5 Bounds in coding theory

Given a q-ary (n, M, d)-code, where n is fixed, the size M is a measure of the efficiency of the code, and the distance d is an indication of its error-correcting capability. It would be nice if both M and d could be as large as possible, but, as we shall see shortly in this chapter, this is not quite possible, and a compromise needs to be struck.

For given q, n and d, we shall discuss some well known upper and lower bounds for the largest possible value of M. In the case where M is actually equal to one of the well known bounds, interesting codes such as perfect codes and MDS codes are obtained. We also discuss certain properties and examples of some of these fascinating families.

5.1 The main coding theory problem

Let C be a q-ary code with parameters (n, M, d). Recall from Chapter 2 that the information rate (or transmission rate) of C is defined to be $\mathcal{R}(C) = (\log_q M)/n$. We also introduce here the notion of the relative minimum distance.

Definition 5.1.1 For a q-ary code C with parameters (n, M, d), the *relative minimum distance* of C is defined to be $\delta(C) = (d-1)/n$.

Remark 5.1.2 The relative minimum distance of C is often defined to be d/n in the literature, but defining it as $(d-1)/n$ leads sometimes to neater formulas (see Remark 5.4.4).

Example 5.1.3 (i) Consider the q-ary code $C = \mathbf{F}_q^n$. It is easy to see that $(n, M, d) = (n, q^n, 1)$ or, alternatively, $[n, k, d] = [n, n, 1]$. Hence,

$$\begin{aligned} \mathcal{R}(C) &= \frac{\log_q(q^n)}{n} = 1, \\ \delta(C) &= 0. \end{aligned}$$

This code has the maximum possible information rate, while its relative minimum distance is 0. As the minimum distance of a code is related closely to its error-correcting capability (cf. Theorem 2.5.10), a low relative minimum distance implies a relatively low error-correcting capability.

(ii) Consider the binary repetition code

$$C = \{\underbrace{00\cdots0}_{n}, \underbrace{11\cdots1}_{n}\}.$$

Clearly, $(n, M, d) = (n, 2, n)$ or, equivalently, C is a binary $[n, 1, n]$-linear code. Hence,

$$\mathcal{R}(C) = \frac{\log_2(2)}{n} = \frac{1}{n} \to 0,$$
$$\delta(C) = \frac{n-1}{n} \to 1,$$

as $n \to \infty$. As this code has the largest possible relative minimum distance, it has excellent error-correcting potential. However, this is achieved at the cost of very low efficiency, as reflected in the low information rate.

(iii) There is a family of binary linear codes (called Hamming codes – see Section 5.3.1) with parameters $(n, M, d) = (2^r - 1, 2^{n-r}, 3)$ or, equivalently, $[n, k, d] = [2^r - 1, 2^r - 1 - r, 3]$, for all integers $r \geq 2$. When $r \to \infty$, we have

$$\mathcal{R}(C) = \frac{\log_2(2^{n-r})}{n} = \frac{2^r - 1 - r}{2^r - 1} \to 1,$$
$$\delta(C) = \frac{2}{n} \to 0.$$

Again, while this family of codes has good information rates asymptotically, the relative minimum distances tend to zero, implying asymptotically bad error-correcting capabilities.

The previous examples should make it clear that a compromise between the transmission rate and the quality of error-correction is necessary.

Definition 5.1.4 For a given code alphabet A of size q (with $q > 1$) and given values of n and d, let $A_q(n, d)$ denote the largest possible size M for which there exists an (n, M, d)-code over A. Thus,

$$A_q(n, d) = \max\{M : \text{there exists an } (n, M, d)\text{-code over } A\}.$$

Any (n, M, d)-code C that has the maximum size, that is, for which $M = A_q(n, d)$, is called an *optimal code*.

Remark 5.1.5 (i) Note that $A_q(n, d)$ depends only on the size of A, n and d. It is independent of A.

(ii) The numbers $A_q(n, d)$ play a central role in coding theory, and much effort has been made in determining their values. In fact, the problem of determining the values of $A_q(n, d)$ is sometimes known as the *main coding theory problem*.

Instead of considering all codes, we may restrict ourselves to linear codes and obtain the following definition:

Definition 5.1.6 For a given prime power q and given values of n and d, let $B_q(n, d)$ denote the largest possible size q^k for which there exists an $[n, k, d]$-code over $\mathbf{F}_q$. Thus,

$$B_q(n, d) = \max\{q^k : \text{there exists an } [n, k, d]\text{-code over } \mathbf{F}_q\}.$$

While it is, in general, rather difficult to determine the exact values of $A_q(n, d)$ and $B_q(n, d)$, there are some properties that afford easy proofs.

Theorem 5.1.7 *Let $q \geq 2$ be a prime power. Then*

(i) $B_q(n, d) \leq A_q(n, d) \leq q^n$ *for all* $1 \leq d \leq n$;
(ii) $B_q(n, 1) = A_q(n, 1) = q^n$;
(iii) $B_q(n, n) = A_q(n, n) = q$.

Proof. The first inequality in (i) is obvious from the definitions, while the second one is clear since any (n, M, d)-code over $\mathbf{F}_q$, being a nonempty subset of $\mathbf{F}_q^n$, must have $M \leq q^n$.

To show (ii), note that $\mathbf{F}_q^n$ is an $[n, n, 1]$-linear code, and hence an $(n, q^n, 1)$-code, over $\mathbf{F}_q$, so $q^n \leq B_q(n, 1) \leq q^n$; i.e., $B_q(n, 1) = A_q(n, 1) = q^n$.

For (iii), let C be an (n, M, n)-code over $\mathbf{F}_q$. Since the codewords are of length n, and the distance between two distinct codewords is $\geq n$, it follows that the distance between two distinct codewords is actually n. This means that two distinct codewords must differ at all the coordinates. Therefore, at each coordinate, all the M words must take different values, so $M \leq q$, implying $B_q(n, n) \leq A_q(n, n) \leq q$. The repetition code of length n, i.e., $\{(a, a, \ldots, a) : a \in \mathbf{F}_q\}$, is an $[n, 1, n]$-linear code, and hence an (n, q, n)-code, over $\mathbf{F}_q$, so $B_q(n, n) = A_q(n, n) = q$. □

In the case of binary codes, there are additional elementary results on $A_2(n, d)$ and $B_2(n, d)$. Before we discuss them, we need to introduce the

notion of the extended code, which is a useful concept in its own right. For a binary linear code, its extended code is obtained by adding a parity-check coordinate. This idea can be generalized to codes over any finite field.

Definition 5.1.8 For any code C over $\mathbf{F}_q$, the *extended code of* C, denoted by $\overline{C}$, is defined to be

$$\overline{C} = \left\{ \left(c_1, \ldots, c_n, -\sum_{i=1}^{n} c_i \right) \ : \ (c_1, \ldots, c_n) \in C \right\}.$$

When $q = 2$, the extra coordinate $-\sum_{i=1}^{n} c_i = \sum_{i=1}^{n} c_i$ added to the codeword $(c_1, \ldots, c_n)$ is called the *parity-check* coordinate.

Theorem 5.1.9 *If C is an (n, M, d)-code over $\mathbf{F}_q$, then $\overline{C}$ is an $(n+1, M, d')$-code over $\mathbf{F}_q$, with $d \leq d' \leq d+1$. If C is linear, then so is $\overline{C}$. Moreover, when C is linear,*

$$\left(\begin{array}{c|c} & 0 \\ H & \vdots \\ & 0 \\ \hline 1 \cdots 1 & 1 \end{array}\right)$$

is a parity-check matrix of $\overline{C}$ if H is a parity-check matrix of C.

The proof is straightforward, so it is left to the reader (Exercise 5.3).

Example 5.1.10 (i) Consider the binary linear code $C_1 = \{000, 110, 011, 101\}$. It has parameters $[3, 2, 2]$. The extended code

$$\overline{C_1} = \{0000, 1100, 0110, 1010\}$$

is a binary $[4, 2, 2]$-linear code.

(ii) Consider the binary linear code $C_2 = \{000, 111, 011, 100\}$. It has parameters $[3, 2, 1]$. The extended code

$$\overline{C_2} = \{0000, 1111, 0110, 1001\}$$

is a binary $[4, 2, 2]$-linear code.

This example shows that the minimum distance $d(\overline{C})$ can achieve both $d(C)$ and $d(C) + 1$. Example 5.1.10(ii) is an illustration of the following fact.

Theorem 5.1.11 *Suppose d is odd.*

(i) *Then a binary (n, M, d)-code exists if and only if a binary $(n+1, M, d+1)$-code exists. Therefore, if d is odd, $A_2(n+1, d+1) = A_2(n, d)$.*

(ii) *Similarly, a binary* $[n, k, d]$*-linear code exists if and only if a binary* $[n+1, k, d+1]$*-linear code exists, so* $B_2(n+1, d+1) = B_2(n, d)$.

Proof. For (i), the latter statement follows immediately from the previous one, so we only prove the earlier statement.

Suppose that there is a binary (n, M, d)-code C, where d is odd. Then, from Theorem 5.1.9, $\overline{C}$ is an $(n+1, M, d')$-code with $d \leq d' \leq d+1$.

Note that $\mathrm{wt}(\mathbf{x}')$ is even for all $\mathbf{x}' \in \overline{C}$. Therefore, Lemma 4.3.5 and Corollary 4.3.4 show that $d(\mathbf{x}', \mathbf{y}')$ is even for all $\mathbf{x}', \mathbf{y}' \in \overline{C}$, so d' is even. Since d is odd and $d \leq d' \leq d+1$, it follows that $d' = d+1$.

We have therefore shown that, if there is a binary (n, M, d)-code C, then $\overline{C}$ is a binary $(n+1, M, d+1)$-code.

Next, we suppose that there exists a binary $(n+1, M, d+1)$-code D, where d is odd. Choose codewords $\mathbf{x}$ and $\mathbf{y}$ in D such that $d(\mathbf{x}, \mathbf{y}) = d+1$. In other words, $\mathbf{x}$ and $\mathbf{y}$ differ at $d+1 \geq 2$ coordinates. Choose a coordinate where $\mathbf{x}$ and $\mathbf{y}$ differ, and let D' be the code obtained by deleting this coordinate from all the codewords of D. (The code D' is called a *punctured code*; see Theorem 6.1.1(iii).) Then D' is a binary (n, M, d)-code.

For (ii), it suffices to observe that, in the proof of (i), if C is linear, then so is $\overline{C}$; similarly, if D is linear, then so is D'. □

Remark 5.1.12 The last statement in Theorem 5.1.11(i) is equivalent to 'if d is even, then $A_2(n, d) = A_2(n-1, d-1)$'. There is also an analogue for (ii).

While the determination of the exact values of $A_q(n, d)$ and $B_q(n, d)$ can be rather difficult, several well known bounds, both upper and lower ones, do exist. We shall discuss some of them in the following sections.

A list of lower bounds and, in some cases, exact values for $A_2(n, d)$ may be found at the following webpage maintained by Simon Litsyn of Tel Aviv University:

http://www.eng.tau.ac.il/~litsyn/tableand/index.html.

The following website, maintained by Andries E. Brouwer of Technische Universiteit Eindhoven, contains tables that give the best known bounds (upper and lower) on the distance d for q-ary linear codes ($q \leq 9$) of given length and dimension:

http://www.win.tue.nl/~aeb/voorlincod.html.

