}$, and so $\mathrm{w_H}(E) \le \mathrm{w_H}(D) - 1 = d_{r+1} - 1$. It follows that $d_r < d_{r+1}$. $\square$

9.5. Definition. A $[n,k]_q$-code C is said to be *nondegenerate* if C is not contained in a coordinate hyperplane of $\mathbb{F}_q^n$ or equivalently, if $d_k(C) = n$.

9.6. Corollary (Generalized Singleton Bound). *Let C be a $[n,k]_q$-code. Then $d_r(C) \le n-k+r$ for all $1 \le r \le k$.*

Proof. By the monotonicity, $d_k(C) \le n \Rightarrow d_{k-1}(C) \le n-1 \Rightarrow d_{k-2}(C) \le n-2$, and so on. In this way, we obtain $d_r(C) \le n-k+r$ for $1 \le r \le k$. $\square$

9.7. Exercise. Show that if a $[n,k]_q$-code C is MDS, then it is a r-MDS code, i.e., $d_r(C) = n-k+r$ for all $r = 1,2,\ldots k$.

9.8. Theorem (Wei Duality Theorem). *Let C be a $[n,k]_q$-code and let $d_r = d_r(C)$ and $d_s^\perp = d_s(C^\perp)$ for $r = 1,2,\ldots,k$ and $s = 1,2,\ldots,n-k$. Then*

$$\{d_1^\perp, d_2^\perp, \ldots, d_{n-k}^\perp\} = \{1,2,\ldots,n\} \setminus \{n+1-d_1, n+1-d_2, \ldots, n+1-d_k\}$$

For a proof of the above theorem, one may refer to the original paper of Wei [18]. In general, determining the weight hierarchy of a code is difficult. In a remarkable paper [10], Heijnen and Pellikaan determined the complete weight hierarchy of Reed-Muller codes. For a more streamlined proof of their result, and in fact, a more general result, we refer to the recent article [2].

References

[1] E. F. Assmus Jr. and J. D. Key, *Designs and their Codes*, Cambridge: Cambridge University Press, 1992.

[2] P. Beelen and M. Datta, Generalized Hamming weights of affine cartesian codes, `arXiv:1706.02114v2 [math.AG]` , 2017.

[3] T. P. Berger and P. Charpin, "The automorphism group of generalized Reed-Muller codes", *Discrete Math.*, vol. 117, pp. 1–17, 1993.

[4] K. Bogart, D.Goldberg and J.Gordon, An elementary proof of the MacWilliams theorem on equivalence of codes, *Inform.and Control* **37** (1978), 19–22.

[5] C. Carvalho, Applications of results from commutative algebra to the study of certain evaluation codes, *Course Notes* of the CIMPA School on Algebraic Methods in Coding Theory, São Paulo, Brazil, July 2017. [Available: `https://www.ime.usp.br/~cimpars/notes/sc4_01.pdf`

[6] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, 3rd ed., Springer-Verlag, New York, 1993.

[7] P. Delsarte, J. M. Goethals, and F. J. MacWilliams, "Generalized Reed-Muller codes and their relatives", *Inf. Control*, vol. 16, pp. 403–442, 1974.

[8] S. R. Ghorpade and G. Lachaud, Hyperplane sections of Grassmannians and the number of MDS linear codes, *Finite Fields Appl.* **7** (2001), 468–506.

[9] S. R. Ghorpade and K. V. Kaipa, Automorphism groups of Grassmann codes, *Finite Fields Appl.* **23** (2013), 80–102.

[10] P. Heijnen and R. Pellikaan, Generalized Hamming weights of q-ary Reed-Muller codes, *IEEE Trans. Inform. Theory* **44** (1998), 181–196.

[11] J.-R. Joly, "Équations et variétés algébraiques sur un corps fini", *Enseign. Math.* **19** (1973), 1–117.

[12] R. Knörr and W. Willems, "The automorphism groups of generalized Reed-Muller codes", *Astérisque*, vol. 181–182, pp. 195–207, 1990.

[13] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*, North Holland, Amsterdam, 1977.

[14] V. Pless and W.C. Huffmann (Eds.), *Handbook of Coding Theory*, North Holland, Amsterdam, 1998.

[15] N. J. A. Sloane, Error-Correcting Codes and Invariant Theory: New Applications of a Nineteenth-Century Technique, *Amer. Math. Monnthly* **84** (1977), 82–107.

[16] M. A. Tsfasman, S. G. Vlăduţ, and D. Nogin, *Algebraic Geometric Codes: Basic Notions*, Amer. Math. Soc., Providence, RI, 2012.

[17] J. H. van Lint, *Introduction to Coding Theory*, Third edition, Springer-Verlag, Berlin, 1999.

[18] V. K. Wei, Generalized Hamming weights for linear codes, *IEEE Trans. Inform. Theory* **37** (1991), 1412–1418.

Department of Mathematics, Indian Institute of Technology Bombay,
Powai, Mumbai 400076, India.
E-mail address: srg@math.iitb.ac.in
URL: http://www.math.iitb.ac.in/∼srg