Coding Theory

Peter Müller

June 1, 2017

Contents

1 Introduction

1.1 ISBN and EAN numbers as examples of error detecting codes

ToDo: To be written

1.2 Why codes?

Assume that we want to transmit (or store) a message which is a stream of letters A, B, C and D over a binary channel which is only using the symbols (bits) 0 and 1. As there are exactly four length 2 words with symbols 0 and 1, we can send 00, 01, 10 and 11 for A, B, C, and D, respectively. For instance, the word DAD will be sent as 110011. If one of the bits gets wrongly transmitted, for instance the first one, so that we receive 010011, we would erroneously decode the message as BAD.

Assume that it happens rarely that a bit gets wrongly transmitted. In order to detect single errors, we could add a parity check bit, that is we add a bit such that the number of 1's is even. The following table shows in the left column the original length 2 words, and in the right column the actually sent word, henceforth called code word:

00	$\to$	000	We detect a single error, but cannot correct it. For instance, if
01	$\to$	011	we receive 010, the correct word could have been 000 or 011.
10	$\to$	101	
11	$\to$	110	

In order to correct a single error, we could add more redundancy. A naive but working idea is to just repeat the original word 3 times:

00	$\to$	000000	Now any two code words differ in at least 3 positions. So if we
01	$\to$	010101	receive a word where one error happened, for instance 011101,
10	$\to$	101010	then we know that 010101 was the intended code word, because
11	$\to$	111111	no other code word differs in at most one position from 011101.

Next one could ask if we can save some bandwidth, and achieve the same result with code words of length 5. Indeed, that is possible:

00	$\to$	00000	Again, any two code words differ in at least 3 positions. Thus
01	$\to$	01011	we can correct a single transmission error.
10	$\to$	10111	
11	$\to$	11100	

The obvious next question is whether we can do that for suitable code words of length 4. We will soon learn how to systematically treat such questions. For later reference we record this question.

Question 1.1. *Are there* 4 *words of length* 4 *using the symbols* 0 *and* 1 *which differ in at least* 3 *positions?*

1.3 Some definitions and easy properties for block codes

The motivation from the previous section can be generalized in various directions. First, we may use code words built of more symbols than just 0 and 1. Also, if there is a high probability that symbols in sent code words get wrong, it might be necessary that we can correct more than one error. These considerations lead to the basic definition of a block code and the Hamming distance.

Definition 1.2 (Block code)**.** Let F be a finite set of size ≥ 2, and $n \in \mathbb{N}$. A *block code* is a subset $C \subseteq F^n$.

We will use the term *code* as a synonym for block code.

Definition 1.3 (Hamming distance)**.** For $x = (x_1\, x_2 \ldots x_n)$ and $y = (y_1\, y_2 \ldots y_n)$ in F^n let the *Hamming distance* $d(x, y)$ be the number of positions where x and y differ. So $d(x, y)$ is the number of indices i such that $x_i \neq y_i$.

Definition 1.4 (Minimum distance)**.** Let $C \subseteq F^n$ be a code. Then the *minimum distance* $d(C)$ of C is the minimum of $d(c, c')$ for distinct $c, c' \in C$. We set $d(C) = \infty$ if $|C| \leq 1$.

Remark 1.5. The minimum distance of a code C is the largest integer d such that $d(c, c') \geq d$ for every pair of distinct elements $c, c' \in C$. If $|C| \leq 1$, then there are no such pairs, therefore the inequality holds for all $d \in \mathbb{N}$, so we consistently set $d(C) = \infty$ in this case.

We record some simple properties of the Hamming distance.

Lemma 1.6. *For $x, y, z \in F^n$ the following holds.*

(a) $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$.

(b) $d(x, y) = d(y, x)$ (symmetry).

(c) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

The following lemma explains the error correction capabilities of a code. If we use a code C with the property (ii) in the lemma, we can correct up to e wrong symbols. For if $c \in C$ was sent, $x \in F^n$ was received, and at most e symbols got wrong, then c is the unique element from C which differs in at most e positions from x. Thus we correctly decode x to c.

Lemma 1.7. *Let $C \subseteq F^n$ be a code and $e \in \mathbb{N}_0$. Then the following assertions are equivalent.*

(i) $d(C) \geq 2e + 1$.

(ii) For any $x \in F^n$ there is at most one $c \in C$ such that $d(x, c) \leq e$.

Proof. (i) implies (ii) by the triangle inequality.
Conversely, suppose that (i) does not hold, so there are distinct $c, c' \in C$ such that $d(c, c') \leq 2e$. Then it is easy so see (confer Problem 1.15) that there is $x \in F^n$ with $d(c, x) \leq e$ and $d(c', x) \leq e$. □

Many codes which we study in the following have more algebraic structure than just being subsets of F^n. In the following $\mathbb{F}_q$ denotes a finite field with q elements.

Definition 1.8 (Linear code)**.** A *linear code* is a subspace C of the vector space $\mathbb{F}_q^n$. We call it an (n, k, d) code, where $k = \dim C$ and $d = d(C)$.

Definition 1.9 (Weight)**.** For $x \in \mathbb{F}_q^n$ define the *weight* $w(x)$ to be the number of nonzero entries in the vector x, so $w(x) = d(x, 0)$.

Lemma 1.10. *Let $C \subseteq \mathbb{F}_q^n$ be a linear code. Then $d(C)$ equals the minimum of the weights $w(c)$ for $0 \neq c \in C$.*

Proof. For $c, c' \in C$ we have $c - c' \in C$ and $d(c, c') = w(c - c')$, from this the claim follows. □

There are some easy modifications of a code which do not change its essential properties. For instance, let $\sigma \in \mathrm{Sym}(\{1, 2, \ldots, n\})$ be a permutation of the letters $\{1, 2, \ldots, n\}$, and $C \subseteq F^n$ be a code. Then

$$C' = \{(c_{\sigma(1)}\, c_{\sigma(2)}\, \ldots\, c_{\sigma(n)}) \mid (c_1\, c_2\, \ldots\, c_n) \in C\}$$

has the same size and minimum distance as C. There is another modification which does not change size and minimum distance. Let F' be another set of size $|F|$, and for each $i = 1, 2, \ldots, n$ let $\sigma_i : F \to F'$ be a bijection. Set

$$C' = \{(\sigma_1(c_1)\, \sigma_2(c_2)\, \ldots\, \sigma_n(c_n)) \mid (c_1\, c_2\, \ldots\, c_n) \in C\} \subseteq F'^n.$$

Then $|C| = |C'|$ and $d(C) = d(C')$.

Typical applications are $F' = F$, or $F' = \mathbb{Z}/q\mathbb{Z}$, where $q = |F|$, or $F' = \mathbb{F}_q$ if q is a prime power.

Definition 1.11 (Equivalence of codes)**.** Suppose that $|F| = |F'|$ and let $C \subseteq F_n$ and $C' \subseteq F'_n$ be codes such that C' arises from C by applying a sequence of the two kinds of modifications described above. Then we call C and C' *equivalent*.

Remark 1.12. It is easy to see that the equivalence of codes indeed is an equivalence relation. Note, however, that if $F = F' = \mathbb{F}_q$ and $C \subseteq \mathbb{F}_q^n$ is a linear code, then a code $C' \subseteq \mathbb{F}_q^n$ which is equivalent to C need not be linear. Linearity of codes is preserved if in the above definition we only allow $\sigma_i : \mathbb{F}_q \to \mathbb{F}_q$ to be maps of the form $\sigma_i(x) = a_i x$ for nonzero elements a_i.

Coding theory essentially consists in the following three aspects:

(a) Given F, n, and d, find codes $C \subseteq F^n$ with $d(C) \geq d$ and $|C|$ as large as possible.

(b) Given F, n, and d, find good upper bounds for the size $|C|$ of codes $C \subseteq F^n$ with $d(C) \geq d$.

(c) Let $C \subseteq F^n$ be a code which can correct up to e errors. Decoding means that given $x \in F^n$, find the (if it exists) the unique code word $c \in C$ with $d(x, c) \leq e$. If $|C|$ is very large, a naive search through all the code words can be too time consuming. For many real world applications, it is important to have efficient decoding algorithms. The lack of fast algorithms can render certain codes with otherwise good parameters useless.

In this course we will be concerned with (a) and (b), and will not touch (c).

1.4 Problems

Problem 1.13. Read how the ISBN-10 numbers are defined and which error recognizing properties they have.

Problem 1.14. The ISBN-13 number is a 13-digit number $a_1 a_2 \dots a_{13}$ with $a_i \in \{0, 1, \dots, 9\}$ such that

$$(a_1 + a_3 + \cdots + a_{13}) + 3(a_2 + a_4 + \cdots + a_{12}) \equiv 0 \pmod{10}.$$

Analyze which switches of two consecutive digits are recognized, and which are not.

Problem 1.15. For $u, v \in F^n$ suppose that $d(u, v) \leq 2e$ for some $e \in \mathbb{N}$. Show that there is an element $w \in F^n$ such that $d(u, w) \leq e$ and $d(v, w) \leq e$.

Problem 1.16. Let $C \subseteq F^n$ be a code with $d(C) = n$. Show that $q = |F| \geq |C| = m$, and that C is equivalent to $C' = \{(i\, i \dots i) \mid i = 0, 1, \dots, m-1\} \subseteq F'^n$, where $F' = \{0, 1, \dots, q-1\}$.

2 Simple bounds and codes attaining these bounds

2.1 The Hamming bound

Definition 2.1 (Hamming ball)**.** For $r \in \mathbb{R}$ and $v \in F^n$ we define the *Hamming ball* with center v and radius r by $B_r(v) = \{x \in F^n \mid d(x,v) \leq r\}$.

Lemma 2.2 (Volume of Hamming ball)**.** *For all $v \in F^n$ and $m \in \mathbb{N}_0$ we have $|B_m(v)| = \sum_{i=0}^{m} \binom{n}{i}(q-1)^i$.*

Proof. For each $0 \leq i \leq m$ we count the number of $x \in F^n$ with $d(x,v) = i$. There are $\binom{n}{i}$ possibilities to pick the i positions in x which differ from v, and for each of these i positions we have $q-1$ choices for the entry of x being different from the one of v. So this number is $\binom{n}{i}(q-1)^i$. □

The following theorem is called the *Hamming bound* or the *sphere packing bound*.

Theorem 2.3 (Hamming bound)**.** *Let $C \subseteq F^n$ be a code with $d(C) \geq 2e+1$ for $e \in \mathbb{N}_0$. Then*

$$|C| \leq \frac{q^n}{\sum_{i=0}^{e} \binom{n}{i}(q-1)^i}$$

where $q = |F|$.

Proof. By the triangle inequality, the balls $B_e(c)$ around the code words $c \in C$ are pairwise disjoint, hence

$$q^n = |F^n| \geq \sum_{c \in C} |B_e(c)| = |C| \sum_{i=0}^{e} \binom{n}{i}(q-1)^i,$$

and the claim follows. □

Example 2.4. In Question 1.1 we asked whether if $q = |F| = 2$, there is a code $C \subseteq F^4$ with $d(C) \geq 3$ and $|C| = 4$. The Hamming bound (with $e = 1$) gives $|C| \leq \frac{2^4}{1+4} < 4$ and therefore $|C| \leq 3$. So the answer is negative.

An important class of codes are those where the Hamming bound is sharp. The proof of the Hamming bound shows that in this case F^n is a disjoint union of the balls $B_e(c)$, $c \in C$.

Definition 2.5 (Perfect code)**.** A code $C \subseteq F^n$ is a *perfect code*, if F^n is the disjoint union of the balls $B_e(c)$ around the code words $c \in C$ for some integer $1 \leq e \leq n$. In order to emphasize the radius e, one calls it also an *e errors correcting perfect code*.

Remark 2.6. The definition excludes the possibilities $e = 0$ and $e = n$. For $e = 0$, we would obtain the uninteresting case $C = F^n$, and for $e = n$ the code C would consist of a single element.

Lemma 2.7. *Let $C \subseteq F^n$ be an e errors correcting perfect code. Then $d(C) = 2e + 1$.*

Proof. We have $d(C) \geq 2e + 1$ by Lemma 1.7. We need to show that $d(C) \leq 2e + 1$. Pick some $c \in C$. As $e < n$, there is an element $x \in F^n$ such that $d(x, c) = e + 1$. Since C is perfect, there is a code word $c' \in C$ with $d(x, c') \leq e$. In particular $c \neq c'$. The triangle inequality yields $d(c, c') \leq d(x, c) + d(x, c') \leq e + 1 + e = 2e + 1$, and the claim follows. □

The following section describes infinite families of 1 error correcting perfect linear codes.

2.2 Hamming codes

The code from the following theorem is called a *Hamming code*.

Theorem 2.8 (Hamming code)**.** *Let $\mathbb{F}_q$ be the finite field with q elements and $2 \leq m \in \mathbb{N}$. For each 1-dimensional subspace of $\mathbb{F}_q^m$ pick a nonzero vector v_i. Let n be the number of these vectors. Set*

$$C = \{(a_1\, a_2\, \ldots\, a_n) \in \mathbb{F}_q^n \mid a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0\}.$$

Then $n = \frac{q^m - 1}{q - 1}$ and C is a perfect linear code with $d(C) = 3$ and of dimension $n - m$.

Proof. Every 1-dimensional subspace of $\mathbb{F}_q^m$ contains exactly $q - 1$ nonzero elements, and every nonzero element from $\mathbb{F}_q^m$ is contained in exactly one 1-dimensional subspace. Thus $\mathbb{F}_q^m$ has exactly $n = \frac{q^m - 1}{q - 1}$ 1-dimensional subspaces. We consider the linear map

$$\varphi : \mathbb{F}_q^n \to \mathbb{F}_q^m,\ (a_1\, a_2\, \ldots\, a_n) \mapsto a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

The map φ is surjective, because any element $x \in \mathbb{F}_q^m$ is a scalar multiple $x = \lambda v_i$

for some i, so the vector from $\mathbb{F}_q^n$ with λ in the i-th position and 0 elsewhere is mapped to x.
By definition, C is the kernel of φ. The dimension formula gives $\dim C = \dim \mathbb{F}_q^n - \dim \mathbb{F}_q^m = n - m$.
We next show that $d(C) \geq 3$. Suppose this does hold, hence $d(C) \leq 2$. Since C is a linear code, there is an element $0 \neq c = (a_1\, a_2\, \ldots\, a_n) \in C$ with $w(c) \leq 2$. Thus there are different indices i and j such that $a_k = 0$ for all k different from i and j. Thus $a_i v_i + a_j v_j = 0$. Since v_i and v_j are linearly independent, we obtain $a_i = a_j = 0$, so $c = 0$, a contradiction.
In order to show that actually $d(C) = 3$, note that $v_1 + v_2$ is a nonzero vector which is not contained in $\langle v_1 \rangle$, nor in $\langle v_2 \rangle$. Thus $v_1 + v_2 = -\lambda v_i$ for some nonzero scalar λ and some index $i \geq 3$. But then $(1\,1\,0\,\ldots\,0\,\lambda\,0\,\ldots\,0) \in C$, where λ is in the i-th position.
The Hamming bound yields

$$|C| \leq \frac{q^n}{1 + n(q-1)} = \frac{q^n}{1 + \frac{q^m - 1}{q-1}(q-1)} = q^{n-m}.$$

But $|C| = q^{\dim C} = q^{n-m}$, so the Hamming bound is sharp, and therefore C is perfect. □

Example 2.9. Take $q = 2$.

(a) For $m = 2$ we may take $v_1 = (0\,1)$, $v_2 = (1\,0)$, $v_3 = (1\,1)$. Then $c = (a_1\, a_2\, a_3) \in C$ if and only if $a_2 + a_3 = 0$ and $a_1 + a_3 = 0$, which is equivalent to $a_1 = a_2 = a_3$. So C is just a (boring) repetition code $C = \{(0\,0\,0), (1\,1\,1)\}$.

(b) More interesting is the case $m = 3$. Then $n = 7$ and $\dim C = 4$. Suppose that we have a long stream of bits which we want to transmit over a binary channel which requires error correction, but where it suffices to correct at most one bit error among say 7 consecutive bits. As $|C| = 2^4$, we have a bijection $\alpha : \mathbb{F}_2^4 \to C \subset \mathbb{F}_2^7$. We can partition the original bit stream in blocks of length 4, identify each of these blocks with an element $x \in \mathbb{F}_2^4$, and send the code word $c = \alpha(x) \in C$.

Under our assumption about the error rate for the bits, we can correct wrongly transmitted code words, and uniquely reconstruct the original word x.

The price we pay is that instead of sending the original blocks of 4 bits, we send blocks of 7 bits. We see that we have complete error correction, by enlarging the message by a factor of $7/3$. This is better than in case

(a), where we have to enlarge by a factor 3, or in the more clever last example from Section 1.2, which still requires the factor 5/2.

2.3 Perfect codes, part 1

A necessary condition for the existence of an e error correcting perfect code $C \subseteq F^n$, where $q = |F|$, is that

$$\sum_{i=0}^{e} \binom{n}{i} (q-1)^i \text{ divides } q^n. \tag{1}$$

Recall that we are only interested in the cases $1 \leq e \leq n-1$. Since the minimum distance of C is $2e+1 \leq n$, we have $e \leq (n-1)/2$. Up to today it is not possible to classify the possibilities for (1). If q is a prime, then (1) is clearly equivalent to

$$\sum_{i=0}^{e} \binom{n}{i} (q-1)^i = q^m \tag{2}$$

for some integer m. In Problem 2.42 we see that for this conclusion it suffices to assume that q is a prime power.

We discuss a few cases of the conditions (1) and (2).

(a) $n = 2e+1$. I do not know if (1) has solutions besides $q = 2$, but it is not hard to show that a perfect code with $n = 2e+1$ exists only for $q = 2$, see Problem 2.40. Up to equivalence the only example is $C = \{(0\,0\,\ldots\,0), (1\,1\,\ldots\,1)\}$.

(b) $e = 1$. Possible solutions are $n = (q^m-1)/(q-1)$ as we have seen for the Hamming codes. Here q need not be a prime power. However, it is not known if there are perfect codes when q is not a prime power and $n = (q^m-1)/(q-1)$.

However, (1) has many more solutions if q is not a prime power. For instance, $n = (16^m-1)/5$ and $q = 6$ solves (1). Later we will learn another necessary condition for perfect codes which does not hold for this particular example.

(c) $e = 2$, $n \neq 2e+1$. If q is a prime power, then the only solutions seem to be $(q,n) = (3,11)$ and $(q,n) = (2,90)$, but there is no proof for that. The first case indeed belongs to a perfect code, the ternary Golay code which we construct in Section 6.2. The second solution $q = 2$, $n = 90$ does not arise from a perfect code, see Problem 2.43.

If q is not a prime power, then one finds a few more solutions, namely $(q,n) = (15,11)$, $(21,52)$, and $(46,93)$. In Section 3.6 we will see that there are no perfect codes with these parameters.

(d) $e = 3$, $n \neq 2e+1$. Probably the only solution of (1) is $q = 2$, $n = 23$. Anyway, there is a perfect with these parameters, the binary Golay code which we study

in Section 6.1. In fact this is one of the most fascinating codes with many connections to other topics like sporadic simple groups, Steiner systems, and dense high dimensional Euclidean sphere packings.

(e) $e \geq 4$, $n \neq 2e+1$. It is unknown whether (1) has any solutions. However, with the methods from 3.6 one can show (which we don't do) that there are no perfect codes with $4 \leq e$ and $2e+1 < n$.

(f) One might wonder if (1) can be discussed at least for $q = 2$. So the question is which sums $\sum_{i=0}^{e} \binom{n}{i}$ are powers of 2. It seems that this question is beyond present techniques in number theory.

2.4 The Singleton bound

Probably the simplest upper bound on code sizes is

Theorem 2.10 (Singleton bound). *Let $C \subseteq F^n$ be a code with minimum distance d. Then*

$$|C| \leq |F|^{n-d+1}.$$

Proof. Let $\phi : C \to F^{n-d+1}$ defined by cutting off the last $d-1$ coordinates, that is $(a_1, a_2, \ldots, a_n)$ is mapped to $(a_1, a_2, \ldots, a_{n-d+1})$. If $\boldsymbol{c}, \boldsymbol{c}' \in C$ are distinct, then $d(\boldsymbol{c}, \boldsymbol{c}') \geq d$, so $\phi(\boldsymbol{c})$ and $\phi(\boldsymbol{c}')$ still differ in at least one position. Hence ϕ is injective. The claim follows from $|C| = |\phi(C)| \leq |F^{n-d+1}|$. □

Remark 2.11. Codes which achieve the Singleton bound are called *MDS (maximum distance separable) codes*. Not very much is known about MDS codes. In the next section we will describe a large family of linear MDS codes.

2.5 Reed-Solomon codes

The codes in the following theorem are called *Reed-Solomon codes*.

