Notes on Finite Fields and Coding Theory

Naina Praveen

August, 2020

Contents

Acknowledgements

These notes have been written for my summer project on Finite Fields and Coding Theory, under the guidance of Prof. Inder Bir Passi in the year 2020. The work is expository in nature. These notes serve as an introduction to Coding Theory, while also providing a basic background to Finite Field theory in the first half.

I have followed [Con13] Keith Conrad's and [Wal20] Michel Waldschmidt's lecture notes as primary sources for the first section on Finite Field theory. I have used [LP84] Rudolf Lidl's book on Applied abstract Algebra and [Hal] J.I Hall's notes as primary sources for the second section on Coding Theory.

I am grateful to Prof. Passi for allowing me to undertake this project under him despite the given circumstances, and for all the guidance and knowledge he provided me with, through the course of our weekly sessions.

1 Finite Fields

1.1 Construction

Consider the ring $\mathbb{Z}$. Let p be a prime belonging to $\mathbb{Z}$. Then, $\mathbb{Z}$ factored by the maximal ideal generated by p, is a finite field of p elements. A finite field must have characteristic p [LP13] We will use $\mathbb{F}_p$ and $\mathbb{Z}/(p)$ interchangeably.

Proposition 1.1.1. *Any finite field must have finite power order.*

Proof. Let F be our finite field. Consider the homomorphism from $\Psi : \mathbb{Z} \to F$ defined as

$$n \mapsto \underbrace{1+1+\cdots+1}_{n \quad \text{times}},$$

where, $1 \in F$. Since F has finite cardinality, and $\mathbb{Z}$ has infinite cardinality, this map cannot be one-one, and hence the kernel of ψ cannot be $\{0\}$. Thus, the kernel must be a non-zero ideal in $\mathbb{Z}$. Let $ker(\psi) = (m)$ for some $m \in \mathbb{Z}$. Since any subring of F must be an integral domain, m must be a prime [Con13]. Then, $ker(\psi) = (p)$ for some prime p in $\mathbb{Z}$.

By First Isomorphism Theorem,

$$\mathbb{Z}/(p) \simeq Im(\psi)$$

$$i.e., \mathbb{Z}/(p) \hookrightarrow F$$

Then, we can view F as a vector space over $\mathbb{Z}/(p)$ and its dimension must be finite, as F is finite. Let n be the dimension of F over $\mathbb{Z}/(p)$.
Let $\{e_1, e_2, ..., e_n\}$ be a basis of F over $\mathbb{Z}/(p)$. Then every element in F can be expressed as $c_1e_1 + ... + c_ne_n$, where $c_i \in \mathbb{Z}/(p)$. Hence, the number of elements in F are p^n □

Proposition 1.1.2. *If F is a finite field, then the group $F^\times$ is cyclic.*

Proof. If $|F| = q$, then $|F^\times| = q - 1$.

If m is the maximal order of an element in the group $F^\times$, then $m|q-1$, by Langrange's theorem. In a finite abelian group, the order of all elements divide the maximal order [Con13]. This implies that every $g \in F^\times$ satisfies $g^m = 1$.

Therefore, all the elements in $F^\times$ are roots of the equation $X^m - 1 = 0$. The number of roots to this equation can be at most m. This would mean that $q - 1 \leq m$. However, since $m|q-1$, we have $m \leq q-1$. Therefore, we have $m = q-1$ and thus $F^\times$ is cyclic. □

Proposition 1.1.3. *Every finite field is isomorphic to $\mathbb{F}_p[X]/(\pi(x))$ for some prime p and some monic irreducible $\pi(x)$ in $\mathbb{F}_p[X]$*

Proof. Let F be an arbitrary finite field. Using Proposition 1.1.1, $\mathbb{F}_p$ is embedded in F, and F must have order p^n, for some prime p and some positive integer n.

From Proposition 1.1.2 $F^\times$ must be cyclic. Let γ be the generator of this group. Now, consider the evaluation homomorphism $ev_\gamma : \mathbb{F}_p[X] \to F$ that maps $f(X) \mapsto f(\gamma)$. This homomorphism is onto, as every element in F is either 0 or a power of γ, and we have $ev_\gamma(0) = 0$ and $ev_\gamma(X^r) = \left(ev_\gamma\right)^r = \gamma^r$, for $1 \leq r \leq p^n - 1$.