5.2 Lower bounds

We discuss two well known lower bounds: the sphere-covering bound (for $A_q(n, d)$) and the Gilbert–Varshamov bound (for $B_q(n, d)$).

5.2.1 Sphere-covering bound

Definition 5.2.1 Let A be an alphabet of size q, where $q > 1$. For any vector $\mathbf{u} \in A^n$ and any integer $r \geq 0$, the *sphere* of radius r and centre $\mathbf{u}$, denoted $S_A(\mathbf{u}, r)$, is the set $\{\mathbf{v} \in A^n : d(\mathbf{u}, \mathbf{v}) \leq r\}$.

Definition 5.2.2 For a given integer $q > 1$, a positive integer n and an integer $r \geq 0$, define $V_q^n(r)$ to be

$$V_q^n(r) = \begin{cases} \binom{n}{0} + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{r}(q-1)^r & \text{if } 0 \leq r \leq n \\ q^n & \text{if } n \leq r. \end{cases}$$

Lemma 5.2.3 *For all integers $r \geq 0$, a sphere of radius r in A^n contains exactly $V_q^n(r)$ vectors, where A is an alphabet of size $q > 1$.*

Proof. Fix a vector $\mathbf{u} \in A^n$. We determine the number of vectors $\mathbf{v} \in A^n$ such that $d(\mathbf{u}, \mathbf{v}) = m$; i.e., the number of vectors in A^n of distance exactly m from $\mathbf{u}$. The number of ways in which to choose the m coordinates where $\mathbf{v}$ differs from $\mathbf{u}$ is given by $\binom{n}{m}$. For each coordinate, we have $q-1$ choices for that coordinate in $\mathbf{v}$. Therefore, the total number of vectors of distance m from $\mathbf{u}$ is given by $\binom{n}{m}(q-1)^m$. For $0 \leq r \leq n$, Lemma 5.2.3 now follows.

When $r \geq n$, note that $S_A(\mathbf{u}, r) = A^n$, hence it contains $V_q^n(r) = q^n$ vectors. □

We are now ready to state and prove the sphere-covering bound.

Theorem 5.2.4 (Sphere-covering bound.) *For an integer $q > 1$ and integers n, d such that $1 \leq d \leq n$, we have*

$$\frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i} \leq A_q(n, d).$$

Proof. Let $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M\}$ be an optimal (n, M, d)-code over A with $|A| = q$, so $M = A_q(n, d)$. Since C has the maximum size, there can be no word in A^n whose distance from every codeword in C is at least d. If there were such a word, we could simply include it in C, and thereby obtain an $(n, M+1, d)$-code.

Therefore, for every vector $\mathbf{x}$ in A^n, there is at least one codeword $\mathbf{c}_i$ in C such that $d(\mathbf{x}, \mathbf{c}_i)$ is at most $d-1$; i.e., $\mathbf{x} \in S_A(\mathbf{c}_i, d-1)$. Hence, every word in A^n is contained in at least one of the spheres $S_A(\mathbf{c}_i, d-1)$. In other words,

$$A^n \subseteq \bigcup_{i=1}^{M} S_A(\mathbf{c}_i, d-1).$$

(For this reason, we say that the spheres $S_A(\mathbf{c}_i, d-1)$ $(1 \leq i \leq M)$ cover A^n, hence the name 'sphere-covering' bound.)

Since $|A^n| = q^n$ and $|S_A(\mathbf{c}_i, d-1)| = V_q^n(d-1)$ for any i, we have

$$q^n \leq M \cdot V_q^n(d-1),$$

implying that

$$\frac{q^n}{V_q^n(d-1)} \leq M = A_q(n, d).$$

□

Some examples of the lower bounds for $A_q(n, d)$ given by the sphere-covering bound are found in Tables 5.2–5.4 (see Example 5.5.5).

The following example illustrates how $A_q(n, d)$ may be found in some special cases. In the example, the lower bound is given by the sphere-covering bound. Then a combinatorial argument shows that the lower bound must also be an upper bound for $A_q(n, d)$, hence yielding the exact value of $A_q(n, d)$.

Example 5.2.5 We prove that $A_2(5, 4) = 2$.

The sphere-covering bound shows that $A_2(5, 4) \geq 2$.

By Theorem 5.1.11, we see that $A_2(5, 4) = A_2(4, 3)$, so we next show that $A_2(4, 3) \leq 2$. Let C be a binary $(4, M, 3)$-code and let (x_1, x_2, x_3, x_4) be a codeword in C. Since $d(C) = 3$, the other codewords in C must be of the following forms:

$$(x_1, \overline{x_2}, \overline{x_3}, \overline{x_4}), \quad (\overline{x_1}, x_2, \overline{x_3}, \overline{x_4}), \quad (\overline{x_1}, \overline{x_2}, x_3, \overline{x_4}),$$
$$(\overline{x_1}, \overline{x_2}, \overline{x_3}, x_4), \quad (\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4}),$$

where $\overline{x_i}$ is defined by

$$\overline{x_i} = \begin{cases} 1 & \text{if } x_i = 0 \\ 0 & \text{if } x_i = 1. \end{cases}$$

However, no pair of these five words are of distance 3 (or more) apart, and so only one of them can be included in C. Hence, $M \leq 2$, implying that $A_2(4, 3) \leq 2$. Therefore, $A_2(5, 4) = A_2(4, 3) = 2$.

5.2.2 Gilbert–Varshamov bound

The Gilbert–Varshamov bound is a lower bound for $B_q(n, d)$ (i.e., for linear codes) known since the 1950s. There is also an asymptotic version of the Gilbert–Varshamov bound, which concerns infinite sequences of codes whose lengths tend to infinity. However, we shall not discuss this asymptotic result here. The interested reader may refer to Chap. 17, Theorem 30 of ref. [13]. For a long time, the asymptotic Gilbert–Varshamov bound was the best lower bound known to be attainable by an infinite family of linear codes, so it became a sort of benchmark for judging the 'goodness' of an infinite sequence of linear codes. Between 1977 and 1982, V. D. Goppa constructed algebraic-geometry codes using algebraic curves over finite fields with many rational points. A major breakthrough in coding theory was achieved shortly after these discoveries, when it was shown that there are sequences of algebraic-geometry codes that perform better than the asymptotic Gilbert–Varshamov bound for certain sufficiently large q.

Theorem 5.2.6 (Gilbert–Varshamov bound.) *Let n, k and d be integers satisfying $2 \le d \le n$ and $1 \le k \le n$. If*

$$\sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i < q^{n-k}, \tag{5.1}$$

then there exists an $[n, k]$-linear code over $\mathbf{F}_q$ with minimum distance at least d.

Proof. We shall show that, if (5.1) holds, then there exists an $(n-k) \times n$ matrix H over $\mathbf{F}_q$ such that every $d-1$ columns of H are linearly independent.

We construct H as follows. Let $\mathbf{c}_j$ denote the jth column of H.

Let $\mathbf{c}_1$ be any nonzero vector in $\mathbf{F}_q^{n-k}$. Let $\mathbf{c}_2$ be any vector not in the span of $\mathbf{c}_1$. For any $2 \le j \le n$, let $\mathbf{c}_j$ be any vector that is not in the linear span of $d-2$ (or fewer) of the vectors $\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}$.

Note that the number of vectors in the linear span of $d-2$ or fewer of $\mathbf{c}_1, \ldots, \mathbf{c}_{j-1}$ $(2 \le j \le n)$ is given by

$$\sum_{i=0}^{d-2} \binom{j-1}{i} (q-1)^i \le \sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i < q^{n-k}.$$

Hence, the vector $\mathbf{c}_j$ $(2 \le j \le n)$ can always be found.

The matrix H constructed in this manner is an $(n-k) \times n$ matrix, and any $d-1$ of its columns are linearly independent. The null space of H is a linear code over $\mathbf{F}_q$ of length n, of distance at least d, and of dimension at least k.

By turning to a k-dimensional subspace, we obtain a linear code of the desired type. □

Corollary 5.2.7 *For a prime power $q > 1$ and integers n, d such that $2 \leq d \leq n$, we have*

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

Proof. Put

$$k = n - \lceil \log_q(V_q^{n-1}(d-2)+1)\rceil.$$

Then (5.1) is satisfied and thus there exists a q-ary $[n,k,d_1]$-linear code with $d_1 \geq d$ by Theorem 5.2.6. By changing certain $d_1 - d$ fixed coordinates to 0, we obtain a q-ary $[n,k,d]$-linear code (see also Theorem 6.1.1(iv)). Our result follows from the fact that $B_q(n,d) \geq q^k$. □

5.3 Hamming bound and perfect codes

The first upper bound for $A_q(n,d)$ that we will discuss is the Hamming bound, also known as the sphere-packing bound.

Theorem 5.3.1 (Hamming or sphere-packing bound.) *For an integer $q > 1$ and integers n, d such that $1 \leq d \leq n$, we have*

$$A_q(n,d) \leq \frac{q^n}{\sum_{i=0}^{\lfloor (d-1)/2\rfloor} \binom{n}{i}(q-1)^i}.$$

Proof. Let $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_M\}$ be an optimal (n, M, d)-code over A (with $|A| = q$), so $M = A_q(n,d)$. Let $e = \lfloor (d-1)/2 \rfloor$; then the packing spheres $S_A(\mathbf{c}_i, e)$ are disjoint. Hence, we have

$$\bigcup_{i=1}^{M} S_A(\mathbf{c}_i, e) \subseteq A^n,$$

where the union on the left hand side is a disjoint union. Since $|A^n| = q^n$ and $|S_A(\mathbf{c}_i, e)| = V_q^n(e)$ for any i, we have

$$M \cdot V_q^n(e) \leq q^n,$$

implying that

$$A_q(n,d) = M \leq \frac{q^n}{V_q^n(e)} = \frac{q^n}{V_q^n(\lfloor (d-1)/2 \rfloor)}.$$

This completes the proof. □

Definition 5.3.2 A q-ary code that attains the Hamming (or sphere-packing) bound, i.e., one which has $q^n \Big/ \left(\sum_{i=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{i}(q-1)^i\right)$ codewords, is called a *perfect* code.

Some of the earliest known codes, such as the Hamming codes and the Golay codes, are perfect codes.

5.3.1 Binary Hamming codes

Hamming codes were discovered by R. W. Hamming and M. J. E. Golay. They form an important class of codes – they have interesting properties and are easy to encode and decode.

While Hamming codes are defined over all finite fields $\mathbf{F}_q$, we begin by discussing specifically the binary Hamming codes. These codes form a special case of the general q-ary Hamming codes, but because they can be described more simply than the general q-ary Hamming codes, and because they are arguably the most interesting Hamming codes, it is worthwhile discussing them separately from the other Hamming codes.

Definition 5.3.3 Let $r \geq 2$. A binary linear code of length $n = 2^r - 1$, with parity-check matrix H whose columns consist of all the nonzero vectors of $\mathbf{F}_2^r$, is called a *binary Hamming code* of length $2^r - 1$. It is denoted by Ham$(r, 2)$.

Remark 5.3.4 (i) The order of the columns of H has not been fixed in Definition 5.3.3. Hence, for each $r \geq 2$, the binary Hamming code Ham$(r, 2)$ is only well defined up to equivalence of codes.

(ii) Note that the rows of H are linearly independent since H contains all the r columns of weight 1 words. Hence, H is indeed a parity-check matrix.