Theorem 2.12 (Reed-Solomon code). *Let $1 \leq d \leq n$ be integers, and $\mathbb{F}_q$ be a finite field with $q \geq n$ elements. Pick n distinct elements $x_1, x_2, \ldots, x_n$ from $\mathbb{F}_q$. Set*

$$C = \{(f(x_1)\, f(x_2)\, \ldots\, f(x_n)) \mid f \in \mathbb{F}_q[X] \text{ of degree } \leq n-d\} \subseteq \mathbb{F}_q^n.$$

Then C is an MDS code with $d(C) = d$ and $\dim C = n - d + 1$.

Proof. Let V be the vector space of the polynomials $f \in \mathbb{F}_q[X]$ of degree $\leq n-d$. Clearly $\dim V = n-d+1$. Consider the linear map $\varphi : V \to C$,

$f \mapsto (f(x_1)\, f(x_2)\, \ldots\, f(x_n))$. Suppose that $f \neq 0$. As $\deg f \leq n - d$, we know that f has at most $n - d$ roots. Thus there are at least d indices i such that $f(x_i) \neq 0$. Thus $w(\varphi(f)) \geq d$, and therefore $d(C) \geq d$. Furthermore, as $d \geq 1$, we see that φ has trivial kernel, so it is an injective map. Thus $\dim C = \dim V = n - d + 1$. The Singleton bound says $n - d + 1 = \dim C \leq n - d(C) + 1$, so $d(C) \leq d$. Above we saw that $d(C) \geq d$, so for our code the Singleton bound is sharp. □

2.6 The Plotkin bound

The next bound, compared to the Singleton bound and the Hamming bound, works particularly well for codes with a large minimum distance.

Theorem 2.13 (Plotkin bound)**.** *Let $C \subseteq F^n$ be a code with $q = |F|$ and $d = d(C) > n(1 - \frac{1}{q})$. Then*

$$|C| \leq \frac{d}{d - n(1 - \frac{1}{q})}.$$

Proof. Set $m = |C|$ and $M = \sum_{u,v \in C} d(u, v)$. As there are $m(m-1)$ pairs u, v in C with $u \neq v$, and $d(u, v) \geq d$ for each such pair, we get

$$M \geq d \cdot m \cdot (m - 1). \tag{3}$$

We compute the contribution of the first position of the code words to M. For each $i \in F$ let a_i be the number of code words $c \in C$ which have the first entry i. Then

$$\sum_{i \in F} a_i = |C| = m. \tag{4}$$

The contribution to M from the first position is

$$M_1 = \sum_{i \in F} a_i (m - a_i).$$

We compute

$$\begin{aligned} M_1 &= \sum_{i \in F} a_i (m - a_i) \\ &\leq \sum_{i \in F} a_i (m - a_i) + \sum_{i \in F} (a_i - \frac{m}{q})^2 \\ &= \sum_{i \in F} \left(a_i m - 2 a_i \frac{m}{q} + \left(\frac{m}{q} \right)^2 \right) \\ &= m^2 (1 - \frac{1}{q}), \end{aligned}$$

where we used (4) in the last step.
This consideration holds for all n positions, hence $M \leq n \cdot m^2(1 - \frac{1}{q})$ and therefore

$$d \cdot m \cdot (m-1) \leq M \leq n \cdot m^2 \cdot (1 - \frac{1}{q}).$$

From this the claim follows. □

Remark 2.14. The proof of the Plotkin bound shows that equality holds if and only all distinct code words $u, v \in C$ have the same distance $d = d(u, v)$, and if in every position each symbol appears the same number of times.

The Plotkin bound has an interesting application to MDS codes. First we need an easy observation.

Lemma 2.15. *Let $C \subseteq F^n$ be an MDS code with $n \geq 2$. For $i \in F$ define*

$$C_i = \{(a_1\, a_2\, \ldots\, a_{n-1}) \mid (a_1\, a_2\, \ldots\, a_n) \in C \text{ and } a_n = i\} \subseteq F^{n-1}.$$

Then C_i is an MDS code with $d(C_i) = d(C)$.

Proof. Set $q = |F|$. Clearly $d(C_i) \geq d(C)$ for all i, and $|C| = \sum_{i \in F} |C_i|$. Furthermore, $|C| = q^{n-d(C)+1}$, because C is an MDS code. The Singleton bound, applied to each of the C_i, gives

$$|C| = \sum_{i \in F} |C_i| \leq \sum_{i \in F} q^{n-1-d(C_i)+1} \leq \sum_{i \in F} q^{n-1-d(C)+1} = q^{n-d(C)+1} = |C|.$$

Thus we have equality everywhere. In particular, $d(C_i) = d(C)$ and $|C_i| = q^{n-1-d(C_i)+1}$ for all i. So C_i is an MDS code with $d(C_i) = d(C)$. □

We learned a huge family of MDS codes, the Reed-Solomon codes. These have the important restriction $q \geq n$. One might ask if there are different constructions which allow small values of q, preferably $q = 2$. Unfortunately, that is not the case.

Theorem 2.16. *Let $C \subseteq F^n$ be an MDS code with $d(C) < n$. Then $|F| \geq d(C)$.*

Proof. Set $d = d(C)$ and $q = |F|$. Repeated application of the previous lemma yields an MDS code $C' \subseteq F^{d+1}$ with $d(C') = d$. If $q \geq d$ then we are done. Thus we may assume that $q \leq d$. Then $d - (d+1)(1 - \frac{1}{q}) = \frac{d-q+1}{q} > 0$, so we may

apply the Plotkin bound to C'. Note that $|C'| = q^{(d+1)-d+1} = q^2$, so

$$q^2 \leq \frac{d}{d-(d+1)(1-\frac{1}{q})}.$$

This yields

$$\frac{q-d}{d-q+1} \geq 0.$$

We assumed $q \leq d$. So the denominator is positive, and we get $q \geq d$. □

Remark 2.17. The idea to apply the Plotkin bound to the MDS code of size q^2 is due to Dominik Barth. My original proof was more complicated.

2.7 Simplex codes

We now present a series of codes which meet the Plotkin bound. The code constructed in the following theorem is called the *simplex code*, because any two distinct elements have the same distance.

Theorem 2.18 (Simplex code)**.** *Let q be a prime power, $n = \frac{q^m-1}{q-1}$, and $A \in \mathbb{F}_q^{m\times n}$ be the matrix where the columns are representatives of the n one-dimensional subspaces of $\mathbb{F}_q^m$. Let $C \subseteq \mathbb{F}_q^n$ be the linear code generated by the rows of A.*
Then $\dim C = m$, and $d(u,v) = q^{m-1}$ for any distinct $u, v \in C$. In particular, $d(C) = q^{m-1}$.

Proof. A contains, up to scalar multiples, all vectors from $\mathbb{F}_q^m$. Therefore the column rank of A is m. By definition, $\dim C$ is the row rank of A, which is m as we know from linear algebra.
It remains to prove the assertion about the distances. As C is a linear code, we only need to show that $w(c) = q^{m-1}$ for each nonzero $c \in C$. Let $a_1, a_2, \ldots, a_m$ be the rows of A, and write $a_i = (a_{i1}\, a_{i2}\, \ldots\, a_{in})$. Let

$$c = \sum_{i=1}^{m} \lambda_i a_i$$

be a nonzero element from C, where $\lambda_i \in \mathbb{F}_q$. Then $w(c)$ is the number of indices $1 \leq j \leq n$ such that $\sum_{i=1}^m \lambda_i a_{ij} \neq 0$. The linear equation $\sum_{i=1}^m \lambda_i x_i = 0$ has q^{m-1} solutions. Thus there are exactly $q^m - q^{m-1}$ vectors $(x_1\, x_2\, \ldots\, x_m)^t \in \mathbb{F}_q^m$ such that $\sum_{i=1}^m \lambda_i x_i \neq 0$. As the matrix A contains from each one-dimensional subspace of $\mathbb{F}_q^m$ precisely one nonzero element, we see that $w(c) = \frac{q^m-q^{m-1}}{q-1} = q^{m-1}$. □

Remark 2.19. (a) The simplex codes meet the Plotkin bound.

(b) Note that $d(C) = q^{m-1}$ is not much smaller than $n = 1 + q + q^2 + \cdots + q^{m-1}$. So compared to the Hamming code, the simplex code can be used in situations where transmission errors arise with a high frequency.

2.8 The Griesmer bound

The proof of the Singleton bound relied on deleting positions of a code. Here we refine this technique for linear codes.

Definition 2.20 (Residual code)**.** Let $C \subseteq \mathbb{F}_q^n$ be a linear code, and $c = (c_1\, c_2\, \ldots\, c_n) \in C$ be a code word with positive weight $w = w(c)$. Let $i_1, i_2, \ldots, i_{n-w}$ be the indices i such that $c_i \neq 0$. Then

$$\mathrm{Res}(C, c) = \{(v_{i_1}\, v_{i_2}\, \ldots\, v_{i_{n-w}}) \mid (v_1\, v_2\, \ldots\, v_n) \in C\} \subseteq \mathbb{F}_q^{n-w}$$

is called the *residual code* with respect to c.

Recall that a linear code $C \subseteq \mathbb{F}_q^n$ is called an $(n, \dim C, d(C))$ code.

Lemma 2.21. *Let $C \subseteq \mathbb{F}_q^n$ be a linear (n, k, d) code. Pick $c \in C$ with $0 < w = w(c) < \frac{q}{q-1}d$. Then $\mathrm{Res}(C, c)$ is a linear $(n - w, k - 1, d')$ code with $d' \geq d - w + \frac{w}{q}$.*

Proof. By replacing C with an equivalent linear code we may assume that $c = (\underbrace{1\,1\,\ldots\,1}_{w}\,\underbrace{0\,0\,\ldots\,0}_{n-w})$. Let $\varphi : C \to \mathrm{Res}(C, c)$ be the linear map sending $v = (v_1\, v_2\, \ldots\, v_n)$ to $\varphi(v) = (v_{w+1}\, v_{w+2}\, \ldots\, v_n)$.
Pick $v \in C$. By the pigeonhole principle, there is an element $\alpha \in \mathbb{F}_q$ which appears at least $\frac{w}{q}$ times among the first w entries $v_1, v_2, \ldots, v_w$ of v. Thus $v - \alpha c$ has at most $w - \frac{w}{q}$ nonzero entries in the first w positions, and exactly $w(\varphi(v))$ nonzero entries in the remaining $n - w$ positions. Thus

$$w(v - \alpha c) \leq w - \frac{w}{q} + w(\varphi(v)).$$

If $v \neq \alpha c$, then $w(v - \alpha c) \geq d$. We obtain

$$w(\varphi(v)) \geq d - w + \frac{w}{q} \quad \text{if } v \neq \alpha c.$$

Any nonzero element from $\mathrm{Res}(C, c)$ has the form $\varphi(v)$ for some $v \in C$. From $\varphi(v) \neq 0$ and $\varphi(c) = 0$ we obtain $v \neq \alpha c$, so $d(\mathrm{Res}(C, c)) \geq d - w + \frac{w}{q}$ as claimed.

It remains to show that $\dim \operatorname{Res}(C,c) = \dim C - 1$. Note that c is in the kernel of φ. Note that by the assumption on w we have $d - w + \frac{w}{q} > 0$, so $w(\varphi(v)) > 0$ if $v \neq \alpha c$. This shows that the kernel of φ is the one-dimensional space $\langle c \rangle$, and the claim follows from the dimension formula. □

Corollary 2.22. *If there is a linear (n,k,d) over $\mathbb{F}_q$ with $k \geq 1$, then there is a linear $(n-d, k-1, d')$ code over $\mathbb{F}_q$ with $d' \geq \frac{d}{q}$.*

Proof. Apply the previous lemma for a nonzero code word c with minimum weight $w(c) = d$. □

In the following theorem $\lceil x \rceil$ is the ceiling function, so $\lceil x \rceil$ is the smallest integer m with $m \geq x$.

Theorem 2.23 (Griesmer bound)**.** *For any linear (n,k,d) code over $\mathbb{F}_q$ the following holds:*

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

Proof. We prove the theorem by induction for k. If $k = 1$ then the inequality claims $d \leq n$, which holds of course.
Thus assume that $k \geq 2$. By the previous corollary we know that there is a linear $(n-d, k-1, d')$ code with $d' \geq \frac{d}{q}$. Applying the induction hypothesis to this code yields

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

From $d' \geq \frac{d}{q}$ we obtain

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

which is equivalent to the claimed inequality. □

Remark 2.24.
(a) For linear codes, the Griesmer bound is always at least as good as the Singleton bound. See Problem 2.47.

(b) The simplex codes not only attain the Plotkin bound, but also the Griesmer bound.

(c) The ternary Golay code, which is a perfect $(11,6,5)$ code over $\mathbb{F}_3$, also attains the Griesmer bound.

(d) There is no known generalization of the Griesmer bound to non-linear codes.

2.9 Hadamard codes

Definition 2.25 (Hadamard matrix)**.** For $n \in \mathbb{N}$ a *Hadamard matrix* is a matrix $H \in \mathbb{Q}^{n\times n}$ with all entries in $\{-1, 1\}$ such that $HH^t = nI_n$, where I_n is the identity matrix.

It is easy to see that the property of being a Hadamard matrix does not change by multiplying rows and columns by -1. In particular, if there is a Hadamard matrix of size n, then there is also one of this size where the entries in the first row and column are 1.

Definition 2.26 (Normalized Hadamard matrix)**.** A Hadamard matrix is *normalized*, if the entries in the first row and column all are 1.

An easy construction of Hadamard matrices is due to Sylvester:

Theorem 2.27 (Sylvester)**.** *If $H \in \mathbb{Q}^{n\times n}$ is a Hadamard matrix, then*

$$\begin{pmatrix} H & H \\ H & -H \end{pmatrix} \in \mathbb{Q}^{2n\times 2n}$$

is a Hadamard matrix as well.

Proof. This follows from

$$\begin{pmatrix} H & H \\ H & -H \end{pmatrix}\begin{pmatrix} H & H \\ H & -H \end{pmatrix}^t = \begin{pmatrix} H & H \\ H & -H \end{pmatrix}\begin{pmatrix} H^t & H^t \\ H^t & -H^t \end{pmatrix} = \begin{pmatrix} 2HH^t & 0 \\ 0 & 2HH^t \end{pmatrix} = 2nI_{2n}$$

□

As $H = (1)$ is a (rather trivial) Hadamard matrix, we see that there are Hadamard matrices of size 2^m for all $m \in \mathbb{N}_0$. Sylvester's construction can easily be generalized. If there are Hadamard matrices of sizes m and n, then there is a Hadamard matrix of size mn, see Problem 2.48.

Next we show a restriction on the sizes of Hadamard matrix.

Theorem 2.28. *If n is the size of a Hadamard matrix, then $n = 1$, 2, or $n \equiv 0 \pmod 4$.*

Proof. Let H be a Hadamard matrix of size $n \geq 3$. As remarked, we may assume that H is normalized, so the first row contains only the entry 1. By permuting columns we may assume that the first three rows of H look as follows:

$$\begin{pmatrix} \underbrace{1 & 1 & \cdots & 1}_{i} & \underbrace{1 & 1 & \cdots & 1}_{j} & \underbrace{1 & 1 & \cdots & 1}_{k} & \underbrace{1 & 1 & \cdots & 1}_{\ell} \\ 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 & -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 & 1 & 1 & \cdots & 1 & -1 & -1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

The pairwise orthogonality of the rows yields

$$\begin{aligned} i + j - k - \ell &= 0 \\ i - j + k - \ell &= 0 \\ i - j - k + \ell &= 0. \end{aligned}$$

From that we get $i = j = k = \ell$, so $n = i + j + k + \ell = 4i$. □

Remark 2.29. There is the conjecture that there is a Hadamard matrix of size n whenever $n \equiv 0 \pmod 4$. The smallest such n for which the existence of a Hadamard matrix is open is 668.

As a preparation for the next theorem, we provide a lemma about the structure of squares in finite fields of odd order.

Lemma 2.30. *Let $\mathbb{F}_q$ be a finite field of odd order q. Let $\chi : \mathbb{F}_q \to \mathbb{Q}$ be the Legendre function, defined by*

$$\chi(i) = \begin{cases} 0, & \text{if } i = 0 \\ 1, & \text{if } 0 \neq i \text{ is a square in } \mathbb{F}_q \\ -1, & \text{if } i \text{ is a non-square in } \mathbb{F}_q. \end{cases}$$

Then the following holds.

(a) $\chi(ij) = \chi(i)\chi(j)$ *for all* $i, j \in \mathbb{F}_q$.

(b) $\sum_{i \in \mathbb{F}_q} \chi(i) = 0$.

(c) $\chi(-1) = -1$ *if* $q \equiv 3 \pmod 4$.

Proof. The map $\alpha : x \mapsto x^2$ is a homomorphism from the multiplicative group $\mathbb{F}_q^\star$ to itself. The kernel of α is $\{-1, 1\}$, so the size of the image S, which is the group of non-zero squares, is $(q-1)/2$. Since S is a subgroup of index 2 in $\mathbb{F}_q^\star$, we get (a). By definition, χ assumes the value 1 on S and -1 on the $(q-1)/2$ elements from $\mathbb{F}_q^\star \setminus S$. This implies (b).
Now suppose that $q \equiv 3 \pmod 4$. Then $(q-1)/2$, the order of S, is odd. Therefore S does not contain the involution -1, and we obtain (c). □

Theorem 2.31 (Paley)**.** *Let q be a prime power such that $q \equiv 3 \pmod 4$. Then there is a Hadamard matrix of size $q+1$.*

Proof. Let $\mathbb{F}_q$ be the finite field with q elements. We use $\mathbb{F}_q$ to number the rows and columns of a matrix $Q \in \mathbb{Q}^{q \times q}$. For $i, j \in \mathbb{F}_q$, let $Q[i,j]$ be the entry of Q in position (i,j).
We use the definitions and results from the previous lemma. Define the matrix $Q \in \mathbb{Q}^{q \times q}$ by $Q[i,j] = \chi(i-j)$. We calculate QQ^t:

$$\begin{aligned}
(QQ^t)[r,s] &= \sum_{k \in \mathbb{F}_q} Q[r,k]Q^t[k,s] \\
&= \sum_{k \in \mathbb{F}_q} Q[r,k]Q[s,k] \\
&= \sum_{k \in \mathbb{F}_q} \chi(r-k)\chi(s-k) \\
&= \sum_{k \in \mathbb{F}_q} \chi(r-s+k)\chi(k) && k \text{ replaced by } s-k \\
&= \sum_{0 \neq k \in \mathbb{F}_q} \chi(r-s+k)\chi(k) \\
&= \sum_{0 \neq k \in \mathbb{F}_q} \chi(r-s+1/k)\chi(1/k) && k \text{ replaced with } 1/k \\
&= \sum_{0 \neq k \in \mathbb{F}_q} \chi(r-s+1/k)\chi(k) && \text{since } \chi(1/k) = \chi(k) \\
&= \sum_{0 \neq k \in \mathbb{F}_q} \chi(k(r-s)+1) && \text{since } \chi \text{ is multiplicative}
\end{aligned}$$

If $r = s$, then the sum is $q-1$, so $(QQ^t)[r,r] = q-1$. If $r \neq s$, then $k(r-s)+1$ runs through all elements from $\mathbb{F}_q$ except 1. Thus the sum is $-\chi(1) = -1$, because $\sum_{i \in \mathbb{F}_q} \chi(i) = 0$. So $(QQ^t)[r,s] = -1$ for $r \neq s$. Set $e = (1\,1\,\ldots\,1) \in \mathbb{Q}^{1 \times q}$. Then all entries of the matrix $e^t e$ are 1. We see that

$$QQ^t = qI_q - e^t e.$$

From $Q[r,s] = \chi(r-s) = \chi(-1)\chi(s-r) = \chi(-1)Q[s,r]$ and $\chi(-1) = -1$ we

furthermore get

$$Q^t = -Q.$$

Define

$$H = \begin{pmatrix} 1 & e \\ e^t & Q - I_q \end{pmatrix}.$$

Then

$$HH^t = \begin{pmatrix} 1 & e \\ e^t & Q - I_q \end{pmatrix} \begin{pmatrix} 1 & e \\ e^t & Q^t - I_q \end{pmatrix} = \begin{pmatrix} q+1 & eQ^t \\ Qe^t & e^te + QQ^t - Q - Q^t + I_q \end{pmatrix}.$$

Part (b) of the previous lemma shows that the sums of the elements in each column and row of Q vanishes. Thus eQ^t, Qe^t are the 0-vectors. Together with $Q^t = -Q$, and $QQ^t = qI_q - e^te$ we obtain $HH^t = (q+1)I_{q+1}$. □

Theorem 2.32. *If there is a Hadamard matrix of size $n > 1$, then there is a code $C \subseteq F^{n-1}$ with $|F| = 2$ of size n and minimum distance $n/2$. In fact, the distance of any distinct code words is $n/2$.*

Proof. Let H be a normalized Hadamard matrix of size n. Let $u = (u_1\, u_2\, \ldots\, u_n)$ and $(v_1\, v_2\, \ldots\, v_n)$ be two distinct code words. Let a and b be the number of indices i where $u_i = v_i$ and $u_i \neq v_i$, respectively. Then $0 = uv^t = \sum_i u_i v_i = a - b$, so $a = b = n/2$. On the other hand, $d(u, v) = b = n/2$. The code of length $n-1$ is the set of rows of H, where we cut off the first element (which is 1 in all rows). □

Remark 2.33. From $n = \frac{\frac{n}{2}}{\frac{n}{2} - (n-1)(1-\frac{1}{2})}$ we see that the Hadamard codes assume the Plotkin bound.