Then, by First Isomorphism theorem,

$$\mathbb{F}_p[X]/ker(ev_\gamma) \simeq F$$

As F is a field, $ker(ev_\gamma)$ must be a maximal ideal. $\mathbb{F}_p[X]$ is a PID, as $\mathbb{F}_p$ is a field. Therefore, the $ker(ev_\gamma)$ must be an ideal generated by an irreducible $\pi(x)$ in $\mathbb{F}_p[X]$. □

So far we have proved that all finite field must have prime power order. Now we ask, do finite fields of order p^n exist for every p^n? If so, then how do we construct them? From Proposition 1.1.3, we see that we can construct a finite field of order p^n from $\mathbb{F}_p$ by adjoining a root of an irreducible polynomial of degree n over $\mathbb{F}_p$[1]. So the question is, does there exist an irreducible polynomial of degree n over $\mathbb{F}_p$ for every n?

1.2 Finite Fields as Splitting fields

We shall see that every finite field is the splitting field of a polynomial. We shall use q to denote p^n.

Lemma 1.2.1. *A field of order q is a splitting field over $\mathbf{F}_p$ of $X^q - X$.*

Proof. Let F be our field of q elements. As $F^\times$ is cyclic and has order $q-1$, every element $g \in F^\times$ satisfies $g^{q-1} = 1$, which means they satisfy $g^q = g$. This follows from the fact that $a^{|G|} = 1_G$ in a finite group G. Thus, every element in F is a root of the polynomial $X^q - X$. From Proposition 1.1.1, we know that $\mathbb{F}_p$ is embedded in F. Then, we see that the derivative of $X^q - X$ in F is -1. This implies that the polynomial does not share any roots with its derivative and hence has q distinct roots. Since F contains q elements, all of which are root of the equation $X^q - X = 0$, F is the splitting field for the polynomial $X^q - X$ over $\mathbb{F}_p$. □

Theorem 1.2.2. *[LP13] For every prime p and every positive integer n, there exists a finite field with p^n elements. Any finite field with $q = p^n$ elements is isomorphic to the splitting field of $X^q - X$ over $\mathbb{F}_p$.*

Proof. Consider $X^q - X$ over $\mathbb{F}_p$. Let F be its splitting field over $\mathbb{F}_p$. Let G be the set of all roots of $X^q - X$ in F. Then, $G = \{a \in F | a^q - a = 0\}$. Note, G is a subfield of F as
i) $1 \in G$,
ii) for $x, y \in G$, $(x-y)^q = x^q - y^q$
iii) for $x, y \in G$, $(xy)^q = x^q y^q = xy$
iv) for $x, y \in G$ and $y \neq 0$ we have $((xy)^{-1})^q = x^q y^{-q} = xy^{-1}$, so $xy^{-1} \in G$.

As F is the splitting field of $X^q - X$, F contains all its roots. The derivative of $X^q - X$ over $\mathbb{F}_p$ is $qX^{q-1} - 1 = -1$. Since $X^q - X$ does not share a root with its derivative, it has q distinct roots. Therefore, the cardinality of G must be q. Now that we have proved the existence of such a field, we shall prove that it is unique up to isomorphism. Let F be a field of cardinality $q = p^n$. Then by Lemma 1.2.1, it is the splitting field of $X^q - X$. The result now follows from the uniqueness (up to isomorphism) of splitting fields. □

[1]Given a field K and an irreducible polynomial $p(X)$ over K, $K[X]/(p(X)) \simeq K(\alpha)$, where α is a root of $p(X)$. More on this can be found in [LP13]

Proposition 1.2.3. *A subfield of $\mathbb{F}_{p^n}$ has order p^d where $d|n$ and there exists one such subfield for every $d|n$.*

Proof. From Proposition 1.1.1, we know that $\mathbb{F}_p$ is contained in $\mathbb{F}_{p^n}$. Let F be a subfield of $\mathbb{F}_{p^n}$ of cardinality p^d. We have $\mathbb{F}_p \subset F \subset \mathbb{F}_{p^n}$. Then d denotes the dimension of F over $\mathbb{F}_p$. So, we have $[\mathbb{F}_{p^n} : \mathbb{F}_p] = [\mathbb{F}_{p^n} : F][F : \mathbb{F}_p]$. As $[\mathbb{F}_{p^n} : \mathbb{F}_p] = n$, we see that $d|n$.
As $|F| = p^d$, we have $|F^\times| = p^d - 1$. Every $g \in F^\times$ satisfies $g^{p^d-1} = 1$, and hence satisfies $g^{p^d} - g = 0$. This is also applicable for 0. Thus, every element in F is a root of $X^{p^d} - X$. Since $X^{p^d} - X$ can have at most p^d roots and the cardinality of F is p^d, F is the splitting field for $X^{p^d} - X$ and these subfields are unique up to isomorphism.
Now, we must show that given $d|n$, there exists a subfield of $\mathbb{F}_{p^n}$ of cardinality p^d. For this, consider $S = \{a \in \mathbb{F}_{p^n} | a^{p^d} = a\}$. All the elements in S are roots of the polynomial $X^{p^d} - X$. S is a field in the same way G is a field in Theorem 1.2.2. We must show that the cardinality of S is p^d. Note, every root of $X^{p^d} - X = 0$ is also a root of $X^{p^n} - X = 0$. This is because if we have $X^{p^d} = X$, then on raising both sides to p^d, we have $X^{p^{2d}} = X$. Let $n = md$. Then, repeating this m times, we have $X^{p^{md}} = X^{p^n} = X$. Therefore all the roots of $X^{p^d} - X$ lie in $\mathbb{F}_{p^n}$. $X^{p^d} - X$ does not share a common root with its derivative, so it has p^d distinct roots. Since all the elements in S are roots of $X^{p^d} - X$, the cardinality of S must be p^d. □