Example 5.3.5 Ham$(3, 2)$: A Hamming code of length 7 with a parity-check matrix

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

Proposition 5.3.6 (Properties of the binary Hamming codes.)

(i) *All the binary Hamming codes of a given length are equivalent.*
(ii) *The dimension of* Ham$(r, 2)$ *is* $k = 2^r - 1 - r$.
(iii) *The distance of* Ham$(r, 2)$ *is* $d = 3$, *hence* Ham$(r, 2)$ *is exactly single-error-correcting.*
(iv) *Binary Hamming codes are perfect codes.*

Proof. (i) For a given length, any parity-check matrix can be obtained from another by a permutation of the columns, hence the corresponding binary Hamming codes are equivalent.

(ii) Since H, a parity-check matrix for Ham$(r, 2)$, is an $r \times (2^r - 1)$ matrix, the dimension of Ham$(r, 2)$ is $2^r - 1 - r$.

(iii) Since no two columns of H are equal, any two columns of H are linearly independent. On the other hand, H contains the columns $(100\ldots0)^{\mathrm{T}}$, $(010\ldots0)^{\mathrm{T}}$ and $(110\ldots0)^{\mathrm{T}}$, which form a linearly dependent set. Hence, by Corollary 4.5.7, the distance of Ham$(r, 2)$ is equal to 3. It then follows from Theorem 2.5.10 that Ham$(r, 2)$ is single-error-correcting.

(iv) It can be verified easily that Ham$(r, 2)$ satisfies the Hamming bound and is hence a perfect code. □

Decoding with a binary Hamming code

Since Ham$(r, 2)$ is perfect single-error-correcting, the coset leaders are precisely the 2^r $(= n + 1)$ vectors of length n of weight ≤ 1. Let $\mathbf{e}_j$ denote the vector with 1 in the jth coordinate and 0 elsewhere. Then the syndrome of $\mathbf{e}_j$ is just $\mathbf{e}_j H^{\mathrm{T}}$, i.e., the transpose of the jth column of H.

Hence, if the columns of H are arranged in the order of increasing binary numbers (i.e., the jth column of H is just the binary representation of j; see Exercise 4.43), the decoding is given by:

Step 1: When $\mathbf{w}$ is received, calculate its syndrome $S(\mathbf{w}) = \mathbf{w}H^{\mathrm{T}}$.

Step 2: If $S(\mathbf{w}) = \mathbf{0}$, assume $\mathbf{w}$ was the codeword sent.

Step 3: If $S(\mathbf{w}) \neq \mathbf{0}$, then $S(\mathbf{w})$ is the binary representation of j, for some $1 \leq j \leq 2^r - 1$. *Assuming a single error*, the word $\mathbf{e}_j$ gives the error, so we take the sent word to be $\mathbf{w} - \mathbf{e}_j$ (or, equivalently, $\mathbf{w} + \mathbf{e}_j$).

Example 5.3.7 We construct a syndrome look-up table for the Hamming code given in Example 5.3.5, and use it to decode $\mathbf{w} = 1001001$ (see Table 5.1).

The syndrome is $\mathbf{w}H^{\mathrm{T}} = 010$, which gives the coset leader $\mathbf{e}_2 = 0100000$. We can then decode $\mathbf{w}$ as $\mathbf{w} - \mathbf{e}_2 = \mathbf{w} + \mathbf{e}_2 = 1101001$.

Definition 5.3.8 The dual of the binary Hamming code Ham$(r, 2)$ is called a binary *simplex code*. It is sometimes denoted by $S(r, 2)$.

Some of the properties of the simplex codes are contained in Exercise 5.19.

Table 5.1.

Coset leader $\mathbf{u}$	Syndrome $S(\mathbf{u})$
0000000	000
1000000	001
0100000	010
0010000	011
0001000	100
0000100	101
0000010	110
0000001	111

Definition 5.3.9 The *extended binary Hamming code*, denoted $\overline{\text{Ham}(r, 2)}$, is the code obtained from $\text{Ham}(r, 2)$ by adding a parity-check coordinate.

Proposition 5.3.10 (Properties of the extended binary Hamming codes.)

(i) $\overline{\text{Ham}(r, 2)}$ *is a binary* $[2^r, 2^r - 1 - r, 4]$*-linear code.*
(ii) *A parity-check matrix* $\overline{H}$ *for* $\overline{\text{Ham}(r, 2)}$ *is*

$$\overline{H} = \left(\begin{array}{c|c} & 0 \\ H & \vdots \\ & 0 \\ \hline 1 \cdots 1 & 1 \end{array}\right),$$

where H *is a parity-check matrix for* $\text{Ham}(r, 2)$.

Proposition 5.3.10 follows immediately from Theorem 5.1.9 and the proof of Theorem 5.1.11.

Remark 5.3.11 The rate of transmission for $\overline{\text{Ham}(r, 2)}$ is slower than that of $\text{Ham}(r, 2)$, but the extended code is better suited for incomplete decoding.

Example 5.3.12 Let $r = 3$ and take

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

Note that every codeword is made up of 8 bits and recall that the syndrome of the error vector $\mathbf{e}_j$ is just the transpose of the jth column of $\overline{H}$. Assuming that as

few errors as possible have occurred, the incomplete decoding works as follows. Suppose the received vector is $\mathbf{w}$, so its syndrome is $S(\mathbf{w}) = \mathbf{w}\overline{H}^{\mathrm{T}}$. Suppose it is $S(\mathbf{w}) = (s_1, s_2, s_3, s_4)$. Then $S(\mathbf{w})$ must fall into one of the following four categories:

(i) $s_4 = 0$ and $(s_1, s_2, s_3) = \mathbf{0}$. In this case, $S(\mathbf{w}) = \mathbf{0}$, so $\mathbf{w} \in \overline{\mathrm{Ham}(3, 2)}$. We may therefore assume that there are no errors.

(ii) $s_4 = 0$ and $(s_1, s_2, s_3) \neq \mathbf{0}$. Since $S(\mathbf{w}) \neq \mathbf{0}$, at least one error must have occurred. If exactly one error occurs and it occurs in the jth bit, then the error vector is $\mathbf{e}_j$, so $S(\mathbf{w}) = S(\mathbf{e}_j)$, which is the transpose of the jth column of $\overline{H}$. An inspection of $\overline{H}$ shows immediately that the last coordinate (the one corresponding to s_4) of every column is 1, contradicting the fact that $s_4 = 0$. Hence, the assumption that exactly one error has occurred is flawed, and we may assume at least two errors have occurred and seek retransmission.

(iii) $s_4 = 1$ and $(s_1, s_2, s_3) = \mathbf{0}$. Again, since $S(\mathbf{w}) \neq \mathbf{0}$, at least one error has occurred. It is easy to see that $S(\mathbf{w}) = S(\mathbf{e}_8)$, so we may assume a single error in the last coordinate, i.e., the parity-check coordinate.

(iv) $s_4 = 1$ and $(s_1, s_2, s_3) \neq \mathbf{0}$. As before, it is easy to check that $S(\mathbf{w})$ must coincide with the transpose of one of the first seven columns of $\overline{H}$, say the jth column. Hence, $S(\mathbf{w}) = S(\mathbf{e}_j)$, and we may assume a single error in the jth coordinate. In fact, given the way the columns of $\overline{H}$ are arranged, j is the number whose binary representation is (s_1, s_2, s_3).

5.3.2 q-ary Hamming codes

Let $q \geq 2$ be any prime power. Note that any nonzero vector $\mathbf{v} \in \mathbf{F}_q^r$ generates a subspace $< \mathbf{v} >$ of dimension 1. Furthermore, for $\mathbf{v}, \mathbf{w} \in \mathbf{F}_q^r \backslash \{\mathbf{0}\}$, $< \mathbf{v} >= < \mathbf{w} >$ if and only if there is a nonzero scalar $\lambda \in \mathbf{F}_q \backslash \{0\}$ such that $\mathbf{v} = \lambda\mathbf{w}$. Therefore, there are exactly $(q^r - 1)/(q - 1)$ distinct subspaces of dimension 1 in $\mathbf{F}_q^r$.

Definition 5.3.13 Let $r \geq 2$. A q-ary linear code, whose parity-check matrix H has the property that the columns of H are made up of precisely one nonzero vector from each vector subspace of dimension 1 of $\mathbf{F}_q^r$, is called a *q-ary Hamming code*, often denoted as $\mathrm{Ham}(r, q)$.

It is an easy exercise to show that, when $q = 2$, the code defined here is the same as the binary Hamming code defined earlier.

Remark 5.3.14 An easy way to write down a parity-check matrix for Ham(r, q) is to list as columns all the nonzero r-tuples in $\mathbf{F}_q^r$ whose first nonzero entry is 1.

Proposition 5.3.15 (Properties of the q-ary Hamming codes.)

(i) Ham(r, q) *is a* $[(q^r - 1)/(q - 1), (q^r - 1)/(q - 1) - r, 3]$*-code.*
(ii) Ham(r, q) *is a perfect exactly single-error-correcting code.*

The proof of Proposition 5.3.15 resembles that of Proposition 5.3.6, so we leave it as an exercise to the reader (Exercise 5.17).

Decoding with a q-ary Hamming code

Since Ham(r, q) is a perfect single-error-correcting code, the coset leaders, other than $\mathbf{0}$, are exactly the vectors of weight 1. A typical coset leader is then denoted by $\mathbf{e}_{j,b}$ $(1 \leq j \leq n, b \in \mathbf{F}_q \backslash \{0\})$ – the vector whose jth coordinate is b and the other coordinates are 0. Note that

$$S(\mathbf{e}_{j,b}) = b\mathbf{c}_j^{\mathrm{T}},$$

where $\mathbf{c}_j$ denotes the jth column of H.

Decoding works as follows:

Step 1: Given a received word $\mathbf{w}$, calculate $S(\mathbf{w}) = \mathbf{w}H^{\mathrm{T}}$.

Step 2: If $S(\mathbf{w}) = \mathbf{0}$, then assume no errors.

Step 3: If $S(\mathbf{w}) \neq \mathbf{0}$, then find the unique $\mathbf{e}_{j,b}$ such that $S(\mathbf{w}) = S(\mathbf{e}_{j,b})$. The received word is then taken to be $\mathbf{w} - \mathbf{e}_{j,b}$.

Definition 5.3.16 The dual of the q-ary Hamming code Ham(r, q) is called a q-ary *simplex code*. It is sometimes denoted by $S(r, q)$.

The reader may refer to Exercise 5.19 for some of the properties of the q-ary simplex codes.

5.3.3 Golay codes

The Golay codes were discovered by M. J. E. Golay in the late 1940s. The (unextended) Golay codes are examples of perfect codes. It turns out that the Golay codes are essentially unique in the sense that binary or ternary codes with the same parameters as them can be shown to be equivalent to them.

Binary Golay codes

Definition 5.3.17 Let G be the 12×24 matrix

$$G = (I_{12}|A),$$

where I_{12} is the 12×12 identity matrix and A is the 12×12 matrix

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

The binary linear code with generator matrix G is called the *extended binary Golay code* and will be denoted by G_{24}.

Remark 5.3.18 (i) The Voyager 1 and 2 spacecraft were launched towards Jupiter and Saturn in 1977. This code was used in the encoding and decoding of the general science and engineering (GSE) data for the missions.

(ii) It is also common to call any code that is equivalent to the linear code with generator matrix G an extended binary Golay code.

