MTH 328: THEORY OF ERROR CORRECTING CODES

Abstract. Class notes for the course MTH328: Coding theory during January semester of 2020. The course policy can be found on the course webpage
www.iiserpune.ac.in/~kaipa/teaching/MTH328.

1. Introduction

When information is transmitted over a communication channel, the information gets corrupted by noise entering the channel. For example:

(1) the data transmitted by Chandrayaan through outer space (which is the communication channel here), gets corrupted by thermal noise
(2) The information written on a CD (compact disk) is corrupted by finger-marks and scratches.
(3) Similarly, the data transmitted from your cell phone to the cell-tower also encounters many sources of noise.

Therefore, in modern communication systems we cannot afford to transmit the raw data. What we transmit is the data with some redundancy built in. The purpose of adding redundancy is to be able to correct the errors picked up during transmission. We take the following point of view.

- There is a set $\mathcal{F}$ of finite size q (which we call the alphabet). The input and output to the channel is $\mathcal{F}^n$ (the Cartesian product of $\mathcal{F}$ with itself n times) – equivalently, the set of words (strings) of length n with letters from $\mathcal{F}$.
- There is a set of messages $\mathcal{M}$. We denote the size of this set by M.
- The encoder is a one-to-one function from $\mathcal{M} \to \mathcal{F}^n$. The image of this map is the code C, and its elements are called codewords. Clearly C also has size M.
- There is a decoder at the receiving end of the channel: when a codeword c is transmitted and a corrupted word c' is received, the job of the decoder is to give its best estimate for what c could have been.

The purpose of encoding is to use the 'larger space' $\mathcal{F}^n$ to spread-out the message words. In order to quantify this, we need a metric $d : \mathcal{F}^n \times \mathcal{F}^n \to \mathbb{R}$. With respect to such a metric d, let

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

The number d_C is called the *minimum distance* of C and the associated number $t_C := (d_C - 1)/2$ is called the *packing radius* of C. Consider the collection of M balls of radius t_C centered at each codeword. Check using triangle inequality for the metric d (see HW) that these balls are disjoint. Next, suppose the nature of the channel and noise is such that the distance between a transmitted codeword c and the received (corrupted version) word c' is at most t_C, then the decoder has no trouble in identifying c correctly, because c' is in the ball centred at c, and the balls are disjoint. This is the *main promise of error-correcting codes.*

The most common metric on $\mathcal{F}^n$ is the *Hamming metric* (named after Richard Hamming, who also created the first error correcting codes). In this course we will use only the Hamming metric on $\mathcal{F}^n$.

$$d(x,y) = \#\{1 \leq i \leq n : x_i \neq y_i\} \tag{1}$$

There are three conditions for a function $d : X \times X \to \mathbb{R}$ to be a metric on a set X: i) $d(x,y) \geq 0$ with equality if and only of $x = y$, ii) $d(x,y) = d(y,x)$, and iii) the triangle inequality $d(x,y) + d(y,z) \geq d(x,z)$. The properties i) and ii) are obvious for the Hamming metric. See HW for the triangle inequality.

We give some basic examples of codes. When we say that C is an $[n, M, d]$ code or $[n, M, d]_q$ code we mean that C is a subset of $\mathcal{F}^n$ of size M, and the alphabet size is $|\mathcal{F}| = q$. The minimum distance of the code is d. The parameter n is called the *length* of the code.

(1) Suppose the set of messages is $\mathcal{M} = \mathcal{F}$. Let t be a natural number and $C \subset \mathcal{F}^{2t+1}$. The encoding is $a \mapsto aa \dots a$ (the letter a is repeated $2t+1$ times). If at most t bit errors happen then the decoder can decode with 100% accuracy. This is a *repetition code* with parameters $[2t+1, q, 2t+1]_q$. This is the simplest example of how building redundancy into the data allows us to correct errors. Clearly it is also very inefficient, in the sense that the rate of transmission is $2t+1$ times slower than transmitting raw data.

(2) Let $\mathcal{F} = \mathbb{F}_2$ the binary field (recall $\mathbb{F}_2 = \{0, 1\}$ with $1 \cdot 1 = 1$ and $1 + 1 = 0$ etc). Let $\mathcal{M} = \mathcal{F}^7$. Let $\mathcal{C} \subset \mathcal{F}^8$ defined by the encoding $(x_1, \dots, x_7) \mapsto (x_1, \dots, x_7, x_1 + x_2 + \dots + x_7)$. This is the *ASCII code* with parameters $[8, 128, 2]_2$. Its minimum distance is 2 (See HW 1). It cannot correct any errors but it can detect 1 error : if exactly one of the symbols $x_1, \dots, x_8$ gets corrupted (i.e. 0 changes to 1 or vice-versa) then the recieved word will not be a codeword.

(3) The ISBN code is a string $x_1 x_2 \dots x_9 x_{10}$ (with some hyphens which we will ignore). Here $x_1, \dots, x_9 \in \{0, 1, \dots, 9\}$ and $x_{10} \in \{0, 1, \dots, 9, 10\}$. The letter x_{10} is determined in terms of $x_1, \dots, x_9$ by the formula $x_{10} = \sum_{i=1}^{9} i x_i \bmod 11$. Also, if x_{10} works out to be 10 mod 11 then the letter X is used for x_{10} instead of $x_{10} = 10$. For example $0550 - 10206 - X$ is a valid ISBN code, because $\sum_{i=1}^{9} i x_i = 98 = 10 \bmod 11$. The ISBN code has minimum distance 2. It has parameters $[10, 10^9, 2]_1 1$. It is designed to detect one error. It can also detect a double-error created by the transposition of two digits (See HW).

(4) Let $\mathcal{F} = \mathbb{F}_q$ be a finite field. (We will see that the size of a finite field is a prime power. Moreover for each prime power there is a unique field of that size up-to isomorphism). Let the message space be $\mathcal{M} = \mathcal{F}^k$, i.e. the vector space $\mathbb{F}_q^k$. We will represent message words as $(a_0, a_1, \dots, a_{k-1})$. Let us fix n distinct elements $x_1, x_2, \dots, x_n$ of $\mathbb{F}_q$. The code $C \subset \mathbb{F}_q^n$ is defined by the encoder

$$(a_0, \dots, a_{k-1}) \mapsto (p(x_1), p(x_2), \dots, p(x_n)),$$

where $p(X)$ is the polynomial $p(X) = a_{k-1}X^{k-1} + a_{k-2}X^{k-2} + \dots + a_1 X + a_0$. In other words we take the components of the message as the coefficients of a degree at most $k-1$ polynomial and evaluate the polynomial at $x_1, \dots, x_n$. This is a *Reed-Solomon* code which is used in CD/DVDs as well as data storage at Facebook. The minimum distance of this code is $n - k + 1$ (see HW). Thus the parameters of this code are $[n, q^k, n-k+1]_q$.

2. Assignment 1

(1) Show that ball of radius $r \leq t_C$ centred at codewords are pairwise disjoint, but if $r > t_C$ we can find codewords c, c' such that balls of radius r centred at these codewords intersect. Thus t_C is the largest integer r such that balls of radius r centred at each $c \in C$, are pairwise disjoint.

(2) Show that the minimum distance of the four codes (repetition, ASCII, ISBN, Reed-Solomon,) is as asserted.

(3) Let x, y, z be words of length n from some alphabet $\mathcal{F}$. Consider the triangle inequality for the Hamming metric $d(x, y) \leq d(x, z) + d(z, y)$.
a) Prove the triangle inequality.
b)* Given x, y describe and count the number of words z for which equality holds in the triangle inequality.

(4) Show that the ISBN code can detect one error, and that it can also detect a double-error created by the transposition of two digits.

Determine x in the following valid ISBNs.

$$0 - 13 - 1x9139 - 9, \qquad 0 - 02 - 32xx80 - 0$$

(5) (optional)* Let $\mathbb{F} = \mathbb{F}_2$ the binary field. Let x, y, and z be words in $\mathbb{F}^n$ that form an "equilateral triangle," that is, $d(x, y) = d(y, z) = d(z, x) = 2t$. Show that there is exactly one word $v \in \mathbb{F}^n$ such that $d(x, v) = d(y, v) = d(z, v) = t$.

3. Introduction continued

Recall that the code C is the image of an injective encoding map $\imath : \mathcal{M} \hookrightarrow \mathcal{F}^n$. The set $\mathcal{M}$ of messages and hence the code C has size M. The ambient space $\mathcal{F}^n$ has size q^n. It is convenient to consider logarithmic versions of sizes: we say that the ambient space $\mathcal{F}^n$ has $\log_q q^n = n$ bits of information, and the set of messages requires $\log_q M$ bits of information. The *rate, also called information rate or transmission rate* of a $[n, M, d]_q$ code C is defined to $(\log_q M)/n$. It is a measure of efficiency of the code.

The *relative minimum distance* of C is d_C/n. Since C can correct $(d_C - 1)/2$ errors, it follows that the relative minimum distance d_C/n is a measure of the *error correcting capacity* of C.

A good code is one which has a high rate and high relative minimum distance. These two requirements are in general conflicting, in the sense that it is easy to achieve one of them high at the expense of lowering the other. We will study the trade-off between these two quantities, in some detail later in this course. Note that both the quantities, relative minimum distance and rate of transmission are numbers in the interval $[0, 1]$. To a code C we can assign a point $(d_C/n, \log_q M/n)$ in the unit square $[0, 1] \times [0, 1]$.

The size of a sphere of radius r centred at $x \in \mathcal{F}^n$ is clearly

$$S_q(n, r) = \binom{n}{r}(q-1)^r.$$

The size or volume of a ball of radius r centred at $x \in \mathcal{F}^n$ is therefore

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

By definition of a ball, it follows that (for fixed n and q) $V_q(n, r)$ is an increasing function of r for $0 \leq r \leq n$. Note that $V_q(n, r) = q^n$ for $r \geq n$. On the other hand $S_q(n, r)$ is not an increasing function of r on all of $0 \leq r \leq n$. For $0 \leq x \leq 1$, the fraction $(\log_q S_q(n, \lfloor xn \rfloor))/n$ has a limit $H_q(x)$ as $n \to \infty$. We will study this function $H_q : [0, 1] \to [0, 1]$ (known as the *Hilbert-Shannon* q-ary entropy function) later in this course. We just mention here that for $q = 2$ we have

$$H_2(x) = -\left(x \log_2(x) + (1-x)\log_2(1-x)\right).$$

The distance of a word $x \in \mathcal{F}^n$ from a code $C \subset \mathcal{F}^n$ is $d(x, C) = \min\{d(x, c) : c \in C\}$. The *covering radius* ρ_C of a code $C \subset \mathcal{F}^n$ is defined to be the largest possible value of $d(x, C)$:

$$\rho_C = \max\{d(x, C) : x \in \mathcal{F}^n\}.$$

It is easy to see that (see HW) ρ_C is the smallest integer r such that the union of balls of radius r centred at each codeword is the whole space $\mathcal{F}^n$. In particular

$$MV_q(n, \rho_C) \geq q^n.$$

Earlier we have seen that the collection of balls of radius t_C centred at each codeword are non-overlapping. In particular

$$MV_q(n, t_C) \leq q^n. \tag{2}$$

This is known as the *Sphere Packing bound* or *Hamming bound* on the size of a code. In terms of the trade-off between the rate and relative minimum distance of a code that we mentioned earlier, we can interpret this bound as saying that for a $[n, M, d]_q$ code C, the rate $n^{-1}\log_q M \leq 1 - \log_q(V_q(n, t_C))/n$.

Since $V_q(n, r)$ is a non-decreasing function of r, and

$$V_q(n, \rho_C) \geq q^n/M \geq V_q(n, t_C),$$

it follows that $t_C \leq \rho_C$. We say a code is *perfect* if $t_C = \rho_C$ or equivalently if $MV_q(n, t_C) = q^n$. For a perfect code d_C is always an odd integer (see HW).

We will now mention three interesting and hard/important problems in Coding theory.

- *The topic of perfect codes presents many interesting combinatorial problems, some of which are still unsolved.* For example it is a hard unsolved problem to determine if there exist perfect codes C with $t_C = 2$ over an alphabet of composite (non prime power) size q.

- Given a code $C \subset \mathcal{F}^n$, we saw that the collection of balls of diameter at most $d_C - 1$ centred at the codewords, are non-overlapping. This gave the Sphere-Packing bound (2). Similarly if we draw a 'box' around each codeword $c \in C$, where the box consists of all $x \in \mathcal{F}^n$ which agree with c in the first $n - d_C + 1$ coordinates, then this collection of boxes is non-overlapping: indeed if the boxes around c and c' have a common member x, then c and c' agree in the first $n - d_C + 1$ coordinates and hence $d(c, c') \leq d_C - 1$ contradicting the definition of d_C. Clearly each box has size $q^{d_C - 1}$. Therefore, the non-overlapping property of this collection of boxes gives the inequality $Mq^{d_C-1} \leq q^n$ or

$$M \leq q^{n-d_C+1}, \tag{3}$$

  known as the *Singleton bound* on the size of a code. Again, in terms of the trade-off between the rate and relative minimum distance of a code that we mentioned earlier, we can interpret this bound as saying that for a $[n, M, d]_q$ code C, the rate $n^{-1}\log_q M \leq 1 + 1/n + d_C/n$. Codes achieving the Singleton bound are known as *MDS codes* (short for Maximum Distance Separable Codes). *The topic of MDS codes opens some fascinating problems in combinatorics.*

  A code $C \subset \mathcal{F}^n$ is called linear is $\mathcal{F}$ is a finite field $\mathbb{F}_q$ of size q and C is a linear subspace of $\mathbb{F}_q^n$. If the dimension of this subspace is k then the size of the code is $M = q^k$. If $k < q$, there is no known examples of linear MDS code of size q^k and length $n > q + 1$. *A famous conjecture in coding theory/combinatorics/finite algebraic geometry* known as the *MDS conjecture* is that there is do not exist linear MDS codes of length $n > q+1$ if the size q^k satisfies $k < q$. (There are some exceptions if q is a power of 2).

- The trade-off between the rate and relative minimum distance of a code is captured by the quantity $A_q(n, d)$ defined to be the largest value of M for which there exists a $[n, M, d]_q$ code. In other words $\log_q A_q(n, d)/n$ is the largest rate possible for a code with minimum distance d. The problem of determining $A_q(n, d)$ for general n and d is not a tractable problem because of *combinatorial explosion.* A central problem in coding theory is to obtain

good upper and lower bounds on $A_q(n,d)$. There is a more tractable and highly interesting version of this problem: For a given relative minimum distance $x \in [0,1]$ we consider the sequence $(\log_q A_q(n, \lfloor xn \rfloor))/n$. While this sequence may not have a limit as $n \to \infty$, it does have a lim sup and lim inf because the sequence takes values in $[0,1]$ and hence is bounded. The asymptotic information rate function is defined to be:

$$\alpha_q(x) = \limsup_{n\to\infty} \frac{\log_q A_q(n, \lfloor xn \rfloor)}{n}.$$

The exact value of this function $\alpha_q : [0,1] \to [0,1]$ is not known. It is known to be continuous and monotone increasing, but we do not know if it is differentiable. Its graph is conjectured to be $\cup$-convex. *Again, it is an important and interesting problem in coding to obtain good upper and lower bounds for this function.* Many of the good bounds on $\alpha(x)$ are expressed in terms of the entropy function $H_q(x)$.

4. Hamming codes: coin weighing problems and IPL betting problem

The Coin Weighing Problem: There are 12 coins of which at most one of them is fake. The fake coin, if it exists, is either slightly lighter or slightly heavier than the standard weight. We have to decide a) if there is a fake coin, b) if yes, then identify it among the 12 coins, c) determine if it is lighter or heavier than the standard weight. Show that, we can use the balance 3 times to achieve this.