Example 2.34. (a) By Paley's theorem for the prime $p = 11$ there is a Hadamard matrix of size 12. So the previous theorem yields a binary code $C \subseteq \{0,1\}^{11}$ of size 12 and minimum distance $d(C) = 6$.

In view of Corollary 2.22 one might wonder if for arbitrary codes $C \subseteq F^n$ there is a code $C' \subseteq F^{n-d(C)}$ such that $|C'| \geq \frac{|C|}{q}$ and $d(C') \geq \frac{d(C)}{q}$.

The example just constructed shows that this is not the case. Otherwise there were a binary code $C' \subseteq \{0,1\}^5$ with $d(C') \geq 3$ and $|C'| \geq 6$. But the Hamming bound $|C'| \leq \frac{2^5}{1+5} < 6$ shows that such a code does not exist.

(b) One might wonder if the following holds: If the Griesmer bound denies the existence of a linear (n, k, d) code, so there is no non-linear code of

size q^k in $\mathbb{F}_q^n$ and minimum distance $d(C)$. However, that does not hold: As in (a), now for the prime $p = 19$, we obtain a binary code $C \subseteq \{0,1\}^{19}$ of size 20 and minimum distance $d(C) = 10$. In particular, a subset of C has size 2^4 and still minimum distance 10. However, the Griesmer bound shows that for a linear $(19, k, 10)$ code we have $k \leq 3$.

2.10 Reed-Muller codes

We now construct a family of binary linear codes using an idea similar to the one for the Reed-Solomon codes.

Theorem 2.35 (Reed-Muller codes)**.** *Fix $0 \leq r \leq m$. Let $P_{r,m}$ be the set of polynomials in $\mathbb{F}_2[X_1, X_2, \ldots, X_m]$ of total degree $\leq r$ and of degree ≤ 1 with respect to each variable. Let $v_1, v_2, \ldots, v_n$ be the $n = 2^m$ elements from $\mathbb{F}_2^m$. Then*

$$C = \{(f(v_1)\, f(v_2)\, \ldots\, f(v_n)) \mid f \in P_{r,m}\}$$

is a linear $(2^m, \sum_{i=0}^{r} \binom{m}{i}, 2^{m-r})$ code over $\mathbb{F}_2$, the so-called Reed-Muller code $RM(r, m)$.

Proof. We prove the theorem by induction on $m + r$. There is nothing to show if $m + r = 0$.
Now suppose that $m + r \geq 1$. Since $P_{r,m}$ is a vector space, and $C \subseteq \mathbb{F}_2^n$ is a homomorphic image of $P_{r,m}$, we see that C is a linear code. Pick $0 \neq f \in P_{r,m}$, and set $c = (f(v_1)\, f(v_2)\, \ldots\, f(v_n)) \in C$. Note that

$$w(c) = |\{(x_1\, x_2\, \ldots\, x_m) \in \mathbb{F}_2^m \mid f(x_1, x_2, \ldots, x_m) \neq 0\}|.$$

Isolate X_m and write $f = g(X_1, X_2, \ldots, X_{m-1}) + X_m h(X_1, X_2, \ldots, X_{m-1})$. Note that $g \in P_{r,m-1}$ and $h \in P_{r-1,m-1}$. Distinguishing between the cases $x_m = 0$ and $x_m = 1$ we see that

$$\begin{aligned} w(c) = \quad & |\{(x_1\, x_2\, \ldots\, x_{m-1}) \in \mathbb{F}_2^{m-1} \mid g(x_1, x_2, \ldots, x_{m-1}) \neq 0\}| \\ & + |\{(x_1\, x_2\, \ldots\, x_{m-1}) \in \mathbb{F}_2^{m-1} \mid (g+h)(x_1, x_2, \ldots, x_{m-1}) \neq 0\}|. \end{aligned}$$

We show that $w(c) \geq 2^{m-r}$ by looking at the contributions of the two summands. If $g = h$, then $0 \neq g = h \in P_{r-1,m-1}$, so by the induction hypothesis the first summand is $\geq 2^{(m-1)-(r-1)} = 2^{m-r}$, hence $w(c) \geq 2^{m-r}$.
Now assume that $g \neq h$. If $g = 0$, then $0 \neq g + h = h \in P_{r-1,m-1}$, so by the induction hypothesis the second summand is $\geq 2^{(m-1)-(r-1)}$, so $w(c) \geq 2^{m-r}$. If however $g \neq 0$ and $g \neq h$, then g and $g + h$ are nonzero elements from $P_{r,m-1}$, so by the induction hypothesis both summands are $\geq 2^{m-1-r}$, and again $w(c) \geq 2^{m-r}$.

Clearly the monomials $\{X_1^{e_1} X_2^{e_2} \dots X_m^{e_m} \mid 0 \le e_k \le 1, \sum e_k \le r\}$ are a basis of $P_{r,m}$. The number of these polynomials with $\sum e_k = i$ is $\binom{m}{i}$. Thus $\dim P_{r,m} = \sum_{i=0}^{r} \binom{m}{i}$. Since $w(f(v)) > 0$ for $0 \neq f \in P_{r,m}$, we see that the linear map $P_{r,m} \to C$ is bijective, so $\dim C = \dim P_{r,m}$. □

Remark 2.36. (a) A historically important example is the Reed-Muller code $C = \mathrm{RM}(1,5)$. Note that this is a linear $(32,6,16)$ code, so $|C| = 64$. This code was used in the Mariner 6 and 7 Mars missions in 1969 to send images. Each pixel allowed for 64 shades of gray, so the raw data of each pixel was a 6-bit word $(a_0\, a_1 \dots a_5)$. Note that $P_{1,5}$ are the linear polynomials in 5 variables. Accordingly, this raw word was encoded and sent as the 32-bit code word with entries $a_0 + \sum_{i \in S} a_i$, where S runs through the 32 subsets of $\{1,2,3,4,5\}$. This encoding scheme was easy to implement in hardware. Also, for this and other Reed-Muller codes, there exist fast decoding methods.

(b) Given the alphabet size 2, the length 32, and minimum distance at least 16, one can show that 64 is indeed an upper bound for the size of the code, see Problem 2.49.

(c) The construction of the Reed-Muller codes can be generalized to arbitrary finite fields $\mathbb{F}_q$ as alphabets. For this take $0 \le r \le m(q-1)$, and replace $P_{r,m}$ by the set of polynomials in $\mathbb{F}_q[X_1, \dots, X_m]$ of total degree $\le r$ and degree $\le q-1$ in each variable, and $v_1, \dots, v_n$, $n = q^m$, by the elements from $\mathbb{F}_q^m$. Note that for $m = 1$ we recover the Reed-Solomon codes of length q.

2.11 Comparison of some bounds

Table 1 compares some of the bounds for small parameters of n, d, and q. We see that for general codes, there are cases where each of the Hamming bound, the Singleton bound, and the Plotkin bound gives the best result. So none of these bounds is superseded by some other bound.

For comparison, we have included the Griesmer bound for linear codes. We see that there are a few cases where the Hamming bound gives a better result. Note that in the listed cases, the Plotkin bound actually isn't worse than the Griesmer bound. For instance for $n = 7$, $q = 5$, $d = 6$, the Plotkin bound gives $|C| \le 15$, while the Griesmer bound gives $|C| \le 5$. But if C is linear, then $|C|$ is a power of 5, so $|C| \le 15$ implies $|C| \le 5$ as well.

	$n=5, q=2$			$n=5, q=3$			$n=6, q=4$			$n=7, q=5$			$n=8, q=6$		
d	2	3	4	2	3	4	2	3	5	2	3	6	2	3	7
Hamming	32	**5**	5	243	**22**	22	4096	**215**	26	78125	**2693**	214	1679616	**40966**	216
Singleton	**16**	8	4	**81**	27	9	**1024**	256	16	**15625**	3125	25	**279936**	46656	36
Plotkin		6	**2**			**6**			**10**			**15**			**21**
Griesmer	16	4	2	81	27	3	1024	256	4	15625	3125	5			

Table 1: Comparison of some bounds

2.12 The Gilbert–Varshamov bounds

We have seen several constructions of codes. They all have a major shortcoming: Given a finite set F of size q, and positive numbers δ and ρ, then only finitely many of the codes C we learned so far fulfill $\frac{d(C)}{n} \geq \delta$ and $\frac{\log_q|C|}{n} \geq \rho$. Note that if C is linear, then $\dim C = \log_q|C|$. Using the existence theorems in this section, one can show that for any $0 \leq \delta < 1 - \frac{1}{q}$ there is $\rho > 0$ such that there are infinitely many inequivalent codes C such that $\frac{d(C)}{n} \geq \delta$ and $\frac{\log_q|C|}{n} \geq \rho$. (See Problem 2.52 for the case $\delta > 1 - \frac{1}{q}$.) So there are many more codes with interesting parameters.

Theorem 2.37 (Gilbert). *For $1 \leq d \leq n$ and $|F| = q$ there exists a code $C \subseteq F^n$ with $d(C) = d$ and*

$$|C| \geq \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i}.$$

Proof. We show that if $d(C) \geq d$ and $|C| < \frac{q^n}{\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i}$ for $C \subseteq F^n$, then there is $c' \in F^n \setminus C$ such that $d(C \cup \{c'\}) \geq d$. So starting with two elements from F^n with distance d, we can keep adding suitable elements until we reach a code C as claimed in the theorem.
It remains to prove the existence of the element c'. Recall that $\sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i$ is the volume of the Hamming ball with radius $d-1$. Thus the inequality $|C| \sum_{i=0}^{d-1} \binom{n}{i}(q-1)^i < q^n$ shows that the balls with radius $d-1$ around the elements from C do not cover F^n. Thus there is an element $c' \in F^n$ which has distance $> d-1$ from all code words in C. □

If F is a field, then an easy modification of the proof allows to assume that C is a linear code. However, in the linear case we can use a different argument and obtain a slightly better lower bound.

Lemma 2.38. *For $1 \leq d \leq n$ and a prime power q assume that*

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

Then there are n elements in $\mathbb{F}_q^{n-k}$ such that any $d-1$ of them are linearly independent.

Proof. Let $v_1, v_2, \ldots, v_m \in \mathbb{F}_q^{n-k}$ be m elements such that any $d-1$ of them are linearly independent. We count the number of linear combinations of these vectors where at most $d-2$ of the coefficients are non-zero. If there are exactly i nonzero coefficients, then there are $\binom{m}{i}(q-1)^i$ possibilities. Thus if $\sum_{i=0}^{d-2} \binom{m}{i}(q-1)^i < q^{n-k}$, then there is an element $v \in \mathbb{F}_q^{n-k}$ which is not a linear combination of $d-2$ of the vectors $v_1, v_2, \ldots, v_m$. Thus any $d-1$ of the $m+1$ vectors $v_1, v_2, \ldots, v_m, v$ are linearly independent.
Thus as long as $m \leq n-1$, we can keep adding vectors. We obtain the claim. □

Theorem 2.39 (Varshamov)**.** *For $1 \leq d \leq n$ and a prime power q there exists a linear code $C \subseteq \mathbb{F}_q^n$ such that $d(C) \geq d$ and*

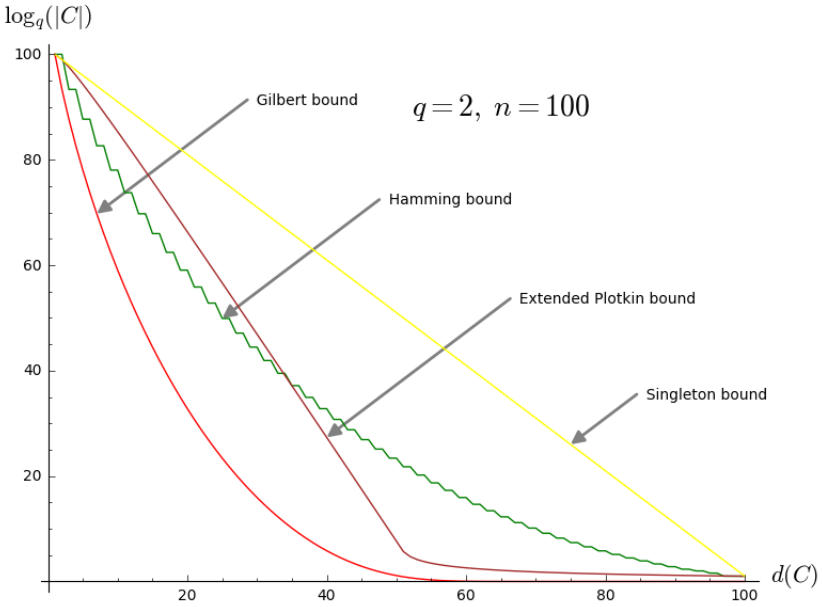
$$\dim C \geq n - \left\lceil \log_q \left(1 + \sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i \right) \right\rceil.$$

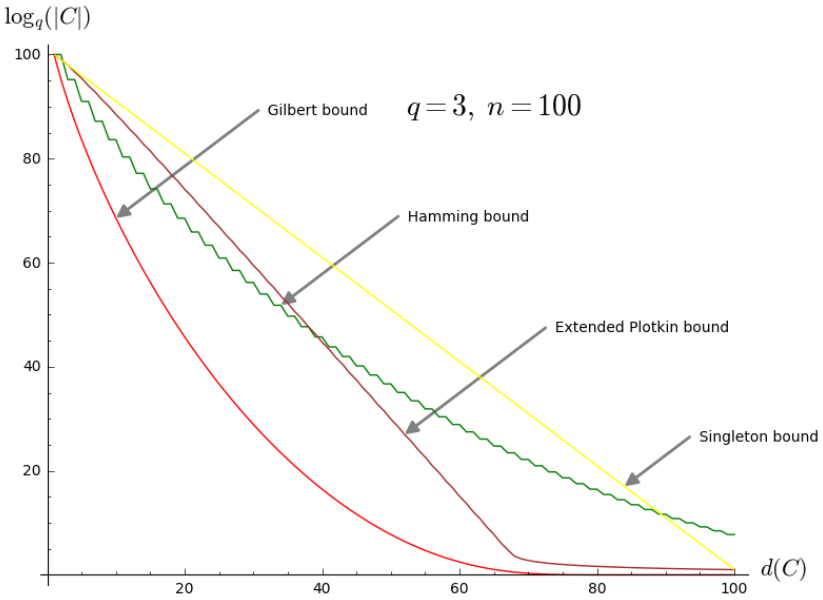
Proof. Set $k = n - \left\lceil \log_q \left(1 + \sum_{i=0}^{d-2} \binom{n-1}{i}(q-1)^i\right)\right\rceil$. Then $k \leq n - \log_q \left(1 + \sum_{i=0}^{d-2} \binom{n-1}{i}(q-1)^i\right)$ and therefore $1 + \sum_{i=0}^{d-2} \binom{n-1}{i}(q-1)^i \leq q^{n-k}$, so $\sum_{i=0}^{d-2} \binom{n-1}{i}(q-1)^i < q^{n-k}$.
By the previous lemma there exist n vectors $v_1, v_2, \ldots, v_n \in \mathbb{F}_q^{n-k}$ such that no $d-1$ of them are linearly independent. Then

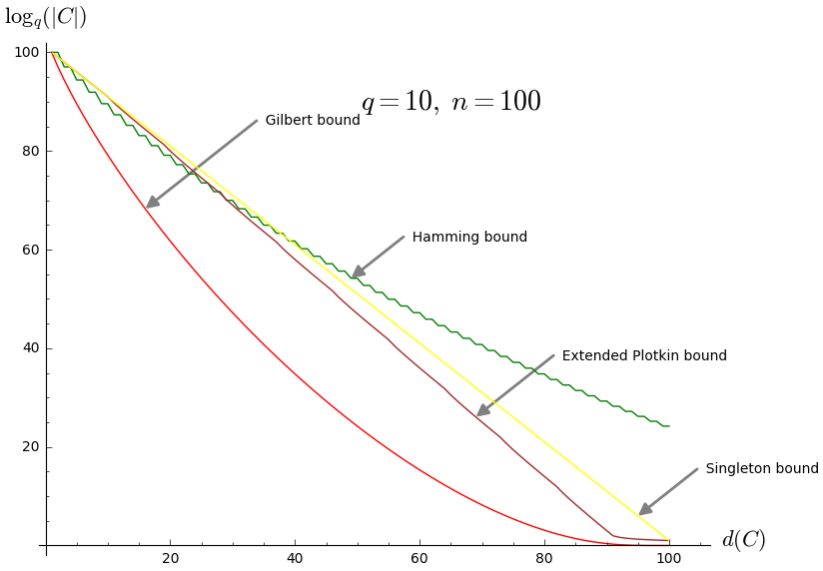
$$C = \{(c_1\, c_2\, \ldots\, c_c) \mid c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0\}$$

has the required properties, for clearly $d(C) \geq d$, and as kernel of the linear map $\mathbb{F}_q^n \to \mathbb{F}_q^{n-k}$, $(c_1\, c_2\, \ldots\, c_c) \mapsto c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$ we have $\dim C \geq n - (n-k) = k$. □

The following graphics show some of our upper bounds and the Gilbert lower bound for $n = 100$ and $q = 2$, 3 and 10.


$\log_q(|C|)$
100
80
60
40
20
Gilbert bound
$q = 2,\ n = 100$
Hamming bound
Extended Plotkin bound
Singleton bound
20
40
60
80
100
$d(C)$


$\log_q(|C|)$
100
80
60
40
20
Gilbert bound
$q = 3,\ n = 100$
Hamming bound
Extended Plotkin bound
Singleton bound
20
40
60
80
100
$d(C)$



2.13 Problems

Problem 2.40. Describe the perfect codes $C \subseteq F^n$ witch $d(C) = n$.
(Hint: Use Problem 1.16.)

Problem 2.41. For $q = 2$ and $e = 3$ determine all possibilities of n, such that $\sum_{i=0}^{e} \binom{n}{i}(q-1)^i$ divides q^n.

Problem 2.42. Suppose that $C \subseteq F^n$ is a perfect code, where $q = |F|$ is a prime power. Show that $|C|$ is a power of q.
(Hint: Show that $q-1$ divides $\frac{q^n}{|C|} - 1$ and use the following easy fact from elementary number theory. If k, ℓ, a are positive integers with $a \geq 2$, such that $a^k - 1$ divides $a^\ell - 1$, then k divides ℓ.)

Problem 2.43. $C \subseteq F^n$ be a perfect code with minimum distance $d(C) = 2e + 1$, where $|F| = 2$. Show that $e + 1$ divides $n + 1$.
(Hint: This does not follow from the sphere packing condition. A possible approach is as follows: Write $F = \{0, 1\}$ and suppose that $(00\ldots0) \in C$. Let B be the set of elements from F^n which are 1 in the first e positions, and have exactly one 1 in the remaining $n - e$ positions. Study how the elements from B are distributed on the balls $B_e(c)$, $c \in C$.)

Problem 2.44. (a) Show that $n = 11$, $d = 5$, $q = 15$ is a solution of the sphere packing condition for perfect codes.

(b) Show that nevertheless there is no perfect code with these parameters.
(Hint: Use a suitable bound.)

Problem 2.45. Let $x_1, x_2, \ldots, x_n$ be distinct elements of the finite field $\mathbb{F}_q$, so $q \geq n$. Pick $1 \leq d \leq n$.

(a) Show that the rows of

$$\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ x_1^2 & \cdots & x_n^2 \\ \vdots & & \vdots \\ x_1^{n-d} & \cdots & x_n^{n-d} \end{pmatrix}$$

generate an $(n, n-d+1, d)$ Reed-Solomon code.

(b) Show that the rows of

$$\begin{pmatrix} 1 & \cdots & 1 & 0 \\ x_1 & \cdots & x_n & 0 \\ x_1^2 & \cdots & x_n^2 & 0 \\ \vdots & & \vdots & \\ x_1^{n-d} & \cdots & x_n^{n-d} & 1 \end{pmatrix}$$

generate an $(n+1, n-d+1, d+1)$ MDS-code. So there are MDS codes with $q < n$.

(c) Now let q be a power of 2. Show that the rows of

$$\begin{pmatrix} 1 & \cdots & 1 & 0 & 0 \\ x_1 & \cdots & x_n & 1 & 0 \\ x_1^2 & \cdots & x_n^2 & 0 & 1 \end{pmatrix}$$

generate an $(n+2, 3, n)$ MDS-code.