Proposition 5.3.19 (Properties of the extended binary Golay code.)

(i) *The length of* G_{24} *is* 24 *and its dimension is* 12.
(ii) *A parity-check matrix for* G_{24} *is the* 12×24 *matrix*

$$H = (A|I_{12}).$$

(iii) *The code* G_{24} *is self-dual, i.e.,* $G_{24}^{\perp} = G_{24}$.
(iv) *Another parity-check matrix for* G_{24} *is the* 12×24 *matrix*

$$H' = (I_{12}|A)(=G).$$

(v) *Another generator matrix for* G_{24} *is the* 12×24 *matrix*

$$G' = (A|I_{12})(=H).$$

(vi) *The weight of every codeword in G_{24} is a multiple of* 4.
(vii) *The code G_{24} has no codeword of weight* 4*, so the distance of G_{24} is* $d = 8$.
(viii) *The code G_{24} is an exactly three-error-correcting code.*

Proof. (i) This is clear from the definition.

(ii) This follows from Theorem 4.5.9.

(iii) Note that the rows of G are orthogonal; i.e., if $\mathbf{r}_i$ and $\mathbf{r}_j$ are any two rows of G, then $\mathbf{r}_i \cdot \mathbf{r}_j = 0$. This implies that $G_{24} \subseteq G_{24}^{\perp}$. On the other hand, since both G_{24} and $G_{24}^{\perp}$ have dimension 12, we must have $G_{24} = G_{24}^{\perp}$.

(iv) A parity-check matrix of G_{24} is a generator matrix of $G_{24}^{\perp} = G_{24}$, and G is one such matrix.

(v) A generator matrix of G_{24} is a parity-check matrix of $G_{24}^{\perp} = G_{24}$, and H is one such matrix.

(vi) Let $\mathbf{v}$ be a codeword in G_{24}. We want to show that $\mathrm{wt}(\mathbf{v})$ is a multiple of 4. Note that $\mathbf{v}$ is a linear combination of the rows of G. Let $\mathbf{r}_i$ denote the ith row of G.

First, suppose $\mathbf{v}$ is one of the rows of G. Since the rows of G have weight 8 or 12, the weight of $\mathbf{v}$ is a multiple of 4.

Next, let $\mathbf{v}$ be the sum $\mathbf{v} = \mathbf{r}_i + \mathbf{r}_j$ of two different rows of G. Since G_{24} is self-dual, Exercise 4.22(d) shows that the weight of $\mathbf{v}$ is divisible by 4.

We then continue by induction to finish the proof.

(vii) Note that the last row of G is a codeword of weight 8. This fact, together with statement (vi) of this proposition, implies that $d = 4$ or 8.

Suppose G_{24} contains a nonzero codeword $\mathbf{v}$ with $\mathrm{wt}(\mathbf{v}) = 4$. Write $\mathbf{v}$ as $(\mathbf{v}_1, \mathbf{v}_2)$, where $\mathbf{v}_1$ is the vector (of length 12) made up of the first 12 coordinates of $\mathbf{v}$, and $\mathbf{v}_2$ is the vector (also of length 12) made up of the last 12 coordinates of $\mathbf{v}$. Then one of the following situations must occur:

Case (1) $\mathrm{wt}(\mathbf{v}_1) = 0$ and $\mathrm{wt}(\mathbf{v}_2) = 4$. This cannot possibly happen since, by looking at the generator matrix G, the only such word is $\mathbf{0}$, which is of weight 0.

Case (2) $\mathrm{wt}(\mathbf{v}_1) = 1$ and $\mathrm{wt}(\mathbf{v}_2) = 3$. In this case, again by looking at G, $\mathbf{v}$ must be one of the rows of G, which is again a contradiction.

Case (3) $\mathrm{wt}(\mathbf{v}_1) = 2$ and $\mathrm{wt}(\mathbf{v}_2) = 2$. Then $\mathbf{v}$ is the sum of two of the rows of G. It is easy to check that none of such sums would give $\mathrm{wt}(\mathbf{v}_2) = 2$.

Case (4) $\mathrm{wt}(\mathbf{v}_1) = 3$ and $\mathrm{wt}(\mathbf{v}_2) = 1$. Since G' is a generator matrix, $\mathbf{v}$ must be one of the rows of G', which clearly gives a contradiction.

Case (5) $\mathrm{wt}(\mathbf{v}_1) = 4$ and $\mathrm{wt}(\mathbf{v}_2) = 0$. This case is similar to case (1), using G' instead of G.

Since we obtain contradictions in all these cases, $d = 4$ is impossible. Thus, $d = 8$.

(viii) This follows from statement (vii) above and Theorem 2.5.10. □

Definition 5.3.20 Let $\hat{G}$ be the 12×23 matrix

$$\hat{G} = (I_{12}|\hat{A}),$$

where I_{12} is the 12×12 identity matrix and $\hat{A}$ is the 12×11 matrix obtained from the matrix A by deleting the last column of A. The binary linear code with generator matrix $\hat{G}$ is called the *binary Golay code* and will be denoted by G_{23}.

Remark 5.3.21 Alternatively, the binary Golay code can be defined as the code obtained from G_{24} by deleting the last coordinate of every codeword.

Proposition 5.3.22 (Properties of the binary Golay code.)

(i) *The length of* G_{23} *is* 23 *and its dimension is* 12.
(ii) *A parity-check matrix for* G_{23} *is the* 11×23 *matrix*

$$\hat{H} = (\hat{A}^{\mathrm{T}}|I_{11}).$$

(iii) *The extended code of* G_{23} *is* G_{24}.
(iv) *The distance of* G_{23} *is* $d = 7$.
(v) *The code* G_{23} *is a perfect exactly three-error-correcting code.*

The proof is left as an exercise to the reader (see Exercise 5.24).

Ternary Golay codes

Definition 5.3.23 The *extended ternary Golay code*, denoted by G_{12}, is the ternary linear code with generator matrix $G = (I_6|B)$, where B is the 6×6 matrix

$$B = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 & 2 & 1 \\ 1 & 1 & 0 & 1 & 2 & 2 \\ 1 & 2 & 1 & 0 & 1 & 2 \\ 1 & 2 & 2 & 1 & 0 & 1 \\ 1 & 1 & 2 & 2 & 1 & 0 \end{pmatrix}.$$

Remark 5.3.24 Any linear code that is equivalent to the above code is also called an extended ternary Golay code.

By mimicking the method used in Proposition 5.3.19, it is possible to check that G_{12} is a self-dual ternary [12, 6, 6]-code (see Exercise 5.28).

Definition 5.3.25 The *ternary Golay code* G_{11} is the code obtained by puncturing G_{12} in the last coordinate.

One can verify that G_{11} satisfies the Hamming bound and is hence a perfect ternary [11, 6, 5]-code (see Exercise 5.29).

5.3.4 Some remarks on perfect codes

The following codes are obviously perfect codes and are called *trivial perfect codes*:

(i) the linear code $C = \mathbf{F}_q^n$ ($d = 1$);
(ii) any C with $|C| = 1$ ($d = \infty$);
(iii) binary repetition codes of odd lengths consisting of two codewords at distance n from each other ($d = n$).

In the earlier subsections, we have seen that the Hamming codes and the Golay codes are examples of nontrivial perfect codes. Various constructions of nonlinear perfect codes with the same parameters as the q-ary Hamming codes have also been found.

In fact, the following result is true.

Theorem 5.3.26 (Van Lint and Tietäväinen.) *When $q \geq 2$ is a prime power, a nontrivial perfect code over $\mathbf{F}_q$ must have the same parameters as one of the Hamming or Golay codes.*

This result was obtained by Tietäväinen [22, 23] with considerable contribution from van Lint [12]. A proof may be found in Chap. 6 of ref. [13]. This result was also independently proved by Zinov'ev and Leont'ev [25].

5.4 Singleton bound and MDS codes

In this section, we discuss an upper bound for $A_q(n, d)$ due to Singleton [20].

Theorem 5.4.1 (Singleton bound.) *For any integer $q > 1$, any positive integer n and any integer d such that $1 \leq d \leq n$, we have*

$$A_q(n, d) \leq q^{n-d+1}.$$

In particular, when q is a prime power, the parameters $[n, k, d]$ of any linear code over $\mathbf{F}_q$ satisfy

$$k + d \leq n + 1.$$

Proof. We first note that the final statement of Theorem 5.4.1 follows from the previous one since, by definition of $A_q(n, d)$, $q^k \leq A_q(n, d)$.

To prove that $A_q(n, d) \leq q^{n-d+1}$, consider an (n, M, d)-code C over an alphabet A of size q, where $M = A_q(n, d)$. Delete the last $d - 1$ coordinates from all the codewords of C. Since the distance of C is d, after deleting the last $d - 1$ coordinates from all the codewords, the remaining words (of length $n - d + 1$) are still all distinct. The maximum number of words of length $n - d + 1$ is q^{n-d+1}, so $A_q(n, d) = M \leq q^{n-d+1}$. □

Remark 5.4.2 The following is another easy direct proof for the inequality $k + d \leq n + 1$ in the case of an $[n, k, d]$-linear code C:

Given any parity-check matrix H for C, the row rank, and hence the rank, of H is, by definition, $n - k$. Therefore, any $n - k + 1$ columns of H form a linearly dependent set. By Theorem 4.5.6(ii), $d \leq n - k + 1$.

Definition 5.4.3 A linear code with parameters $[n, k, d]$ such that $k+d = n+1$ is called a *maximum distance separable* (*MDS*) code.

Remark 5.4.4 An alternative way to state the Singleton bound is: for any q-ary code C, we have

$$\mathcal{R}(C) + \delta(C) \leq 1.$$

(In this situation, we see that our choice of the definition of the relative minimum distance $\delta(C)$ gives a neater inequality than if $\delta(C)$ is defined to be d/n.) A linear code C is MDS if and only if $\mathcal{R}(C) + \delta(C) = 1$.

One of the interesting properties of MDS codes is the following.

Theorem 5.4.5 *Let C be a linear code over $\mathbf{F}_q$ with parameters $[n, k, d]$. Let G, H be a generator matrix and a parity-check matrix, respectively, for C. Then, the following statements are equivalent:*

(i) *C is an MDS code*;
(ii) *every set of $n - k$ columns of H is linearly independent*;
(iii) *every set of k columns of G is linearly independent*;
(iv) *$C^\perp$ is an MDS code.*

Proof. The equivalence of (i) and (ii) follows directly from Corollary 4.5.7, with $d = n - k + 1$.

Since G is a parity-check matrix for $C^{\perp}$, (iii) and (iv) are also equivalent by Corollary 4.5.7.

Next, we prove that (i) implies (iv).

Recall that H is a generator matrix for $C^{\perp}$, so the length of $C^{\perp}$ is n and the dimension is $n - k$. To show that $C^{\perp}$ is MDS, we need to show that the minimum distance d' is $k + 1$.

Suppose $d' \leq k$. Then there is a word $\mathbf{c} \in C^{\perp}$ with at most k nonzero entries (and hence at least $n - k$ zero coordinates). Permuting the coordinates does not change the weight of the words, so we may assume that the last $n - k$ coordinates of $\mathbf{c}$ are 0.