If we replace the number of coins from 12 to a general number n, then show that we need to use the balance $r = \lceil \log_3(2n+3) \rceil$ times.

The IPL Betting problem: Suppose there are t IPL matches in a tournament. The matches are expressed as Team A vs Team B. A bet consists of predicting the outcome of all the t matches as a word in $\mathcal{F}^t$ where $\mathcal{F} = \{0,1,2\}$. Here the outcomes $0, 1$ and 2 stand for a draw, Team A wins, and Team B wins respectively. A first prize in the betting goes to someone who predicts all the outcomes correctly. A second prize to someone who predicts all but one outcomes correctly. Let $f(t)$ denote the least number $f(t)$ of bets required to guarantee at least a second prize.
Find the value of $f(t)$ for values of t of the form $(3^r-1)/2$ for some integer $r \geq 2$; i.e. for $t = 4, 13, 40, 121, \ldots$.

The answer to both these problems is based on a family of perfect codes of minimum distance 3 over the finite field $\mathbb{F}_3$ (with 3 elements), known as the *Hamming codes*, which we next discuss.

For the betting problem, consider the space $\mathcal{F}^t$ consisting of words of length t from the alphabet $\mathcal{F} = \{0,1,2\}$. The result of the tournament is a word in $\mathcal{F}^t$. We assume every word of $\mathcal{F}^t$ is a possible outcome of the tournament. The bets placed by a person consists of a code $C \subset \mathcal{F}^t$. If this person wants to be guaranteed of at least a second prize, then we must have $d(x, C) \leq 1$ for each $x \in \mathcal{F}^t$. (For every possible outcome x, there should be a codeword $c \in C$ differing from x in at most one coordinate). Thus $\mathcal{F}^t$ is the union of balls of radius 1 centered at each codeword. Each such ball has size $V_3(t,1) = 1 + 2t$. If $|C| = M$, then the union of the balls has size at most $M(1+2t)$. Thus

$$3^t = |\mathcal{F}^t| \leq M(1+2t).$$

This gives $M \geq 3^t/(1+2t)$. In other words, the person cannot be guaranteed of at least a second prize if $M < 3^t/(1+2t)$. We now show that whenever $3^t/(1+2t)$ is an integer, there is indeed a code C of size $M = 3^t/(1+2t)$ with $d(x, C) \leq 1$ for all $x \in \mathcal{F}^t$. (The covering radius $\rho_C = 1$). Since 3 is a prime number $3^t/(1+2t)$ is an integer if and only if it is of the form 3^{t-r}, i.e. $t = (3^r-1)/2$ for some $r \leq t$.

We think of the alphabet $\mathcal{F} = \{0,1,2\}$ as the field $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$. Let S be the subset of $\mathbb{F}_3^r$ consisting of all nonzero vectors whose first nonzero coordinate is 1. Note that the set of nonzero vectors of $\mathbb{F}_3^r$ has size $3^r - 1$ and between the pair v and $-v$ exactly one of them has 1 for the value of the first nonzero coordinate. This shows that $|S| = (3^r-1)/2 = t$. Also note that each one dimensional subspace of $\mathbb{F}^3$ has 3 elements consisting of $\{0, v, -v\}$ for a unique $v \in S$. Thus we may also think of S as the set of one-dimensional subspaces of $\mathbb{F}_3^r$ (equivalently the set of lines through the origin of $\mathbb{F}_3^r$). Consider the matrix H of size $r \times (3^r-1)/2$ whose columns consist of some ordering of the S. We define $C \subset \mathbb{F}_3^t$ to be the null space of H, i.e. $C = \{y \in \mathbb{F}_3^t : Hy = 0\}$. Since the $r \times r$ identity matrix appears as a sub-matrix of H, we see that H has full rank r, and

hence $\dim(C) = t - r$ by the rank-nullity theorem. In particular $|C| = 3^{t-r}$ as required. In order to show that $d(x, C) \leq 1$ for all $x \in \mathbb{F}_3^t$, we will need the following lemma:

Lemma 1. *Let $\mathcal{C} \subset \mathbb{F}_q^n$ be a linear code given as the null space of a $r \times n$ matrix H of full rank r. Let $x \in \mathbb{F}_q^n$. Then*

$$d(x, \mathcal{C}) = min\{\mu : Hx \text{ can be expressed as a linear combination of } \mu \text{ columns of } H\}.$$

Proof. Let $x \in \mathbb{F}_q^n$, and let $c \in \mathcal{C}$. Consider $e = x - c$. Note that the distance of x from c is the number μ of nonzero entries of e. Since c is in the null space of H, we have $Hx = He$ is a linear combination of μ columns of H. Conversely if Hx is a linear combination of μ columns of H we can write $Hx = He$ where e has μ nonzero entries. It then follows that $c = x - e \in \mathcal{C}$ and $d(x, c) = \mu$. This establishes that $d(x, C)$ is the minimum number of columns of H such that Hx is in the span of these columns. □

Returning to the code $C \subset \mathbb{F}_3^t$, we note that for any $x \in \mathbb{F}_3^t$, either Hx or $-Hx$ is in S. Therefore, by the lemma above $d(x, C) \leq 1$ for all $x \in \mathbb{F}_3^t$.

Next, we discuss the coin weighing problem with n coins. Let the coins be labeled as $C_1, \ldots, C_n$. Each time we use the balance we have to put an equal number of coins on the left pan and the right pan of the balance. Consider the vector $(x_1, \ldots, x_n) \in \mathbb{F}_3^n$ where $x_j = -1$ if the coin C_j is put on the left pan, and $x_j = 1$ if C_j is put on the right pan. We put $x_j = 0$ if C_j is not weighed. Suppose we use the balance r times. Consider the matrix G of size $r \times n$ whose i-th row consists of the vector $(x_1, \ldots, x_n)$ associated with the i-th use of the balance. The reading of the balance can be expressed as an element of $y \in \mathbb{F}_3$: if the left pan is lighter we put $y = -1$, and if the right pan is lighter we $y = +1$ and if the pans are of equal weight, we put $y = 0$. Let $y = (y_1, \ldots, y_r) \in \mathbb{F}_3^r$ be the result of all the r uses of the balance. We now make an observation. If there is no fake coin then y is the zero vector. If C_j is fake and lighter, then y is the j-th column of G. If C_j is fake and heavier, then y is the j-th column of G multiplied by -1. Therefore, the coin weighing problem can be solved using the balance r times, provided that no column of G is the zero vector, and $v \neq \pm u$ for any two columns of G. In short no two columns of G are linearly dependent. We have already seen that the largest subset of $\mathbb{F}_3^r$ with the property that no two elements of S are linearly dependent is the set S above of size $(3^r - 1)/2$. However, the matrix H above does not have the property that each row sum of H is zero (which reflects the fact that we need to put an equal number of coins on each pan of the balance). A simple exercise shows that removing the column $(0, \ldots, 0, 1)$ from H gives a $r \times (3^r - 3)/2$ matrix G all whose row sums are zero. The property of H that no two columns are dependent is not disturbed. Thus the minimum number of uses of the balance for n coins is the least integer r such that $(3^r - 3)/2 \geq n$, i.e. $r = \lceil (2n + 3) \rceil$. For example if $n = 12$ then $r = 3$.

Note that the optimality consideration in both problems forces linear algebra over $\mathbb{F}_3$ into the picture.

5. Assignment 2

(1) Show that the matrix H has all row sums zero, except the last row which has row sum 1.

(2) Show that the repetition code of length $2t + 1$ above is a perfect code when $q = 2$.

(3) Show that the minimum distance of a perfect code is an odd integer.

(4) Show that in every linear code over $\mathbb{F}_2$, either all codewords have even Hamming weight or exactly half of the codewords have even Hamming weight. The Hamming weight of a vector is its distance from the zero vector. (Hint : Let $\mathcal{C} \subset \mathbb{F}_2^n$. Consider the linear map $\mathbb{F}_2^n \to \mathbb{F}_2$ given by $x \mapsto x_1 + x_2 + \cdots + x_n$.)

(5) Given a word $x \in \mathcal{F}^n$ where $|\mathcal{F}| = q$, what is the average distance of a word y from x ?

6. Linear Codes -I

For linear codes the alphabet $\mathcal{F}$ is a finite field. A finite field is any field whose size q is finite. As we show below, the size of such a field is always of the form $q = p^m$ where p is a prime number and m is a positive integer. It can be shown that for each prime power integer q there is (up to isomorphism) only one field of size q. Such a field is denoted $\mathbb{F}_q$.

A linear code C is a linear subspace of the vector space $\mathbb{F}_q^n$. If $0 \leq k \leq n$ is the dimension of C as a vector space over $\mathbb{F}_q$, then C has size q^m: this is because if $\vec{b_1}, \ldots, \vec{b_k}$ is a basis for the vector space C, then the map $\mathbb{F}_q^k \to C$ given $(a_1, \ldots, a_k) \mapsto a_1\vec{b_1} + \ldots a_k\vec{b_k}$ is a vector space isomorphism. In particular, the rate of a linear code is $\frac{\log_q |C|}{n} = k/n$.

For $x \in \mathbb{F}_q^n$ let $\mathrm{wt}(x)$ denote the number of nonzero entries of x. Note that $d(x, y) = \mathrm{wt}(x - y)$. We also note that

$$\{x - y : x, y \in C, x \neq y\} = C \setminus \{0\}.$$

Therefore

$$d_C = \min\{\mathrm{wt}(x - y) : x, y \in C, x \neq y\} = \min\{\mathrm{wt}(x) : x \in C \setminus \{0\}\}.$$

6.1. **Finite Fields-I.** Let n be a positive integer. Recall that for integers m, n the division with remainder of m by n consists of writing $m = qn + r$ for a unique $0 \leq r \leq n - 1$. Partition the set of integers $\mathbb{Z}$ into n parts $\mathbb{Z} = A_0 \coprod A_1 \coprod \cdots \coprod A_{n-1}$ where A_i consists of those integers which leave a remainder i when divided by n. There are natural operations $+$ and $\cdot$ on the set $\{A_0, \ldots, A_{n-1}\}$. If $x \in A_i$ and $y \in A_j$ then $x + y \in A_k$ and $xy \in A_\ell$ for unique k, ℓ which do not depend on which elements x of A_i and y of A_j are chosen. In this way we can define $A_i + A_j$ and $A_i \cdot A_j$. Clearly

$$A_i + A_j = A_j + A_i,\ (A_i + A_j) + A_k = A_i + (A_j + A_k),\ A_i + A_0 = A_i,\ A_i + A_{n-i} = A_0.$$

Thus $\{A_0, \ldots, A_{n-1}\}$ is an abelian group with respect to $+$. Also

$$A_i \cdot A_j = A_j \cdot A_i,\ (A_i \cdot A_j) \cdot A_k = A_i \cdot (A_j \cdot A_k),\ A_i \cdot A_1 = A_i,$$

and

$$A_i \cdot \cdot (A_j + A_k) = A_i \cdot A_j + A_i \cdot A_k.$$

Thus $\{A_0, \ldots, A_{n-1}\}$ with the operations of $+, \cdot$ is a commutative ring. We usually just denote A_i as $i \bmod n$ and the set (or ring) $\{A_0, \ldots, A_{n-1}\}$ is denoted $\mathbb{Z}/n\mathbb{Z}$, the ring of integers mod n. If $j \neq 0 \bmod n$, then the function $f_j : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ defined by $x \bmod n \mapsto jx \bmod n$ is a group homomorphism of the group $(\mathbb{Z}/n\mathbb{Z}, +)$. If $j \bmod n$ has a multiplicative inverse $j' \bmod n$ then we note that f_j is surjective (because $x \bmod n = f_j(j'x \bmod n)$. Conversely, if f_j is surjective, then we note that the element $j' \bmod n$ such that $f_j(j') = 1 \bmod n$ is a multiplicative inverse of $j \bmod n$. Since the set $\mathbb{Z}/n\mathbb{Z}$ is finite, the map f_j is surjective if and only if it is injective. Moreover, f_j is injective if and only if the kernel of the homomorphism f_j is $\{0 \bmod n\}$. The kernel consists of those $x \bmod n$ for which $jx = 0 \bmod n$. If $m = n/\gcd\{j, n\}$ then, it is clear that the $\ker(f_j) = \{0, m, 2m, \ldots, (\gcd\{j, n\} - 1)m\}$. In particular f_j is injective and hence $j \bmod n$ has a multiplicative inverse if and only if j is relatively prime to n. We recall that the number $\#\{1 \leq j \leq n : j \text{ is relatively prime to } n\ \}$ is denoted by the symbol $\phi(n)$ and the function ϕ is called the Euler ϕ function or sometimes the totient function.

The condition for $\mathbb{Z}/n\mathbb{Z}$ to be a field is that each j mod n except 0 mod n must have a multiplicative inverse. In other words each $1 \leq j \leq n-1$ must be relatively prime to n. This is the case if and only if n is a prime number. When n is a prime number p we will also denote the field $\mathbb{Z}/p\mathbb{Z}$ as $\mathbb{F}_p$.

Now suppose F is a field which has a finite number of elements. Let 1_F denote the multiplicative identity of F and consider the function from $\imath : \mathbb{Z} \to F$ defined by $\imath(n) = 1_F + \cdots + 1_F$ (n times) when n is positive, and in case n is negative we define $\imath(n) = -\imath(|n|)$. It is easy to check that $\imath$ is a ring homomorphism from $(\mathbb{Z}, +1)$ to $(F, +)$, that is $\imath(x+y) = \imath(x) + \imath(y)$ and $\imath(xy) = \imath(x) \cdot \imath(y)$. It is well known that any subgroup of $(\mathbb{Z}, +)$ is of the form $m\mathbb{Z}$ for some $m \in \{0, 1, 2, \dots\}$. Also $m \neq 0$ for otherwise $\imath$ will be injective, contradicting the finiteness of $\mathbb{F}$. Suppose $1 < d < m$ is a divisor of m, then $\imath(d), \imath(md)$ are both nonzero (because m is the least positive integer for which $\imath(m) = 0_F$). However $\imath(d) \cdot \imath(m/d) = 0_F$. Since the product of two nonzero elements of a field is never 0, it follows that no such divisor d exists, i.e. m is a prime number p. We also note that the homomorphism $\imath$ also induces injective field homomorphism of $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ into $\mathbb{F}$. In other words the subfield of $\mathbb{F}$ generated by 0_F and 1_F is $\mathbb{F}_p$. This is called the prime subfield of $\mathbb{F}$.

Thinking of $(\mathbb{F}, +1)$ as an abelian group, and the field $\mathbb{F}_p \subset \mathbb{F}$ as scalars, we may view the operation $y \mapsto xy$ (as defined in the field $\mathbb{F}$) as multiplying elements y of $(\mathbb{F}, +)$ by scalars $x \in \mathbb{F}_p$. This scalar multiplication satisfies the conditions required for a $\mathbb{F}_p$ vector space structure of the abelian group $(\mathbb{F}, +)$:

$$1_{\mathbb{F}_p} \cdot y = y, \;\; x \cdot (y+z) = x \cdot y + x \cdot z, \; (x+x') \cdot y = x \cdot y + x' \cdot y, \; (xx') \cdot y = x \cdot (x' \cdot y).$$

Therefore $\mathbb{F}$ is a vector space over $\mathbb{F}_p$. Suppose there exists a set of m linearly independent elements of $\mathbb{F}$ over $\mathbb{F}_p$, then, as noted above the span of this set has size p^m. In particular $m \leq \log_p |\mathbb{F}|$, and hence $\mathbb{F}$ is finite dimensional as a vector space over $\mathbb{F}_p$. Let k be this dimension. As above, the size $\mathbb{F}$ is p^k. This establishes that the size of a finite field $\mathbb{F}$ is p^k for a unique prime p and natural number k, and $\mathbb{F}_p$ is the prime subfield of $\mathbb{F}$. The prime p is also called the *characteristic* of the field $\mathbb{F}$.