Proposition 1.2.4. *Let $f(X) \in \mathbb{F}_q[X]$ be an irreducible polynomial over $\mathbb{F}_q$ of degree m. Then $f(X)$ divides $X^{q^n} - X$ if and only if m divides n.*

Proof. Suppose $f(X)$ divides $X^{q^n} - X$. Let α be a root of $f(X)$ in a splitting field over $\mathbb{F}_q$. Then, α satisfies $\alpha^{q^n} = \alpha$ and so $\alpha \in \mathbb{F}_{q^n}$. So, $\mathbb{F}_q(\alpha)$ is a subfield of $\mathbb{F}_{q^n}$ and as $f(X)$ is an irreducible of degree m, we have $[\mathbb{F}_q(\alpha) : \mathbb{F}_q] = m$. As $[\mathbb{F}_{q^n} : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}(\alpha)][\mathbb{F}(\alpha) : \mathbb{F}_q]$, we have that $m|n$.
Conversely, let $m|n$. Then, from Proposition 1.2.3, we have that $\mathbb{F}_{q^m}$ is a subfield of $\mathbb{F}_{q^n}$. Let α be a root of $f(X)$. As $f(X)$ is an irreducible of degree m and $[\mathbb{F}(\alpha) : \mathbb{F}_q] = m$, from Theorem 1.2.2, $\mathbb{F}(\alpha) \simeq \mathbb{F}_{q^m}$. Thus, α satisfies $\alpha^{q^m} = \alpha$. If $n = mk$, then raising $\alpha^{q^m} = \alpha$ to q^m, k times, we find that α also satisfies $\alpha^{q^n} = \alpha$ and is thus an element of $\mathbb{F}_{q^n}$. This implies that α is a root of $X^{q^n} - X = 0$ and $f(X)$ divides $X^{q^n} - X = 0$. □

1.3 Galois theory of Finite Fields

Here, we will be introduced to the Frobenius automorphism and see how Galois Theory applies to Finite fields. As $\mathbb{F}_{q^n}$ is the splitting of the polynomial $X^{q^n} - X$ over $\mathbb{F}_q$ which is separable, it is a Galois extension of $\mathbb{F}_q$. Thus, any finite extension over $\mathbb{F}_q$ is a Galois extension.

Proposition 1.3.1. *Let F be a field of characteristic p.*
(a) The map

$$\begin{aligned} \sigma_q : & F \to F \\ & x \mapsto x^q \end{aligned}$$

is an endomorphism of F.
(b) If F is finite,then σ_q is surjective, hence is an automorphism of the field F.

Proof. This is a morphism of fields as in F,

$$\sigma_q(x+y) = (x+y)^q = x^q + y^q = \sigma_q(x) + \sigma_q(y)$$

and

$$\sigma(xy) = (xy)^q = x^q y^q = \sigma_q(x)\sigma_q(y)$$

It is injective as every field homomorphism is injective. If F is finite, it is surjective because it is injective. Hence σ_q is an automorphism of a finite field F. □

This map is also called the Frobenius automorphism, that maps an element to its q^{th} power.