Problem 2.46. Euler's thirty-six officers problem from 1782 asks the following:

> Suppose that there are six regiments, each of these consisting of six officers of different ranks. Can we arrange these 36 officers in a 6×6 square such that each regiment and each rank appears exactly once in each row and column?

(a) Show that there is a solution if and only if there is an MDS-code $C \subseteq F^4$ with $|C| = 36$ and $|F| = 6$.

(b) Show that there is no such MDS-code, so Euler's question is negative. (Remark: There is no known easy proof without using a computer.)

Problem 2.47. Show that for linear codes, the Griesmer bound is at least as strong as the Singleton bound.

Problem 2.48. Suppose that there are Hadamard matrices of sizes m and n. Show that there is a Hadamard matrix of size mn.
(Hint: For $H = (h_{ij})$ and H' consider the block matrix where the block in position (i, j) is $h_{ij}H'$.)

Problem 2.49. Let $C \subseteq F^n$ be a code, and $|F| = q$.

(a) Show that for each $0 \leq i \leq n$ there is a code $C' \subseteq F^{n-i}$ with $|C'| \geq \frac{|C|}{q^i}$ and $d(C') \geq d(C)$.
(Hint: Fix i positions. Show that a pattern at these positions occurs at least $\frac{|C|}{q^i}$ times.)

(b) (Extended Plotkin bound) Set $d = d(C)$. Then for any i with $0 \leq i \leq n$ and $d - (n-i)(1-\frac{1}{q}) > 0$ the following holds: $|C| \leq \frac{dq^i}{d-(n-i)(1-\frac{1}{q})}$.

(c) Show that the Reed-Muller code $\mathrm{RM}(1,5)$, a linear $(32, 6, 16)$ code, is the largest possible binary (not necessarily linear) code of length 32 and minimum distance ≥ 16.

Problem 2.50. Let $A_2(n, d)$ be the largest size of a binary code of length n and with minimum distance $\geq d$.

(a) Show that $A_2(n, 2e) = A_2(n-1, 2e-1)$.

(b) Can (a) help in applying the Plotkin (or other) bound?

Problem 2.51. Set $F = \{0, 1, 2, 3, 4, 5\}$. Suppose that there is a perfect code $C \subseteq F^7$ with $d(C) = 3$. Let $C' \subseteq F^4$ be the set of those tuples $(u\,v\,w\,x)$ such that $(0\,0\,0\,u\,v\,w\,x) \in C$. Show that C' is an MDS code with $|C'| = 36$ and $d(C') = 3$.
(Hint: Show that each 5-tuple $(f_1\,f_2\,\ldots\,f_5)$ can be completed to a code word $(f_1\,f_2\,\ldots\,f_7) \in C$.)

Problem 2.52. Set $q = |F|$ and pick δ, ρ with $\delta > 1 - \frac{1}{q}$ and $\rho > 0$. Show that there are only finitely many $n \in \mathbb{N}$ for which there is a code $C \subseteq F^n$ such that $d(C) \geq \delta n$ and $\log_q(|C|) \geq \rho n$.

3 Duality

3.1 The dual code

Let F be a field. We consider the standard symmetric bilinear form on F^n, which is defined by

$$\langle \boldsymbol{u} | \boldsymbol{v} \rangle = \sum_{i=1}^{n} u_i v_i,$$

where $\boldsymbol{u} = (u_1\, u_2\, \ldots\, u_n)$ and $\boldsymbol{v} = (v_1\, v_2\, \ldots\, v_n)$. Note that this form is not degenerate, that means if $\boldsymbol{v} \neq 0$, then there is a vector $\boldsymbol{u}$ with $\langle \boldsymbol{u}|\boldsymbol{v}\rangle \neq 0$.

For a subset C of F^n set

$$C^\perp = \{\boldsymbol{u} \mid \boldsymbol{u} \in F^n, \langle \boldsymbol{u}|\boldsymbol{c}\rangle = 0 \text{ for all } \boldsymbol{c} \in C\}.$$

Clearly $C^\perp$ is a subspace of F^n. The most important case for us is when C is a subspace of F^n. If $F = \mathbb{R}$, then we know from linear algebra that then $F^n = C \oplus C^\perp$. Here, however, we will be mainly interested in the case that F is a finite field. Then it need not be true anymore that $C \cap C^\perp = \{0\}$. In fact, we will see later that codes $C \leq \mathbb{F}_q^n$ with $C = C^\perp$ will play a prominent role. If $C \leq \mathbb{F}_q^n$ is a code, then we will call $C^\perp$ the *dual code*.

Lemma 3.1. *Let $C \leq F^n$ be a subspace. Then*

(a) $\dim C + \dim C^\perp = n$.

(b) $(C^\perp)^\perp = C$.

Proof. Set $k = \dim C$, and let M be the $(k \times n)$–matrix whose rows are a basis of C. Then by definition $C^\perp$ consist of those vectors $\boldsymbol{v}$ where $M\boldsymbol{v}^t = 0$. So $C^\perp$ is the kernel of the linear map $F^n \to F^k$, $\boldsymbol{v} \mapsto M\boldsymbol{v}^t$. This map is surjective, because M has rank k. The dimension formula gives (a).
By definition $C \subseteq (C^\perp)^\perp$. On the other hand $\dim(C^\perp)^\perp = n - \dim C^\perp = n - (n - \dim C) = \dim C$. This yields (b). □

Remark 3.2. We see that by definition, the dual code of a Hamming code is a simplex code.

3.2 Linear characters of finite groups

Let $(G, +)$ be a (not necessarily abelian) group, and let $\mathbb{C}^\star$ be the multiplicative group of the complex numbers. A *character* is a homomorphism $\chi : G \to \mathbb{C}^\star$, so $\chi(g+h) = \chi(g)\chi(h)$ for all $g, h \in G$. We say that G is trivial if $\chi(g) = 1$ for all $g \in G$. Note that $\chi(0) = \chi(0+0) = \chi(0)\chi(0)$, hence $\chi(0) = 1$. Somewhat less obvious is

Lemma 3.3. *Let $\chi : G \to \mathbb{C}^\star$ be a character of the finite group G. Then*

(a) $\chi(-g) = \overline{\chi(g)}$ *for all* $g \in G$.

(b)

$$\sum_{g\in G}\chi(g) = \begin{cases} |G|, & \text{if } \chi \text{ is trivial} \\ 0, & \text{if } \chi \text{ is non-trivial} \end{cases}$$

Proof. (a) From $1 = \chi(0) = \chi(-g+g) = \chi(-g)\chi(g)$ we get $\chi(-g) = 1/\chi(g)$. As G is a finite group, there is $n \in \mathbb{N}$ such that $ng = 0$. So $1 = \chi(ng) = \chi(g)^n$. We see that $\chi(g)$ is an n-th root of unity, so in particular $\chi(g)\overline{\chi(g)} = 1$ and the claim follows.
(b) The assertion is clear if χ is trivial. Thus assume that χ is non-trivial. So there is $h \in G$ which $\chi(h) \neq 1$. Set $S = \sum_{g\in G}\chi(g)$. If g runs through G, then $h+g$ runs trough G as well. We obtain

$$S = \sum_{g\in G}\chi(h+g) = \sum_{g\in G}\chi(h)\chi(g) = \chi(h)S,$$

hence $S = 0$ as $\chi(h) \neq 1$. □

Lemma 3.4. *Let F be the ring $\mathbb{Z}/q\mathbb{Z}$ for $q \geq 2$, or a finite field.*
Then $(F,+)$ has a non-trivial character χ with the following additional property: For each $0 \neq b \in F$ there is $a \in F$ with $\chi(ba) \neq 1$.

Proof. Suppose that $F = \mathbb{Z}/q\mathbb{Z}$. Let $\zeta = e^{2\pi i/q}$ be a primitive q-th root of unity. Then for $g = j + q\mathbb{Z}$ set $\chi(g) = \zeta^j$. This is well-defined, and defines a nontrivial character on F with $\chi(g) = 1$ if and only if $g = 0$. So in particular $\chi(ba) \neq 1$ for $b \neq 0$ and $a = 1$.
Next suppose that F is a finite field. Let p be the characteristic of F, so F contains the prime field $\mathbb{F}_p$. By (a), there is a non-trivial character χ' on $(\mathbb{F}_p,+)$. Now F is a vector space over $\mathbb{F}_p$, so there is a surjective linear map $\psi : F \to \mathbb{F}_p$. Then $\chi = \chi' \circ \psi$ is a non-trivial character on F. As χ is not trivial, there is $c \in F$ with $\chi(c) \neq 1$. So $\chi(ba) \neq 1$ for $b \neq 0$ and $a = b^{-1}c$. □

Remark 3.5. (a) In group theory, a character as defined here is called a *linear character*. In general a character is defined as $\chi(g) = \mathrm{trace}(\phi(g))$, where $\phi : G \to \mathrm{GL}_n(\mathbb{C})$ is a homomorphism. Note that in this case χ need not be a homomorphism anymore. For finite abelian groups, one can show that any character is built from linear characters, so there is little point in considering non-linear characters.

(b) In number theory characters of finite abelian groups play a prominent role, for instance in the proof of Dirichlet's theorem about primes in

arithmetic progressions.

(c) Any finite abelian group of order ≥ 2 has a non-trivial linear character, see Problem 3.41. However, we cannot drop the assumption that G is abelian: As $G/\ker\chi$ is isomorphic to a subgroup of the abelian group $\mathbb{C}^\star$, we see that $\ker\chi$ contains the derived subgroup (syn. commutator subgroup) G'. So if $G = G'$, which happens for instance for the alternating group Alt_5, there is no non-trivial linear character.

(d) If χ is a non-trivial character of the finite field F, then of course the additional property is automatically satisfied. This is not true if F is not a field. Take e.g. $F = \mathbb{Z}/4\mathbb{Z}$, and define $\chi(j + 4\mathbb{Z}) = (-1)^j$. Then $\chi(2a) = 1$ for all $a \in F$.

In the next three sections F will be a finite field $\mathbb{F}_q$, or the residue ring $\mathbb{Z}/q\mathbb{Z}$ for some $q \geq 2$, respectively. Suppose that $F = \mathbb{Z}/q\mathbb{Z}$. We consider F^n as an additive group, and like in the case that F is a field we define the bi-additive function $F^n \times F^n \to F$ which sends the pair $\boldsymbol{u} = (u_1\, u_2\, \ldots\, u_n)$, $\boldsymbol{v} = (v_1\, v_2\, \ldots\, v_n)$ to $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = \sum u_i v_i$. For $\boldsymbol{v} \in F^n$, let $w(\boldsymbol{v})$ be the number of non-zero components v_i. So $d(\boldsymbol{u}, \boldsymbol{v}) = w(\boldsymbol{u} - \boldsymbol{v})$.

The main technical lemma to be used in the next three sections is

Lemma 3.6. *Let z be a polynomial variable over $\mathbb{C}$, and F be the residue class ring $\mathbb{Z}/q\mathbb{Z}$ for some $q \geq 2$, or a finite field of order q. Then $(F, +)$ has a character χ such that for all $n \in \mathbb{N}$ and $\boldsymbol{u} \in F^n$ the following holds:*

$$\sum_{\boldsymbol{v}\in F^n} z^{w(\boldsymbol{v})}\chi(\langle \boldsymbol{u}|\boldsymbol{v}\rangle) = (1-z)^{w(\boldsymbol{u})}(1+(q-1)z)^{n-w(\boldsymbol{u})}.$$

Proof. Let χ be a character of $(F, +)$ such that for each $0 \neq b \in F$ there is $a \in F$ with $\chi(ba) \neq 1$. Such a character exists by Lemma 3.4.
For $a \in F$ set $w(a) = 1$ if $a \neq 0$, and $w(a) = 0$ otherwise. Thus $w(\boldsymbol{v}) = w(v_1) + w(v_2) + \cdots + w(v_n)$ for $\boldsymbol{v} = (v_1\, v_2\, \ldots\, v_n) \in F^n$.
Write $\boldsymbol{u} = (u_1\, u_2\, \ldots\, u_n)$, so $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = u_1v_1 + u_2v_2 + \cdots + u_nv_n$. We compute

$$\begin{aligned}\sum_{\boldsymbol{v}\in F^n} z^{w(\boldsymbol{v})}\chi(\langle \boldsymbol{u}|\boldsymbol{v}\rangle) &= \sum_{v_1,v_2,\ldots,v_n\in F} \chi(u_1v_1 + u_2v_2 + \cdots + u_nv_n)z^{w(v_1)+w(v_2)+\cdots+w(v_n)} \\ &= \sum_{v_1,v_2,\ldots,v_n\in F} (\chi(u_1v_1)z^{w(v_1)})(\chi(u_2v_2)z^{w(v_2)})\ldots(\chi(u_nv_n)z^{w(v_n)}) \\ &= \prod_{i=1}^{n}\sum_{a\in F}\chi(u_ia)z^{w(a)}\end{aligned}$$

It remains to compute the terms $\sum_{a\in F}\chi(u_ia)z^{w(a)}$. Recall that $w(a) = 1$ if $a \neq 0$,

and $w(0) = 0$. This yields

$$\sum_{a \in F} \chi(u_i a) z^{w(a)} = 1 - z + \sum_{a \in F} \chi(u_i a).$$

Note that $F \to F$, $a \mapsto u_i a$ is additive, so $a \mapsto \chi_i(a) = \chi(u_i a)$ is a character of $(\mathbb{F}_q, +)$. By the choice of χ we see that χ_i is non-trivial if and only if $u_i \neq 0$. Lemma 3.3 then yields

$$\sum_{a \in F} \chi(u_i a) z^{w(a)} = \begin{cases} 1 - z, & \text{if } u_i \neq 0 \\ 1 + (q-1)z, & \text{if } u_i = 0. \end{cases}$$

So in the product, the factor $1 - z$ appears $w(\boldsymbol{u})$ times, and the remaining $n - w(\boldsymbol{u})$ factors equal $1 + (q-1)z$. □

Let $C \subseteq F^n$ be a code. For $\boldsymbol{x} \in F^n$ and $0 \leq i \leq n$ we let $A_i(\boldsymbol{x})$ be the number of code words $\boldsymbol{c} \in C$ with $d(\boldsymbol{x}, \boldsymbol{c}) = i$. Several results in the following sections depend on

Lemma 3.7. *Let z be a polynomial variable over $\mathbb{C}$, and F be the residue class ring $\mathbb{Z}/q\mathbb{Z}$ for some $q \geq 2$, or a finite field of order q. Let χ be a character as in the previous lemma.*
Let $C \subseteq F^n$ be a code, $\boldsymbol{x} \in F^n$, and $\hat{A}_i(\boldsymbol{x})$ be the coefficient of z^i in

$$\sum_{i=0}^{n} A_i(\boldsymbol{x})(1-z)^i (1 + (q-1)z)^{n-i}.$$

Then

$$\hat{A}_i(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{v} \in F^n \\ w(\boldsymbol{v}) = i}} \left(\overline{\chi(\langle \boldsymbol{x} | \boldsymbol{v} \rangle)} \sum_{\boldsymbol{c} \in C} \chi(\langle \boldsymbol{c} | \boldsymbol{v} \rangle) \right).$$

Proof. As F is a ring, we have $d(\boldsymbol{c}, \boldsymbol{x}) = w(\boldsymbol{c} - \boldsymbol{x})$. Clearly

$$\sum_{i=0}^{n} A_i(\boldsymbol{x})(1-z)^i (1 + (q-1)z)^{n-i} = \sum_{\boldsymbol{c} \in C} (1-z)^{w(\boldsymbol{c}-\boldsymbol{x})} (1 + (q-1)z)^{n - w(\boldsymbol{c}-\boldsymbol{x})}.$$

Using the previous lemma, the bi-additivity of the form $\langle \cdot | \cdot \rangle$, and $\chi(-g) = \overline{\chi(g)}$,

we compute the right hand side:

$$\begin{aligned}\sum_{c\in C}(1-z)^{w(c-x)}(1+(q-1)z)^{n-w(c-x)} &= \sum_{c\in C}\sum_{v\in F^n} z^{w(v)}\chi(\langle c-x|v\rangle)\\ &= \sum_{c\in C}\sum_{v\in F^n} z^{w(v)}\chi(\langle c|v\rangle)\chi(-\langle x|v\rangle)\\ &= \sum_{v\in F^n} z^{w(v)}\left(\overline{\chi(\langle x|v\rangle)}\sum_{c\in C}\chi(\langle c|v\rangle)\right)\end{aligned}$$

The coefficient of z^i arises from summing over those $v \in F^n$ with $w(v) = i$. The claim follows. □

3.3 The MacWilliams identity

Let $C \leq \mathbb{F}_q^n$ be a linear code. We set $A_i = A_i(\mathbf{0})$, so A_i is the number of code words of weight i. One calls the sequence $A_0, A_1, \ldots, A_n$ the *weight distribution* of C. Note that $A_0 = 1$, $A_0 + A_1 + \cdots + A_n = |C|$, and $A_1 = A_2 = \ldots A_{d-1} = 0$ for $d = d(C)$.

Let z be a variable over $\mathbb{C}$. For nice codes it will turn out that the generating function (which is a polynomial of degree at most n in this case) $A_C(z) = \sum_{i=1}^n A_i z^i$ often has a nice closed form. $A_C(z)$ is called the *weight enumerator* of C. Clearly

$$A_C(z) = \sum_{i=1}^{n} A_i z^i = \sum_{c\in C} z^{w(c)}.$$

The weight enumerator of a code has several important applications. In this section we will prove the surprising result by MacWilliams that for linear codes C, the weight enumerator of $C^\perp$ can be computed in terms of the weight enumerator of C. For another application of weight enumerators see Problem ???.

Before stating and proving the MacWilliams Identity, we show how characters can be used to check if a vector lies in the dual of a code.

Lemma 3.8. *Let $C \leq \mathbb{F}_q^n$ be a linear code, and $\chi : \mathbb{F}_q \to \mathbb{C}^\star$ a non-trivial character. Then, for all $v \in \mathbb{F}_q^n$,*

$$\sum_{c\in C}\chi(\langle c|v\rangle) = \begin{cases}|C|, & v \in C^\perp\\ 0, & v \notin C^\perp\end{cases}$$

Proof. The assertion is clear if $v \in C^\perp$. Thus assume that $v \notin C^\perp$. Then there is $c \in C$ with $\langle c|v\rangle \neq 0$, and by linearity of the scalar product in the first component, we see that $\psi : C \to \mathbb{F}_q$, $\psi(c) = \langle c|v\rangle$, is surjective. But then $\chi \circ \psi : C \to \mathbb{C}^\star$, $c \mapsto \chi(\langle c|v\rangle)$, is a non-trivial character, and the claim follows from Lemma 3.3. □

Theorem 3.9 (MacWilliams). *Let $C \leq \mathbb{F}_q^n$ be a linear code with weight enumerator $A(z) = \sum_{i=0}^{n} A_i z^i$. Then the weight enumerator $A^\perp(z)$ of the dual code $C^\perp$ is given by*

$$A^\perp(z) = \frac{1}{|C|} \sum_{i=0}^{n} A_i (1-z)^i (1+(q-1)z)^{n-i}.$$

Proof. We apply Lemma 3.7 with $\boldsymbol{x} = \mathbf{0}$, hence $\chi(\langle \boldsymbol{x}|\boldsymbol{v}\rangle) = 1$. So the coefficient of z^i in

$$\sum_{i=0}^{n} A_i (1-z)^i (1+(q-1)z)^{n-i}$$

is

$$\sum_{\substack{\boldsymbol{v}\in F^n \\ w(\boldsymbol{v})=i}} \sum_{\boldsymbol{c}\in C} \chi(\langle \boldsymbol{c}|\boldsymbol{v}\rangle).$$

By the previous lemma, this inner sum is $|C|$ if $\boldsymbol{v} \in C^\perp$, and 0 otherwise. So this sum simplifies to

$$|C| \sum_{\substack{\boldsymbol{v}\in C^\perp \\ w(\boldsymbol{v})=i}} = |C| A_i^\perp,$$

where $A_0^\perp, A_1^\perp, \ldots, A_n^\perp$ is the weight distribution of $C^\perp$. □

Remark 3.10. The MacWilliams Identity can also be written as

$$A^\perp(z) = \frac{(1+(q-1)z)^n}{|C|} A\left(\frac{1-z}{1+(q-1)z}\right).$$

Example 3.11. The MacWilliams identity can be used for instance to compute the weight enumerator of the Hamming codes. Set $n = \frac{q^m-1}{q-1}$, and let C be the $(n, n-m, 3)$ Hamming code over $\mathbb{F}_q$. Then we know that the dual code $C^\perp$ is the (n, m, q^{m-1}) Simplex code. Let $A(z)$ and $A^\perp(z)$ be the weight enumerators of C and $C^\perp$, respectively. Since all nonzero codewords in $C^\perp$ have weight q^{m-1}, we have

$$A^\perp(z) = 1 + (q^m - 1) z^{q^{m-1}}.$$

The MacWilliams identity and $C = (C^\perp)^\perp$ then yield

$$\begin{aligned} A(z) &= (A^\perp)^\perp(z) \\ &= \frac{(1+(q-1)z)^n}{q^m}\left(1+(q^m-1)\left(\frac{1-z}{1+(q-1)z}\right)^{q^{m-1}}\right) \\ &= \frac{1}{q^m}\left((1+(q-1)z)^{\frac{q^m-1}{q-1}} + (q^m-1)(1-z)^{q^{m-1}}(1+(q-1)z)^{\frac{q^{m-1}-1}{q-1}}\right). \end{aligned}$$

Expanding in powers of z then yields very messy expressions for the coefficients A_i. Without duality and the relation to the simplex code, it would certainly have been very difficult to compute the weight enumerator of a Hamming code.