6.2. **Linear Codes continued.** Let $C \subset \mathbb{F}_q^n$ be a linear code of dimension k. Let G be a $k \times n$ matrix whose rows from a basis for C. Such a matrix is called a *generator matrix* for the code C.

Any other generator matrix of C is of the form PG where P is an invertible $k \times k$ matrix over $\mathbb{F}_q$. (see HW).

Since the rows of G are linearly independent (by definition of G), it follows that G has full rank k. By rank-nullity theorem the kernel or null space of G is a $n-k$ dimensional subspace of $\mathbb{F}_q^n$. This subspace of $\mathbb{F}_q^n$ (when regarded as a code) is called the dual code of C and denoted as $C^\perp$. Note that $C^\perp$ depends only on C and not the choice of the generator matrix G.

If H is a $(n-k) \times n$ matrix whose rows form a basis for $C^\perp$ (i.e. H is a generator matrix for $C^\perp$) then H has rank $n-k$ and hence nullity k. Since $C \subset \ker(H)$ it follows that $C = \ker(H)$. In general for a k-dimensional linear code $C \subset \mathbb{F}_q^n$, any $(n-k) \times n$ matrix H of rank $n-k$ whose kernel is C, is called a *parity check* matrix of C. As before, if H, H' are parity check matrices of C, then

$H' = QH$ for some invertible $(n-k)\times(n-k)$ matrix Q. Often, linear codes are presented by parity check matrices instead of generator matrices. For example, the Hamming codes in the previous section. Also, the $[8,7]_2$ ASCII code has parity check [1111111|1] and the $[10,9]_{11}$ (extended) ISBN code has parity check [123456789|10].

We have seen in Lemma 1 that the distance of a word $x \in \mathbb{F}_q^n$ from a linear code $C \subset \mathbb{F}_q^n$ can be expressed in terms of a parity check matrix H of C.

$$\text{syn} : \mathbb{F}_q^n \to \mathbb{F}_q^{n-k} \quad \text{defined by } \text{syn}(x) = Hx. \tag{4}$$

The vector $Hx \in \mathbb{F}_q^{n-k}$ is called the *syndrome* of x. The distance of x from C is the smallest integer j such that Hx can be expressed as a linear combination of j columns of H. Since the linear map syn above is surjective, we get a characterization of the covering radius $\rho(C)$ in terms of H :

Lemma 2. *Let $C \subset \mathbb{F}_q^n$ be a k-dimensional linear code with parity check matrix H. The covering radius $\rho(C)$ is the least integer j such that every vector of $\mathbb{F}_q^{n-k}$ is in the span of some (variable set of) j columns of H.*

We now characterize the minimum distance d_C in terms of the parity check matrix H :

Lemma 3. $d(C) - 1 = \max\{j : \text{ any } j \text{ columns of } H \text{ are linearly independent}\}$.

Proof. Let Δ denote the right side of the above equation. There exist some j columns of H which are linearly dependent if and only if there is a codeword $x \in C$ of weight at most j. (This is because a relation of linear dependence in the columns of H looks like $Hx = 0$ i.e. x is a codeword). Thus, every set of j columns of H is linearly independent if and only if there is no codeword of weight at most j. This is equivalent to $d(C) > j$. Thus Δ is the largest j satisfying $d(C) > j$ which is $d(C) - 1$. □

The q-ary Hamming codes. For a finite field $\mathbb{F}_q$ and an integer $r \geq 2$, the Hamming code is a linear code of length $(q^r-1)/(q-1)$ and dimension $(q^r-1)/(q-1) - r$. It is defined by a parity check matrix H of size $r \times (q^r-1)/(q-1)$ constructed as follows: Let $S \subset \mathbb{F}_q^r$ be the set of all non-zero vectors whose first non-zero entry is 1. We may decompose

$$S = S_1 \amalg S_2 \amalg \cdots \amalg S_r,$$

where S_i is the subset of S consisting of vectors whose first nonzero entry occurs in the i-place. Clearly $|S_i| = q^{r-i}$ and hence $|S| = 1 + q + \cdots + q^{r-1} = (q^r-1)/(q-1)$.

We may also view S as the set of of one-dimensional subspaces of $\mathbb{F}_q^r$. On the set $\mathbb{F}_q^r \setminus \{\vec{0}\}$ of size $q^r - 1$, put an equivalence relation $v \sim w$ if and only if w is a nonzero scalar multiple of v (equivalently v and w generate the same one dimensional subspace of $\mathbb{F}_q^r$). Clearly each equivalence class has size $(q-1)$. Thus the set of equivalence classes has size $(q^r-1)/(q-1)$. Also each equivalence class has a unique representative in the set S. The columns of H are taken to be some ordering of the set S.

Note that any two columns of H are independent, but given any two columns there is a third column which is a linear combination of the first two. Thus $d_C = 3$ by Lemma 3, and $\rho_C = 1$ by

Lemma 2. In this case the packing radius $\lfloor (d(\mathcal{C})-1)/2 \rfloor$ equals the covering radius $\rho(C)$ (which are both equal to 1). Thus the Hamming codes are perfect codes. We summarize:

Proposition 4. *The* $[(q^r-1)/(q-1), (q^r-1)/(q-1)-r]_q$ *Hamming code has minimum distance* 3*, covering radius* 1 *and is a perfect code.*

The dual $C^\perp$ of the Hamming codes are called the *simplex codes*. They have the interesting property that all nonzero codewords of $C^\perp$ have the same weight q^{r-1}. (See HW)

7. Assignment 3

(1) Verify that the function $\imath : \mathbb{Z} \to \mathbb{F}$ respects addition and multiplication.

(2) Let G be a subgroup of the additive group $(\mathbb{Z}, +)$. If $G \neq \{0\}$, let n be the least positive member of G. Show that $G = n\mathbb{Z}$. Thus every subgroup of $(\mathbb{Z}, +)$ is of the form $n\mathbb{Z}$ for some $n \in \{0, 1, 2, \dots\}$.

(3) Prove that $\phi(n) = n \prod_{p|n}(1 - p^{-1})$.

(4) Let C be the q-ary Hamming code of length $(q^r - 1)/(q-1)$ and dimension $(q^r - 1)/(q-1) - r$. Show that every nonzero codeword of the dual code $C^\perp$ has the same weight q^{r-1}.

(5) Let C be an $[n, q^k]_q$ linear code with generator matrix G. If G does not have a column of 0's, then the prove that the average weight of a codeword of C is $n(1 - q^{-1})$.

(6) * Let $C \subset \mathbb{F}_q^n$ be a k-dimensional linear code with a generator matrix G. Show that d_C is the largest integer j such that every $k \times (n - j + 1)$ sub-matrix of G has rank k.

8. Hamming isometries

For any set S, the set $\mathrm{Perm}(S)$ of bijections of S is a group with respect to composition of functions. The identity element is the transformation $\mathrm{id} : S \to S$ defined by $\mathrm{id}(s) = s$ for all $s \in S$. For $f \in \mathrm{Perm}(S)$, the set theoretic inverse function f^{-1} is the inverse in the group $\mathrm{Perm}(S)$. If S is a metric space with metric $d : S \times S \to \mathbb{R}$, then the subset of $\mathrm{Perm}(S)$ defined as $\mathrm{Isom}(S) = \{f \in \mathrm{Perm}(S) : d(f(x), f(y)) = d(x, y)\}$ is easily seen to be a subgroup of $\mathrm{Perm}(S)$:

(1) clearly $\mathrm{id} : S \to S$ is in $\mathrm{Isom}(S)$,
(2) if f is an isometry and $x, y \in S$, let $u = f(x)$ and $v = f(y)$ and let $g = f^{-1}$. Since $f \in \mathrm{Isom}(S)$, we have:
$$d(x, y) = d(f(x), f(y)) = d(u, v) = d(g(x), g(y)),$$
which shows that $g = f^{-1} \in \mathrm{Isom}(S)$.
(3) if $f, h \in \mathrm{Isom}(S)$, then
$$d(f(h(x)), f(h(y))) = d(h(x), h(y)) = d(x, y),$$
which shows that $f \circ h \in \mathrm{Isom}(S)$.

This completes the verification of $\mathrm{Isom}(S)$ being a subgroup of $\mathrm{Perm}(S)$.

Let $\mathcal{F}$ be any alphabet and consider the space $\mathcal{F}^n$ with the Hamming metric. Some elements of $\mathrm{Isom}(\mathcal{F}^n)$ are easy to find.

- For $\sigma \in \mathrm{Perm}\{1, \ldots, n\}$, the transformation
$$T_\sigma : (x_1, \ldots, x_n) \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(n)}),$$
is an isometry. This is because $d(x, y)$ being the number of coordinates in which x and y differ does not depend on the ordering of the coordinates. Note that $\sigma \to T_\sigma$ is an injective group homomorphism from $\mathrm{Perm}\{1, \ldots, n\} \to \mathrm{Isom}(\mathcal{F}^n)$, and hence $\{T_\sigma : \sigma \in \mathrm{Perm}\{1, \ldots, n\}\}$ is a copy of $\mathrm{Perm}\{1, \ldots, n\}$ in $\mathrm{Isom}(\mathcal{F}^n)$
- If $\theta_1, \ldots, \theta_n \in \mathrm{Perm}(\mathcal{F})$, let $\vec{\theta}$ denote the n-tuple $(\theta_1, \ldots, \theta_n)$. The transformation
$$T_{\vec{\theta}} : (x_1, \ldots, x_n) \mapsto (\theta_1(x_1), \ldots, \theta_n(x_n)),$$
is an isometry. This is because if $x' = T_{\vec{\theta}}(x)$ and $y' = T_{\vec{\theta}}(y)$ then x and y agree in the i-th coordinate if and only if x' and y' agree in the i-th coordinate. We note that $\vec{\theta} \to T_{\vec{\theta}}$ is an injective group homomorphism from direct product $\mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})$ to $\mathrm{Isom}(\mathcal{F}^n)$, and hence $\{T_{\vec{\theta}} : \theta \in \mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})\}$ is a copy of $\mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})$ in $\mathrm{Isom}(\mathcal{F}^n)$.
- Since $\mathrm{Isom}(\mathcal{F}^n)$ is a group, we can compose the above two examples: for $\sigma \in \mathrm{Perm}\{1, \ldots, n\}$ and $\vec{\theta}$ as above, the composition $T_{\sigma,\vec{\theta}} = T_{\vec{\theta}} \circ T_\sigma$ is an isometry.
$$T_{\sigma;\vec{\theta}} : (x_1, \ldots, x_n) \mapsto (\theta_1(x_{\sigma(1)}), \ldots, \theta_n(x_{\sigma(n)})).$$

Let $G \subset \mathrm{Isom}(\mathcal{F}^n)$ consist of all such transformations $T_{\sigma,\vec{\theta}}$. We now check that G is a subgroup of $\mathrm{Isom}(\mathcal{F}^n)$. First, we note that

$$T_\sigma \circ T_{\vec{\theta}} \circ T_\sigma^{-1} = T_{\vec{\theta}'} \text{ where } \vec{\theta}' = (\theta_{\sigma(1)}, \ldots, \theta_{\sigma(n)}). \tag{5}$$

(1) Clearly $\mathrm{id} : \mathcal{F}^n \to \mathcal{F}^n$ is in G because $\mathrm{id} = T_{\sigma;\vec{\theta}}$ for $\sigma = \mathrm{id} : \{1, \ldots, n\} \to \{1, \ldots, n\}$ and $\theta_1 = \cdots = \theta_n = \mathrm{id} : \mathcal{F} \to \mathcal{F}$.

(2) let $\sigma, \tau \in \mathrm{Perm}\{1,\ldots,n\}$ and let $\theta_1,\ldots,\theta_n,\beta_1,\ldots,\beta_n \in \mathrm{Perm}(\mathcal{F})$. We will show that $T_{\sigma,\vec{\theta}} \circ T_{\tau,\vec{\beta}} = T_{\sigma\circ\tau,\vec{\theta}\circ\vec{\beta'}}$ where $\vec{\beta'} = (\beta_{\sigma(1)},\ldots,\beta_{\sigma(n)})$:

$$T_{\sigma,\vec{\theta}} \circ T_{\tau,\vec{\beta}} = T_{\vec{\theta}} \circ T_\sigma \circ T_{\vec{\beta}} \circ T_\tau = T_{\vec{\theta}} \circ (T_\sigma \circ T_{\vec{\beta}} \circ T_{\sigma^{-1}}) \circ T_{\sigma\circ\tau} = T_{\vec{\theta}\circ\vec{\beta'}} \circ T_{\sigma\circ\tau} = T_{(\sigma\circ\tau),(\vec{\theta}\circ\vec{\beta'})},$$

(3) If we take τ and $\vec{\beta}$ above to be $\tau = \sigma^{-1}$ and $\vec{\beta} = (\theta^{-1}_{\sigma^{-1}(1)},\ldots,\theta^{-1}_{\sigma^{-1}(n)})$, we see that $T_{\tau,\vec{\beta}} = T^{-1}_{\sigma,\vec{\theta}}$

Theorem 5. *The subgroup G of $\mathrm{Isom}(\mathcal{F}^n)$ above is the whole group $\mathrm{Isom}(\mathcal{F}^n)$. In other words*

$$\mathrm{Isom}(\mathcal{F}^n) = \{T_{\sigma,\vec{\theta}} : \sigma \in \mathrm{Perm}\{1,\ldots,n\},\ \vec{\theta} \in \mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})\}.$$

We postpone the proof of this theorem to discuss the linear case which is a bit simpler. Let $\mathcal{F} = \mathbb{F}_q$ and we now consider the group of linear Hamming-isometries of $\mathbb{F}_q^n$, the subgroup of $\mathrm{Isom}(\mathbb{F}_q^n)$ which are of the form $x \mapsto xM$ for some $M \in GL_n(\mathbb{F}_q)$ (the group of $n \times n$ invertible matrices with entries from $\mathbb{F}_q$). Recall we are writing elements of $\mathbb{F}_q^n$ as row vectors here. For $\sigma \in \mathrm{Perm}\{1,\ldots,n\}$ we note that

$$T_\sigma x = xP_\sigma, \quad \text{where } (P_\sigma)_{ij} = \delta_{i,\sigma(i)},$$

and hence T_σ's are linear. The *permutation matrix* P_σ is obtained by permuting the columns of the $n \times n$ identity matrix by the permutation σ, i.e. the j-th column of P_σ is the $\sigma(j)$-th column of the identity matrix. Note that $\sigma \mapsto P_\sigma$ is an injective group homomorphism from $\mathrm{Perm}\{1,\ldots,n\} \to GL_n(\mathbb{F}_q)$, and hence $\{P_\sigma : \sigma \in \mathrm{Perm}\{1,\ldots,n\}\}$ is a copy of $\mathrm{Perm}\{1,\ldots,n\}$ in $GL_n(\mathbb{F}_q)$.