Write H as $H = \left(A|H'\right)$, where A is some $(n - k) \times k$ matrix and H' is a square $(n - k) \times (n - k)$ matrix. Since the columns of H' are linearly independent (for (i) and (ii) are equivalent), H' is invertible. Hence, the rows of H' are linearly independent. The only way to obtain 0 in all the last $n - k$ coordinates (such as for $\mathbf{c}$) is to use the 0-linear combination of the rows of H' (by linear independence). Therefore, the entire word $\mathbf{c}$ is the all-zero word $\mathbf{0}$. Consequently, $d' \geq k + 1$. Together with the Singleton bound, it now follows that $d' = k + 1$.

Since $(C^{\perp})^{\perp} = C$, the above also shows that (iv) implies (i). This completes the proof of the theorem. □

Definition 5.4.6 An MDS code C over $\mathbf{F}_q$ is *trivial* if and only if C satisfies one of the following:

(i) $C = \mathbf{F}_q^n$;
(ii) C is equivalent to the code generated by $\mathbf{1} = (1, \ldots, 1)$; or
(iii) C is equivalent to the dual of the code generated by $\mathbf{1}$.

Otherwise, C is said to be *nontrivial*.

Remark 5.4.7 When $q = 2$, the only MDS codes are the trivial ones. This fact follows easily by considering the generator matrix in standard form (see Exercise 5.32).

An interesting family of examples of MDS codes is given by the (generalized) Reed–Solomon codes. For more details, see Chapters 8 and 9. Some other examples may also be found in the exercises at the end of this chapter.

5.5 Plotkin bound

The next upper bound for $A_q(n, d)$ that we will discuss is the Plotkin bound, which holds for codes for which d is large relative to n. It often gives a tighter upper bound than many of the other upper bounds, though it is only applicable to a comparatively smaller range of values of d. The proof we give for the Plotkin bound makes use of the following well known Cauchy–Schwarz inequality.

Lemma 5.5.1 (Cauchy–Schwarz inequality.) *Let* $\{a_1, \ldots, a_m\}$ *and* $\{b_1, \ldots, b_m\}$ *be any two sets of real numbers. Then*

$$\left(\sum_{r=1}^{m} a_r b_r\right)^2 = \left(\sum_{r=1}^{m} a_r^2\right)\left(\sum_{s=1}^{m} b_s^2\right) - \sum_{r=1}^{m}\sum_{s=1}^{m}(a_r b_s - a_s b_r)^2/2$$

Consequently,

$$\left(\sum_{r=1}^{m} a_r b_r\right)^2 \leq \left(\sum_{r=1}^{m} a_r^2\right)\left(\sum_{r=1}^{m} b_r^2\right).$$

For more details on the Cauchy–Schwarz inequality, see, for example, ref. [9].

Theorem 5.5.2 (Plotkin bound.) *Let* $q > 1$ *be an integer and suppose that* n, d *satisfy* $rn < d$, *where* $r = 1 - q^{-1}$. *Then,*

$$A_q(n, d) \leq \left\lfloor \frac{d}{d - rn} \right\rfloor.$$

Proof. Let C be an (n, M, d)-code over an alphabet A of size q. Let

$$T = \sum_{\mathbf{c}\in C}\sum_{\mathbf{c}'\in C} d(\mathbf{c}, \mathbf{c}').$$

Since $d \leq d(\mathbf{c}, \mathbf{c}')$ for $\mathbf{c}, \mathbf{c}' \in C$ such that $\mathbf{c} \neq \mathbf{c}'$, it follows that

$$M(M-1)d \leq T. \tag{5.2}$$

Now let $\mathcal{A}$ be the $M \times n$ array whose rows are made up of the M codewords in C. For $1 \leq i \leq n$ and $a \in A$, let $n_{i,a}$ denote the number of entries in the ith column of $\mathcal{A}$ that are equal to a. Hence, $\sum_{a\in A} n_{i,a} = M$ for every $1 \leq i \leq n$. Consequently, writing $\mathbf{c} = (c_1, \ldots, c_n)$ and $\mathbf{c}' = (c_1', \ldots, c_n')$, we have

$$T = \sum_{i=1}^{n}\left(\sum_{\mathbf{c}\in C}\sum_{\mathbf{c}'\in C} d(c_i, c_i')\right) = \sum_{i=1}^{n}\sum_{a\in A} n_{i,a}(M - n_{i,a}) = M^2 n - \sum_{i=1}^{n}\sum_{a\in A} n_{i,a}^2.$$

Applying Lemma 5.5.1, with $m = q$ and $a_1 = \cdots = a_q = 1$, it follows that

$$T \leq M^2 n - \sum_{i=1}^{n} q^{-1} \left(\sum_{a \in A} n_{i,a} \right)^2 = M^2 r n. \tag{5.3}$$

The Plotkin bound now follows from (5.2) and (5.3). □

In fact, when $q = 2$, a more refined version of the Plotkin bound is available.

Theorem 5.5.3 (Plotkin bound for binary codes.)
(i) *When d is even,*

$$A_2(n, d) \leq \begin{cases} 2\lfloor d/(2d - n) \rfloor & \text{for } n < 2d \\ 4d & \text{for } n = 2d. \end{cases}$$

(ii) *When d is odd,*

$$A_2(n, d) \leq \begin{cases} 2\lfloor (d + 1)/(2d + 1 - n) \rfloor & \text{for } n < 2d + 1 \\ 4d + 4 & \text{for } n = 2d + 1. \end{cases}$$

We leave the proof of Theorem 5.5.3 as an exercise (Exercise 5.30).

Example 5.5.4 To illustrate that Theorem 5.5.3 gives a more refined bound than Theorem 5.5.2, note that Theorem 5.5.2 gives $A_2(8, 5) \leq 5$, $A_2(8, 6) \leq 3$, $A_2(12, 7) \leq 7$ and $A_2(11, 8) \leq 3$, whereas Theorem 5.5.3 gives $A_2(8, 5) \leq 4$, $A_2(8, 6) \leq 2$, $A_2(12, 7) \leq 4$ and $A_2(11, 8) \leq 2$.

Example 5.5.5 In Tables 5.2–5.4, we list the sphere-covering lower bound and compare the Hamming, Singleton and Plotkin upper bounds for $A_2(n, d)$, with $d = 3, 5, 7$ and $d \leq n \leq 12$. In cases where the Plotkin bound is not applicable, the entry is marked '–'.

5.6 Nonlinear codes

Whereas most of this book focuses on linear codes, there are several families of (binary) nonlinear codes that are well known and important in coding theory. We provide a brief introduction to some of them in this section.

Table 5.2. Bounds for $A_2(n, 3)$.

n	Sphere-covering	Hamming	Singleton	Plotkin
3	2	2	2	2
4	2	3	4	2
5	2	5	8	4
6	3	9	16	8
7	5	16	32	16
8	7	28	64	–
9	12	51	128	–
10	19	93	256	–
11	31	170	512	–
12	52	315	1024	–

Table 5.3. Bounds for $A_2(n, 5)$.

n	Sphere-covering	Hamming	Singleton	Plotkin
5	2	2	2	2
6	2	2	4	2
7	2	4	8	2
8	2	6	16	4
9	2	11	32	6
10	3	18	64	12
11	4	30	128	24
12	6	51	256	–

Table 5.4. Bounds for $A_2(n, 7)$.

n	Sphere-covering	Hamming	Singleton	Plotkin
7	2	2	2	2
8	2	2	4	2
9	2	3	8	2
10	2	5	16	2
11	2	8	32	4
12	2	13	64	4

5.6.1 Hadamard matrix codes

Definition 5.6.1 A *Hadamard matrix* H_n is an $n \times n$ integer matrix whose entries are 1 or -1 and which satisfies $H_n H_n^{\mathrm{T}} = nI_n$, where I_n is the identity matrix.

When such a Hadamard matrix exists, then either $n = 1, 2$ or n is a multiple of 4. The existence of Hadamard matrices is known for many n; for example when n is a power of 2 (these are called Sylvester matrices), and when $n = p^m + 1$, where p is a prime and n is divisible by 4 (this is called the Paley construction). The construction of Sylvester matrices is easy. We begin with $H_1 = (1)$ and use the observation that, whenever H_n is a Hadamard matrix of order n, the matrix

$$H_{2n} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}$$

is a Hadamard matrix of order $2n$.

The existence of a Hadamard matrix H_n implies the existence of binary nonlinear codes of the following parameters:

$$\begin{aligned} (n, 2\lfloor d/(2d-n)\rfloor, d) &\quad \text{for } d \text{ even and } d \le n < 2d; \\ (2d, 4d, d) &\quad \text{for } d \text{ even}; \\ (n, 2\lfloor (d+1)/(2d+1-n)\rfloor, d) &\quad \text{for } d \text{ odd and } d \le n < 2d+1; \\ (2d+1, 4d+4, d) &\quad \text{for } d \text{ odd}. \end{aligned}$$

These codes were constructed by Levenshtein [10]. By the Plotkin bound, they are optimal.

5.6.2 Nordstrom–Robinson code

It can be shown that there cannot be any binary linear codes of parameters [16, 8, 6] (see Exercise 5.35). However, there does exist a binary nonlinear code, called the *Nordstrom–Robinson code*, of parameters $(16, 2^8, 6)$. It was discovered by Nordstrom and Robinson [16] (when Nordstrom was still a high school student!) and later independently by Semakov and Zinov'ev [19]. One construction of this famous code is as follows.

Rearrange the columns of the extended binary Golay code so that the new code (also called G_{24}) contains the word $111111110\cdots0$, and let G denote a generator matrix for this new G_{24}. Since $d(G_{24}) = 8 > 7$, Theorem 4.5.6(i) shows that the first seven columns of G are linearly independent. One can then show that each of the 2^7 possible vectors in $\mathbf{F}_2^7$ appears as the first seven coordinates of some codeword in G_{24}. In fact, each of them appears in exactly

$2^{12}/2^7 = 32$ codewords of G_{24}. Now collect all those words in G_{24} whose first seven coordinates are either all 0 or are made up of six 0s and one 1. There are altogether $8 \times 32 = 256 = 2^8$ of them.

The Nordstrom–Robinson code is obtained by deleting the first eight coordinates from these 2^8 vectors. It can be shown that this code has minimum distance 6 and is nonlinear.

5.6.3 Preparata codes

For $m \geq 2$, Preparata codes are binary nonlinear codes with the parameters $(2^{2m}, 2^{2^{2m}-4m}, 6)$.

There are several different ways to construct the Preparata codes; one way is as follows.

Write the vectors of $\mathbf{F}_2^{2^{2m}}$ in the form $(\mathbf{u}, \mathbf{v})$, where $\mathbf{u}, \mathbf{v} \in \mathbf{F}_2^{2^{2m-1}}$. Label the coordinate positions of these vectors in $\mathbf{F}_2^{2^{2m-1}}$ by the elements of $\mathbf{F}_{2^{2m-1}}$, with the first coordinate position corresponding to 0. For $\alpha \in \mathbf{F}_{2^{2m-1}}$, denote the entry at the αth coordinate of $\mathbf{u}$, $\mathbf{v}$ by u_α, v_α, respectively.

Definition 5.6.2 For $m \geq 2$, the *Preparata code* $P(m)$ of length 2^{2m} consists of all the codewords $(\mathbf{u}, \mathbf{v})$, where $\mathbf{u}, \mathbf{v} \in \mathbf{F}_2^{2^{2m-1}}$, satisfying the following conditions:

(i) both $\mathbf{u}$ and $\mathbf{v}$ are of even Hamming weight;
(ii) $\sum_{u_\alpha=1} \alpha = \sum_{v_\alpha=1} \alpha$;
(iii) $\sum_{u_\alpha=1} \alpha^3 + \left(\sum_{u_\alpha=1} \alpha\right)^3 = \sum_{v_\alpha=1} \alpha^3$.