3.4 The Linear Programming Bound

The MacWilliams Identity has a surprising consequence. As the weight enumerator of the dual code $C^\perp$ of a linear code C of course has non-negative coefficients, we obtain that the coefficients of $\sum_{i=0}^n A_i(1-z)^i(1+(q-1)z)^{n-i}$ are non-negative. This yields linear inequalities of the numbers A_i, which are non-negative as well. If $d = d(C)$, then in addition to $A_0 = 1$ we have $A_1 = A_2 = \cdots = A_{d-1} = 0$. These restrictions allow to bound $|C| = A_0 + A_1 + \cdots + A_n$ from above. It turns out that the bounds one obtains this way in general are much better than the upper bounds we have developed so far.

What is even more surprising is that we can obtain the same bounds for non-linear codes, despite the fact that for non-linear codes there are no good notions of duality.

In order to do so, we introduce the *distance distribution* and the *distance enumerator* polynomial.

Definition 3.12. Let F be a finite set of order ≥ 2, and $C \subseteq F^n$ be a code. For $0 \leq i \leq n$ let $|C|B_i$ be the number of pairs $\boldsymbol{c}_1, \boldsymbol{c}_2 \in C$ with $d(\boldsymbol{c}_1, \boldsymbol{c}_2) = i$. Then $B_0, B_1, \ldots, B_n$ is the *distance distribution* of C, and the *distance enumerator* of C is the polynomial

$$B(z) = \sum_{i=0}^n B_i z^i = \frac{1}{|C|}\sum_{\boldsymbol{c}_1,\boldsymbol{c}_2 \in C} z^{d(\boldsymbol{c}_1,\boldsymbol{c}_2)}.$$

Remark 3.13. We record a few simple facts, some of which we use without further notice:

(a) $B_0 = 1$ and $|C| = B_0 + B_1 + \cdots + B_n$.

(b) $B_i = \frac{1}{|C|} \sum_{\boldsymbol{x} \in C} A_i(\boldsymbol{x})$.

(c) If C is linear, then $A_i = B_i$ for all i, and $A(z) = B(z)$. Note that for non-linear codes, B_i need not be integral.

Theorem 3.14 (Delsarte). *Let $C \subseteq F^n$ be a code with distance enumerator $B(z) = \sum_{i=0}^{n} B_i z^i$ and $|F| = q \geq 2$. Then the coefficients of*

$$\sum_{i=0}^{n} B_i (1-z)^i (1+(q-1)z)^{n-i}$$

are non-negative.

Proof. Without loss of generality we may assume that $F = \mathbb{Z}/q\mathbb{Z}$. Let $\hat{B}_i$ be the coefficient of z^i in $\sum_{i=0}^{n} B_i (1-z)^i (1+(q-1)z)^{n-i}$. From $|C| B_i = \sum_{\boldsymbol{x} \in C} A_i(\boldsymbol{x})$ we get $|C| \hat{B}_i = \sum_{\boldsymbol{x} \in C} \hat{A}_i(\boldsymbol{x})$, so

$$\begin{aligned}
|C| \hat{B}_i &= \sum_{\boldsymbol{x} \in C} \hat{A}_i(\boldsymbol{x}) \\
&= \sum_{\boldsymbol{x} \in C} \sum_{\substack{\boldsymbol{v} \in F^n \\ w(\boldsymbol{v}) = i}} \left(\overline{\chi(\langle \boldsymbol{x} | \boldsymbol{v} \rangle)} \sum_{\boldsymbol{c} \in C} \chi(\langle \boldsymbol{c} | \boldsymbol{v} \rangle) \right) \qquad \text{(by Lemma 3.7)} \\
&= \sum_{\substack{\boldsymbol{v} \in F^n \\ w(\boldsymbol{v}) = i}} \left(\sum_{\boldsymbol{x} \in C} \overline{\chi(\langle \boldsymbol{x} | \boldsymbol{v} \rangle)} \sum_{\boldsymbol{c} \in C} \chi(\langle \boldsymbol{c} | \boldsymbol{v} \rangle) \right) \\
&= \sum_{\substack{\boldsymbol{v} \in F^n \\ w(\boldsymbol{v}) = i}} \left| \sum_{\boldsymbol{c} \in C} \chi(\langle \boldsymbol{c} | \boldsymbol{v} \rangle) \right|^2 \\
&\geq 0
\end{aligned}$$

□

For later use we record an important corollary of the previous proof

Lemma 3.15. *Let $C \subseteq F^n$ be a code with distance enumerator $B(z) = \sum_{i=0}^{n} B_i z^i$ and $|F| = q \geq 2$. If the coefficient of z^i in*

$$\sum_{i=0}^{n} B_i (1-z)^i (1+(q-1)z)^{n-i}$$

is 0, then for each $\boldsymbol{x} \in F^n$ the coefficient of z^i in

$$\sum_{i=0}^{n} A_i(\boldsymbol{x})(1-z)^i(1+(q-1)z)^{n-i}$$

is 0 too.

Proof. In the notation of the proof of the previous theorem assume that $\hat{B}_i = 0$. The proof then shows that

$$\sum_{\boldsymbol{c} \in C} \chi(\langle \boldsymbol{c} | \boldsymbol{v} \rangle) = 0$$

for each $\boldsymbol{v} \in F^n$ with $w(\boldsymbol{v}) = i$. The claim then follows from Lemma 3.7. □

We can slightly rephrase Delsarte's Theorem. To do so, let $K_k(i)$ be the coefficient of z^k in $(1-z)^i(1+(q-1)z)^{n-i}$. Then Delsarte's Theorem says that $\sum_{i=0}^{n} B_i K_k(i) \geq 0$ for all $0 \leq k \leq n$. Note $B_1 = B_2 = \cdots = B_{d-1} = 0$ for $d = d(C)$. Together with $B_0 = 1$ and $|C| = B_0 + B_1 + \cdots + B_n$ we get

Theorem 3.16 (Linear Programming Bound). *$C \subseteq F^n$ be a code with $d = d(C)$, $|F| = q$, and $K_k(i)$ be the coefficient of z^k in $(1-z)^i(1+(q-1)z)^{n-i}$. Choose $b_d, b_{d+1}, \ldots, b_n \in \mathbb{R}$ such that $b = b_d + b_{d+1} + \cdots + b_n$ is maximal subject to the following restrictions:*

$$\begin{array}{ll} b_i \geq 0 & \text{for all } d \leq i \leq n, \text{ and} \\ \sum_{i=d}^{n} b_i K_k(i) \geq -K_k(0) & \text{for all } 0 \leq k \leq n. \end{array}$$

Then $|C| \leq b + 1$.

The problem here is that it is difficult to determine the maximum b. In fact, while one can compute numerically this maximum for small values of n and q, it is unknown in general. However, one can compute more easily upper bounds for the maximum. This is based on the easy concept of duality of linear programs.

In the following we consider real column vectors. For $a, b \in \mathbb{R}^n$ we write $a \geq b$ if $a - b$ has no negative entries.

Definition 3.17 (Linear Program). Let $m, n \in \mathbb{N}$, $A \in \mathbb{R}^{m \times n}$ be a real $m \times n$ matrix, $\boldsymbol{b} \in \mathbb{R}^m$, and $\boldsymbol{c} \in \mathbb{R}^n$. A linear programs asks for the maximum of $\boldsymbol{c}^t \boldsymbol{x}$ for $\boldsymbol{x} \in \mathbb{R}^n$ with $\boldsymbol{x} \geq \boldsymbol{0}$ and $A\boldsymbol{x} \leq \boldsymbol{b}$.

It may happen that there is no solution $\boldsymbol{x}$ at all. In this case one says that the linear program is *infeasible*. It may also happen that there is no upper bound for $\boldsymbol{c}^t \boldsymbol{x}$ for all solutions $\boldsymbol{x}$, so the maximum is ∞. In this case the linear program is said to

unbounded.

In our case, it is easy to see that the previous theorem yields a linear program. Of course it is feasible, because for instance $b_d = b_{d+1} = \cdots = b_n = 0$ is a solution, and $K_k(0) \geq 0$.

It is less clear that the linear program is bounded. To show that, and in order to obtain upper bounds for $\boldsymbol{c}^t\boldsymbol{x}$, one considers the so-called *dual linear program*:

Theorem 3.18 (Weak Duality Theorem)**.** *With the notation from the previous definition, suppose that there is $\boldsymbol{y} \in \mathbb{R}^m$ with $\boldsymbol{y} \geq \boldsymbol{0}$ and $A^t\boldsymbol{y} \geq \boldsymbol{c}$. Then $\boldsymbol{c}^t\boldsymbol{x} \leq b^t\boldsymbol{y}$ for each solution $\boldsymbol{x}$ of the original linear program.*

Proof. This follows immediately from

$$\boldsymbol{c}^t\boldsymbol{x} = \boldsymbol{x}^t\boldsymbol{c} \leq \boldsymbol{x}^t(A^t\boldsymbol{y}) = (A\boldsymbol{x})^t\boldsymbol{y} \leq b^t\boldsymbol{y}.$$

□

By a suitable (yet very technical) choice of the vector $\boldsymbol{y}$, one can use this theorem for instance to retrieve the Singleton bound from Theorem 3.16. So this Theorem is at least as strong as the Singleton bound. Other (more complicated) choices of $\boldsymbol{y}$ yield the Hamming bound or the Plotkin bound. This already indicates that the linear programming bound, despite its difficulty to actually use it, is a very strong bound.

By applying the Weak Duality Theorem 3.18 in order to bound the maximum b in Theorem 3.16 one does not loose anything. This follows from the Strong Duality Theorem: If the linear program in Definition 3.17 is feasible and bounded, then there exists a solution $\boldsymbol{y}$ in Theorem 3.18. Choose $\boldsymbol{y}$ such that $b^t\boldsymbol{y}$ is minimal, and choose $\boldsymbol{x}$ such that $\boldsymbol{c}^t\boldsymbol{x}$ is maximal. Then $\boldsymbol{c}^t\boldsymbol{x} = b^t\boldsymbol{y}$. For a proof see any book about linear programming or linear optimization.

3.5 The Covering Radius

Let $C \subseteq F^n$ be a code. The *covering radius* $r(C)$ of C is the smallest $r \geq 0$ such that the Hamming balls of radius r around the code words cover F^n. Clearly, $\lfloor \frac{d(C)-1}{2} \rfloor \leq r(C) \leq n$. Also, if C is perfect with minimum distance $2e+1$, then $r(C) = e$.

Definition 3.19 (MacWilliams transformation)**.** Let $n \in \mathbb{N}$ and $q \in \mathbb{C}$ be fixed. For $A(z) = a_0 + a_1 z + \cdots + a_n z^n \in \mathbb{C}[z]$ we define the *MacWilliams transformation* $\hat{A}(z)$ by $\hat{A}(z) = \sum_{i=0}^{n} a_i (1-z)^i (1+(q-1)z)^{n-i}$. Note that $A \mapsto \hat{A}$ is a linear map from the vector space of polynomials of degree $\leq n$ to itself.

Lemma 3.20. *$\hat{\hat{A}}(z) = q^n A(z)$. In other words, if*

$$\sum_{i=0}^{n} a_i (1-z)^i (1+(q-1)z)^{n-i} = \sum_{i=0}^{n} b_i z^i,$$

then

$$\sum_{i=0}^{n} b_i (1-z)^i (1+(q-1)z)^{n-i} = q^n \sum_{i=0}^{n} a_i z^i.$$

Proof. Clearly $\hat{A}(z) = (1+(q-1)z)^n A(\rho(z))$ with $\rho(z) = \frac{1-z}{1+(q-1)z}$. Now $\rho(\rho(z)) = z$, so

$$\hat{\hat{A}}(z) = (1+(q-1)z)^n \hat{A}(\rho(z)) = (1+(q-1)z)^n (1+(q-1)\rho(z))^n A(z) = q^n A(z).$$

□

Lemma 3.21. *Suppose that $q \neq 0$. Let $I \subseteq \{0,1,\ldots,n\}$ and $a_i \in \mathbb{C}$ for $i \in I$. Suppose that not all a_i vanish. Then at least one of the coefficients of $z^0, z^1, \ldots, z^{|I|-1}$ in $\sum_{i\in I} a_i (1-z)^i (1+(q-1)z)^{n-i}$ does not vanish.*

Proof. Suppose wrong. Then

$$\sum_{i\in I} a_i (1-z)^i (1+(q-1)z)^{n-i} = \sum_{k=m}^{n} b_k z^k$$

with $m = |I|$ and $b_k \in \mathbb{C}$. Apply the MacWilliams transformation to obtain

$$q^n \sum_{i\in I} a_i z^i = \sum_{k=m}^{n} b_k (1-z)^k (1+(q-1)z)^{n-k}.$$

We see that 1 is an at least m–fold zero of $\sum_{i\in I} a_i z^i$, so 0 is a root of multiplicity $\geq m$ of

$$\sum_{i\in I} a_i (z+1)^i = \sum_{k=0}^{n} (z^k \sum_{i\in I} \binom{i}{k} a_i).$$

We obtain

$$\sum_{i\in I} \binom{i}{k} a_i = 0 \text{ for all } 0 \leq k \leq m-1.$$

So the $m \times m$ matrix with entry $\binom{i}{k}$ in position (k,i) is singular. Therefore there exist a non-zero sequence $c_0, c_1, \ldots, c_{m-1}$ with

$$\sum_{k=0}^{m-1} \binom{i}{k} c_k = 0 \text{ for all } i \in I.$$

Consider the polynomial

$$f(x) = \sum_{k=0}^{m-1} c_k \binom{x}{k},$$

where $\binom{x}{k} = x(x-1)\cdots(x-k+1)/k!$. Then $f(x)$ has degree at most $m-1$. On the other hand, $f(x)$ has the m roots $i \in I$. Hence $f(x)$ is the zero polynomial. But the polynomials $\binom{x}{k}$ have pairwise distinct degrees, so they are linearly independent. This implies $c_k = 0$ for all k, a contradiction. □

Theorem 3.22. *Let $C \subseteq F^n$ be a code with distance enumerator $B(z) = \sum_{i=0}^n B_i z^i$ and $|F| = q \geq 2$. Let r be the number of non-zero coefficients of $z, z^2, \ldots, z^n$ in*

$$\sum_{i=0}^{n} B_i (1-z)^i (1+(q-1)z)^{n-i}.$$

Then $r(C) \leq r$.

Proof. Pick $\boldsymbol{x} \in F^n$. Again, let $A_i(\boldsymbol{x})$ be the number of $\boldsymbol{c} \in C$ with $d(\boldsymbol{x}, \boldsymbol{c}) = i$. Define $\hat{A}_i(\boldsymbol{x})$ via

$$\sum_{i=0}^{n} \hat{A}_i(\boldsymbol{x}) z^i = \sum_{i=0}^{n} A_i(\boldsymbol{x}) (1-z)^i (1+(q-1)z)^{n-i}.$$

Then, by Lemma 3.20,

$$\sum_{i=0}^{n} A_i(\boldsymbol{x}) z^i = \frac{1}{q^n} \sum_{i=0}^{n} \hat{A}_i(\boldsymbol{x}) (1-z)^i (1+(q-1)z)^{n-i}.$$

By Lemma 3.15, at most r of the numbers $\hat{A}_i(\boldsymbol{x})$, $1 \leq i \leq r$, are non-zero. So including $\hat{A}_0(\boldsymbol{x})$, at most $r+1$ of these numbers are non-zero. Of course, not all $\hat{A}_i(\boldsymbol{x})$ vanish. So we may apply Lemma 3.21 which tells us that at least one of the coefficients $A_0(\boldsymbol{x}), A_1(\boldsymbol{x}), \ldots, A_r(\boldsymbol{x})$ does not vanish. But that means that there is a code word $\boldsymbol{c} \in C$ with $d(\boldsymbol{x}, \boldsymbol{c}) \leq r$. As $\boldsymbol{x}$ was arbitrary, the claim follows. □

Corollary 3.23. *Let $C \leq \mathbb{F}_q^n$ be a linear code, and let r be the number of different weights which appear in the dual code $C^\perp$. Then $r(C) \leq r$.*

Proof. By MacWilliams Theorem 3.9, the sequence of the coefficients of $\sum_{i=0}^n B_i (1-z)^i (1+(q-1)z)^{n-i}$ is, up to a constant factor $\neq 0$, the weight distribution of the dual code $C^\perp$. The claim then follows from the previous theorem. □

3.6 Perfect Codes, Part 2 (Lloyd's Theorem)

As we have seen in Section 2.3, it has been impossible so far to use the sphere packing condition in order to determine the potential parameters of a perfect code. Using the techniques from the previous section, it is possible to prove a more technical, but yet more useful condition on the parameters of a perfect code.

This criterion is most easily expressed in terms of the Krawtchouk polynomials.

Definition 3.24. For $0 \leq k \leq n$ and $q \in \mathbb{C}$ let

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

be the k-th *Krawtchouk polynomial* over $\mathbb{C}$.

If q and n are fixed, the we simply write $K_k(x)$ instead of $K_k(x; n, q)$. Here $\binom{x}{j}$ is the polynomial $x(x-1)\dots(x-j+1)/j!$.

Lemma 3.25. *Suppose that $q \neq 0$. Then the following holds:*

(a) $K_k(x)$ has degree k, with leading term $\frac{(-q)^k}{k!}$.

(b) $(1-z)^i(1+(q-1)z)^{n-i} = \sum_{k=0}^{\infty} K_k(i) z^k$.