The transformation $T_{\vec{\theta}}$ is linear if and only if each $\theta_i \in \mathrm{Perm}(\mathbb{F}_q)$ is a linear bijection of $\mathbb{F}_q$. To see this let $e_1,\ldots e_n$ be the standard basis of $\mathbb{F}_q^n$ given by the columns of the $n \times n$ identity matrix. Since

$$(\theta_1(0),\ldots,\theta_{j-1}(0),\theta_j(c),\theta_{j+1}(0),\ldots,\theta_n(0)) = T_{\vec{\theta}}(ce_j) = cT_{\vec{\theta}}(e_j) = c(\theta_1(0),\ldots,\theta_{j-1}(0),\theta_j(1),\theta_{j+1}(0),\ldots,\theta_n($$

Taking $c = 0$, we get $\theta_i(0) = 0$ for all i. Next, taking $c = 1$, we get $\theta_j(c) = c\lambda_j$ where $\lambda_j = \theta_j(1)$. This shows that $T_{\vec{\theta}}x = xD_{\vec{\lambda}}$ for a diagonal matrix $D_{\vec{\lambda}} = \mathrm{diag}(\lambda_1,\ldots,\lambda_n)$ where $\lambda_1,\ldots,\lambda_n \in \mathbb{F}_q \setminus \{0\}$. We denote $T_{\vec{\theta}}$ by $T_{\vec{\lambda}}$. The composition $T_{\sigma,\vec{\lambda}} = T_{\vec{\lambda}} \circ T_\sigma$ can be written as

$$T_{\sigma,\vec{\lambda}}x = xM, \quad \text{where } M_{ij} = \lambda_i\delta_{i,\sigma(i)}.$$

We can also write $M = P_\sigma D_{\vec{\lambda}}$. A $n \times n$ *monomial matrix* is a product $P_\sigma \cdot D_{\vec{\lambda}}$ of a permutation matrix and an invertible diagonal matrix. The equation (5) shows that the set $n \times n$ monomial matrices $\mathrm{Mon}_n(\mathbb{F}_q)$ is a subgroup of $GL_n(\mathbb{F}_q)$.

Theorem 6. *The group of linear Hamming isometries of $\mathbb{F}_q^n$ is the group $Mon_n(\mathbb{F}_q)$ of $n \times n$ monomial matrices.*

Proof. Let φ be a linear isometry of $\mathbb{F}_q^n$. Note that $\mathrm{wt}(\varphi(x)) = \mathrm{wt}(x)$ for any vector $x \in \mathbb{F}_q^n$. In particular $\varphi(e_j) = \lambda_j e_{j'}$ for some $\lambda_j \in \mathbb{F}_q \setminus \{0\}$ and some $j' \in \{1,\ldots,n\}$. For $i \neq j$ we have

$$2 = \mathrm{wt}(e_i - e_j) = \mathrm{wt}(\varphi(e_i - e_j)) = \mathrm{wt}(\lambda_i e'_i - \lambda_j e'_j).$$

If $e_{i'} = e_{j'}$ then the last quantity in the equation above would be at most 1. Therefore $e'_i \neq e'_j$ if $i \neq j$. Thus there exists $\sigma \in \mathrm{Perm}\{1,\ldots,n\}$ such that $e_{j'} = e_{\sigma(j)}$. We now have

$$\varphi(x) = \varphi(\sum_{j=1}^n x_j e_j) = \sum_{j=1}^n x_j \varphi(e_j) = \sum_{j=1}^n x_j \lambda_j e_{\sigma(j)} = \sum_{j=1}^n x_{\sigma^{-1}(j)} \lambda_{\sigma^{-1}(j)} e_j = xM,$$

where M is the monomial matrix $P_\tau D_{\vec{\beta}} = D_{\vec{\lambda}} P_\tau$ where$\tau = \sigma^{-1}$ and $\beta_j = \lambda_{\sigma^{-1}(j)}$. This completes the proof that any linear Hamming isometry of $\mathbb{F}_q^n$ is a monomial transformation. □

Proof. (Proof of Theorem 5) We may assume $\mathcal{F} = \mathbb{Z}/q\mathbb{Z}$. The weight of a word $x \in (\mathbb{Z}/q\mathbb{Z})^n$ will denote the number of non-zero components of x. For each $a \in \mathcal{F}$ the translations $x \mapsto x + a$ are permutations of $\mathcal{F}$. For $\vec{a} = (a_1, \ldots, a_n)$ let $T_a(\vec{x}) = \vec{x} + \vec{a}$. Given $\varphi \in \mathrm{Isom}(\mathcal{F}^n)$, let $\varphi(\vec{0}) = \vec{a}$. It suffices to consider $T_{-a} \circ \phi$ instead of ϕ. Thus we may assume $\varphi(\vec{0}) = \vec{0}$, and hence $\mathrm{wt}(\varphi(x)) = \mathrm{wt}(x)$ for all $x \in \mathcal{F}^n$. We now proceed as above: since φ permutes the set of weight 1 words, let $\varphi(e_i) = \lambda_i e_{i'}$ and $\varphi(e_j) = \lambda_j e_{j'}$, where $i \neq j$ and $\lambda_i, \lambda_j \neq 0$. Since $d(e_i, e_j) = 2 = d(\lambda_i e_{i'}, \lambda_j e_{j'}) = d(e_{i'}, e_{j'})$ it follows that $i' \neq j'$. Thus there is a $\sigma \in \mathrm{Perm}\{1, \ldots, n\}$ such that $i' = \sigma(i)$ for all i. Next we note that for $c \neq 1$, we have $d(ce_i, e_i) = 1 = d(\varphi(ce_i), e_{\sigma(i)})$ and $\varphi(ce_i)$ is also a weight 1 word. This forces $\varphi(ce_i) = c' e_{\sigma(i)}$. For $c \neq a$, the fact that $d(ce_i, ae_i) = 1 = d(c' e_{\sigma(i)}, a' e_{\sigma(i)}) = 1 - \delta_{a',c'}$ forces $\varphi(ce_i) = \theta_i(c) e_{\sigma(i)}$ for some element $\theta_i \in \mathrm{Perm}(\mathcal{F})$. Let $\psi = T_{\sigma^{-1}} \circ T_{\vec{\theta}} \in G$. Note that $\psi^{-1}\varphi$ fixes each weight 1 word. In order to show that $\varphi \in G$, it suffices to show that $\psi^{-1}\varphi \in G$. Therefore, we may assume that φ itself fixes each weight 1 word. We leave it as an exercise (See HW) to show that φ must then fix all words, i.e. φ is the trivial element of G. □

(optional) Structure of the group $\mathbf{Isom}(\mathcal{F}^n)$.
If A is a group with disjoint subgroups B and C satisfying the following two conditions:

(1) the function $B \times C \to A$ given by $(b, c) \mapsto bc$ is surjective
(2) for each $b \in B$ and $c \in C$ the product $bcb^{-1} \in C$ (i.e. B normalizes C)

then A is called the internal semi-direct product of its subgroups B and C , and denoted $A = C \rtimes B$.

Taking $A = \mathrm{Isom}(\mathcal{F}^n)$, $B = \mathrm{Perm}\{1, \ldots, n\}$ and $C = \mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})$, the condition 1) above as verified in the proof of Theorem 5, and the condition 2) was verified in (5). Thus,

$$\mathrm{Isom}(\mathcal{F}^n) = (\mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})) \rtimes \mathrm{Perm}\{1, \ldots, n\},$$

is the internal semi-direct product of its subgroups $\mathrm{Perm}\{1, \ldots, n\}$ and $\mathrm{Perm}(\mathcal{F}) \times \cdots \times \mathrm{Perm}(\mathcal{F})$. Similarly,

$$\mathrm{Mon}_n(\mathbb{F}_q) = D_n \rtimes \mathrm{Perm}\{1, \ldots, n\},$$

is the internal semi-direct product of its subgroups $\mathrm{Perm}\{1, \ldots, n\}$ and $\mathbb{F}_q^\times \times \cdots \times \mathbb{F}_q^\times$ (the group of diagonal matrices D_n).

The structure of the isometry group can also be stated concisely using the notion of *wreath product* of two groups. $\mathrm{Isom}(\mathcal{F}^n) = \mathrm{Perm}(\mathcal{F})$ wr S_n and similarly $\mathrm{Mon}_n(\mathbb{F}_q) = \mathbb{F}_q^\times$ wr S_n.

9. Assignment 4

(1) Suppose $\mathcal{F} = \mathbb{Z}/q\mathbb{Z}$, and $\varphi \in \mathrm{Isom}(\mathcal{F}^n)$ fixes each weight 1 word, then show that φ fixes every word of $\mathcal{F}^n$.

(2) Suppose $C \subset \mathbb{F}_q^n$ is a k-dimensional linear code with generator matrix $G = [I_k|A_{k\times(n-k)}]$. Show that $H = [-A^t|I_{n-k}]$ is a parity check matrix for C. Here I_k denotes the $k \times k$ identity matrix.

(3) We say two linear codes $C, C' \subset \mathbb{F}_q^n$ are *monomially equivalent* if there is a $M \in \mathrm{Mon}_n(\mathbb{F}_q)$ such that $C' = \{cM : c \in C\}$. In case M is a permutation matrix, we say C, C' are *permutation equivalent*.

Show that upto permutation equivalence I(i.e. replacing C by $C\dot{P}$ where P is a $n \times n$ permutation matrix) every k-dimensional linear code $C \subset \mathbb{F}_q^n$ has a generator matrix of the form $G = [I_k|A_{k\times(n-k)}]$. Moreover, such a generator matrix is unique. In other words, there exists a permutation matrix P such that GP is of the form $GP = [I_k|A]$. Show that C has parity check matrix $[-A^t|I_{n-k}]P$.

(4) Let A be an invertible $n \times n$ matrix over a field $\mathbb{K}$. (We also say $A \in GL_n(\mathbb{K})$). Let $A^{-t} = (A^t)^{-1} = (A^{-1})^t$ denote the inverse transpose of A. Show that $A \mapsto A^{-t}$ is an automorphism (an invertible homomorphism) of $GL_n(\mathbb{K})$.

(5) Let $C \subset \mathbb{F}_q^n$ be a k-dimensional linear code. The linear automorphism group if C is

$$\mathrm{MAut}(C) = \{M \in \mathrm{Mon}_n(\mathbb{F}_q) : cM \in C \text{ for all } c \in C\}.$$

Show that the inverse-transpose automorphism of $GL_n(\mathbb{F}_q)$ carries $\mathrm{MAut}(C)$ to $\mathrm{MAut}(C^\perp)$

10. Quiz 1

In this quiz, C denotes a $[n, M, d]_q$ code, i.e. C is a subset of $\mathcal{F}^n$ of size M and minimum distance d. The alphabet $\mathcal{F}$ has size q.

(1) (5 points) Suppose we receive a word $x \in \mathcal{F}^n$ which equals a codeword of C but for the fact that e of the components of x are erased or not readable. There are no other errors. (This is called an *erasure communication channel*). What is the maximum value of e so that the actual codeword x can be accurately determined ?
Ans: The maximum value of e is $d-1$. Suppose x is received with e components unreadable and there being no errors in the remaining $n-e$ co-ordinates. Let these $n-e$ coordinates be indexed by $I \subset \{1, \ldots, n\}$. We want that there should be no $y \in C$ with $x_{|I} = y_{|I}$. This is guaranteed if $e \leq d-1$ (because no two codewords can agree in more than $n-d$ positions). Conversely, taking a pair of codewords x, y at a distance of d apart and taking $I \subset \{i : x_i = y_i\}$, we see that any $e \geq d$ will not have this guarantee.

(2) (5 points) Suppose $\mathcal{F} = \mathbb{F}_q$, $M = q^k$ and C is linear with generator matrix G. If exactly r of the columns of G are $\vec{0} \in \mathbb{F}_q^k$, then determine the average weight of a codeword of C.
Ans: Done in class. The answer is $(n-r)(1-q^{-1})$.

(3) (6 points) (note C is <u>not assumed</u> to be linear in parts a), c) below)
(a) Suppose C is perfect and q is a prime, then show that M is a power of q.

(b) Let r and s be positive integers. Show that for a prime number p the number $p^r - 1$ divides $p^s - 1$ if and only if r divides s.

(c) Suppose C is perfect and q is a prime power, then show that M is a power of q.

Ans: a) For a perfect code M divides q^n. If q is prime, it follows that M is a prime power.

b) Using division with remainder, we write $s = ar + b$ with $0 \leq b \leq r-1$. It is easy to see that

$$\tfrac{p^s-1}{p^r-1} = p^b \tfrac{p^{ar}-1}{p^r-1} + \tfrac{p^b-1}{p^r-1}$$

is an integer if and only if $b = 0$ i.e. r divides s.
c) Now suppose $q = p^r$. We have $MV_q(n; r) = q^n = p^{nr}$. Therefore, $V_q(n, r) = p^s$ for some s. It suffices to show that r divides s, because it then follows that both $V_q(n, r)$ and M are powers of q. Since

$$V_q(n; r) - 1 = n(q-1) + \cdots + \tbinom{n}{(d-1)/2}(q-1)^{(d-1)/2},$$

is divisible by $(q-1) = p^r - 1$, it follows from part b) that r divides s.

(4) (4 points) Suppose the Singleton bound is achieved by C, i.e. $M = q^{n-d+1}$. Suppose $\mathcal{F} = \mathbb{F}_q$ and C is a k-dimensional linear code with a parity check matrix H. Show that all $(n-k) \times (n-k)$ submatrices of H are invertible.
Ans: see next section

11. Linear MDS codes

We recall that the notation C is a $[n, M, d]_q$ code means that $C \subset \mathcal{F}^n$ has size M and minimum distance d, and the alphabet has size q. We also recall the *Singleton bound* which states that any $[n, M, d]_q$ code satisfies $M \leq q^{n-d+1}$.

An $[n, M, d]_q$ code C is called *Maximum Distance Separable* (in short MDS), if the Singleton bound is achieved for C, i.e. $M = q^{n-d+1}$. Linear MDS codes are particularly interesting

Theorem. *Let $C \subset \mathbb{F}_q^n$ be a k-dimensional linear code. TFAE*

(1) $d(C) = n - k + 1$
(2) all $k \times k$ minors of a generator matrix of C are non-zero
(3) all $(n-k) \times (n-k)$ minors of a parity check matrix of C are non-zero
(4) $d(C^\perp) = k + 1$

Any of these equivalent conditions characterizes linear MDS codes.

Proof. Equality in the Singleton bound $M = q^{n-d+1}$ is equivalent to $k = n - d + 1$ which is the condition 1) above.
1) $\Leftrightarrow$ 3): Let H be a parity check matrix for C. We recall from Lemma 3 that $d(C) - 1 = \max\{j : \text{ any } j \text{ columns of } H \text{ are linearly independent}\}$. Therefore condition 3) is equivalent to $d_C - 1 = n - k$ which is condition 1).

1) $\Leftrightarrow$ 2): Let G be a generator matrix G of C. There exists a zero $k \times k$ minor of G if and only if there exists a non-zero codeword having zero entries on the k coordinates indexed by the columns of the minor concerned (i.e. a codeword of weight at most $n-k$). This is turn is equivalent to $d_C < n - k + 1$, i.e. inequality holding in the Singleton bound. Therefore the condition 2) is equivalent to the condition 1). Applying this to $C^\perp$ we conclude that the condition 3) is equivalent to the condition 4)

4) $\Leftrightarrow$ 2) and 4) $\Leftrightarrow$ 3): Since G is a parity check matrix for $C^\perp$ our proofs of 1) $\Leftrightarrow$ 3) and 1) $\Leftrightarrow$ 2) establishes 4) $\Leftrightarrow$ 2) and 4) $\Leftrightarrow$ 3). □

There are some MDS codes which are sometimes called trivial. They are non-linear but can also be realized as linear codes when the alphabet is $\mathbb{F}_q$:

(1) The whole space $\mathcal{F}^n$ is a a $[n, q^n, 1]_q$ nonlinear MDS code.
(2) ii) the repetition code is a nonlinear $[n, q, n]$ MDS code
(3) iii) the $[n, q^{n-1}, 2]_q$ nonlinear code given by the encoding $(\mathbb{Z}/q\mathbb{Z})^{n-1} \to (\mathbb{Z}/q\mathbb{Z})^n$ given by $(x_1, \ldots, x_{n-1}) \mapsto (x_1, \ldots, x_{n-1}, x_1 + \cdots + x_{n-1})$. Prove that this code has minimum distance 2 and is thus MDS.