It can be shown that $P(m)$ is a subcode of the extended binary Hamming code of the same length (see Chap. 15 of ref. [13] or Sect. 9.4 of ref. [24]).

The first code in this family, with $m = 2$, can be shown to be equivalent to the Nordstrom–Robinson $(16, 2^8, 6)$-code in Section 5.6.2.

5.6.4 Kerdock codes

For $m \geq 2$, the *Kerdock codes* $K(m)$ are binary nonlinear codes with parameters $(2^{2m}, 2^{4m}, 2^{2m-1} - 2^{m-1})$.

The Kerdock code $K(m)$ is constructed as a union of 2^{2m-1} cosets of the Reed–Muller code $\mathcal{R}(1, 2m)$ in $\mathcal{R}(2, 2m)$ (see Section 6.2).

Once again, the first code in this family, with $m = 2$, is equivalent to the Nordstrom–Robinson code. The Kerdock codes form a special case of a more general family of nonlinear codes called the *Delsarte–Goethals codes*.

The weight enumerators of the Kerdock and Preparata codes can be shown to satisfy the MacWilliams identity (see Exercise 4.49), thus giving a 'formal duality' between the Kerdock and Preparata codes. However, this falls beyond the scope of this book, so we will not elaborate further on this formal duality. The interested reader may refer to Chap. 15, Theorem 24 of ref. [13] for more details. This mystery of the formal duality between the Kerdock and Preparata codes was explained when it was shown by Nechaev [15] and Hammons *et al.* [7] that the Kerdock codes can be viewed as linear codes over the ring $\mathbf{Z}_4$, and by Hammons *et al.* [7] that the binary images of the $\mathbf{Z}_4$-dual of the Kerdock codes over $\mathbf{Z}_4$ can be regarded as variants of the Preparata codes.

5.7 Griesmer bound

The next bound we shall discuss is the Griesmer bound, which applies specifically to linear codes.

Let C be a linear code over $\mathbf{F}_q$ with parameters $[n, k]$ and suppose $\mathbf{c}$ is a codeword in C with $\mathrm{wt}(\mathbf{c}) = w$.

Definition 5.7.1 The *support* of $\mathbf{c}$, denoted by $\mathrm{Supp}(\mathbf{c})$, is the set of coordinates at which $\mathbf{c}$ is nonzero.

Definition 5.7.2 The *residual code of C with respect to* $\mathbf{c}$, denoted $\mathrm{Res}(C, \mathbf{c})$, is the code of length $n - w$ obtained from C by puncturing on all the coordinates of $\mathrm{Supp}(\mathbf{c})$.

Note that $w = |\mathrm{Supp}(\mathbf{c})|$.

Lemma 5.7.3 *If C is an $[n, k, d]$-code over $\mathbf{F}_q$ and $\mathbf{c} \in C$ is a codeword of weight d, then* $\mathrm{Res}(C, \mathbf{c})$ *is an $[n-d, k-1, d']$-code, where $d' \geq \lceil d/q \rceil$. Here, $\lceil x \rceil$ is the least integer greater than or equal to x.*

Proof. Without loss of generality, we may replace C by an equivalent code so that $\mathbf{c} = (1, 1, \ldots, 1, 0, 0, \ldots, 0)$, where the first d coordinates are 1 and the other coordinates are 0.

We first note that $\mathrm{Res}(C, \mathbf{c})$ has dimension at most $k-1$. To see this, observe first that $\mathrm{Res}(C, \mathbf{c})$ is a linear code. For every $\mathbf{x} \in \mathbf{F}_q^n$, denote by $\mathbf{x}'$ the vector obtained from $\mathbf{x}$ by deleting the first d coordinates, i.e., by puncturing on the coordinates of $\mathrm{Supp}(\mathbf{c})$. Now, it is easy to see that the map $C \to \mathrm{Res}(C, \mathbf{c})$ given by $\mathbf{x} \mapsto \mathbf{x}'$ is a well defined surjective linear transformation of vector

spaces, whose kernel contains $\mathbf{c}$ and is hence a subspace of C of dimension at least 1. Therefore, $\mathrm{Res}(C, \mathbf{c})$ has dimension at most $k-1$.

We shall show that $\mathrm{Res}(C, \mathbf{c})$ has dimension exactly $k-1$.

Suppose that the dimension is strictly less than $k-1$. Then there is a nonzero codeword $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in C that is not a multiple of $\mathbf{c}$ and that has the property that $v_{d+1} = \cdots = v_n = 0$. Then $\mathbf{v} - v_1\mathbf{c}$ is a nonzero codeword that belongs to C and that has weight strictly less than d, contradicting the definition of d. Hence, $\mathrm{Res}(C, \mathbf{c})$ has dimension $k-1$.

To show that $d' \geq \lceil d/q \rceil$, let $(x_{d+1}, \ldots, x_n)$ be any nonzero codeword of $\mathrm{Res}(C, \mathbf{c})$, and let $\mathbf{x} = (x_1, \ldots, x_d, x_{d+1}, \ldots, x_n)$ be a corresponding word in C. By the pigeonhole principle, there is an $\alpha \in \mathbf{F}_q$ such that at least d/q coordinates of $(x_1, \ldots, x_d)$ are equal to α. Hence,

$$d \leq \mathrm{wt}(\mathbf{x} - \alpha\mathbf{c}) \leq d - \frac{d}{q} + \mathrm{wt}((x_{d+1}, \ldots, x_n)).$$

The inequality $d' \geq \lceil d/q \rceil$ now follows. □

Theorem 5.7.4 (Griesmer bound.) *Let C be a q-ary code of parameters $[n, k, d]$, where $k \geq 1$. Then*

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

Proof. We prove the Griesmer bound by induction on k. Clearly, when $k = 1$, Theorem 5.7.4 holds.

When $k > 1$ and $\mathbf{c} \in C$ is a codeword of minimum weight d, then Lemma 5.7.3 shows that $\mathrm{Res}(C, \mathbf{c})$ is an $[n-d, k-1, d']$-code, where $d' \geq \lceil d/q \rceil$. By the inductive hypothesis, we may assume that the Griesmer bound holds for $\mathrm{Res}(C, \mathbf{c})$, hence

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

Theorem 5.7.4 now follows. □

Example 5.7.5 From Exercise 5.19, the q-ary simplex code $S(r, q)$ has parameters $[(q^r - 1)/(q-1), r, q^{r-1}]$, so it meets the Griesmer bound.

5.8 Linear programming bound

One of the best bounds in existence for $A_q(n, d)$ is one that is based on linear programming techniques. It is due to Delsarte [2]. The bound obtained through this method is often called the linear programming bound.

A family of polynomials, called the *Krawtchouk polynomials*, plays a pivotal role in this theory. Krawtchouk polynomials are also very useful in other areas of coding theory. We give the definition and summarize some properties of these polynomials below.

Definition 5.8.1 For a given q, the *Krawtchouk polynomial* $K_k(x; n)$ is defined to be

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

When there is no ambiguity for n, the notation is often simplified to $K_k(x)$.

Proposition 5.8.2 (Properties of Krawtchouk polynomials.)

(i) *If z is a variable, then* $\sum_{k=0}^{\infty} K_k(x) z^k = (1 + (q-1)z)^{n-x}(1-z)^x$.
(ii) $K_k(x) = \sum_{j=0}^{k} (-1)^j q^{k-j} \binom{n-k+j}{j} \binom{n-x}{k-j}$.
(iii) *$K_k(x)$ is a polynomial of degree k, with leading coefficient $(-q)^k/k!$ and constant term $K_k(0) = \binom{n}{k}(q-1)^k$.*
(iv) (Orthogonality relations.) $\sum_{i=0}^{n} \binom{n}{i} (q-1)^i K_k(i) K_\ell(i) = \delta_{k\ell} \binom{n}{k} (q-1)^k q^n$, *where $\delta_{k\ell}$ is the Kronecker delta function; i.e.,*

$$\delta_{k\ell} = \begin{cases} 1 & \text{if } k = \ell \\ 0 & \text{otherwise.} \end{cases}$$

(v) $(q-1)^i \binom{n}{i} K_k(i) = (q-1)^k \binom{n}{k} K_i(k)$.
(vi) $\sum_{i=0}^{n} K_\ell(i) K_i(k) = \delta_{k\ell} q^n$.
(vii) $\sum_{k=0}^{j} \binom{n-k}{n-j} K_k(x) = q^j \binom{n-x}{j}$.
(viii) *When $q = 2$, we have* $K_i(x) K_j(x) = \sum_{k=0}^{n} \binom{n-k}{(i+j-k)/2} \binom{k}{(i-j+k)/2} K_k(x)$.
(ix) *Every polynomial $f(x)$ of degree r can be expressed as* $f(x) = \sum_{k=0}^{r} f_k K_k(x)$, *where* $f_k = q^{-n} \sum_{i=0}^{n} f(i) K_i(k)$. *(This way of expressing $f(x)$ is called the* Krawtchouk expansion *of $f(x)$.)*

We leave the proof of Proposition 5.8.2 to the reader (see Exercise 5.42).

The linear programming bound gives an upper bound for $A_q(n, d)$; i.e., it applies also to nonlinear codes. Therefore, we will deal with the distance between two distinct codewords and not the weight of each codeword. For the main result in this section, we need the following notion.

Definition 5.8.3 Let A be an alphabet of size q. For C an (n, M)-code over A and for all $0 \leq i \leq n$, let

$$A_i(C) = \frac{1}{M}|\{(\mathbf{u}, \mathbf{v}) \in C \times C \ : \ d(\mathbf{u}, \mathbf{v}) = i\}|.$$

The sequence $\{A_i(C)\}_{i=0}^n$ is called the *distance distribution* of C.

Remark 5.8.4 Note that the distance distribution depends only on the size q of the code alphabet and not on the alphabet itself. To obtain the linear programming bound, it is more convenient to work with the ring $\mathbf{Z}_q$ as the alphabet. Hence, in the discussion below, while we begin with codes over an alphabet A of size q, we pass immediately to codes over $\mathbf{Z}_q$ in the proofs.

Lemma 5.8.5 *Let C be a q-ary code of length n. Then*

$$\sum_{i=0}^{n} A_i(C)K_k(i) \geq 0$$

for all integers $0 \leq k \leq n$.