Proof. (a) As $\binom{x}{j}$ and $\binom{n-x}{k-j}$ have degrees j and $k-j$, we see that the degree of $K_k(x)$ is at most k. The claim then follows from checking that the coefficient of z^k does not vanish:

$$\sum_{j=0}^{k} (-1)^j \frac{1}{j!} \frac{(-1)^{k-j}}{(k-j)!} (q-1)^{k-j} = \frac{(-1)^k}{k!} \sum_{j=0}^{k} \binom{k}{j} (q-1)^{k-j} = \frac{(-q)^k}{k!}$$

(b) We compute

$$\begin{aligned}
\sum_{k=0}^{\infty} K_k(i) z^k &= \sum_{k=0}^{\infty} \sum_{j=0}^{k} (-1)^j \binom{i}{j} \binom{n-i}{k-j} (q-1)^{k-j} z^k \\
&= \sum_{j=0}^{\infty} (-1)^j \binom{i}{j} \sum_{k=j}^{\infty} \binom{n-i}{k-j} (q-1)^{k-j} z^k \\
&= \sum_{j=0}^{\infty} (-1)^j \binom{i}{j} \sum_{k=0}^{\infty} \binom{n-i}{k} (q-1)^{k} z^{k+j} \\
&= \sum_{j=0}^{\infty} (-1)^j \binom{i}{j} z^j (1+(q-1)z)^{n-i} \\
&= (1-z)^i (1+(q-1)z)^{n-i}.
\end{aligned}$$

□

An immediate corollary from part (b) and Lemma 3.6 is

Lemma 3.26. *Let F be the residue class ring $\mathbb{Z}/q\mathbb{Z}$ for some $q \geq 2$. Then $(F, +)$ has a character χ such that for all $n \in \mathbb{N}$, $\boldsymbol{u} \in F^n$ and $0 \leq k \leq n$ the following holds:*

$$\sum_{\substack{\boldsymbol{v} \in F^n \\ w(\boldsymbol{v})=k}} \chi(\langle \boldsymbol{u} | \boldsymbol{v} \rangle) = K_k(w(\boldsymbol{u})).$$

Lemma 3.27. *Let $C \subseteq F^n$ be a perfect code with $|F| = q$ and minimum distance $2e+1$. Set $\Psi_e(x) = \sum_{k=0}^{e} K_k(x)$. Let $B_0, \ldots, B_n$ be the distance distribution of C. Define $\hat{B}_i$ by*

$$\sum_{i=0}^{n} \hat{B}_i z^i = \sum_{i=0}^{n} B_i (1-z)^i (1+(q-1)z)^{n-i}.$$

Then

$$\Psi_e(i) \hat{B}_i = 0 \text{ for } 1 \leq i \leq n.$$

Proof. Let S be the set of $\boldsymbol{v} \in F^n$ with $w(\boldsymbol{v}) \leq e$. As C is perfect, every vector $\boldsymbol{v} \in F^n$ has a unique representation $\boldsymbol{v} = \boldsymbol{s} + \boldsymbol{c}$ with $\boldsymbol{s} \in S$, $\boldsymbol{c} \in C$. In particular, for any fixed $\boldsymbol{u} \in F^n$, we have

$$\sum_{\boldsymbol{v} \in F^n} \chi(\langle \boldsymbol{u} | \boldsymbol{v} \rangle) = \sum_{\boldsymbol{s} \in S, \boldsymbol{c} \in C} \chi(\langle \boldsymbol{u} | \boldsymbol{s} + \boldsymbol{c} \rangle) = \sum_{\boldsymbol{s} \in S, \boldsymbol{c} \in C} \chi(\langle \boldsymbol{u} | \boldsymbol{s} \rangle) \chi(\langle \boldsymbol{u} | \boldsymbol{c} \rangle) = \sum_{\boldsymbol{s} \in S} \chi(\langle \boldsymbol{u} | \boldsymbol{s} \rangle) \sum_{\boldsymbol{c} \in C} \chi(\langle \boldsymbol{u} | \boldsymbol{c} \rangle)$$

The previous Lemma shows that

$$\sum_{s\in S}\chi(\langle u|s\rangle)=\sum_{k=0}^{e}\sum_{\substack{v\in F^n\\ w(v)=k}}\chi(\langle u|v\rangle)=\Psi_e(w(u)).$$

From now on assume that $u\neq 0$. Setting $z=1$ in Lemma 3.6 yields

$$\sum_{v\in F^n}\chi(\langle u|v\rangle)=0.$$

We obtain

$$0=\Psi_e(w(u))\sum_{c\in C}\chi(\langle u|c\rangle), \tag{5}$$

The proof of Theorem 3.14 shows that

$$\hat{B}_i=\frac{1}{|C|}\sum_{\substack{v\in F^n\\ w(v)=i}}|\sum_{c\in C}\chi(\langle v|c\rangle)|^2.$$

Suppose that $\hat{B}_i\neq 0$ for $i\geq 1$. Then there is an element $u\in F^n$ with $w(u)=i$ and $\chi(\langle u|c\rangle)\neq 0$, hence $\Psi_e(w(u))=0$ by (5). □

Lemma 3.28. *$\sum_{k=0}^{e}K_k(x;n,q)=K_e(x-1;n-1,q)$ for all $n\geq 1$, $0\leq e\leq n$.*

Proof. Pick $1\leq i\leq n$. Using generating functions we see that

$$\begin{aligned}\sum_{e=0}^{\infty}(\sum_{k=0}^{e}K_k(i;n,q))z^e&=(1+z+z^2+\dots)\sum_{e=0}^{\infty}K_e(i;n,q)z^e\\&=\frac{1}{1-z}(1-z)^i(1+(q-1)z)^{n-i}\\&=(1-z)^{i-1}(1+(q-1)z)^{(n-1)-(i-1)}\\&=\sum_{e=0}^{\infty}K_e(i-1;n-1,q)z^e.\end{aligned}$$

We see that the polynomials $\sum_{k=0}^{e}K_k(x;n,q)$ and $K_e(x-1;n-1,q)$ coincide in $x=1,2,3,\dots,n$. Furthermore, by Lemma 3.25, they have the same leading term, so their difference has degree $\leq e-1$. But a polynomial of degree $\leq e-1$ with at least $n\geq e$ roots is 0, and the claim follows. □

Theorem 3.29 (LLoyd). *Let $C\subseteq F^n$ be a perfect code with $|F|=q$ and $d(C)=2e+1$. Then $K_e(x-1;n-1,q)$ has e distinct integral roots in the interval $[1;n]$.*

Proof. As C is perfect with minimum distance $2e+1$, we see that $r(C) = e$. By Theorem 3.22, at least e of the elements $\hat{B}_1, \hat{B}_2, \ldots, \hat{B}_n$ from Lemma 3.27 do not vanish. So, Lemma 3.27 tells us that at $\Psi_e(i) = 0$ for at least e integers $1 \leq i \leq n$. But $\Psi_e(x)$ is a polynomial of degree e, hence $\Psi_e(i) = 0$ for exactly e integers $1 \leq i \leq n$, and these are all the roots of Ψ_e. The claim then follows from Lemma 3.28, as $\Psi_e(x) = K_e(x-1; n-1, q)$. □

Remark 3.30. (a) The proof shows that if C is perfect with $d(C) = 2e+1$, then exactly e of the numbers $\hat{B}_1, \hat{B}_2, \ldots, \hat{B}_n$ do not vanish.

(b) It is known that any Krawtchouk polynomial $K_k(x; n, q)$ of degree k has k distinct real roots in the interval $[0; n]$. So the only extra information provided by Lloyd's Theorem is that these roots are integers.

Using another identity of Krawtchouk values, we can state Lloyd's Theorem in another form.

Lemma 3.31. *Set $K_k(x) = K_k(x; n, q)$. Then*

$$(q-1)^i \binom{n}{i} K_k(i) = (q-1)^k \binom{n}{k} K_i(k)$$

for all $0 \leq i, k \leq n$.

Proof. This follows directly from the definition of the Krawtchouk polynomials and the identity $\binom{a}{b} = \frac{a!}{b!(a-b)!}$. Note that in evaluating $K_k(i) = \sum_{j=0}^{k} (-1)^j \binom{i}{j} \binom{n-i}{k-j} (q-1)^{k-j}$, it suffices to sum from $j = 0$ to the minimum of k and i. □

Corollary 3.32. *Let $C \subseteq F^n$ be a perfect code with $|F| = q$ and $d(C) = 2e+1$. Then there are exactly e indices $0 \leq i \leq n-1$ such that the coefficient of z^i in $(1-z)^e(1+(q-1)z)^{n-1-e}$ vanishes.*

Proof. By Lloyd's Theorem, $K_e(i-1; n, q) = 0$ for exactly e integers $1 \leq i \leq n$. The previous lemma shows that $K_e(i-1; n, q) = 0$ if and only if $K_{i-1}(e; n, q) = 0$, and the claim follows. □

3.6.1 Application for $e = 1$

Let $C \subseteq F^n$ be a perfect code with $d(C) = 3$, so $e = 1$ in Lloyd's Theorem. We have

$$\frac{-1}{q} K_1(x-1; n-1, q) = x - n + \frac{n-1}{q}. \tag{6}$$

So the necessary condition in Lloyd's Theorem shows that q divides $n-1$. This condition cannot be obtained from the sphere packing condition, which says that

$$1+n(q-1) \text{ divides } q^n. \tag{7}$$

It is not hard to show that there are infinitely many pairs (q,n) for which (6) holds, but (7) does not hold; and that there are also infinitely many pairs, for which (7) holds, but (6) does not hold. So in a sense these two conditions are independent from each other.

It is unknown if there exist perfect codes C with $d(C) = 3 < n$ for which q is not a prime power. Unfortunately, there are infinitely many pairs (q,n) for which (6) and (7) hold, yet q is not a prime power. For example $n = q+1$ fulfills both conditions, but $q = 6$ is the only non prime power for which the non-existence of a perfect $(q+1, q^{q-1}, 3)$–code is known (see Problem ???).

If q is a prime power, then (7) implies (6): From Problem ??? we know that (7) implies $1+n(q-1) = q^k$ for some $k \geq 1$. Then $n-1 = \frac{q^k-q}{q-1} = q\frac{q^{k-1}-1}{q-1}$, and we get (6).

3.6.2 Application for $e = 2$

In this case $d(C) = 5$, it is again not known if there are non-trivial perfect codes for which q is not a prime power. If however q is a prime power, then a combination of the sphere packing condition and Lloyd's theorem allows us to describe the possible parameters of a perfect code.

Theorem 3.33. *Let $C \subseteq F^n$ be a perfect code with $|F| = q$ and $d(C) = 5 \leq n$. Suppose that q is a prime power. Then either $q = 2$ and $n = 5$, or $q = 3$ and $n = 11$.*

Proof. The sphere packing condition

$$1+n(q-1)+\frac{n(n-1)}{2}(q-1)^2 \text{ divides } q^n$$

together with Problem ??? shows that

$$1+n(q-1)+\frac{n(n-1)}{2}(q-1)^2 = q^k \tag{8}$$

for some $k \geq 2$. We compute

$$\frac{2}{q^2}K_2(x-1; n-1, q) = x^2 + (1-2n+\frac{2n-4}{q})x + c_0$$

with $c_0 = \frac{2}{q^2}(1+n(q-1)+\frac{n(n-1)}{2}(q-1)^2) = 2q^{k-2}$. So Lloyd's Theorem says that

$$x^2 + (1-2n+\frac{2n-4}{q})x + 2q^{k-2} \tag{9}$$

has positive integral roots $x_1 \geq 1$ and $x_2 \geq 1$. Write $q = p^f$ with $f \in \mathbb{N}$ and p a prime. As $x_1 x_2 = 2q^{k-2} = 2p^{(k-2)f}$, we obtain $x_1 = 2p^a$ and $x_2 = p^b$ for some $a, b \geq 0$. Note that $x_1 \neq x_2$ by Lloyd's theorem. We see that either $x_1/x_2 \geq 2$ or $x_1/x_2 \leq 1/2$ unless $x_1 = 2 \cdot 3^a$, $x_2 = 3^{a+1}$. The function $z \mapsto z + 1/z$ is convex for $z > 0$, with a minimum in $z = 1$. This shows that

$$\frac{x_1}{x_2} + \frac{x_2}{x_1} \text{ is } \begin{cases} \geq \frac{5}{2}, & \text{if } \{x_1, x_2\} \neq \{2 \cdot 3^a, 3^{a+1}\} \\ = \frac{13}{6}, & \text{if } \{x_1, x_2\} = \{2 \cdot 3^a, 3^{a+1}\}. \end{cases} \tag{10}$$

Note that

$$\begin{aligned} \frac{x_1}{x_2} + \frac{x_2}{x_1} &= \frac{x_1^2 + x_2^2}{x_1 x_2} \\ &= \frac{(x_1 + x_2)^2 - 2x_1 x_2}{x_1 x_2} \\ &= \frac{(x_1 + x_2)^2}{x_1 x_2} - 2 \\ &= \frac{(1 - 2n + \frac{2n-4}{q})^2}{\frac{2}{q^2}(1 + n(q-1) + \frac{n(n-1)}{2}(q-1)^2)} - 2 \end{aligned}$$

We look at the first case $\frac{x_1}{x_2} + \frac{x_2}{x_1} \geq \frac{5}{2}$. After some calculation we arrive at

$$-3(nq - n + q - 7)(n-2)(q-1) \geq 0.$$

Now $n \geq 5$ and $q \geq 2$, which yields

$$n(q-1) + q \leq 7,$$

hence $n = 5$, $q = 2$.
It remains to look at the case $p = 3$, $\frac{x_1}{x_2} + \frac{x_2}{x_1} = \frac{2}{3} + \frac{3}{2}$. This yields the equation

$$(n^2 q - n^2 - nq - 21n - 6q + 42)(q-1) - 4 = 0.$$

So $q - 1$ divides 4. But q is a power of 3, hence $q = 3$. This yields $4(n-1)(n-11) = 0$, hence $n = 11$ and we are done.
ToDo: Do also the non prime power cases from Section 2.3. □

3.6.3 Application for $e = 3$

The case $e = 3$ and $|F| = q$ being a prime power is, up to a slightly lengthy computation, even easier than the case $e = 2$.

Theorem 3.34. *Let $C \subseteq F^n$ be a perfect code with $|F| = q$ and $d(C) = 7 \leq n$. Suppose that q is a prime power. Then $q = 2$ and $n = 7$ or $n = 23$.*

Proof. We have

$$F(x) = \frac{-K_3(x-1; n-1, q)}{q^3} = x^3 + c_2x^2 + (3n^2 - 6n + 2 - \frac{3(2nq - n - q + 2)(n-3)}{q^2})x + c_0$$

with complicated coefficients c_2, c_0 which we leave to the reader to compute.
Set $a = n - 1 - \frac{n-3}{q}$ and $b = n - \frac{n-2}{q}$. Then

$$\begin{array}{rcrcl} q^3F(a) & = & (n-3)(q-1)(q-2) & \geq & 0 \\ q^3F(b) & = & -(n-2)(2q-1)(q-1) & < & 0 \end{array}$$

First suppose that $q \neq 2$. Then $F(a) > 0$, $F(b) < 0$, so by the intermediate value theorem, there is a real γ with $F(\gamma) = 0$ and $a < \gamma < b$. By Lloyd's Theorem, we have $\gamma \in \mathbb{Z}$. As $0 < b - a = 1 - \frac{1}{q} < 1$, neither a nor b can be integral. So q divides neither $n-3$, nor $n-2$. On the other hand, the coefficients of $F(x)$ need to be integers, so q^2 divides $3(2nq - n - q + 2)(n-3)$. Recall that q is a prime power $q = p^f$. As p^f does not divide $n-3$, we get that p^{f+1} divides $3(2nq - n - q + 2)$, so $q = p^f$ divides $2nq - n - q + 2$, and then q divides $n-2$, a contradiction.
Thus we know that $q = 2$. Here the Lloyd condition is of little help. However, the sphere packing condition shows that $1 + n + \binom{n}{2} + \binom{n}{3} = \frac{(n+1)(n^2-n+6)}{6}$ is a power of 2, a case handled in Problem ???. □

Remark 3.35. (a) The parameters $q = 3$, $n = 11$ and $q = 2$, $n = 23$ in the previous two theorems indeed give rise to perfect codes, the ternary Golay code (see Section 6.2) and the binary Golay code (see Section 6.1), respectively.

(b) The preceding theorems can be extended to any $e \geq 2$, besides the two Golay codes no non-trivial perfect codes show up. This classification result was a joint effort by several authors in the years 1970–1973. See [vL75] for a nice historical survey, and [vL99] for a readable proof in the case $q = 2$, $e \geq 2$.

(c) If one does not require q to be a prime power anymore, it is much more difficult to show that there are no new perfect codes for $e \geq 3$. (As remarked already, the cases $e = 1$ and $e = 2$ are wide open.) In [Bes83] Best treated the cases $e \geq 3$ except for $e = 6$ and $e = 8$. Hong settled the remaining cases $e = 6$ and $e = 8$ in [Hon84].

3.7 Selfdual Codes

Let $C \leq \mathbb{F}_q^n$ be a linear code. Recall that the dual code $C^\perp$ is defined as the set of $\boldsymbol{u} \in F^n$ with $\langle \boldsymbol{u}|\boldsymbol{c}\rangle = 0$ for all $\boldsymbol{c} = 0$. The code C is called *self-dual* if $C = C^\perp$. As $\dim C + \dim C^\perp = n$, we see that self-dual codes can exist only in even dimensions $n = 2m$, and then $\dim C = m$.

If $C \leq \mathbb{F}_2^n$ is a binary self-dual code, then in particular $\langle \boldsymbol{c}|\boldsymbol{c}\rangle = 0$ for all $\boldsymbol{c} \in C$, so c has even weight. As each $\boldsymbol{c} \in C$ has even weight, we conclude that $(1\,1\,\ldots\,1) \in C^\perp = C$. We summarize these two easy observations:

Lemma 3.36. *Let $C \leq \mathbb{F}_2^n$ be a binary self-dual code. Then*

(a) each $c \in C$ has even weight, and

(b) $(1\,1\,\ldots\,1) \in C$.

It has become customary to call a code *doubly-even* if all weights of the code words are divisible by 4. Many interesting self-dual binary codes are doubly even. For later use we state another easy lemma:

Lemma 3.37. *Let $C \leq \mathbb{F}_2^n$ be a binary linear code with $C \leq C^\perp$. Suppose that C is generated by elements of weight divisible by 4. Then C is doubly-even.*

Proof. It suffices to show that if $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}_2^n$ have weights divisible by 4 and $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = 0$, then the weight of $\boldsymbol{u} + \boldsymbol{v}$ is divisible by 4 too.
Let c be the number of positions where $\boldsymbol{u}$ and $\boldsymbol{v}$ have the entry 1. Let a and b be the number of further 1's in $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively. Then 4 divides $w(\boldsymbol{u}) = a + c$ and $w(\boldsymbol{v}) = b + c$, so 4 divides $a + b + 2c = w(\boldsymbol{u} + \boldsymbol{v}) + 2c$. As $\langle \boldsymbol{u}|\boldsymbol{v}\rangle = 0$, we have that c is even, and the claim follows. □

Let $C \leq \mathbb{F}_2^n$ be a binary doubly-even self-dual code. As $(1\,1\,\ldots\,1) \in C$ has weight n, we get $4 \mid n$. A surprising consequence of the MacWilliams identity is that even stronger $8|n$.

Theorem 3.38 (Gleason)**.** *Let $C \leq \mathbb{F}_2^n$ be a binary doubly-even self-dual code. Then 8 divides n.*

Proof. Write $n = 2m$, so $|C| = 2^m$. Let $A(z)$ be the weight enumerator of C. As $C = C^\perp$, MacWilliams' Theorem 3.9 tells us that

$$A(z) = \frac{(1+z)^n}{2^m} A\left(\frac{1-z}{1+z}\right).$$

Set $I = \sqrt{-1}$. As $I^4 = 1$ and all the coefficients of z^i in $A(z)$ vanish if $4 \nmid i$, we

have $A(z) = A(Iz)$. Replacing z with $\frac{1-z}{1+z}$ yields $A(\frac{1-z}{1+z}) = A(I\frac{1-z}{1+z})$, hence

$$A(z) = \frac{(1+z)^n}{2^m} A(\mu(z)) \text{ with } \mu(z) = I\frac{1-z}{1+z}. \tag{11}$$

Note that

$$\mu(0) = I, \quad \mu(I) = 1, \quad \mu(1) = 0.$$

So setting $z = 0$, $z = I$, and $z = 1$ in (11) gives

$$A(0) = \frac{1}{2^m} A(I), \quad A(I) = \frac{(1+I)^n}{2^m} A(1), \quad A(1) = \frac{2^n}{2^m} A(0). \tag{12}$$

Note that $A(0) = 1 \neq 0$, so also $A(I), A(1) \neq 0$. Multiplying the three equations from (12), and noting that $n = 2m$, yields

$$1 = \frac{(1+I)^{2m}}{2^m} = \frac{(2I)^m}{2^m} = I^m.$$

However, the multiplicative order of I is 4, hence $4 \mid m$ and therefore $8 \mid n$. □

Remark 3.39. The usual proof of Gleason's Theorem uses stronger techniques from invariant theory, and gives a finer result about the shape of the weight enumerator of a binary doubly-even self-dual code. See e.g. [Wil99]. We quickly sketch the main ideas:
Let $C \leq \mathbb{F}_q^n$ be a linear code. It is convenient to work with a bivariate version of the weight enumerator $A(x,y) = \sum_{i=0}^n A_i x^i y^{n-i}$. The MacWilliams identity then gives the weight enumerator $\frac{1}{|C|} \sum_{i=0}^n A_i (y-x)^i (y+(q-1)x)^{n-i}$ for the dual code. So if C is self-dual, then

$$A(x,y) = A\left(\frac{y-x}{\sqrt{q}}, \frac{y+(q-1)x}{\sqrt{q}}\right) = A((x\,y)U)$$

with

$$U = \frac{1}{\sqrt{q}} \begin{pmatrix} -1 & q-1 \\ 1 & 1 \end{pmatrix}.$$

If we assume in addition that $q = 2$ and C is doubly even, then $A(x,y) = A(Ix,y)$, which we can express as

$$A(x,y) = A((x\,y)V)$$

with

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

Let $G \leq \mathrm{GL}_2(\mathbb{C})$ be the group generated by U and V. Then $A(x,y) = A((x\,y)g)$ for all $g \in G$. One computes that $(UV)^3 = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta \end{pmatrix}$ where ζ is a primitive 8th root of unity. So $A(x,y) = A(\zeta x, \zeta y)$, which again implies $8 \mid n$. This explains the background of our proof from above, which bypasses these considerations.
Of course knowing that 8 divides n is a rather coarse information about the weight enumerator $A(x,y)$. Using invariant theory one can describe the subring R of $\mathbb{Q}[x,y]$ of those polynomials which are fixed by all $g \in G$. One obtains $R = \mathbb{Q}[F_8, F_{24}]$ with $F_8 = x^8 + 14x^4y^4 + y^8$, $F_{24} = x^4y^4(x^4 - y^4)^4$. Furthermore, F_8 and F_{24} are algebraically independent, so there is a unique polynomial $P(x,y) \in \mathbb{Q}[x,y]$ with $A(x,y) = P(F_8, F_{24})$.
In order to prove these results, one has to study more carefully the group G which happens to have order 192.