If G is a $k \times n$ matrix which generates a linear k-dimensional MDS code C. Let B be the submatrix of G on the first k columns. Since all $k \times k$ minors of G are non-zero, it follows in particular that B is invertible. The matrix $B^{-1}G$ is row equivalent to G and is hence another generator matrix for C. Thus, we may assume wlog that G has the form $[I_k|A_{n-k}]$ for some matrix A of size $k \times (n-k)$.

Lemma 7. *The matrix $G = [I_k|A]$ generates an MDS code if and only if all $j \times j$ minors of A are nonzero for $j = 1, \ldots, k$.*

This result is a corollary of the following Lemma:

Lemma 8. *Let $|x|$ denote the equivalence class of $x \in \mathbb{F}_q$ under the equivalence relation $x \sim -x$ on $\mathbb{F}_q$. For the $k \times n$ matrix $G = [I_k|A]$ consider the sets:*

$$S = \{|det(B)| : B \text{ is a } k \times k \text{ submatrix of } G\}.$$

$$S' = \{|1|\} \cup \{|m| : m \text{ is a } j \times j \text{ minor of } A \text{ for some } j \in \{1, \ldots, k\}\}.$$

We claim $S = S'$.

Proof. Let B be a $k \times k$ submatrix of the matrix. If $B = I_k$ then $|\det(B)| = |1|$. If $B \neq I_k$, let $j \in \{1, \ldots, k\}$ be the number of columns of B taken from A (with the remaining $k - j$ columns taken from I_k). If Let $e_1, \ldots, e_k$ denote the columns of I_k, and let $e_{\mu_1}, \ldots, e_{\mu_j}$ be the columns of I_k which are not columns of B. We may assume $1 \leq \mu_1 < \cdots < \mu_j \leq k$. Let $v_1, \ldots, v_{n-k}$ denote the columns of A. Let $v_{\nu_1}, \ldots, v_{\nu_j}$ be the list of columns of B taken from A. We may assume $1 \leq \nu_1 < \cdots < \nu_j \leq n - k$. Expanding $\det(B)$, we see that $\pm\det(B)$ is the $j \times j$ minor of A on columns $\nu_1, \ldots, \nu_j$ and rows $\mu_1, \ldots, \mu_j$. Clearly this establishes a bijective correspondence between S and S'. □

11.1. **Reed-Solomon Codes - I.**
Let $x_1, \ldots, x_q$ be a listing of the elements of $\mathbb{F}_q$. Let G be the $k \times (q + 1)$ matrix

$$G = \begin{pmatrix} 1 & 1 & \ldots & 1 & 0 \\ x_1 & x_2 & \ldots & x_q & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{k-2} & x_2^{k-2} & \ldots & x_q^{k-2} & 0 \\ x_1^{k-1} & x_2^{k-1} & \ldots & x_q^{k-1} & 1 \end{pmatrix} \tag{6}$$

A Reed-Solomon code of length n and dimension k is the code generated by a $k \times n$ submatrix of G. In order to show that a Reed-Solomon code (or RS-code in short) is MDS, it suffices to show that all $k \times k$ minors of G (and hence of G with some columns deleted) are nonzero. Consider the polynomial $V(X_1, \ldots, X_m) \in \mathbb{Z}[X_1, \ldots, X_m]$ (the ring of polynomials in m indeterminates with integer coefficients) defined by

$$V(X_1, \ldots, X_m) = \det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ X_1 & X_2 & \ldots & X_m \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{m-2} & X_2^{m-2} & \ldots & X_m^{m-2} \\ X_1^{m-1} & X_2^{m-1} & \ldots & X_m^{m-1} \end{pmatrix}$$

Lemma 9. $V(X_1, \ldots, X_m) = \prod_{i<j}(X_j - X_i)$ *in* $\mathbb{Z}[X_1, \ldots, X_m]$

Proof. Let $\mathbb{K} = \mathbb{Q}(X_1, \ldots, X_{m-1})$ denote the field of rational functions in $m - 1$ indeterminates $X_1, \ldots, X_{m-1}$ with $\mathbb{Q}$ coefficients. Let $f_{m-1} \in \mathbb{K}$ be the coefficient of X_m^{m-1} in $V(X_1, \ldots, X_m)$ regarded as an element of $\mathbb{K}[X_m]$. Just by expanding the determinant in the definition of $V(X_1, \ldots, X_m)$ along the last column, it is clear that $f_{m-1} = V(X_1, \ldots, X_{m-1})$. From the elementary fact that a matrix in which two columns are identical, has zero determinant, it follows that $V(X_1, \ldots, X_{m-1}, X_i)$ is the zero element of $\mathbb{K}$ for $1 \leq i \leq m-1$. In particular for $1 \leq i \leq m-1$ each X_i viewed as an element of $\mathbb{K}$ is a root of $V(X_1, \ldots, X_m)$ viewed as an element of $\mathbb{K}[X_m]$ i.e. as a polynomial in X_m with $\mathbb{K}$-coefficients. From that fact that a polynomial $p(X)$ in one variable X over any field F admits $a \in F$ as a root if and only if $X - a$ divides $p(X)$, it follows that $V(X_1, \ldots, X_m) = g \prod_{i<m}(X_m - X_i)$

for some element $g \in \mathbb{K}$. The coefficient of X_m^{m-1} in this expression is precisely g. Therefore $g = f_{m-1} = V(X_1, \ldots, X_{m-1})$. Since $V(X_1, X_2) = \det \left(\begin{smallmatrix} 1 & 1 \\ X_1 & X_2 \end{smallmatrix} \right) = (X_2 - X_1)$. It follows that

$$V(X_1, \ldots, X_m) = \left(\prod_{i<m} (X_m - X_i) \right) \cdot \left(\prod_{i<m-1} (X_{m-1} - X_i) \right) \cdot \cdots \cdot \left(\prod_{i<2} (X_2 - X_i) \right) = \prod_{i<j} (X_j - X_i)$$

□

Returning to the matrix G, we note that a $k \times k$ minor of G either equals $V(a_1, \ldots, a_k)$ or $V(a_1, \ldots, a_{k-1})$ for distinct elements $a_1, \ldots, a_k \in \mathbb{F}_q$. It follows from the lemma above, that all $k \times k$ minors of G are nonzero.

12. Distance Distribution of a code

Let C be a $[n, M, d]_q$ code. For an element $x \in C$ let

$$A_i(C; x) = \#\{y \in C : d(x, y) = i\},$$

denote the number of codewords which are at a distance of i from x. These numbers can be collected into a generating function

$$W_C(Z; x) = \sum_{i=0}^{n} A_i(C; x) Z^i,$$

which may be viewed as an element of the polynomial ring $\mathbb{Z}[Z]$ or $\mathbb{Q}[Z]$. We may average the quantities $A_i(C; x)$ over all codewords $x \in C$ to obtain the (rational) number

$$A_i(C) = \tfrac{1}{M} \sum_{x \in C} A_i(C; x),$$

and similarly the polynomial $W_C(Z) \in \mathbb{Q}[Z]$ defined by

$$W_C(Z) = \sum_{i=0}^{n} A_i(C) Z^i.$$

A code C is called *distance-invariant* if $A_i(C; x)$ is independent of $x \in C$. In this case $A_i(C) = A_i(C; x)$ for any $x \in C$. Clearly, linear code are distance-invariant because $A_i(C; x)$ is independent of $x \in C$ and equals the number of codewords of weight i. We will see that perfect codes and MDS codes (neither assumed to be linear) are distance-invariant.

It will be useful to define some other quantities related to $A_i(C; x)$. Let $I_{j,n}$ denote the collection of subsets of $\{1, \ldots, n\}$ that have size j. Any such a subset can be represented as an increasing sequence $1 \leq i_1 < \cdots < i_j \leq n$. Let

$$A(C; x, I) = \#\{y \in C : y_j \neq x_j \text{ if and only if } j \in I\}.$$

We note that

$$A_i(C; x) = \sum_{I \in I_{i,n}} A(C; x, I).$$

We also define:

$$B(C; x, I) = \#\{y \in C : y_j = x_j \text{ if } j \in I\},$$

and

$$B_i(C; x) = \sum_{I \in I_{i,n}} B(C; x, I).$$

We take the convention that $B(C; x, I) = |C| = M$ if $I = \emptyset$. In particular $B_0(C; x) = M$. The quantities $B_i(C)$ are usually known in literature as the *binomial moments* of the code C.

Proposition 10. *The quantities* $(A_0(C; x), \ldots, A_n(C; x))$ *and* $(B_0(C; x), \ldots, B_n(C; x))$ *are related as*

$$\begin{aligned} B_{n-i}(C; x) &= \textstyle\sum_\ell \binom{n-\ell}{i-\ell} A_\ell(C; x) \\ A_i(C; x) &= \textstyle\sum_\ell B_{n-\ell}(C; x)(-1)^{i-\ell} \binom{n-\ell}{i-\ell} \end{aligned} \tag{7}$$

Proof. Let $X = \{1, \ldots, n\}$. Let $J = X \setminus I$. Let S_J denote the set in right side of the definition of $B(C; x, J)$ above. Clearly

$$B(C; x, J) = |S_J| = \sum_{K \subset I} A(C; x, K),$$

and hence

$$B_{n-i}(C; x) = \sum_{I \in I_{i,n}} \sum_{K \subset I} A(C; x, K) = \sum_{K} A(C; x, K) \tbinom{n-|K|}{i-|K|} = \textstyle\sum_k \tbinom{n-k}{i-k} A_k(C; x)$$

This proves the identity expressing $B_{n-i}(C; x)$ in terms of $A_j(C; x)$ for $j \leq i$. The other identity $A_i(C; x)$ in terms of the $B_j(C; x)$ for $j \geq n - i$ is a consequence of the first identity using the inclusion-exclusion principle from Combinatorics. There are several variants of this principle for example Lemma 11 and Lemma 12.

In order to apply Lemma 11, we take $v, w \in \mathbb{R}^{n+1}$ are given by $\vec{v} = (B_n(C; x), \ldots, B_0(C; x))$ and $\vec{w} = (A_0(C; x), \ldots, A_n(C; x))$ and rewrite the first identity in (7) as $\vec{v} = P\vec{w}$. Lemma 11 gives the second identity in (7):

$$A_i(C; x) = \textstyle\sum_\ell B_{n-\ell}(C; x)(-1)^{i-\ell} \tbinom{n-\ell}{i-\ell}.$$

To obtain this same result using Lemma 12, we take $X = \{1, \ldots, n\}$, $f(I) = A(C; x, I)$ and $f^\dagger(I) = B(C; x, X \setminus I)$ to obtain:

$$A(C; x, I) = \sum_{L \subset I} (-1)^{|I \setminus L|} B(C; x, X \setminus L) \tag{8}$$

Using (8) we get:

$$A_i(C; x) = \sum_{I \in I_{i,n}} \sum_{L \subset I} (-1)^{I \setminus L} B(C; x, B(C; x, X \setminus L) = \sum_{L} B(C; x, B(C; x, X \setminus L) \sum_{\{I : I \supset L, |I| = i\}} (-1)^{I \setminus L}.$$

This further simplifies to

$$A_i(C; x) = \sum_{\ell} B_{n-\ell}(C; x)(-1)^{i-\ell} \tbinom{n-\ell}{i-\ell}.$$

□

Lemma 11. *Let P be $(n+1) \times (n+1)$ lower triangular Pascal matrix whose ij-th entry is $\binom{n+1-j}{i-j}$. Since P is triangular with diagonal entries being 1, it follows that P is invertible. The ij-th entry of P^{-1} is just $(-1)^{i+j}$ times the ij-th entry of P.*

Proof. Let V be the $(n+1)$-dimensional vector space of polynomials in one variable X of degree at most n and real coefficients:

$$V = \{a_0 + a_1 X + \cdots + a_n X^n : a_0, \ldots, a_n \in \mathbb{R}\}.$$

The function $T : V \to V$ given by $T(f(X)) = f(X + 1)$ is a linear transformation. The matrix of T with respect to the basis $\{X^n, X^{n-1}, \ldots, X, 1\}$ is clearly P. The inverse transformation T^{-1} is given by $T^{-1}(f(X)) = f(X - 1)$. Since

$$(X-1)^{n+1-j} = X^{n+1-j} - (n+1-j)X^{n-j} + \cdots + (-1)^{n+1-j} X^0,$$

it follows that the ij-th entry of the matrix representing T^{-1} is $(-1)^{i+j} \binom{n+1-j}{i-j}$. □

Lemma 12. *Let X be a finite set, and 2^X the power set of X. To each function $f : 2^X \to \mathbb{R}$, we define another function $f^\dagger : 2^X \to \mathbb{R}$ by*

$$f^\dagger(K) = \sum_{L \subset K} f(L).$$

We can recover f from $f^\dagger$ by:

$$f(K) = \sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L)$$

Proof. Using the definition of $f^\dagger$ the right hand side is:

$$\sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L) = \sum_{L \subset K} (-1)^{|K \setminus L|} \sum_{J \subset L} f(J) = \sum_{J \subset K} f(J) \sum_{\{L : K \supset L \supset J\}} (-1)^{|K \setminus L|}$$

Let $\bar{K} = K \setminus J$ and let $\bar{L} = L \setminus J$. We have

$$\sum_{\{L : K \supset L \supset J\}} (-1)^{|K \setminus L|} = \sum_{\bar{L} \subset \bar{K}} (-1)^{|\bar{K} \setminus \bar{L}|} = \sum_{i=0}^{|\bar{K}|} \binom{|\bar{K}|}{i} (-1)^i = \delta_{0,|\bar{K}|}.$$

Since $\delta_{0,|\bar{K}|} = \delta_{J,K}$, we get:

$$\sum_{L \subset K} (-1)^{|K \setminus L|} f^\dagger(L) = \sum_{J \subset K} f(J) \delta_{J,K} = f(K)$$

□

Recall that the distance enumerator polynomial $W_C(Z; x)$ of a $[n, M, d]_q$ code with respect to a codeword $x \in C$ was defined to be

$$W_C(Z; x) = \sum_{y \in C} Z^{d(x,y)} = \sum_{i=0}^{n} A_i(C; x) Z^i$$

The next lemma allows us to describe this polynomial in terms of the quantities $B_i(C; x)$

Lemma 13.

$$\begin{aligned} Z^n W_C(1/Z; x) &= \textstyle\sum_{i=0}^n B_i(C; x) \, (Z-1)^i \\ Z^n W_C(1/Z) &= \textstyle\sum_{i=0}^n B_i(C) \, (Z-1)^i \\ W_C(Z) &= \textstyle\sum_{i=0}^n B_{n-i}(C) \, (Z-1)^i Z^{n-i} \end{aligned} \tag{9}$$

Proof. We will prove the first result. The second result follows from the first result by averaging over all $x \in C$. The third result is a restatement of the second result. For a polynomial $f(Z) \in \mathbb{R}[Z]$ we know that if $f_r = \frac{1}{r!} \frac{d^r}{dZ^r}|_{Z=1} f(Z)$ then $f(Z) = \sum f_i (Z-1)^i$. Applying this to $f(Z) = Z^n W_C(1/Z) = \sum_{i=0}^n A_i(C; x) Z^{n-i}$, the quantities

$$f_r = \tfrac{1}{r!} \tfrac{d^r}{dZ^r}|_{Z=1} f(Z) = \textstyle\sum_{i=0}^n A_i(C; x) \binom{n-i}{r} = B_r(C; x)$$