Proof. As mentioned in Remark 5.8.4, we assume C is defined over $\mathbf{Z}_q$. It suffices to show that $M \sum_{i=0}^n A_i(C)K_k(i) \geq 0$, where $M = |C|$. Using Exercise 5.46,

$$M \sum_{i=0}^{n} A_i(C)K_k(i) = \sum_{i=0}^{n} \sum_{\substack{(\mathbf{u},\mathbf{v}) \in C^2 \\ d(\mathbf{u},\mathbf{v})=i}} \sum_{\substack{\mathbf{w} \in \mathbf{Z}_q^n \\ \mathrm{wt}(\mathbf{w})=k}} \zeta^{(\mathbf{u}-\mathbf{v})\cdot\mathbf{w}} = \sum_{\substack{\mathbf{w} \in \mathbf{Z}_q^n \\ \mathrm{wt}(\mathbf{w})=k}} \left| \sum_{\mathbf{u} \in C} \zeta^{\mathbf{u}\cdot\mathbf{w}} \right|^2 \geq 0,$$

where, for $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$, $\mathbf{u}\cdot\mathbf{w} = u_1w_1 + \cdots + u_nw_n$, and ζ is a primitive qth root of unity in $\mathbf{C}$; i.e., $\zeta^q = 1$ but $\zeta^i \neq 1$ for all $0 < i < q$. □

Theorem 5.8.6 (Linear programming bound – version 1.) *For a given integer $q > 1$ and positive integers n and d $(1 \leq d \leq n)$, we have*

$$A_q(n, d) \leq \max \left\{ \sum_{i=0}^{n} A_i : A_0 = 1, A_i = 0 \text{ for } 1 \leq i < d, A_i \geq 0 \text{ for } 0 \leq i \leq n \right.$$
$$\left. \sum_{i=0}^{n} A_i K_k(i) \geq 0 \text{ for } 0 \leq k \leq n \right\}. \tag{5.4}$$

Proof. Let $M = A_q(n, d)$. If C is a q-ary (n, M)-code, its distance distribution $\{A_i(C)\}_{i=0}^n$ satisfies the following conditions:

(i) $A_0(C) = 1$;

(ii) $A_i(C) = 0$ for $1 \leq i < d$;
(iii) $A_i(C) \geq 0$ for all $0 \leq i \leq n$;
(iv) $\sum_{i=0}^{n} A_i(C)K_k(i) \geq 0$ for $0 \leq k \leq n$ (from Lemma 5.8.5);
(v) $M = A_q(n, d) = \sum_{i=0}^{n} A_i(C)$.

Hence, the inequality (5.4) follows immediately. □

The following theorem is the duality theorem of Theorem 5.8.6 in linear programming. It is often more useful than Theorem 5.8.6 because any polynomial $f(x)$ that satisfies Theorem 5.8.7 gives an upper bound for $A_q(n, d)$, while an optimal solution for the linear programming problem in (5.4) is required to give an upper bound for $A_q(n, d)$.

Theorem 5.8.7 (Linear programming bound – version 2.) *Let $q > 1$ be an integer. For positive integers n and d $(1 \leq d \leq n)$, let $f(x) = 1 + \sum_{k=1}^{n} f_k K_k(x)$ be a polynomial such that $f_k \geq 0$ $(1 \leq k \leq n)$ and $f(i) \leq 0$ for $d \leq i \leq n$. Then $A_q(n, d) \leq f(0)$.*

Proof. As in the proof of Theorem 5.8.6, let $M = A_q(n, d)$, let C be a q-ary (n, M)-code and let $\{A_i(C)\}_{i=0}^{n}$ be its distance distribution.

Note that conditions (i), (ii) and (iv) in the proof of Theorem 5.8.6 imply that $K_k(0) \geq -\sum_{i=d}^{n} A_i(C)K_k(i)$ for all $0 \leq k \leq n$. The condition that $f(i) \leq 0$ for $d \leq i \leq n$ implies that $\sum_{i=d}^{n} A_i(C) f(i) \leq 0$, which means that

$$\begin{aligned} f(0) &= 1 + \sum_{k=1}^{n} f_k K_k(0) \\ &\geq 1 - \sum_{k=1}^{n} f_k \sum_{i=d}^{n} A_i(C)K_k(i) \\ &= 1 - \sum_{i=d}^{n} A_i(C) \sum_{k=1}^{n} f_k K_k(i) \\ &= 1 - \sum_{i=d}^{n} A_i(C)(f(i) - 1) \\ &\geq 1 + \sum_{i=d}^{n} A_i(C) \\ &= M = A_q(n, d). \end{aligned}$$

□

To illustrate that the linear programming bound can be better than some other bounds that we have discussed in this chapter, we show in Example 5.8.8

how one can deduce the Singleton bound, the Hamming bound and the Plotkin bound from the linear programming bound.

Example 5.8.8 (i) (Singleton bound.) Let

$$f(x) = q^{n-d+1} \prod_{j=d}^{n} \left(1 - \frac{x}{j}\right).$$

By Proposition 5.8.2(ix), $f(x) = \sum_{k=0}^{n} f_k K_k(x)$, where f_k is given by

$$\begin{aligned} f_k &= \frac{1}{q^n} \sum_{i=0}^{n} f(i) K_i(k) \\ &= \frac{1}{q^{d-1}} \sum_{i=0}^{d-1} \binom{n-i}{n-d+1} K_i(k) \Big/ \binom{n}{d-1} \\ &= \binom{n-k}{d-1} \Big/ \binom{n}{d-1} \geq 0, \end{aligned}$$

where the last equality follows from Proposition 5.8.2(vii). In particular, $f_0 = 1$. Clearly, $f(i) = 0$ for $d \leq i \leq n$.

Hence, by Theorem 5.8.7, it follows that $A_q(n, d) \leq f(0) = q^{n-d+1}$, which is the Singleton bound (cf. Theorem 5.4.1).

(ii) (Hamming bound.) Let $d = 2e + 1$. Let $f(x) = \sum_{k=0}^{n} f_k K_k(x)$, where

$$f_k = \left\{ L_e(k) \Big/ \sum_{i=0}^{e} (q-1)^i \binom{n}{i} \right\}^2 \qquad (0 \leq k \leq n),$$

with $L_e(x) = \sum_{i=0}^{e} K_i(x) = K_e(x-1; n-1)$. (The polynomial $L_e(x)$ is called a *Lloyd polynomial*.) Clearly, $f_k \geq 0$ for all $0 \leq k \leq n$ and $f_0 = 1$. Using Proposition 5.8.2(viii) and (vi), it can be shown that $f(i) = 0$ for $d \leq i \leq n$. Therefore, Theorem 5.8.7 and Proposition 5.8.2(iv) show that

$$A_q(n, d) \leq f(0) = q^n \Big/ \sum_{i=0}^{e} (q-1)^i \binom{n}{i},$$

which is exactly the Hamming bound.

(iii) (Plotkin bound for $A_2(2\ell + 1, \ell + 1)$.) Set $q = 2$, $n = 2\ell + 1$ and $d = \ell + 1$. Take $f_1 = (\ell + 1)/(2\ell + 1)$ and $f_2 = 1/(2\ell + 1)$, so that

$$\begin{aligned} f(x) &= 1 + \frac{\ell+1}{2\ell+1} K_1(x) + \frac{1}{2\ell+1} K_2(x) \\ &= 1 + \frac{\ell+1}{2\ell+1} (2\ell + 1 - 2x) + \frac{1}{2\ell+1} (2x^2 - 2(2\ell+1)x + \ell(2\ell+1)). \end{aligned}$$

Clearly, $f_k \geq 0$ for all $1 \leq k \leq n$, and it is straightforward to verify that $f(i) \leq 0$ for $\ell + 1 = d \leq i \leq n = 2\ell + 1$. (In fact, $f(x)$ is a quadratic polynomial such that $f(\ell + 1) = 0 = f(2\ell + 1)$.)

Hence, by Theorem 5.8.7, it follows that

$$A_2(2\ell + 1, \ell + 1) \leq f(0) = 1 + \frac{\ell + 1}{2\ell + 1}(2\ell + 1) + \frac{1}{2\ell + 1}\ell(2\ell + 1) = 2\ell + 2,$$

which is exactly the Plotkin bound (cf. Theorem 5.5.2). (Note: when ℓ is even, Theorem 5.5.3 in fact gives a better bound.)

Exercises

5.1 Find the size, (minimum) distance, information rate and relative minimum distance of each of the following codes:
- (a) the binary code of all the words of length 3;
- (b) the ternary code consisting of all the words of length 4 whose second and fourth coordinates are 0;
- (c) the code over the alphabet $\mathbf{F}_p$ (p prime) consisting of all the words of length 3 whose first coordinate is $p - 1$ and whose second coordinate is 1;
- (d) the repetition code over the alphabet $\mathbf{F}_p$ (p prime) consisting of the following words of length n: $(0, 0, \ldots, 0), (1, 1, \ldots, 1), \ldots, (p - 1, p - 1, \ldots, p - 1)$.

5.2 For n odd, let C be a self-orthogonal binary $[n, (n - 1)/2]$-code. Show that $\overline{C^{\perp}}$ is a self-dual code. (Note: compare with Exercise 4.26.)

5.3 For any code C over $\mathbf{F}_q$ and any $\epsilon \in \mathbf{F}_q^*$, let

$$\overline{C}_\epsilon = \left\{ \left(c_1, \ldots, c_n, \epsilon \sum_{i=1}^{n} c_i \right) : (c_1, \ldots, c_n) \in C \right\}.$$

(In particular, $\overline{C}_{-1}$ is the extended code $\overline{C}$ of C defined in Definition 5.1.8.)
- (i) If C is an (n, M, d)-code, show that $\overline{C}_\epsilon$ is an $(n + 1, M, d')$-code, where $d \leq d' \leq d + 1$.
- (ii) If C is linear, show that $\overline{C}_\epsilon$ is linear also. Find a parity-check matrix for $\overline{C}_\epsilon$ in terms of a parity-check matrix H of C.

5.4 Without using any of the bounds discussed in this chapter, show that
(a) $A_2(6, 5) = 2$, (b) $A_2(7, 5) = 2$.
(Hint: For (a), first show that $A_2(6, 5) \geq 2$ by producing a code explicitly.

Then try to show that $A_2(6, 5) \leq 2$ using a simple combinatorial argument similar to the one in Example 5.2.5.)

5.5 Find an optimal binary code with $n = 3$ and $d = 2$.

5.6 Prove that $A_q(n, d) \leq q A_q(n-1, d)$.

5.7 For each of the following spheres in $A^n = \mathbf{F}_2^n$, list its elements and compute its volume:
(a) $S_A(110, 4)$, (b) $S_A(1100, 3)$, (c) $S_A(10101, 2)$.

5.8 For each n such that $4 \leq n \leq 12$, compute the Hamming bound and the sphere-covering bound for $A_2(n, 4)$.

5.9 Prove that a $(6, 20, 4)$-code over $\mathbf{F}_7$ cannot be an optimal code.

5.10 Let $q \geq 2$ and $n \geq 2$ be any integers. Show that $A_q(n, 2) = q^{n-1}$.

5.11 Let C be an $[n, k, d]$-code over $\mathbf{F}_q$, where $\gcd(d, q) = 1$. Suppose that all the codewords of C have weight congruent to 0 or d modulo q. Using Exercise 4.30(iv), or otherwise, show the existence of an $[n+1, k, d+1]$-code over $\mathbf{F}_q$.

5.12 Let C be an optimal code over $\mathbf{F}_{11}$ of length 12 and minimum distance 2. Show that C must have a transmission rate of at least $5/6$.

5.13 For positive integers n, M, d and $q > 1$ (with $1 \leq d \leq n$), show that, if $(M-1)\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i < q^n$, then there exists a q-ary (n, M)-code of minimum distance at least d. (Note: this is often known as the *Gilbert–Varshamov bound* for nonlinear codes.)

5.14 Determine whether each of the following codes exists. Justify your answer.
(a) A binary code with parameters $(8, 29, 3)$.
(b) A binary linear code with parameters $(8, 8, 5)$.
(c) A binary linear code with parameters $(8, 5, 5)$.
(d) A binary linear code with parameters $(24, 2^{12}, 8)$.
(e) A perfect binary linear code with parameters $(63, 2^{57}, 3)$.