3.8 Problems

Problem 3.40. Let $C \leq \mathbb{F}_q^n$ be a linear $(n, k, n-k+1)$ code, so C is an MDS code. Show that the dual code $C^\perp$ is an $(n, n-k, k+1)$ code, so it is an MDS code too.
(Hint: Let M be the matrix whose rows are a basis of C. If $C^\perp$ contains an element $v \neq 0$ with $w(v) \leq k$, then there are k columns in M which are linearly dependent. Show that there is a non-trivial linear combination of the rows of M which is 0 in k positions.)

Problem 3.41. Let G be a finite abelian group of order > 1. Show, without using the structure theorem for finite abelian groups, that G has a non-trivial character.

4 Comparing the Bounds

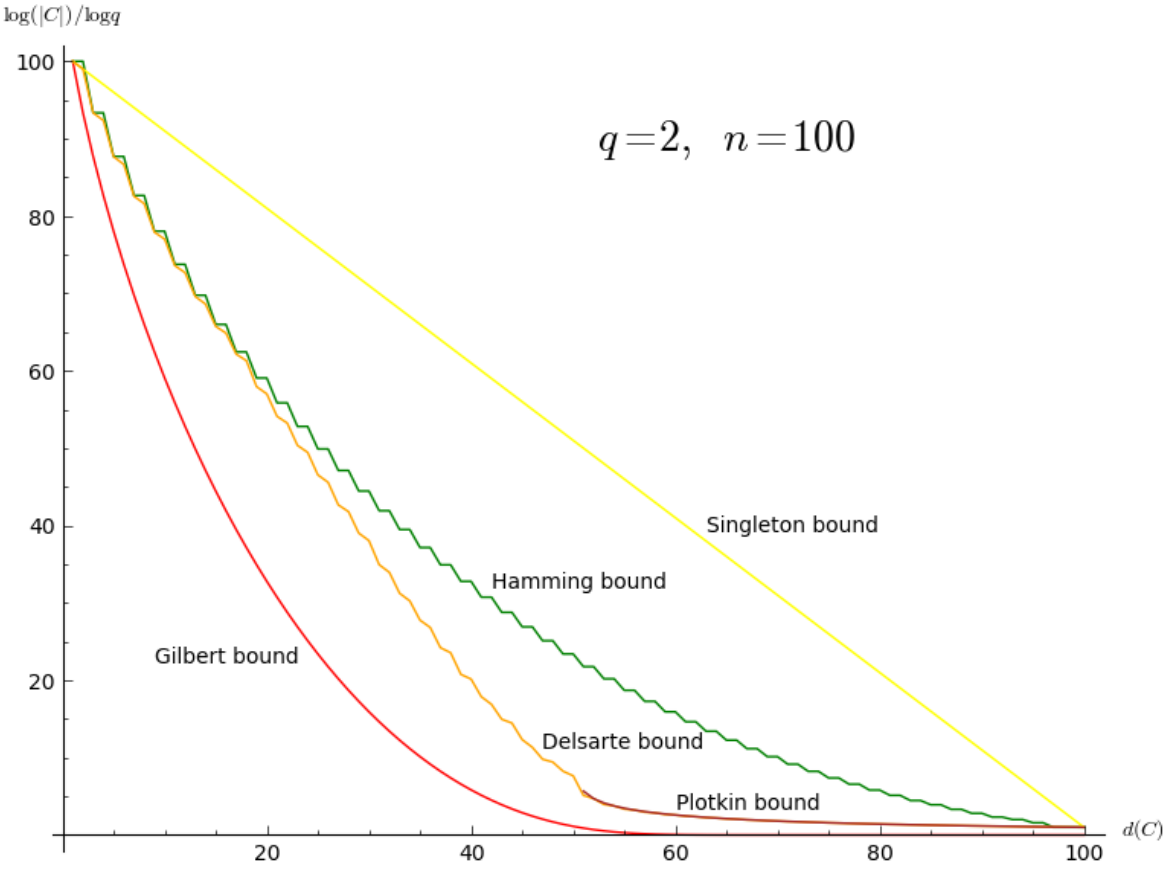
One can show that the linear programming bound is always at least as good as the Singleton bound, the Hamming bound, and the Plotkin bound. For q small compared to n, the Hamming bound is far stronger than the Singleton bound. However, if $q \geq n$ is a prime power, then we found the Reed-Solomon codes C for any $1 \leq d(C) \leq n$ which match the Singleton bound. That indicates that if q is large compared to n, the quality of the Singleton bound improves. In particular, for $q \geq n$ a prime power, it is the best bound.

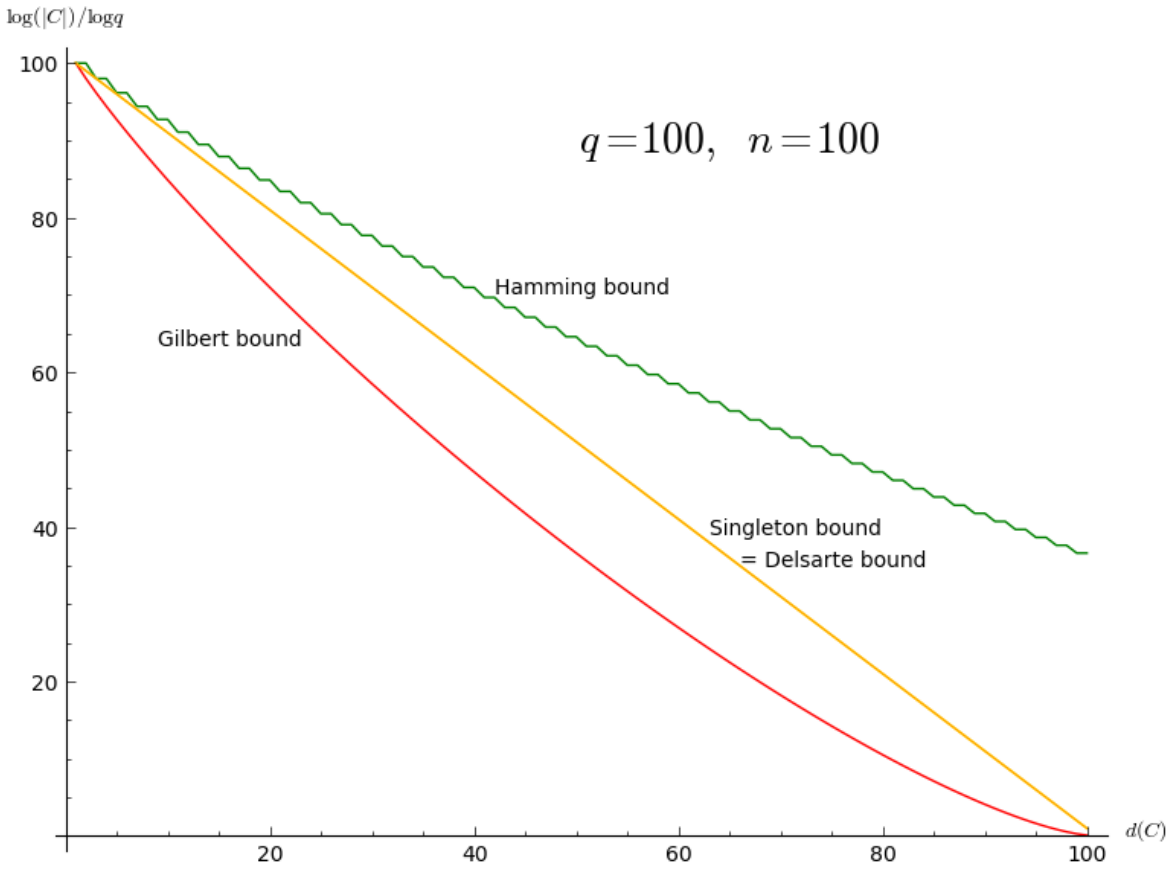
Fix q. If $f(n,d,q)$ is one of the upper or lower bounds by Singleton, Hamming, Plotkin (in the allowed range), Gilbert, or Varshamov, then one can show the following: Fix $0 \leq \delta \leq 1$, and let $n_i \in \mathbb{N}$, d_i be sequences with $0 \leq d_i \leq n_i$, $\lim n_i = \infty$ and $\lim \frac{d_i}{n_i} = \delta$, then $\lim \frac{\log f(n_i, d_i, q)}{n \log q}$ exists and is a continuous function $\hat{f}(\delta)$ depending

on δ. So asymptotically $\hat{f}$ bounds the relative rate of a code in terms of the relative minimum distance.

On can compute these functions $\hat{f}$ for the bounds mentioned above. For the Delsarte bound, where $f(n,d,q)$ is the bound given by solving the associated linear program, $\hat{f}(\delta)$ is not known.

The following displays two examples of the quality of the bounds. We have omitted the Griesmer bound, because it applies only to linear codes. A refinement of the proof of the Plotkin bound yields the Elias bound, which is often the next best bound after the Delsarte bound. As we have not discussed this bound, we don't display it here.





5 Application to Projective Planes

5.1 Codes Generated from Integral Square Matrices

Theorem 5.1. *Let $A \in \mathbb{Z}^{n\times n}$ be a matrix with $\det A \neq 0$, and p be a prime. Let $C \leq \mathbb{F}_p^n$ be the code generated by the rows of A, and let m be maximal with $p^m \mid \det A$. Then $\dim C \geq n - m$.*

Proof. For $X \in \mathbb{Z}^{r\times s}$ let $\bar{X}$ be the natural image of X in $\mathbb{F}_p^{r\times s}$. Conversely, for $\bar{X} \in \mathbb{F}_p^{r\times s}$ let $X \in \mathbb{Z}^{r\times s}$ be any matrix which is mapped to $\bar{X}$.
Set $k = \dim C$. Then $\dim C^\perp = n - k$. Let $\bar{M}$ be the $(n-k) \times n$ matrix whose rows are a basis of $C^\perp$. As $\bar{M}$ has rank $n - k$, we may assume after permuting coordinates in $C^\perp$ and C that the first $n - k$ columns of $\bar{M}$ are linearly independent. Write $\bar{M} = \begin{pmatrix}\bar{U} & \bar{V}\end{pmatrix}$ with a $U \in \mathbb{Z}^{(n-k)\times(n-k)}$ and $V \in \mathbb{Z}^{(n-k)\times k}$. Let $I_k \in \mathbb{Z}^{k\times k}$ be

the identity matrix, and set

$$N = \begin{pmatrix} U & V \\ 0 & I_k \end{pmatrix} \in \mathbb{Z}^{n \times n}.$$

By definition of duality of codes, the first $n-k$ columns in $\bar{A}\bar{N}^t = \overline{AN^t}$ vanish, so the entries in the first $n-k$ columns of AN^t are all divisible by p, so p^{n-k} divides $\det AN^t = \det A \det N$. On the other hand, $\det N = \det U$. As $\bar{U}$ is invertible, we have $\det \bar{U} \neq 0$, so $p \nmid \det U$. We obtain $p^{n-k} \mid \det A$, so $n-k \leq m$ and therefore $\dim C = k \geq n-m$. □

5.2 Incidence Matrices of Finite Projective Planes

A *projective plane* consists of two sets P and L and a subset I of $P \times L$ subject to the following conditions:

(a) For and $p_1, p_2 \in P$ with $p_1 \neq p_2$ there is exactly one $\ell \in L$ with $(p_1, \ell), (p_2, \ell) \in I$.

(b) For and $\ell_1, \ell_2 \in L$ with $\ell_1 \neq \ell_2$ there is exactly one $p \in L$ with $(p, \ell_1), (p, \ell_2) \in I$.

(c) There are four distinct elements Q in P such that for any $\ell \in L$ there are at most 2 elements $q \in Q$ with $(q, \ell) \in I$.

We can make this abstract definition a little more concrete: Identify each $\ell \in L$ with the subset of those $p \in P$ such that $(p, \ell) \in I$, and call these subsets *lines*. Call the elements in P *points*. Then (a) says that any two distinct points lie on exactly one line, (b) says that any two distinct lines have exactly one common point, and (c) says that there is a non-degenerate quadrangle.

From now on we take this point of view, so L is a set of subsets of P with the properties from above.

Lemma 5.2. *Let (P, L) be a projective plane, and ℓ_1, ℓ_2 be two distinct lines. Then there is a point which is neither on ℓ_1, nor on ℓ_2.*

Proof. Let Q be a quadrangle according to (c). If each of it points is on ℓ_1 or ℓ_2, then $q_1, q_2 \in \ell_1$ and $q_3, q_4 \in \ell_2$ with $Q = \{q_1, q_2, q_3, q_4\}$. But then the intersection point of the line through q_1 and q_3 with the line through q_2 and q_4 is neither on ℓ_1, nor on ℓ_2. □

Theorem 5.3. *Each line of a projective plane contains the same number of points.*

Proof. Let ℓ_1 and ℓ_2 be two distinct lines, and p be a point not contained in these lines. For $q \in \ell_1$ let $\phi(q)$ be the intersection of the line ℓ_2 with the line through q and p. The properties of a projective plane then show that ϕ is bijective. □

A *finite projective plane* is a projective plane with P a finite set. As L is a set of subsets of P, we see that L is finite as well. Let $n+1$ be the common number of points on a line of a finite projective plane. We call n the *order* of the projective plane.

Theorem 5.4. *Let (P, L) be a finite projective plane of order n. Then there are $n+1$ lines through each point, and $|P| = |L| = n^2 + n + 1$.*

Proof. Let p be a point. Not all sides of a quadrangle can contain p, so there is a line ℓ not containing p. For each line ℓ' with $p \in \ell'$ let $\phi(\ell')$ be the intersection of ℓ' with l. Then ϕ is a bijection, so there are $n+1$ lines through p.
The $n+1$ lines through p are, up to p, disjoint, and contain n points besides p. Also, every point $\neq p$ is on exactly one line through p. So there are $(n+1)n+1 = n^2+n+1$ points.
The assertion about the number of lines either follows from the same argument, where the lines and points switch their roles, or by this counting argument: Let m the number of lines. We count the triples (p_1, p_2, ℓ) with $p_1, p_2 \in \ell$, $p_1 \neq p_2$ in two ways. One the one hand, this number equals $\binom{n^2+n+1}{2} = (n^2+n+1)\frac{n+1}{2}$. On the other, this number is $m\binom{n+1}{2}$, and the claim follows. □

Definition 5.5. Let (P, L) be a projective plane of order n. Set $N = n^2 + n + 1$. Number the points and lines $P = \{p_1, p_2, \ldots, p_N\}$ and $L = \{\ell_1, \ell_2, \ldots, \ell_N\}$. Let $A \in \mathbb{Z}^{N \times N}$ with entry 1 in position (i, j) if $p_i \in \ell_j$, and 0 otherwise.

Theorem 5.6. *Let A be the incidence matrix of a projective plane of order n. Then*

$$AA^t = A^tA = \begin{pmatrix} n+1 & 1 & \ldots & 1 \\ 1 & n+1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & n+1 \end{pmatrix},$$

a matrix with $n+1$ on the diagonal, and 1 elsewhere. Furthermore, each row and column of A contains the entry 1 exactly $n+1$ times.

Proof. This is just a translation of the properties we developed so far. □

5.3 A Special Case of the Bruck–Ryser Theorem

A major open problem in finite geometry is the question for which $n \in \mathbb{N}$ there is a projective plane of order n. Below we see that if n is a prime power, there is a plane of order n. On the other hand, no projective plane of order not a prime power has been found yet. Up to the combination of theoretical and computational arguments in the proof of the non-existence of planes of order 10 (see Section 5.3.1), the only other non-existence result is

Theorem 5.7 (Bruck-Ryser 1949, [BR49]). *Suppose that $n \equiv 1$ or $2 \pmod 4$, and that $n \neq x^2 + y^2$ for all $x, y \in \mathbb{Z}$. Then there is no projective plane of order n.*

We will not prove this result here, but see Remark 5.11. Instead, we will show how our results in coding theory can be used to get a special case. Namely suppose that $n \equiv 6 \pmod 8$. Then n cannot be sum of two squares, and the Bruck–Ryser Theorem shows that there is no plane of order n. Below we give a direct proof for this.

But before doing so, we quickly show that projective planes exist for each prime power order:

Theorem 5.8. *For every prime power $q > 1$ there is a projective plane of order q.*

Proof. Let V be the vector space $\mathbb{F}_q^3$. Let P and L be the set of 1-dimensional and 2-dimensional subspaces, respectively. We say that $p \in P$ is on the line $\ell \in L$, if $p \in \ell$. Then one easily verifies the properties of a projective plane. Note that the four 1-dimensional spaces $\langle(1\,0\,0)\rangle$, $\langle(0\,1\,0)\rangle$, $\langle(0\,0\,1)\rangle$, $\langle(1\,1\,1)\rangle$ have the property that no 3 of them are contained in a 2–dimensional subspace, so we get a quadrangle.
The number of 1-dimensional subspaces in a 2-dimensional subspace is $(q^2 - 1)/(q-1) = q+1$, so our plane has order q. □

Lemma 5.9. *Let $A \in \mathbb{Z}^{N \times N}$ be the incidence matrix of a projective plane of order n, where $N = n^2 + n + 1$. Define*

$$B = \begin{pmatrix} & & 1 \\ & A & \vdots \\ & & 1 \\ 1 & \dots & 1 \quad n+1 \end{pmatrix} \text{ and the diagonal matrix } D = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix},$$

both in $\mathbb{Z}^{(N+1)\times(N+1)}$. Then $BDB^t = nD$.

Proof. Note that

$$DB^tD^{-1} = \begin{pmatrix} & & & -1 \\ & A^t & & \vdots \\ & & & -1 \\ -1 & \dots & -1 & n+1 \end{pmatrix}.$$

The claim then follows from the form of AA^t given in Theorem 5.6, and the fact that each row and column of A contains the entry 1 exactly $n+1$ times. □

Theorem 5.10. *Let $n \in \mathbb{N}$ with $n \equiv 6 \pmod 8$. Then there is no finite projective plane of order n.*

Proof. (Assmus, see [HW06, Satz 7.4.16]) Suppose that $n \equiv 6 \pmod 8$, and that A is the incidence matrix of a projective plane of order n. Set $N = n^2 + n + 1$. The assumption on n gives

$$N + 1 \equiv 4 \pmod 8. \tag{13}$$

Build $B \in \mathbb{Z}^{(N+1)\times(N+1)}$ as in the previous lemma. Taking determinants in $BDB^t = nD$ shows $(\det B)^2 = n^{N+1}$, hence $\det B = \pm n^{(N+1)/2}$. Let m be maximal with $2^m \mid \det B$. The assumption on n shows that $4 \nmid n$, hence $m = \frac{N+1}{2}$.
Let $\bar{X}$ be the image in $\mathbb{F}_2^{(N+1)\times(N+1)}$ of $X \in \mathbb{Z}^{(N+1)\times(N+1)}$, and $C \leq \mathbb{F}_2^{N+1}$ the binary code generated by the rows of $\bar{B}$. As $\bar{D}$ is the identity matrix and $2 \mid n$, we see that $\bar{B}\bar{B}^t = 0$. That means that $\langle \mathbf{b}_1 | \mathbf{b}_2 \rangle = 0$ for any two rows $\mathbf{b}_1, \mathbf{b}_2$ of B. So $C \leq C^\perp$ and therefore $\dim C \leq \frac{N+1}{2}$. On the other hand, $\dim C \geq N + 1 - m \geq N + 1 - \frac{N+1}{2} = \frac{N+1}{2}$ by Theorem 5.1. So we have equality everywhere, so C is self-dual.
Each row of $\bar{B}$ but the last one contains $n+2$ times the entry 1, while the last row contains $N+1$ times the entry 1. So in each row, the number of 1s is divisible by 4. In particular, C is doubly-even by Lemma 3.37. But then $8 \mid N+1$ by Theorem 3.38, contrary to (13). □

Remark 5.11. All proofs of the Bruck–Ryser Theorem are based on Lemma 5.9 or minor modifications of it. Suppose that n is the order of a projective plane. In the language of quadratic forms, the lemma says that the quadratic forms described by the matrices D and nD are equivalent over the rationals. Using the easier direction of the Hasse–Minkowski Theorem one shows (under the assumption on n modulo 4) that each prime p with $p \equiv 3 \pmod 4$ occurs in n with an even (possibly 0) multiplicity. From elementary number theory, we get that $n = x^2 + y^2$ for $x, y \in \mathbb{Z}$.
There is a modification which avoids the use of the Hasse–Minkowski Theorem, and instead uses Lagrange's Theorem that each positive integer n is a sum

of 4 squares. Using this and elementary calculations, one arrives at $n = u^2 + v^2$ with *rational* u, v. Again, elementary number theory implies that $n = x^2 + y^2$ for suitable integers x, y. That is the usual proof in text books, see e.g. [Hal98]. One might wonder if the coding–theoretic proof of the special case $n \equiv 6 \pmod 8$ can be extended to a full proof of the Bruck–Ryser Theorem. Indeed, while somewhat technical, this is possible. See [Lan83].