□

12.1. **The distance distribution of (non-linear) MDS codes.**
Let C be a $[n, M, d]_q$ code. Let $I \in I_{n-d+1,n}$ i.e., $I \subset \{1, \dots, n\}$ has size $n-d+1$. Consider the function $\pi_I : C \to \mathcal{F}^{n-d+1}$ given by $f(x) = (\dots, x_i, \dots)$ for $i \in I$. This function is injective for otherwise there will be a pair of distinct codewords x, y with $d(x, y) < d$. The function π_I is bijective if and only if $M = q^{n-d+1}$, i.e. if and only if C is MDS. Given $J \subset \{1, \dots, n\}$, if $|J| \leq n-d+1$, then picking $I \in I_{n-d+1}$ such that $I \supset J$ and considering the function π_I we see that for any $x \in C$, the set $S_J = \{y \in C : y_{|J} = x_{|J}\}$ has size $q^{n-d+1-j} = M/q^j$. On the other hand if $|J| > n-d+1$, then picking $I \in I_{n-d+1}$ with $I \subset J$, and considering the function π_I we see that for any $x \in C$, the set $S_J = \{y \in C : y_{|J} = x_{|J}\}$ equals $\{x\}$ and hence has size 1. Hence we get

Theorem 14. *For an $[n, M, d]_q$ MDS code, we have $B(C; x, I) = max\{M/q^{|I|}, 1\}$, which depends only on the parameters n and d and not on x or the code C. Thus*

$$B_i(C) = B_i(C; x) = \tbinom{n}{i} \cdot max\{M/q^i, 1\}.$$

In particular

$$W_C(Z) = \textstyle\sum_{i=0}^{n} B_{n-i}(Z-1)^i = Z^n + \sum_{i=d}^{n} (Z-1)^i \binom{n}{i} (q^{i+1-d} - 1).$$

The quantities $A_i(C)$ are given by $A_0(C) = 1$ and for $i > 0$:

$$A_i(C) = \tbinom{n}{i} \textstyle\sum_{\ell=0}^{i-d} (-1)^\ell \binom{i}{\ell} (q^{i-d-\ell+1} - 1) = \binom{n}{i} (q-1) \sum_{\ell=0}^{i-d} (-1)^\ell \binom{i-1}{\ell} q^{i-d-\ell}.$$

Proof. We have already proved the assertion about $B_i(C)$. Using (7) we get

$$A_i(C; x) = \textstyle\sum_\ell B_{n-\ell}(C; x)(-1)^{i-\ell} \binom{n-\ell}{i-\ell} = \sum_\ell \binom{n}{\ell} \cdot \max\{q^{\ell-d+1}, 1\}(-1)^{i-\ell} \binom{n-\ell}{n-i}.$$

This simplifies to

$$\tfrac{A_i(C;x)}{\binom{n}{i}} = \textstyle\sum_\ell (-1)^{i-\ell} \binom{i}{\ell} \cdot (1 + \max\{q^{\ell-d+1} - 1, 0\}) = \delta_{i,0} + \sum_{\ell=d}^{i} (-1)^{i-\ell} \binom{i}{\ell} (q^{\ell-d+1} - 1).$$

The other form of $A_i(C; x)$ is left as a combinatorial exercise (see HW). Finally we note that $A_i(C; x) = A_i(C)$ because $B_j(C; x) = B_j(C)$ for all j. □

13. The MacWilliams identities for Linear Codes

Let C be a linear code of length n and dimension k over $\mathbb{F}_q$. The next Theorem due to Florence Jessie MacWilliams (1963) shows that the knowledge of $W_C(Z)$ allows to determine $W_{C^\perp}(Z)$ and vice-versa.

Theorem 15. *(MacWilliams identities) . The following are equivalent:*

(1) (Binomial moments form) $B_i(\mathcal{C}^\perp) = q^{n-k-i}\, B_{n-i}(\mathcal{C})$.

(2) (Standard form)

$$W_{C^\perp}(Z) = \tfrac{(1+(q-1)Z)^n}{q^k}\, W_C(\tfrac{1-Z}{1+(q-1)Z})$$

Proof. Let $S_I = \{y \in C : y_{|I} = \vec{0}\}$. If G is a generator matrix of C and G_I is the $k \times |I|$ submatrix of G on the columns indexed by I, then S_I is the kernel of the linear map $\mathbb{F}_q^k \mapsto \mathbb{F}_q^{|I|}$ given by $a \mapsto aG_I$. Thus $|S_I| = q^{k-\mathrm{rank}(G_I)}$.

Let $J = \{1, \ldots, n\} \setminus I$ and consider

$$S_J(C^\perp) = \{y \in C^\perp : y_{|J} = \vec{0}\}.$$

Since $C^\perp = \ker(G)$, we can rewrite this as

$$S_J(C^\perp) = \{\vec{v} \in \mathbb{F}_q^{|I|} : G_I \vec{v} = \vec{0}\} = \ker(G_I).$$

Thus $|S_J(C^\perp)| = q^{|I|-\mathrm{rank}(G_I)}$. Thus , we have shown that

$$|S_I(C)| = q^{k-|I|}\, |S_J(C^\perp)|$$

Since $B(C;I) = |S_I|$ we have shown that $B(C;I) = q^{k-|I|}B(C^\perp;J)$. Summing over $I \in I_{j,n}$ for each j we get

$$B_j(C) = q^{k-j}B_{n-j}(C^\perp)$$

In order to prove the asserted relation between $W_C(Z)$ and $W_{C^\perp}(Z)$ we first recall (from (9)) that:

$$Y^n W_{C^\perp}(1/Y) = \textstyle\sum_{i=0}^n B_{n-i}(C^\perp)\,(Y-1)^{n-i}.$$

Using the fact that $B_{n-i}(C^\perp) = q^{i-k}B_i(C)$ we get

$$Y^n W_{C^\perp}(1/Y) = \sum_{i=0}^n B_i(C) q^{i-k}\,(Y-1)^{n-i} = \tfrac{(Y-1)^n}{q^k} \sum_{i=0}^n B_i(C)\,(\tfrac{q}{Y-1})^i$$

In terms of $Z = 1/Y$ we get:

$$W_{C^\perp}(Z) = \tfrac{(1-Z)^n}{q^k} \sum_{i=0}^n B_i(C)\,(\tfrac{qZ}{1-Z})^i$$

Let T be defined by $T - 1 = \frac{qZ}{1-Z}$, so that the right side of the above equation is

$$\tfrac{(1-Z)^n}{q^k} T^n W_C(1/T) = \tfrac{(1+(q-1)Z)^n}{q^k} W_C(\tfrac{1-Z}{1+(q-1)Z})$$

□

As an application consider the problem of determining $W_C(Z)$ for the $[\frac{q^m-1}{q-1}, q^{\frac{q^m-1}{q-1}-m}, 3]_q$ linear Hamming codes. Since we already know

$$W_{C^\perp}(Z) = 1 + Z^{q^{m-1}}(q^m - 1)$$

we can use MacWilliams identities to obtain:

$$W_C(Z) = \frac{(1+(q-1)Z)^{\frac{q^m-1}{q-1}}}{q^m} W_{C^\perp}(\frac{1-Z}{1+(q-1)Z}) = \frac{(1+(q-1)Z)^{\frac{q^m-1}{q-1}}}{q^m} \left(1 + (q^m-1)(\frac{1-Z}{1+(q-1)Z})^{q^{m-1}}\right)$$

This simplifies to

$$W_C(Z) = \frac{(1+(q-1)Z)^{\frac{q^{m-1}-1}{q-1}}}{q^m} \left((1+(q-1)Z)^{q^{m-1}} + (q^m-1)(1-Z)^{q^{m-1}}\right)$$

For example, let C be the linear $[7, 2^4, 3]_2$ Hamming code. We get:

$$W_C(Z) = \frac{(1+Z)^3}{8} \left((1+Z)^4 + 7(1-Z)^4\right) = \frac{(1+Z)^3}{8} \left(8 - 24Z + 48Z^2 - 24Z^3 + 8Z^4)\right)$$

This simplifies to

$$W_C(Z) = 1 + 7(Z^3 + Z^4) + Z^7$$

14. Assignment 5

(1) Consider the matrix $\begin{pmatrix} 1 & 6 & 2 & 5 & 1 \\ 1 & 4 & 3 & 3 & 6 \\ 1 & 5 & 5 & 1 & 5 \end{pmatrix}$ over $\mathbb{F}_7$. Check that every square submatrix of A is non-singular.

(2) Prove that every 2-dimensional MDS code is monomially equivalent to a Reed-Solomon code.

(3) Show that a linear MDS code of length n and dimension k, with $k \geq q$ exists only if $n = k, k+1$. Describe these codes.

(4) Prove the other formula for $A_i(C)$ for MDS codes C given in Theorem 14.

(5) Let $C_1 \subset \mathbb{F}_q^{n_1}$ and $C_2 \subset \mathbb{F}_q^{n_2}$ be $[n_1, q^{k_1}, d_1]_q$ and $[n_2, q^{k_2}, d_2]_q$ linear codes. Let $C = C_1 \oplus C_2 \subset \mathbb{F}_q^{n_1} \oplus \mathbb{F}_q^{n_2}$. It is a$(k_1 + k_2)$-dimensional linear code of length $n_1 + n_2$.
   (a) Show that $d_C = \min\{d_1, d_2\}$.

   (b) Show that $W_C(Z) = W_{C_1}(Z) \cdot W_{C_2}(Z)$.

(6) Let R denote either the field of real numbers or an arbitrary ring (your choice). Let $V = \{f(X) \in R[X] : \deg(f) \leq n\}$. Note that V has a basis $1, X, \ldots, X^n$. Let $D_r : V \to V$ be the linear map (called the r-th Hasse derivative) defined on the basis elements by $D_r X^m = \binom{m}{r} X^{m-r}$. If $a \in R$ show that

$$f(X) = \sum_{j=1}^{n} a^j (D_j f)_{|X=a} (X-a)^j$$

   (Just expand $X^i = (X - a + a)^i$ in powers of $(X - a)$)

(7) Find the weight enumerator of the binary code whose generator matrix is $\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$.
   (a) directly,
   (b) by using the MacWilliams identity.

(8) Let $C \subset \mathbb{F}_q^n$ be a k-dimensional linear code. The code is being used over a communication channel which corrupts a bit with probability t where $0 \leq t < 1$. For $c \in C$, the probability $P(c \text{ is sent}) = 1/q^k$ is the same for all $c \in C$. If a codeword c is sent and a word y is received, consider the decoding scheme where the decoder returns y if $y \in C$ and returns "error" if $y \notin C$. We say we have a decoding result if decoder does not return "error". Prove that the probability of decoding result is:

$$(1-t)^n W_C(\tfrac{t}{(q-1)(1-t)}).$$

   (Hint: determine the conditional probability $P(c' \text{ is received} \mid c \text{ is sent})$ for $c' \in C$.)

15. Quiz 2 (with answers) Feb 17

(1) (7 points) Let C be a $[n, M, d]_q$ MDS code. Determine $A_d(C)$ and $A_{d+1}(C)$.
Ans: For any code C, the relation between B_{n-i} and A_j's for $j \leq i$ is easy to remember: it follows from the definitions that for $J \subset \{1, \ldots, n\}$ of size $n-i$, we have $B_{n-i}(C; J, x) = \sum_{I \subset \{1,\ldots,n\} \setminus J} A(C; I, x)$. Averaging over $x \in C$ and summing over J, we get:

$$B_{n-i}(C) = \textstyle\sum_{\ell \leq i} A_\ell(C) \binom{n-\ell}{i-\ell}.$$

For MDS codes C, we know $B_\ell(C) = \binom{n}{\ell} M/q^\ell$ for $0 \leq \ell \leq n-d+1$. Taking $i = d, d+1$ we get

$$\begin{aligned} \binom{n}{d} q &= B_{n-d} = A_d(C) + A_0(C)\binom{n}{d} \\ \binom{n}{d+1} q^2 &= B_{n-d-1} = A_{d+1}(C) + A_d(C)\binom{n-d}{1} + A_0(C)\binom{n}{d+1} \end{aligned}$$

Since $A_0(C) = 1$, we get $A_d(C) = (q-1)\binom{n}{d}$. Using this, we get

$$A_{d+1}(C) = \binom{n}{d+1}(q^2-1) - (q-1)\binom{n}{d}(n-d) = \binom{n}{d+1}\left[(q^2-1) - (d+1)(q-1)\right]$$

(2) (7 points) Show that a linear MDS code of length n and dimension k, with $k \geq q$ exists only if $n = k, k+1$. Describe the duals of these codes.
Ans: Suppose there exists an MDS code with $k \geq q$ and $n \geq k+2$. Then there exists such an MDS code with $n = k+2$. Let $G = [I_k|A]$ be a generator matrix for this code with all 1×1 and 2×2 minors of the $k \times 2$ matrix A being non-zero. Suppose the columns of A are $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$. Note that $a_1/b_1, a_2/b_2, \ldots, a_k/b_k$ are distinct elements of $\mathbb{F}_q^\times$ because otherwise a minor of the form $a_i b_j - a_j b_i$ will be zero. Since $|\mathbb{F}_q^\times| = q-1$ we get $k \leq q-1$. This contradiction shows that there is no MDS code of length $n \geq k+2$ if $k \geq q$.

If $n = k$ then $C^\perp = \{0\} \subset \mathbb{F}_q$. If $n = k+1$, then $C^\perp$ is monomially equivalent to the repetition code generated by $1 \times (n+1)$ matrix $[1, 1 \ldots, 1]$.

(3) (7 points) Let C be the linear code over $\mathbb{F}_2$ generated by the matrix $G = [I_6|u|v]$ where $u^T = (0, 0, 1, 1, 1, 1)$ and $v^T = (1, 1, 0, 1, 1, 1)$.
(a) Determine $W_C(Z)$ the weight enumerator polynomial of C. You need not simplify your answer.
(b) Determine the number of elements of C of weight 2.

Ans: The parity check for C and the generator matrix for $C^\perp$ is $\left(\begin{smallmatrix} u^t & 1 & 0 \\ v^t & 0 & 1 \end{smallmatrix}\right)$. Thus $C^\perp$ has 4 codewords with weights $0, 5, 5, 6$. This gives

$$W_{C^\perp}(Z) = 1 + 2Z^5 + Z^6$$

By Macwilliams identity

$$W_C(Z) = \tfrac{(1+Z)^8}{4}\left(1 + 2(\tfrac{1-Z}{1+Z})^5 + (\tfrac{1-Z}{1+Z})^6\right)$$

The coefficient of Z^2 in this polynomial is 7.

16. Mid-Semester Exam

Question 1
a) Prove that a linear code of length n and dimension k has covering radius at most $n-k$.
b) Let $C = \{(f(x_1), \ldots, f(x_n)) : f \in \mathbb{F}_q[X], \deg(f) \leq k-1\}$ be a Reed-Solomon code over $\mathbb{F}_q$ of length n and dimension k, where $x_1, \ldots, x_n$ are n distinct elements of $\mathbb{F}_q$. What is the distance of the word $u = (x_1^k, x_2^k, \ldots, x_n^k)$ from C?
c) Determine the covering radius of the code C in part b).

Question 2 Let C be a binary code with generator matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

a) Decode (i.e. find a nearest codeword to) the following received words.
(1) (1101011)
(2) (0110111)
(3) (0111000)
b) Determine the covering radius of C.
c) Determine the minimum distance of C.