5.15 Write down a parity-check matrix H for a binary Hamming code of length 15, where the jth column of H is the binary representation of j. Then use H to construct a syndrome look-up table and use it to decode the following words:
(a) 01010 01010 01000,
(b) 11100 01110 00111,
(c) 11001 11001 11000.

5.16 (i) Show that there exist no binary linear codes with parameters $[2^m, 2^m - m, 3]$, for any $m \geq 2$.
(ii) Let C be a binary linear code with parameters $[2^m, k, 4]$, for some $m \geq 2$. Show that $k \leq 2^m - m - 1$.

5.17 Prove Proposition 5.3.15.

5.18 (i) Let $n \geq 3$ be an integer. Show that there is an $[n, k, 3]$-code defined over $\mathbf{F}_q$ if and only if $q^{n-k} - 1 \geq (q-1)n$.
(ii) Find the smallest n for which there exists a ternary $[n, 5, 3]$-code.

5.19 (i) Let $\mathbf{v}$ be a nonzero vector in $\mathbf{F}_q^r$. Show that the set of vectors in $\mathbf{F}_q^r$ orthogonal to $\mathbf{v}$, i.e., $\{\mathbf{v}\}^\perp$, forms a subspace of $\mathbf{F}_q^r$ of dimension $r-1$.
(ii) Let G be a generator matrix for the simplex code $S(r, q)$. Show that, for a given nonzero vector $\mathbf{v} \in \mathbf{F}_q^r$, there are exactly $(q^{r-1}-1)/(q-1)$ columns $\mathbf{c}$ of G such that $\mathbf{v} \cdot \mathbf{c} = 0$.
(iii) Using the observation that $S(r, q) = \{\mathbf{v}G : \mathbf{v} \in \mathbf{F}_q^r\}$, or otherwise, show that every nonzero codeword of $S(r, q)$ has weight q^{r-1}. (Hint: Use (ii) to determine the number of coordinates of $\mathbf{v}G$ that are equal to 0.)

5.20 Determine the Hamming weight enumerators of Ham(3, 2) and $S(3, 2)$. Verify that they satisfy the MacWilliams identity (see Exercise 4.49).

5.21 The ternary Hamming code Ham(2, 3) is also known as the *tetracode*.
(i) Show that the tetracode is a self-dual MDS code.
(ii) Without writing down all the elements of Ham(2, 3), determine the weights of all its codewords.
(iii) Determine the Hamming weight enumerator of Ham(2, 3) and show that the MacWilliams identity (see Exercise 4.49) holds for $C = C^\perp = \text{Ham}(2, 3)$.

5.22 Let $\mathcal{G}_6$ denote the hexacode defined in Exercise 4.10(b).
(i) Show that $\mathcal{G}_6$ is a $[6, 3, 4]$-code over $\mathbf{F}_4$. (Hence, $\mathcal{G}_6$ is an MDS quaternary code.)
(ii) Let $\mathcal{G}_6'$ be the code obtained from $\mathcal{G}_6$ by deleting the last coordinate from every codeword. Show that $\mathcal{G}_6'$ is a Hamming code over $\mathbf{F}_4$.

5.23 (i) Show that the all-one vector $(1, 1, \ldots, 1)$ is in the extended binary Golay code G_{24}.
(ii) Deduce from (i) that G_{24} does not have any word of weight 20.

5.24 Prove Proposition 5.3.22.

5.25 (i) Show that every word of weight 4 in $\mathbf{F}_2^{23}$ is of distance 3 from exactly one codeword in the binary Golay code G_{23}.
(ii) Use (i) to count the number of codewords of weight 7 in G_{23}.
(iii) Use (ii) to show that the extended binary Golay code G_{24} contains precisely 759 codewords of weight 8.

5.26 Show that the extended binary Golay code G_{24} has the weight distribution shown in Table 5.5 for its codewords.

5.27 Verify the MacWilliams identity (see Exercise 4.49) with $C = C^\perp = G_{24}$.

Table 5.5.

Weight	0	4	8	12	16	20	24
Number of codewords	1	0	759	2576	759	0	1

5.28 Prove that the extended ternary Golay code G_{12} is a [12, 6, 6]-code.

5.29 Show that the ternary Golay code G_{11} satisfies the Hamming bound.

5.30 Prove Theorem 5.5.3. (Hint: When d is even and $n < 2d$, mimic the proof of Theorem 5.5.2. Divide into the two cases M even and M odd, and maximize the expression $\sum_{i=1}^{n} \sum_{a \in \mathbf{F}_2} n_{i,a}(M - n_{i,a})$ in each case. For the case of even d and $n = 2d$, apply Exercise 5.6, with $q = 2$, and the previous case. When d is odd, apply Theorem 5.1.11 with the result for even d.)

5.31 Let C be the code over $\mathbf{F}_4 = \{0, 1, \alpha, \alpha^2\}$ with generator matrix

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & \alpha & \alpha^2 \end{pmatrix}.$$

(i) Show that C is an MDS code.

(ii) Write down a generator matrix for the dual $C^{\perp}$.

(iii) Show that $C^{\perp}$ is an MDS code.

5.32 Show that the only binary MDS codes are the trivial ones.

5.33 Suppose there is a q-ary MDS code C of length n and dimension k, where $k < n$.

(i) Show that there is also a q-ary MDS code of length $n - 1$ and dimension k.

(ii) For a given $1 \leq i \leq n$, let C_i be the subcode of C consisting of all the codewords with 0 in the ith position, and let D_i be the code obtained by deleting the ith coordinate from every codeword of C_i. Show that D_i is an MDS code. (Hint: You may need to show that there is at least one minimum weight codeword of C with 0 in the ith position.)

5.34 For each n such that $9 \leq n \leq 16$, compare the Singleton, Plotkin and Hamming upper bounds for $A_2(n, 9)$.

5.35 Suppose there exists a binary linear code C of parameters [16, 8, 6].

(i) Let C' be the residual code of C with respect to a codeword of weight 6. Show that C' is a binary linear code of parameters $[10, 7, d']$, where $3 \leq d' \leq 4$.

(ii) Use Exercise 5.32 to show that $d' = 3$.

(iii) Using the Hamming bound, or otherwise, show that such a C' cannot exist.

5.36 A binary (n, M, d)-code C is called a *constant-weight binary code* if there exists an integer w such that $\mathrm{wt}(\mathbf{c}) = w$ for all $\mathbf{c} \in C$. In this case, we say that C is a constant-weight binary $(n, M, d; w)$-code.

(a) Show that the minimum distance of a constant-weight binary code is always even.

(b) Show that a constant-weight binary $(n, M, d; w)$-code satisfies $M \leq \binom{n}{w}$.

(c) Prove that a constant-weight binary $(n, M, d; w)$-code can detect at least one error.

5.37 Let $A_2(n, d, w)$ be the maximum possible number M of codewords in a constant-weight binary $(n, M, d; w)$-code. Show that

(a) $1 \leq A_2(n, d, w) \leq \binom{n}{w}$;

(b) $A_2(n, 2, w) = \binom{n}{w}$;

(c) $A_2(n, d, w) = 1$ for $d > 2w$;

(d) $A_2(n, d, w) = A_2(n, d, n - w)$.

5.38 Use the Griesmer bound to find an upper bound for d for the q-ary linear codes of the following n and k:

(a) $q = 2$, $n = 10$ and $k = 3$;

(b) $q = 3$, $n = 8$ and $k = 4$;

(c) $q = 4$, $n = 10$ and $k = 5$;

(d) $q = 5$, $n = 9$ and $k = 2$.

5.39 For a prime power q and positive integers k and u with $k > u > 0$, the *MacDonald code* $C_{k,u}$ is a q-ary linear code, of parameters $[(q^k - q^u)/(q - 1), k, q^{k-1} - q^{u-1}]$, that has nonzero codewords of only two possible weights: $q^{k-1} - q^{u-1}$ and q^{k-1}. Show that the MacDonald codes attain the Griesmer bound.

5.40 Let C be an $[n, k, d]$-code over $\mathbf{F}_q$ and let $\mathbf{c} \in C$ be a codeword of weight w, where $w < dq/(q - 1)$. Show that the residual code $\mathrm{Res}(C, \mathbf{c})$ is an $[n - w, k - 1, d']$-code, where $d' \geq d - w + \lceil w/q \rceil$.

5.41 Let C be a $[q^2, 4, q^2 - q - 1]$-code over $\mathbf{F}_q$.

(i) By considering $\mathrm{Res}(C, \mathbf{c})$, where $\mathrm{wt}(\mathbf{c}) = q^2 - t$ with $2 \leq t \leq q - 1$, or otherwise, show that the only possible weights of the codewords in C are: $0, q^2 - q - 1, q^2 - q, q^2 - 1$ and q^2.

(ii) Show the existence of a $[q^2 + 1, 4, q^2 - q]$-code over $\mathbf{F}_q$. (Hint: Compare with Exercise 5.11.)

5.42 Prove the properties of the Krawtchouk polynomials listed in Proposition 5.8.2. (Hint: For (ii), use the fact that

$$(1+(q-1)z)^{n-x}(1-z)^x = (1-z)^n\left(1+\frac{qz}{1-z}\right)^{n-x}.$$

For (iv), multiply both sides of the equality by $y^k z^\ell$ and sum over all $k, \ell \geq 0$. For (vii), use (ii). For (viii), use (i) by multiplying two power series together.)

5.43 Show that the Krawtchouk polynomials satisfy the following recurrence relation:

$$(k+1)K_{k+1}(x) = (k+(q-1)(n-k)-qx)\,K_k(x) - (q-1)(n-k+1)K_{k-1}(x).$$

5.44 Show that $K_k(x) = \sum_{j=0}^{k}(-q)^j(q-1)^{k-j}\binom{n-j}{k-j}\binom{x}{j}$.

5.45 Let $q = 2$. Show that:
(a) $K_0(x) = 1$;
(b) $K_1(x) = -2x + n$;
(c) $K_2(x) = 2x^2 - 2nx + \binom{n}{2}$;
(d) $K_3(x) = -4x^3/3 + 2nx^2 - (n^2 - n + 2/3)x + \binom{n}{3}$.

5.46 Let ζ be a primitive qth root of unity in $\mathbf{C}$. Suppose $\mathbf{u} \in \mathbf{Z}_q^n$ is a word of weight i. Show that

$$\sum_{\substack{\mathbf{w}\in\mathbf{Z}_q^n \\ \mathrm{wt}(\mathbf{w})=k}} \zeta^{\mathbf{u}\cdot\mathbf{w}} = K_k(i),$$

where, for $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$, $\mathbf{u}\cdot\mathbf{w} = u_1w_1 + \cdots + u_nw_n$.

5.47 Use the linear programming bound (Theorem 5.8.7) to show that the Hadamard matrix code of parameters $(2d, 4d, d)$, with d even, is an optimal code. (Hint: Use $f(x) = 1 + K_1(x) + \frac{1}{d}K_2(x)$.)

5.48 Let d be such that $2d > n$. Use the linear programming bound (Theorem 5.8.7) to show that $A_2(n,d) \leq 2d/(2d-n)$. Note that this bound is slightly weaker than the Plotkin bound. (Hint: Use $f(x) = 1 + \frac{1}{2d-n}K_1(x)$.)