Theorem 1.3.2. *The Frobenius map generates the Galois group of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$, denoted as $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$.*

Proof. From Proposition 1.3.1, σ_q is an automorphism of $\mathbb{F}_{q^n}$. We also know that every $\alpha \in \mathbb{F}_q$, satisfies $\alpha^q = \alpha$. Hence, φ_q acts as identity on the elements of $\mathbb{F}_q$ and belongs to $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$. We also know that $|Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)| =$
$[\mathbb{F}_{q^n} : \mathbb{F}_q] = n$. To show that $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$ is cyclic, we need to show that the order of σ_q in $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$ is n.
Let σ_q^i denote the composition of σ_q i times, i.e., $\sigma_q^i(x) = x^{q^i}$. In order to find the order of φ_q, we need to find the smallest value of i such that $\sigma_q^i(x) = x^{q^i} = x$. We know that an element of $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$ maps the roots of a polynomial to other roots of the polynomial. If for $i < n$, we had σ_q^i as the identity map, it would imply that the elements are roots of the polynomial $X^{q^i} - X = 0$, which has only q^i roots. However, $\mathbb{F}_{q^n}$ is the splitting field of $X^{q^n} - X = 0$, which has q^n distinct roots. Hence, $i = n$, and the order of σ_q is n in $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$. □

Proposition 1.3.3. *If $\pi(x) \in \mathbb{F}_q[X]$ is an irreducible with degree n, then it is separable and if α is one of its roots in some extension field of $\mathbb{F}_q$, then its full sets of roots are $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{n-1}}$.*

Proof. We know that any finite extension of prime power order is Galois over $\mathbb{F}_q$. Then the field $\mathbb{F}_q(\alpha)$ is a finite extension over $\mathbb{F}_q$ that is Galois. The roots of $\pi(x)$ can be obtained from α on applying $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$ to it.

We know that the $Gal(\mathbb{F}_{q^n}/\mathbb{F}_q)$ is generated by the Frobenius automorphism. Applying the Frobenius automorphism to α we get the elements to be $\alpha, \alpha^q, \alpha^{q^2}, \ldots$. Since $\mathbb{F}_q[\alpha] \simeq \mathbb{F}_q/(\pi(X))$, and from Proposition 1.1.3, $\mathbb{F}_q/(\pi(X))$ has order q^n, thus, so does $\mathbb{F}_q[\alpha]$. Every element in $\mathbb{F}_q[\alpha]$ must satisfy $\alpha^{q^n} = \alpha$. Hence, on application of the Frobenius automorphism n times, we have cycled back to the start and our set of roots then are $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{n-1}}$. Note, these are all distinct as $\pi(X)$ is separable because we are in a Galois extension and $\pi(X)$ is an irreducible with a root in $\mathbb{F}_q[\alpha]$. Thus, the n distinct roots of $\pi(x)$ are $\alpha, \alpha^q, \alpha^{q^2}, \ldots, \alpha^{q^{n-1}}$. □

Theorem 1.3.4 (Galois Theory for Finite Fields)**.** *[Wal20] Let F be a finite field with q elements and K, an extension of degree n. Then, there exists a bijection between the divisors d of n, the subfields E containing F, and the subgroups H of $Gal(K/F)$.*

Proof. From Theorem 1.2.2, we know that the $Gal(K/F)$ is generated by the Frobenius automorphism and is cyclic, of order n. Hence, there exists a bijection between the divisors d of n and the subgroups H of $Gal(K/F)$ (because for a finite cyclic group of order n, every subgroup's order is a divisor of n, and there is exactly one subgroup for each divisor). The order of the element σ_q^d is $\frac{n}{gcd(n,d)}$ which is n/d. So the subgroup H of index d and order n/d is generated by σ_{q^d}. The fixed field of H is defined to be $x \in K$ that satisfy $\sigma(x) = x$ for all $x \in H$. Hence, the fixed field for H would be the unique subfield E with q^d elements, which we know exists for every $d|n$ from Proposition 1.2.3. □

1.4 Roots of Unity over Finite Fields

In this section, we will see first understand roots of unity and see how roots of unity can also be used to describe elements of finite fields. The splitting field of $X^n - 1$ over a field K is called the n^{th} cyclotomic field over K, denoted by $K^{(n)}$. The set of roots of unity are denoted as $E^{(n)}$. The derivative of $X^n - 1$ is nX^{n-1}, and they do not share any common roots. Hence it has n distinct roots and the group $E^{(n)}$ forms a cyclic group. The proof is similar to Proposition 1.1.2. This also implies that $E^{(n)}$ has $\varphi(n)$ generators.
If K has characteristic p, and an integer n such that $n = p^k m$. Then $X^n - 1 = X^{mp^k} - 1 = (X^m - 1)^{p^k}$. Then, $E^{(n)} = E^{(m)}$ and the roots of $X^m - 1$ each occur with multiplicity p^k. Hence, here on, we will consider $(n, p) = 1$.