Question 3 Consider the $3 \times (q+2)$ matrix G below. Do there exist a, b, c such that G generates a linear MDS code of dimension 3 and length $q+2$ over $\mathbb{F}_q$. If yes then determine all such a, b, c. (The answer will be different for even and odd characteristic).

$$G = \begin{bmatrix} a & 0 & 1 & 1 & \ldots & 1 \\ b & 0 & 0 & x_1 & \ldots & x_{q-1} \\ c & 1 & 0 & x_1^2 & \ldots & x_{q-1}^2 \end{bmatrix}$$

Here $\{x_1, \ldots, x_{q-1}\} = \mathbb{F}_q^\times$.

Question 4 For an integer $m \geq 2$, let C be the Hamming code over $\mathbb{F}_2$ of length n and dimension $n-m$ where $n = 2^m - 1$. Recall that the minimum distance of C is 3. Let $A_i(C)$ denote the number of codewords of C of Hamming weight i.
a) Determine $A_3(C)$.
b) Determine (with full proof) $A_n(C)$.
c) Determine $A_{n-1}(C), A_{n-2}(C)$ and $A_{n-3}(C)$

Answers:
Problem 1: a) We know that for a $[n,k]_q$ linear code with parity check matrix H, the covering radius is the least integer j such that any $u \in \mathbb{F}_q^{n-k}$ is expressible as a linear combinations of some j

columns of H. Since H has rank $n-k$, every u can be expressed as a linear combinations of $n-k$ columns of H. Thus the covering radius is at most $n-k$.

b) By part a) the distance of the given word u from C is at most $n-k$. In case it is less than $n-k$ then, there is a codeword $(f(x_1),\dots,f(x_n))$ (where $\deg(f) \le k-1$) such that $(f(x_1)-x_1^k,\dots,f(x_n)-x_n^k)$ has more than k zeros. But then the polynomial $X^k - f(X)$ is a non-zero polynomial with more roots than its degree, which is impossible. Thus the distance of u from the code is exactly $n-k$.

c) by parts a) and b), the covering radius of C is exactly $n-k$.

Problem 2: The parity check matrix is

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

The syndromes of the three received words are a) $(0,0,0)$, b) $(0,0,1)$ and c) $(1,0,1)$. The distance of a received word $\vec{u}$ from the code is the least integer j such that the syndrome Hu can be expressed as a linear combination of j columns of H. (In particular the distance is zero, i.e. $\vec{u} \in C$ if and only if $Hu = 0$). Thus
1) u is a codeword. it is decoded to itself.
2) $d(u,C) = 1$ and u is decoded to the codeword $u - e_7$.
3) $d(u,C) = 1$ and u can be decoded to either $u - e_1$ or $u - e_2$.

b) and c) the covering radius $\rho(C) = 2$ and the minimum distance $d(C) = 2$. This can be read from the matrix H using the characterization of these numbers in terms of H.

Problem 3: The minor on the first three columns is zero if and only if $b = 0$, therefore $b \neq 0$. Also $b - ax$ and $cx - bx^2$ are 3×3 minors of G for each $x \in \mathbb{F}_q^\times$. If $a \neq 0$ then $b - ax = 0$ for $x = b/a$. If $c \neq 0$ then $cx - bx^2 = 0$ for $x = c/b$. Thus $a = c = 0$. Using this we note that $b(x^2 - y^2)$ are 3×3 minors for $x \neq y$. Since $b \neq 0$, this means $x \neq -y$ for all pairs $x \neq y$ in $\mathbb{F}_q^\times$. If q is odd we can take $x = 1 = -y$ to conclude that there are no solutions (a,b,c). If q is even then $x \neq y$ already implies $x \neq -y$ because $y = -y$. Thus all vectors $(0,b,0)$ for $b \neq 0$ are solutions.

Problem 4: a) Codewords of weight 3 of the given code $\mathcal{C}$ are in bijective correspondence with unordered triples of distinct columns of a parity check matrix H, such that the triple is linearly dependent. For H we can take the matrix whose columns are all the nonzero vectors of $\mathbb{F}_2^m$. For every pair of columns of H, there is a unique third column of H such that the triple is dependent. Thus there are $\frac{1}{3}\binom{n}{2} = \frac{n(n-1)}{6}$ such triples.
b) We must show that the sum S of all non-zero vectors (or equivalently all vectors) of $\mathbb{F}_2^m$ is the zero vector. We do this by induction on m. Since $m \ge 2$, the base case is $m = 2$: in this case $S = \binom{0}{1} + \binom{1}{0} + \binom{1}{1} = \binom{0}{0}$. Assume the result is true for $2 \le m \le \ell - 1$. When $m = \ell$, it is easy to see that

$$S(\ell) = \binom{S(\ell-1)}{0} + \binom{S(\ell-1)}{1\cdot(2^{\ell-1})} = \binom{2S(\ell-1)}{2^{\ell-1}} \equiv \vec{0}$$

c) Suppose there is a codeword c of Hamming weight $n-j$ where $j \in \{1,2,3\}$. Then $0 = Hc = S - S'$, where S is the sum of all columns of H, and S' is a sum of some j columns of H. Since $S = 0$ as shown in part b), we get $S' = 0$. This shows that codewords of weight $n-j$ are in bijective correspondence with codewords of weight j. Thus there are no codewords of weight $n-1, n-2$ and there are $n(n-1)/6$ codewords of weight $n-3$.

17. Finite Fields -II

For any field F, the set of non-zero elements of F denoted $F^\times$ is an abelian group with respect to multiplication. For a finite field $\mathbb{F}_q$ the group $\mathbb{F}_q^\times$ is actually a cyclic group of order $q-1$. In fact we will prove a stronger result:

Theorem 16. *Let F be an arbitrary field and let G be a finite subgroup of the multiplicative group $F^\times$. The group G is cyclic.*

Proof. Let $n = |G|$. It suffices to show that there exists an element of order n in G. Let

$$\psi(d) = \#\{g \in G : \text{ order of } g \text{ is d}\}.$$

Since the order of any element of G is a divisor d of n, we note that $\sum_{d|n} \psi(d) = |G| = n$.
Let C_n denote the cyclic group of order n. It is an elementary exercise to show that

$$\phi(d) = \#\{g \in C_n : \text{ order of } g \text{ is } d\} = \#\{j : 1 \le j \le d : \gcd(j,d) = 1\}.$$

Again $\sum_{d|n} \phi(d) = |C_n| = n$. Thus

$$\sum_{d|n} (\phi(d) - \psi(d)) = 0.$$

Note that for $g \in G$, the order of g is a divisor of d if and only g is a root of the polynomial $f(X) = X^d - 1$ over F. Since F is a field, $f(X)$ has at most d roots. In case there exists an element $g \in G$ of order d, we note that $f(X) = X^d - 1$ has exactly d roots given by $\langle g \rangle := \{1, g, g^2, \ldots, g^{d-1}\}$: the subgroup of G generated by g. Thus the set of elements of order d in G is the set of elemets of order d in $\langle g \rangle$ which is $\phi(d)$ because $\langle g \rangle$ is cyclic of order d. Thus $\psi(d) > 0$ implies $\psi(d) = \phi(d)$. In other words either $\psi(d) = 0$ or $\psi(d) = \phi(d)$. It follows that each term in the left side of the identity $\sum_{d|n}(\phi(d) - \psi(d)) = 0$, is non-negative. Therefore, $\phi(d) = \psi(d)$ for all $d|n$. In particular $\psi(n) = \phi(n) > 0$ as was to be shown. □

As an application, we have:

Lemma 17.

$$\sum_{x \in F^\times} x^j = \begin{cases} 0 & \text{if } (q-1) \nmid j \\ -1 & \text{if } (q-1) \mid j \end{cases}$$

Proof. Let S denotes the left hand side of this identity. Let g be a generator for the cyclic group $\mathbb{F}_q^\times$. We have $S = \sum_{i=0}^{q-2} g^{ji}$. Therefore

$$S(g^j - 1) = g^{j(q-1)} - 1 = 0,$$

because $g^{q-1} = 1$. If $(q-1) \nmid j$ then $g^j - 1 \neq 0$ and hence $S = 0$. On the other hand if $(q-1) \mid j$ then $S = \sum_{i=0}^{q-2} g^{ji} = \sum_{i=0}^{q-2} 1 = q - 1 = -1$. □

A variant of the above lemma is as follows:

Lemma 18. *Suppose $0 \le m \le q-2$ and following the convention that $0^0 = 1$ (in $\mathbb{F}_q$) we have:*

$$S_m = \sum_{x \in \mathbb{F}_q} x^m = 0, \qquad 0 \le m \le q-2. \tag{10}$$

Since $\mathbb{F}_q^\times$ is cyclic, we have $x^{q-1} = 1$ for all $x \in \mathbb{F}_q^\times$ and hence $x^q = x$ for all $x \in \mathbb{F}_q$. This shows that

$$X^q - X = \prod_{x \in \mathbb{F}_q} (X - x), \qquad \text{in } \mathbb{F}_q[X] \tag{11}$$

Theorem 19. *Suppose F is a field of characteristic a prime number p. The function $\phi : F \to F$ given by $\phi(x) = x^p$ is an injective homomorphism of fields. It is known as the Frobenius map. If F is finite, then ϕ is an automorphism of F. More generally, if $im(\phi) = F$ then ϕ is an automorphism.*

Proof. Clearly $\phi(xy) = \phi(x)\phi(y)$, $\phi(1) = 1$ $\phi(0) = 0$. It remains to check that $\phi(x+y) = \phi(x)+\phi(y)$. Since

$$\phi(x + y) - \phi(x) - \phi(y) = \sum_{i=1}^{p-1} \tbinom{p}{i} x^i y^{p-i},$$

our result follows by observing that for $1 \le i \le p-1$, the integer $\binom{p}{i} = p!/(i!(p-i)!)$ has p in the numerator but not in the denominator. Thus $\binom{p}{i} = 0$ in F when $1 \le i \le p-1$. □

Theorem 20. *Let p be a prime number. For each q which is a power of p, there exists a field $\mathbb{F}_q$ of size q which is unique upto isomorphism.*

It will take some time to prove this theorem. First we consider the polynomial ring $F[X]$ where F is an arbitrary field. We will first explore some properties of $F[X]$ using analogies with the ring $\mathbb{Z}$. So let R denote either of the rings $\mathbb{Z}$ or $F[X]$. In R we have

Euclidean division with remainder. Given $a, b \in R$ with $a \neq 0$, we can write $b = aq + r$ for unique $q, r \in R$ with the property that the 'remainder' r satisfies the condition: either $r = 0$ or $0 < \text{size}(r) < \text{size}(a)$.

Here the 'size' is defined on $R \setminus 0$ as $|a|$ if $R = \mathbb{Z}$, and if $R = F[X]$ then $2^{\deg(a)}$ works as a size function. We could use $\deg(a)$ itself as a size function, but the degree of a product is the sum of the degrees, where as in $\mathbb{Z}$ the size of a product is the product of the sizes. Therefore, to maintain uniformity between $\mathbb{Z}$ and $F[X]$ we choose the exponent of deg (the base could be any positive number for example 2). The Euclidean long division is still true with this definition of size.

We say $c \in R$ is irreducible if c cannot be written as $c = ab$ for any a, b of positive size. If c is reducible, then the size of the factors a, b of $c = ab$ is strictly less than the size of c. Therefore we can always find a factorization $c = f_1^{a_1} f_2^{a_2} \dots f_r^{a_r}$ with f_i's irreducible and a_i non-negative integers. We will show that any c can be uniquely factored as $c = u f_1^{a_1} f_2^{a_2} \dots f_r^{a_r}$ where

- f_i's are i) positive irreducibles if $R = \mathbb{Z}$, and ii) monic irreducibles if $R = F[X]$ (recall that $g(X)$ is monic if the coefficient of the highest power of X in $g(X)$ is 1).
- $u \in U$ – the set of *units* of R, i.e. those elements of R which have a multiplicative inverse. We note that $U = \{\pm 1\}$ for $R = \mathbb{Z}$ and $U = F^\times$ for $R = F[X]$.

We note that for $R = Z, F[X]$ the set of elements of size 1 is precisely the set of units U. To establish uniqueness of factorizations, we first define a notion closely related to irreducibility: we say an element $a \in R$ is prime if $a|bc$ implies $a|b$ or $a|c$. Next we show that the two notions are identical in R. Suppose a is prime and $a = bc$. Then $a|bc$ and hence a divides one of b and c, say b. Writing $b = ab'$ we get $a(1 - b'c) = 0$. Since a product of two non-zero elements in R (either $\mathbb{Z}$

or $F[X]$) is non-zero, it follows that either $a = 0$ or $b'c = 1$. Since $a \neq 0$, we must have $c \in U$ (note that U is preciesly the set of size 0 elements of $R \setminus \{0\}$), and hence a is irreducible. The reverse direction that if a is irreducible, then it is prime requires a different idea :

Let $S \subset R$ satisfying the following two properies

- S is closed with respect to addition,
- if $a \in S$ and $r \in R$ then $ar \in S$.

Such a subset $S \subset R$ is called an *ideal* of R. Let d be an element of S of least size, then $S = d \cdot R$. To see this suppose there exists $a \in S$ with $d \nmid a$, then writing $a = dq + r$ with $0 < \text{size}(r) < d$, it follows that $r \in S$ contradicting the fact that no element of S has size smaller than the size of d.

Suppose a is irreducible and $a|bc$ with $a \nmid b$. Let $S = \{ax + by : x, y \in R\}$. This is an ideal of $R$and hence $S = dR$ for some $d \in R$. Since $a, b \in S$, we have $d|a$ and $d|b$. Since a is irreducible, we must have $d = au$ or $d = u$ for some unit u. Since $d|b$ and $a \nmid b$, it follows that d is a unit and hence $S = dR = R$. In particular $1 \in S$. i.e. there exist $x, y \in R$ with $ax + by = 1$. Multiplying this with c, we get $acx + bc = c$. Since $a|bc$, it follows that $a|(acx + bc)$ and hence $a|c$.

Returning to the uniqueness of factorizations in R, suppose $h = u\Pi_{i=1}^{r} f_i^{a_i} = u'\Pi_{j=1}^{s} g_j^{b_j}$ are two factorizations into irreducibles of some $h \in R$. Since f_i's are ireducible and $f_i|\Pi_{j=1}^{s} g_j^{b_j}$. It follows that $f_i|g_j$ for some j. Since g_j is irreducible, it follows that $\{f_1, \ldots, f_r\} \subset \{g_1, \ldots, g_s\}$. Arguing similarly, we get $\{g_1, \ldots, g_r\} \subset \{f_1, \ldots, f_s\}$. It follows that $r = s$, $\{f_1, \ldots, f_r\} = \{g_1, \ldots, g_s\}$, and the corresponding exponents are identical, and $u = u'$.

We now recall the ring $\mathbb{Z}/n\mathbb{Z}$ (integers mod n) where n is a positive integer. More generally, for $f \in R \setminus \{0\}$, let $R/(f)$ be the set of equivalence classes of R under the equivalence relation, $x \sim y$ if $f|(y - x)$. For $x \in R$, we denote by $[x]$ the class of x in $R/(f)$. As a simple exercise check that the operation $[x] + [y] := [x + y]$ and $[x][y] := [xy]$ where $x \in [x]$ and $y \in [y]$, are independent of the choices of x, y, and thus are well-defined. Thus $R/(f)$ is itself is a ring the given addition and multiplication. This is summarized by saying that the map from $R \to R/(f)$ taking x to its class $[x]$ is a ring homomorphism.