5.3.1 A Remark on Projective Planes of Order 10

The Bruck–Ryser Theorem does not apply to the order $n = 10$, as $10 = 1^2 + 3^2$. Still, much of the proof of Theorem 5.10 carries over. Let C' and C be the binary codes generated by the rows of A and B, respectively. Then again C is doubly-even and self–dual. So the MacWilliams identity puts severe restrictions on the weight enumerator of C, and thus on C' too. However, one does not achieve a contradiction by theoretical arguments only. Instead massive computer calculations eventually show that such a code C' cannot arise from a plane of order 10. Up to today, there is no reproducible account of this non–existence proof – one has to believe in the correctness of poorly documented computer calculations. See [Hal98] and [KÖ06] for theoretical results, and [Lam91] for an overview of the computation.

6 The Golay Codes

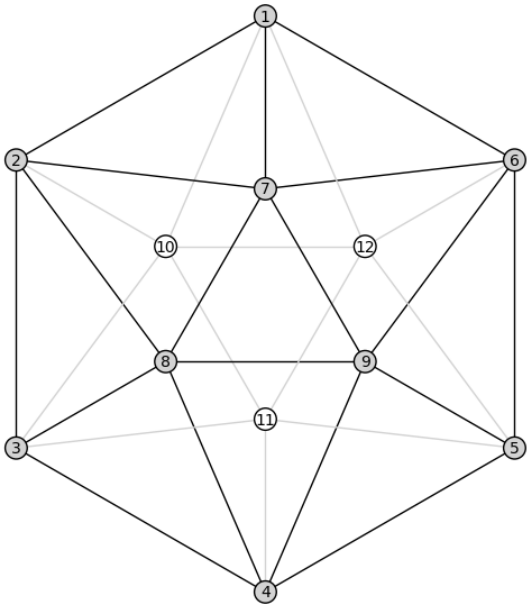
6.1 The Binary Golay Code

Previously we had seen from the sphere packing condition and Lloyd's Theorem that there is a potential perfect binary $[23, 12, 7]$–code. In this section we will construct such a code, the *binary Golay code* G_{23}. While there are direct constructions of this code (for instance as a cyclic code or as a quadratic residue code), we prefer to obtain this code from the equally important *extended binary Golay code* G_{24}, which is a $[24, 12, 8]$–code.

In doing so, we number the vertices of an icosahedron from 1 to 12. Next we define a matrix $A \in \mathbb{F}_2^{12\times 12}$ via

$$A_{ij} = \begin{cases} 0, & i \neq j \text{ and } i, j \text{ are connected by an edge} \\ 1, & \textit{otherwise} \end{cases}$$

With the numbering as in



we obtain the first four rows of the matrix A, where we drop the 0s for better readability:

$$A = \begin{pmatrix} 1 & & 1 & 1 & 1 & & & 1 & 1 & & 1 & \\ & 1 & & 1 & 1 & 1 & & & 1 & & 1 & 1 \\ 1 & & 1 & & 1 & 1 & 1 & & 1 & & & 1 \\ 1 & 1 & & 1 & & 1 & 1 & & & 1 & & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Lemma 6.1. *(a) $w(a) = 7$ for each row a of A.*

(b) Any two distinct row of A differ in 6 or 10 positions.

(c) $\langle a|a'\rangle = 0$ for any two distinct rows of A, and $AA^t = A^2 = I$, where I denotes the identity matrix.

Proof. (a) This holds, because every vertex is not connected by an edge with $12 - 5 = 7$ vertices. In particular, $\langle a|a\rangle = 1$ for each row a of A.
(b) Any pair of distinct vertices can be mapped by a symmetry of the icosahedron to the pair $(1,2)$, $(1,3)$, or $(1,4)$. So we only have to compare the first row of A with the second, third and fourth row.
(c) As in (b), we see that a and a' have a common 1 in 2 or 4 positions, hence $\langle a|a'\rangle = 0$. We get $AA^t = I$ from this and (a), and then $A^2 = I$ because A is symmetric. □

Set

$$B = \begin{pmatrix} I & A \end{pmatrix} \in \mathbb{F}_2^{12\times 24}$$

and let $G_{24} = C \leq \mathbb{F}_2^{24}$ be the binary code generated by the rows of B.

Theorem 6.2. *The extended binary Golay code G_{24} is a self–dual, doubly-even $[24, 12, 8]$–code.*

Proof. Clearly B has rank 12, so $\dim C = 12$. From $AA^t = I$ we get that the rows of B are pairwise orthogonal, so $C \leq C^\perp$. But $\dim C + \dim C^\perp = 24$ and $\dim C = 12$, hence $C = C^\perp$. Each row of B has weight 8, so C is doubly-even by Lemma 3.37. Of course $d(C) \leq 8$. Suppose that $d(C) \leq 4$. As the weight of the code words of C are divisible by 4, we get $d(C) = 4$. Pick $c \in C$ with $w(c) = 4$. The elements of C have the form $uB = (u\, uA)$, where u runs through $\mathbb{F}_2^{12}$. Hence $c = (u\, uA)$ for a suitable u. We claim that $c' = (uA\, u) \in C$. Set $v = uA$. From $(v\, vA) \in C$ and $A^2 = I$ we obtain $c' = (uA\, u) \in C$. From $4 = w(c) = w(u) + w(uA)$ we obtain that $w(u) \leq 2$ and $w(uA) \leq 2$. So by probably passing to c', we may and do assume that $w(u) \leq 2$.
If $w(u) = 0$ then $u = 0$, hence $c = 0$, a contradiction. If $w(u) = 1$, then c is a row of B, hence $w(c) = 8$, again contrary to $w(c) = 4$. Finally assume that $w(u) = 2$. Then c is the sum of two distinct rows of B. So there are two distinct rows a and a' of A with $2 = w(a + a') = d(a, a')$. However, $d(a, a') \geq 6$ by the lemma. □

One can easily determine the weight distribution A_i of $C = G_{24}$ without computing all the 2^{12} code words:

Theorem 6.3. *The weight enumerator of the extended binary Golay code G_{24} is $1 + 759z^8 + 2576z^{12} + 759z^{16} + z^{24}$.*

Proof. Let A_i be the number of code words of weight i. As C is doubly-even, we have $A_i = 0$ unless $4 \mid i$. Furthermore, the all 1 vector is in C (it is the sum of all rows of B, or see Lemma 3.36), so $A_i = A_{24-i}$. Next, $A_4 = 0$ as $d(G_{24}) = 8$. Thus the weight enumerator has the form $A(z) = 1 + A_8(z^8 + z^{16}) + A_{12}z^{12} + z^{24}$. From $2^{12} = A(1)$ we get

$$2A_8 + A_{12} = 2^{12} - 2. \tag{14}$$

The MacWilliams identity and $A(z) = A^\perp(z)$ gives

$$2^{12}A(z) = 2^{12}A^\perp(z) = (1+z)^{24}A(\frac{1-z}{1+z}).$$

Pick $\omega \in \mathbb{C}$ with $\omega^2 + \omega + 1 = 0$. Then $\omega^3 = 1$, so $\omega^8 + \omega^{16} = \omega^2 + \omega = -1$, hence

$$A(\omega) = 2 - A_8 + A_{12}.$$

Set $\beta = \frac{1-\omega}{1+\omega}$. Then $\beta^2 = \frac{1-2\omega+\omega^2-4(1+\omega+\omega^2)}{(1+\omega)^2} = \frac{-3(1+\omega)^2}{(1+\omega)^2} = -3$. Furthermore,

$(1+\omega)^{24} = (-\omega^2)^{24} = 1$. This yields

$$(1+\omega)^{24} A(\frac{1-\omega}{1+\omega}) = A(\beta) = (1+3^{12}) + (3^4+3^8)A_8 + 3^6 A_{12},$$

hence

$$2^{12}(2 - A_8 + A_{12}) = (1+3^{12}) + (3^4+3^8)A_8 + 3^6 A_{12}.$$

This, together with (14) eventually yields $A_8 = 759$, $A_{12} = 2576$, and the claim follows. □

Remark 6.4. The extended binary Golay code G_{24} has many exciting connections to other areas of mathematics. For instance, we can consider its automorphism group G, which is the group of those permutations of the coordinates which map code words to code words. Using our definition of the code, it is easy to see that there are at least 120 automorphisms of the code, see Problem ???. Actually, the group G of all automorphisms is quite big. It acts 5–fold transitively on the coordinate positions, and the pointwise stabilizer of 5 points has order 48. Thus $|G| = 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 48 = 244823040$. Furthermore, one can show that G is simple, it is called the Mathieu group M_{24}, one of the 26 sporadic simple groups.
The code G_{24} is connected to other fascinating combinatorial objects, the *Steiner systems*. Let $2 \leq t < k < n$ be integers. A Steiner system $S(t,k,n)$ is a set S of k–element subsets of a set M of size n, such that each t–element subset of M is a subset of exactly one set from S. For instance, a projective plane of order n gives an $S(2, n+1, n^2+n+1)$ Steiner system, because any two distinct points are contained in exactly one line (which has size $n+1$).
It is an open problem if there are Steiner systems $S(t,k,n)$ with $t \geq 6$. However, there are a few with $t = 5$. The code G_{24} allows to construct a $S(5,8,24)$ Steiner system: For each of the 759 code words $c = (c_1\, c_2\, \ldots, c_{24})$ of weight 8 associate the 8–element subset of $M = \{1, 2, \ldots, 24\}$ consisting of the indices i with $c_i = 1$. It is not hard to see that any 5–element subset of M is contained in exactly one of these 8–element subsets. See Problem ???.
Another interesting construction, a slight modification of the preimage in $\mathbb{Z}^{24}$ of G_{24} under the natural map $\mathbb{Z}^{24} \to \mathbb{F}_2^{24}$, gives the *Leech lattice*. It is an exceptionally dense lattice in $\mathbb{R}^{24}$ whose automorphism group yields further interesting sporadic simple groups, the three Conway groups.
For this and many further connections see [CS99].

Lemma 6.5. *Let C be a binary $[24, 12, 8]$ code, and C' be the code obtained form C by deleting a coordinate position. Then C' is a perfect $[23, 12, 7]$–code.*

Proof. Deleting one coordinate position lowers the weight by at most 1, hence $d(C') \geq 7$, so the balls of radius 3 around the code words of C' are disjoint. However, $2^{12}(1 + \binom{23}{1} + \binom{23}{2} + \binom{23}{3}) = 2^{23}$, so these balls disjointly cover $\mathbb{F}_2^{23}$, hence C' is perfect with covering radius 3, and therefore $d(C') = 2 \cdot 3 + 1 = 7$. □

Let G_{23} be the code derived from G_{24} by deleting a coordinate position. The lemma then tells us that G_{23} is a perfect $[23, 12, 7]$-code. This code is called a *binary Golay code*. It looks as if G_{23} would seriously depend on which position gets deleted. However, one can actually show that up to isomorphism there is only one binary $[23, 12, 7]$-code. For this reason one speaks of *the* binary Golay code. As to the uniqueness of the $[24, 12, 8]$-code, or even stronger of the $(24, 2^{12}, 8)$-code, see Problem ???.

Remark 6.6. As all the code words in G_{24} have even weight, one can recover G_{24} from G_{23} by appending a parity check bit. Historically G_{23} was known and used before G_{24}. This explains the *extended* in the name of G_{24}.
Besides the beautiful mathematical properties, the binary Golay codes (G_{24} more than G_{23}) had been used in the past for the NASA deep space missions, and is used nowadays in radio communication.

Remark 6.7. There is the following uniqueness result for the extended binary Golay code: Let $C \subseteq \mathbb{F}_2^{24}$ be a not necessarily linear code with $|C| = 2^{12}$ and $d(C) = 8$. Assume that $0 \in C$. Then, up to a permutation of coordinates, $C = G_{24}$. The essential step in showing this is to prove that C is linear. See Problem ??? for this.

6.2 The Ternary Golay Code

In this section we construct the two ternary Golay codes.

Lemma 6.8. *Set*

$$A = \begin{pmatrix} 0 & 1 & -1 & -1 & 1 \\ 1 & 0 & 1 & -1 & -1 \\ -1 & 1 & 0 & 1 & -1 \\ -1 & -1 & 1 & 0 & 1 \\ 1 & -1 & -1 & 1 & 0 \end{pmatrix} \in \mathbb{F}_3^{5\times 5} \text{ and } C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ & & & & & -1 \\ & & & & & -1 \\ & & A & & & -1 \\ & & & & & -1 \\ & & & & & -1 \end{pmatrix} \in \mathbb{F}_3^{6\times 6}.$$

Then $CC^t = -I_6$.

Proof. Let a_i be the i-th row of A. Then $\langle a_i|a_j\rangle = 1$ if $i = j$, and $\langle a_i|a_j\rangle = -1$ for $i \neq j$. As each row in A arises from rotating the preceding row to the right by one step, we may assume $i = 1$ in verifying the cases. Together with the observation that the entries of each row of A sum up to 0, we get that $CC^t = -I_6$. □

Remark 6.9. The matrix A is the modulo 3 reduction of a so-called *Paley matrix*, which in general is defined as follows: Index the rows and columns of A with the finite field $\mathbb{F}_q$. For $i, j \in \mathbb{F}_q$ define the entry $A_{i,j}$ of $A \in \mathbb{Z}^{q\times q}$ in position (i, j) by

$$A_{i,j} = \begin{cases} 0, & \text{if } i = j \\ 1, & \text{if } 0 \neq i - j \text{ is a square in } \mathbb{F}_q \\ -1, & \text{if } i - j \text{ is not a square in } \mathbb{F}_q. \end{cases}$$

Then $AA^t = qI_q - J_q$, where I_q is the all 1 matrix.
Extending to the matrix C as above (still over $\mathbb{Z}$), one gets $CC^t = qI_{q+1}$. Such a matrix C, with 0 on the diagonal, -1 or 1 off the diagonal, and CC^t a multiple of the identity matrix, is called a *conference matrix*. For more about this see [vLW01].

With the matrix C from the lemma, we define

$$B = \begin{pmatrix} I_6 & C \end{pmatrix} \in \mathbb{F}_3^{6\times 12}$$

and let $G_{12} \leq \mathbb{F}_3^{12}$ be the ternary code generated by the rows of B. We call G_{12} the *extended ternary Golay code*.

Theorem 6.10. *The extended ternary Golay code G_{12} is a self–dual $[12, 6, 6]$–code.*

Proof. Length and dimension of G_{12} are clear. The previous lemma implies $BB^t = 0$, so the rows of B generate a self-dual code. From $\langle c|c\rangle = 0$ for each $c \in G_{12}$ and $i^2 = 1$ for each $0 \neq i \in \mathbb{F}_3$ we see that the weights of the code words of G_{12} are divisible by 3. So in order to show that $d(C) = 6$ we need to show that there are no words of weight 3. Suppose that there is $c \in G_{12}$ with $w(c) = 3$. Write $c = (u\, uC)$ with $u \in \mathbb{F}_3^6$. Then $3 = w(u) + w(uC)$. Clearly $w(u) = 0$ and $w(u) = 1$ are not possible. Suppose that $w(u) = 2$. Then $w(uC) = 1$, and uC is a linear combination of two rows of C. But this implies $w(uC) \geq 2$, because, each row of C has a unique 0 entry in a different position. Finally assume $w(u) = 3$, then $w(uC) = 0$, so $uC = 0$. But $CC^t = -I_6$ implies that C is invertible, hence $u = 0$, a contradiction again. □

Let G_{11} be the code obtained from deleting a coordinate position of G_{12}. As in the binary case, we see that G_{11} is a perfect $[11, 6, 5]$-code, the *ternary Golay code*. Again, one can show that up to isomorphism there is only one $(11, 3^6, 5)$-code and only one

$(12, 3^6, 6)$-code. This, however, is much more difficult to prove than in the binary case.

6.3 Covering Codes and Virtakallio's Discovery of the Ternary Golay Code

It was long believed that the ternary Golay code was first constructed by Golay in 1949. However, it turned out much later that the first description of this code, essentially by listing $243 = 3^5$ codewords and explaining how to obtain the remaining $2 \cdot 3^5$ words, was given by the Finnish football pool specialist Juhani Virtakallio in 1947 in the soccer magazine Veikkaaja! Every football match has three possible outcomes, a home win, a home loss, or a draw. In the 1940s, there was a Finnish football pool where one had to bet the outcome of 12 matches. As the correct prediction of 10 matches was rewarded with a reasonable high prize, Virtakallio tried to find a system with a minimal number of bets such that at least one of them predicts correctly the outcome of at least 10 matches. He assumed that for one of the matches the teams are of such a different strength such that one can be sure to make a correct guess. So that one needs a system for the remaining 11 games. Phrased mathematically, one has the following problem after identifying the three possible outcomes of a match with the elements of $\mathbb{F}_3$. Find a set $C \subseteq \mathbb{F}_3^{11}$ of minimal cardinality such that for each $\boldsymbol{v} \in \mathbb{F}_3^{11}$ there is a $\boldsymbol{c} \in C$ with $d(\boldsymbol{c}, \boldsymbol{v}) \leq 2$. In other hands, the balls of radius 2 around the elements $\boldsymbol{c} \in C$ should cover all of $\mathbb{F}_3^{11}$. As the volume of such a ball is $\sum_{i=0}^{2} \binom{11}{i} 2^i = 243 = 3^5$, one gets $|C|3^5 \geq 3^{11}$, hence $|C| \geq 3^6$.

Of course, this bound is sharp if and only if C is a perfect $(11, 3^6, 5)$-code. Surprisingly, Virtakallio found a system with $|C| = 3^6$, the ternary Golay code! The way he described his code indicates that he didn't know much about mathematics, which makes it even more mysterious that he could find this code!

This example is a special case of a so-called *covering code*, arising from a question which is kind of opposite to the error correcting codes: Set $|F| = q$, and let $K_q(n, e)$ be the minimal size of a set $C \subseteq F^n$ such that the balls of radius e around $\boldsymbol{c} \in C$ cover F^n. So the perfectness of the ternary Golay codes gives $K_3(11, 2) = 3^6 = 729$. Virtakallio's achievement is even more surprising considering that as of 2011, the values $K_3(n, 2)$ were unknown for $6 \leq n \leq 10$. For instance, one only knows that $15 \leq K_3(6, 2) \leq 17$ (see [Kér11]).

See [CHLL97] for a whole book devoted to covering codes. It also contains more details about Virtakallio's exciting discovery.

7 Goppa Codes

to be written

7.1 Classical Goppa Codes

to be written

7.2 Algebraic Curves

to be written

7.3 Geometric Goppa Codes

to be written

8 Appendix: Some Tools from Algebra

to be written

Index

References

[Bes83] M. R. Best, *A contribution to the nonexistence of perfect codes*, Ph.D. thesis, Mathematisch Centrum, Amsterdam (1983).

[BR49] R. H. Bruck, H. J. Ryser, *The nonexistence of certain finite projective planes*, Canadian J. Math. (1949), **1**, 88–93.

[CHLL97] G. Cohen, I. Honkala, S. Litsyn, A. Lobstein, *Covering Codes*, vol. 54 of *North-Holland Mathematical Library*, North-Holland Publishing Co., Amsterdam (1997).

[CS99] J. H. Conway, N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, vol. 290 of *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, New York, 3rd edn. (1999), With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.

[Hal98] M. Hall, Jr., *Combinatorial Theory*, Wiley Classics Library, John Wiley & Sons Inc., New York, 2nd edn. (1998), A Wiley-Interscience Publication.

[Hon84] Y. Hong, *On the nonexistence of unknown perfect* 6- *and* 8-*codes in Hamming schemes* $H(n, q)$ *with* q *arbitrary*, Osaka J. Math. (1984), **21**(3), 687–700.

[HW06] B. Huppert, W. Willems, *Lineare Algebra.*, Wiesbaden: Teubner (2006).

[Kér11] G. Kéri, *Tables for bounds on covering codes*, http://www.sztaki.hu/~keri/codes/index.htm (2011).

[KÖ06] P. Kaski, P. R. J. Östergård, *Classification Algorithms for Codes and Designs*, vol. 15 of *Algorithms and Computation in Mathematics*, Springer-Verlag, Berlin (2006), With 1 DVD-ROM (Windows, Macintosh and UNIX).

[Lam91] C. W. H. Lam, *The search for a finite projective plane of order* 10, Amer. Math. Monthly (1991), **98**(4), 305–318.

[Lan83] E. S. Lander, *Symmetric Designs: An Algebraic Approach*, vol. 74 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge (1983).

[vL75] J. H. van Lint, *A survey of perfect codes*, Rocky Mountain J. Math. (1975), **5**, 199–224.

[vL99] J. H. van Lint, *Introduction to Coding Theory*, vol. 86 of *Graduate Texts in Mathematics*, Springer-Verlag, Berlin, 3rd edn. (1999).

[vLW01] J. H. van Lint, R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 2nd edn. (2001).

[Wil99] W. Willems, *Codierungstheorie.*, Berlin: de Gruyter (1999).