Lemma 21. *If $f \in R$ is irreducible, then the ring $R/(f)$ constructed above is a field.*

Proof. We must show that any $[g] \neq [0]$ has a multiplicative inverse $[h]$ in $R/(f)$. This is equivalent to showing that if $f \nmid g$ then there exists $h \in R$ with $gh = 1 + af$ for some $a \in R$. This follows from the fact (established above) that for f, g such that f is irreducible and $f \nmid g$, the set $S = \{xf + yg : x, y \in R\} = R$. □

Now, let K_0 be a base field of size $q = p^m$ where p is a prime. A good example to keep in mind is just $m = 1$ i.e. $K_0 = \mathbb{F}_p$. Let $f(X) \in K_0[X]$ is an irreducible polynomial of degree n, and let $K = K_0[X]/(f)$ be the field constructed above. Note that the map $\imath : K_0 \to K$ that takes $a \in K_0$ to $[a] \in K$ is an isomorphism of K_0 onto a subfield of K. In this way we regard K_0 as a subfield of K. It is easy to check that the classes of $1, X, \ldots, X^{n-1}$ (where $n = \deg(f)$) span K as a vector space over K_0 and are also linearly independent over K_0. Thus the dimension of K over K_0 is n. This shows that K is a finite field of size q^n containing K_0. We will show that there always exists such

an irreducible polynomial $f(X) \in K_0[X]$ for each natural number n, thus showing the existence of fields of size q^n for each n.

Lemma 22. *Let $K_0 \subset K = K_0[X]/(f)$ be as above, let θ denote the class of X in K. We have*

$$f(X) = (X-\theta)(X-\theta^q)\dots(X-\theta^{q^{n-1}}), \quad \text{in } K[X]$$

Proof. Let $\psi : K_0 \to K_0$ be the automorphism $\psi(x) = x^q$. (If $q = p^m$ then $\psi = \phi^m$). Let $f(X) = \sum_{i=0}^n a_i X^i$ with $a_i \in K_0$. Since $\psi(a_i) = a_i$, we note that

$$0 = \psi(0) = \psi(f(\theta)) = \psi(\sum a_i\theta^i) = \sum a_i(\psi(\theta))^i,$$

which shows that $\psi(\theta), \psi^2(\theta), \psi^3(\theta), \dots$ are roots of $f \in K[X]$. There is a smallest positive integer such that $\psi^r(\theta) = \theta$. It follows from the fact that $a^{q^n} = a$ for all $a \in K$, that ψ^n is the identity map on K. Therefore $r \leq n$. The fact that $\psi^r(\theta) = \theta$ implies that $\psi^r(a) = a$ for all $a \in K$ (because a can be written as $a_0 + a_1\theta + a_2\theta^2 + \dots a_{n-1}\theta^{n-1}$) and hence all the q^n elements of K are roots of $X^{q^r} - X$. This shows that $r = n$. □

Theorem 23. *Let $h(X) \in K_0[X]$ be the polynomial $X^{q^n} - X$.*

- *every irreducible factor f of h satisfies $\deg(f)|n$.*
- *every irreducible $g(X) \in K_0[X]$ of degree $d|n$ is a factor of $h(X)$.*

Proof. Let $f \in K_0[X]$ be an irreducible factor of $h(X)$ of degree d. Let $K = K_0[X]/(f)$ and let θ denote the class of X in K. Since $f|h$ and $f(X) = \prod_{i=0}^{d-1}(X - \psi^i(\theta))$ in $K[X]$ as proved in the lemma above, it follows that θ is a root of h, i.e. $\psi^n(\theta) = \theta$. We have noted in the previous lemma that d is the least r such that $\psi^r(\theta) = \theta$. It follows that $d|n$.

Let $g(X) \in K_0[X]$ be irreducible of degree $d|n$. Since $d|n$, we know that $X^{q^d} - X$ divides $X^{q^n} - X$ in $K_0[X]$. Therefore, in order to show that $g|h$, it suffices to show that $g|(X^{q^d} - X)$ in $K_0[X]$. Let $L = K_0[X]/(g)$. We note that $X^{q^d} - X = \prod_{a\in L}(X - a)$ in $L[X]$, and $g(X) = \prod_{i=0}^{d-1}(X - \psi^i(\theta))$ in $L[X]$. Therefore $g|(X^{q^d} - X)$ in $L[X]$. It is easy to check that if $a(X), b(X)$ are polynomials over a field F, such that $a(X)|b(X)$ in $F'[X]$ for some extension field F' of F, then in fact $a(X)|b(X)$ in $F[X]$. Therefore, $g|X^{q^d} - X$ in $K_0[X]$. □

For an irreducible factor $f(X)$ of $h(X) \in K_0[X]$ as above, we claim that the exponent of f in the unique factorization of $h(X)$ is 1. Suppose not. Then $(X-\theta)^2|h(X)$ in $K[X]$ where θ denotes the class of X in $K = K_0[X]/(f)$. Writing $h(X) = (X-\theta)^2 h_1(X)$ in $K[X]$, we note that the derivative

$$h'(X) = 2(X-\theta)h_1 + (X-\theta)^2 h_1'(X),$$

also has θ as a root over K. However $h'(X) = q^n X^{q^n-1} - 1 = -1$ has no roots over K. This contradiction shows that $f^2 \nmid h$. We have now established that

Lemma 24.

$$X^{q^n} - X = \prod_{d|n} \prod_{\{f\in K_0[X]: f \text{ is monic irreducible of degree } d\}} f(X)$$

Let a_d denote the number of irreducible polynomials of degree d in $K_0[X]$. It follows from the result above that

$$\sum_{d|n} d a_d = q^n.$$

The Mobius inversion formula from Combinatorics/elementary number theory shows that

Theorem 25 (Gauss)**.**

$$na_n = \sum_{d|n} \mu(d) q^{n/d}.$$

Here the Mobius function $\mu(d)$ is defined as follows. If $d = p_1^{a_1} \dots p_r^{a_r}$ is the prime factorization of d, then $\mu(d) = 0$ if some $a_i > 1$, and $\mu(d) = (-1)^r$ if all $a_i = 1$. We also take $\mu(1) = 1$. It remains to show that $a_n > 0$. We write the right side of the above formula for na_n as $A - B$ where $A = \sum q^{n/d}$ with d running over products of an even number of distinct prime factors of n and $B = \sum q^{n/d}$ with d running over products of an odd number of distinct prime factors of n. Recall that the base q representation of an integer m is of the form $m = a_0 + a_1 q + \cdots + a_r q^r$ with $0 \leq a_i \leq q-1$. If $a_r = 1$ then m is strictly greater than any integer $b_0 + b_1 q + \cdots + b_{r-1} q^{r-1}$. In the case at hand $A = a_0 + a_1 q + \cdots + q^n$ and $B = b_0 + b_1 q + \cdots + b_{n-1} q^{n-1}$ (with a_i's and b_i's in $\{0, 1\}$) which shows that $A > B$.

17.1. **Uniqueness of finite fields upto isomorphism.** Let K_1 and K_2 be fields of size q^n containing $K_0 = \mathbb{F}_q$. By the previous lemma, there does exist a monic irreducible polynomial $f(X) \in K_0[X]$ of degree n and $f(X) | (X^{q^n} - X)$. Since $(X^{q^n} - X) = \prod_{a \in K_1} (X - a) = \prod_{b \in K_2} (X - b)$, we know that there exist roots $\alpha_i \in K_i$ Consider the homomorphisms $F[X] \mapsto K_i$ given by $X \mapsto \alpha_i$. The kernel of these homomorphisms is an ideal of $F[X]$ and hence of the form $g(X) \cdot F[X]$ for some monic $g(X)$. Note $g(X) \neq 1$ because K_i is finite. Since the irreducible polynomial $f(X)$ is in the kernel and $g | f$ it follows that $g = f$. The induced homorphisms $F[X]/(f) \to K_i$ are then injective and hence bijective (because both sides of the homomorphism have the same size q^n). Thus K_1 and K_2 are both isomorphic to $F[X]/(f)$.

18. Duals of Reed-Solomon codes

We now turn to study the duals of Reed-Solomon codes. Let

$$c_k(x) = \begin{cases} (1, x, \ldots, x^{k-1})^T & \text{if } x \in \mathbb{F}_q \\ (0, 0, \ldots, 0, 1)^T & \text{if } x = \infty \end{cases}.$$

For D an ordered subset $(x_1, \ldots, x_n)$ of $\mathbb{F}_q \cup \infty$, we define the generalized Reed-Solomon code (GRS code in short) of dimension k associated with the evaluation set D and column multipliers $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n) \in (\mathbb{F}_q^\times)^n$ to be the code generated by the matrix

$$G_k(D; \vec{\lambda}) = [\lambda_1 c_k(x_1) | \lambda_2 c_k(x_2) | \ldots | \lambda_n c_k(x_n)].$$

Recall that Reed-Solomon codes are exactly GRS codes satisfying $\vec{\lambda} = (1, 1, \ldots, 1)$.

Let us begin with RS codes of length $q+1$ and length q. If D has size $q+1$ (i.e. $D = (x_1, \ldots, x_{q+1})$ is an ordering of $\mathbb{F}_q \cup \infty$) and $\vec{\lambda} = 1$ we denote the matrix $G_k(D; \vec{\lambda})$ as just

$$G_k = \begin{pmatrix} 1 & 1 & \ldots & 1 & 0 \\ x_1 & x_2 & \ldots & x_q & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{k-2} & x_2^{k-2} & \ldots & x_q^{k-2} & 0 \\ x_1^{k-1} & x_2^{k-1} & \ldots & x_q^{k-1} & 1 \end{pmatrix}.$$

Lemma 26. *For $1 \le k \le q$ we have $G_k \cdot G_{q+1-k}^T = 0_{k \times (q+1-k)}$ and hence G_{q+1-k} is the parity check matrix for the code generated by G_k. Thus the dual of the RS code of length $(q+1)$ is a RS code of length $(q+1)$.*

The dual of a k-dimensional RS code with evaluation set $D = \mathbb{F}_q$ is a $(q-k)$-dimensional RS code with evaluation set $D = \mathbb{F}_q$.

Proof. Let a_{ij} denote the (ij)-th entry of $G_k \cdot G_{q+1-k}^T$. For $1 \le i \le k$ and $1 \le j \le q+1-k$ we have $0 \le i+j-2 \le q-1$ with $i+j-2 = q-1$ if and only if $(i,j) = (k, q+1-k)$. If $(i,j) = (k, q+1-k)$ then $a_{ij} = 1 + \sum_{x \in \mathbb{F}_q^\times} x^{q-1} = 1 - 1 = 0$. If $(i,j) \neq (k, q+1-k)$ then, by (10) we have $a_{ij} = S_{i+j-2} = 0$ because $1 \le i+j-2 \le q-2$.

Let M_k be the matrix $[c_k(x_1) | \ldots | c_k(x_q)]$. Let b_{ij} denote the (ij)-th entry of $M_k M_{q-k}^T$. For $1 \le i \le k$ and $1 \le j \le q-k$ we have $0 \le i+j-2 \le q-2$ and hence $b_{ij} = S_{i+j-2} = 0$ by (10). □

Theorem. *The dual of the GRS code $C_k(D, \vec{\lambda})$ is the GRS code $C_{n-k}(D, \vec{\lambda}')$ where*

$$\lambda_j' \lambda_j = \mu_j := \begin{cases} 1/\left(\Pi_{x \in D \setminus \{\infty, x_j\}}(x_j - x)\right) & \text{if } x_j \neq \infty \\ -1 & \text{if } x_j = \infty \end{cases}$$

Proof. Let a_{ij} denote the (ij)-th entry of $G_k(D; \vec{\lambda}) \cdot G_{n-k}(D; \vec{\lambda}')$. We must show that $a_{ij} = 0$ for all $1 \le i \le k$ and $1 \le j \le n-k$. First suppose $\infty \notin D$. In this case we must show that

$$\sum_{j=1}^{n} \mu_j x_j^m = 0, \quad 0 \le m \le n-2$$

because a_{ij} is the left side of this identity for $m = (i + j - 2)$. It follows from the next lemma that

$$X^m = \sum_{j=1}^{n} x_j^m \mu_j \prod_{1 \le i \le n, i \ne j} (X - x_i), \quad 0 \le m \le n-1.$$

In particular for $0 \le m \le n-2$, we note that the coefficient of X^{n-1} in 0 on the left side, and $\sum_{j=1}^{n} \mu_j x_j^m$ on the right side. This shows that

$$\sum_{j=1}^{n} \mu_j x_j^m = 0, \quad 0 \le m \le n-2$$

as was to be shown. Taking $m = n-1$, we also note for later use that

$$1 = \sum_{j=1}^{n} \mu_j x_j^{n-1} \tag{12}$$

We now turn to the case when $\infty \in D$. If C is a linear code with generator matrix G and parity check matrix H, and M is a permutation matrix, then the code C' generated by GM clearly has parity check matrix HM: this follows by noting that a permutation matrix is orthogonal (i.e $M^T = M^{-1}$), or simply by observing that the dot product of two row vectors v, w is unaffected if we permute the coordinates of both v and w by the same permutation. Thus, we may suppose that $i = n$, i.e. $x_n = \infty$. Let $\mu_j^{-1} = \prod_{\ell \ne j}(x_j - x_\ell)$ for $1 \le j \le n-1$ and let $\mu_n = -1$ From the theorem about duality of RS codes when no x_i is ∞, we know that the product $G_{k-1}((x_1, \ldots, x_{n-1}))\cdot G_{(n-1)-(k-1)}((x_1, \ldots, x_{n-1}), (\mu_1, \ldots, \mu_{n-1}))^T$ is zero. In long form this can be written as

$$\sum_{j=1}^{n-1} \mu_j x_j^\ell = 0 \quad \text{for } \ 0 \le \ell \le n-3.$$

Note especially that the above identity holds for all $\ell \le n-3$. This means that the product

$$G_k(\vec{x}) G_{n-k}(\vec{x}, \vec{\mu})^T = \begin{bmatrix} G_{k-1}((x_1, \ldots, x_{n-1})) & \vec{0} \\ x_1^{k-1} \ldots x_{n-1}^{k-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} G_{(n-1)-(k-1)}((x_1, \ldots, x_{n-1}), (\mu_1, \ldots, \mu_{n-1}))^T \\ (0, \ldots, 0, \mu_n) \end{bmatrix}$$

equals

$$\begin{bmatrix} 0_{(k-1)\times(n-k)} \\ 0_{1\times(n-k-1)} \,|\, \mu_n + \sum_{j=1}^{n-1} \mu_j x_j^{n-2} \end{bmatrix}$$

Since $\mu_n = -1$, we have $\mu_n + \sum_{j=1}^{n-1} \mu_j x_j^{n-2} = 0$ by (12) (for $n-1$ points instead of n points). Since $G_{n-k}(\vec{x}, \vec{\mu})$ has the correct size $(n-k) \times n$ and rank $(n-k)$, we conclude that it is a parity check matrix for $C_k(\vec{x})$. □

Lemma 27. (Lagrange interpolation) *Let K be an arbitrary field. If $(x_1, y_1), \ldots, (x_n, y_n)$ are n points of K^2 with distinct x-coordinates, then there is a unique polynomial $f(X)$ of degree at most $n-1$, with coefficients in K, such that $f(x_i) = y_i$.*

Proof. Existence is given by the explicit formula $f(X) = \sum_{j=1}^{n} y_j E_j(X)$ where

$$E_j(X) = \prod_{\ell \ne j} \frac{X - x_\ell}{x_j - x_\ell}$$

Note that each $\deg(E_j(X)) = n-1$ and hence $\deg(f(X)) \leq n-1$. If $g(X)$ is another polynomial of degree at most $n-1$ satisfying $g(x_j) = y_j$, then $f(X) - g(X)$ has degree at most $n-1$ and n distinct roots, hence it is the zero polynomial. $\square$