Defintion 1.4.1. If K has characteristic p and n is an integer not divisible by p, and ζ a primitive n^{th} root of unity over K, then,

$$\Phi_n(x) = \prod_{\substack{s=1 \\ (s,n)=1}}^{n} (x - \zeta^s)$$

where $\Phi_n(X)$ is the n^{th} cyclotomic polynomial over K and has degree $\varphi(n)$.

Proposition 1.4.2. *[Neu13] Over $\mathbb{Z}$, we have*

$$X^n - 1 = \prod_{\substack{1 \le d \le n \\ d|n}} \Phi_d(X)$$

Proof. Suppose ζ be a root of $\Phi_d(X)$ where $d|n$. This implies that ζ is a d^{th} root of unity. Let q be such that $n = qd$. Then, $\zeta^n = (\zeta^d)^q = 1^q = 1$. Thus, ζ is also a root of $X^n - 1$. So, for every $d|n$, the d^{th} root of unity is also a n^{th} root of unity.
Now, let ζ be a root of $X^n - 1$, i.e., ζ is a n^{th} root of unity. Let the order of ζ be k. This implies that ζ is a k^{th} root of unity as it satisfies $X^k - 1$. This means it is a root of $\Phi_k(X)$. Since the n^{th} roots of unity form a group of n elements, the order of ζ must divide the order of the group i.e., n. This implies that k is a divisor of n. Thus, ζ is a root of $\Phi_d(X)$ for some $d|n$.
Now, we have shown that $X^n - 1$ and $\prod_{d|n} \Phi_d(X)$ share all their roots. Since $\prod_{d|n} \Phi_d(X)$ is a product of a collection of monic polynomials, its product will also be monic. As $X^n - 1$ is also monic, they must be equal. □

Theorem 1.4.3. *[Neu13] The cyclotomic polynomial $\Phi_n(X)$ has coefficients in $\mathbb{Z}$ and is irreducible over $\mathbb{Z}[X]$.*

Proof. If $n = p$ where p is some prime. We then know that $\Phi_p(X) = \frac{X^p-1}{X-1}$. Let $X = Y + 1$. Then we have

$$\Phi_p(Y+1) = \frac{(Y+1)^p - 1}{Y+1-1} = \frac{(Y+1)^p - 1}{Y}$$

By binomial expansion, we have

$$Y^{p-1} +^p C_{p-1}Y^{p-2} + ... +^p C_1$$

Each ${}^pC_i$ is divisible by p, p^2 does not divide p and p does not divide 1. hence, by Eisenstein's criterion, $\Phi_p(Y+1) = \Phi_p(X)$ is irreducible over $\mathbb{Z}[X]$.
Now, assume that n is not a prime. Assume $\Phi_n(X)$ is reducible.
Let $f(X) \in \mathbb{Z}[X]$ be an irreducible divisor of $\Phi_n(X)$. Then, let $g(X) \in \mathbb{Z}[X]$ be such that

$$\Phi_n(X) = f(X)g(X) \tag{1}$$

. Since $\Phi_n(X)$ is monic, $f(X)$ and $g(X)$ must be monic too. Note, from Proposition 1.4.2, as $X^n - 1 = \prod_{d|n} \Phi_d(X)$, and $f(X)$ divides $\Phi_n(X)$, we have that $f(X)$ divides $X^n - 1$ i.e., $X^n - 1 = \Phi_n(X)d(X)$, for some $d(X) \in \mathbb{Z}[X]$. From (1), we have

$$X^n - 1 = f(X)g(X)d(X) \tag{2}$$

Let ζ be a root of $f(X)$ in $\mathbb{C}$. Since $f(X)$ is a factor of $\Phi_n(X)$, ζ must be a primitive root of unity. Let p be a prime such that $(p, n) = 1$. We need to show that ζ^p is a root of f as well. For the above prime p and a primitive root of unity, ζ^p is also a primitive n^{th} root of unity, which means that $\Phi_n(\zeta^p) = 0$. Now, assume that ζ^p is *not* a root of $f(X)$. Then, it must be a root of $g(X)$. If ζ^p is a root of $g(X)$, then ζ is a root of $g(X^p)$. Since $f(X)$ is irreducible, it is the minimal polynomial of ζ. This implies that $f(X) \mid g(X^p)$. Then,

$$g(X^p) = f(X)h(X) \tag{3}$$

for some $h(X) \in \mathbb{Z}[X]$. Now, consider the morphism of reduction modulo p. In $\mathbb{F}_p[X]$, we have $g(X^p) = (g(X))^p$. Let $\tilde{f}(X)$ and $\tilde{h}(X)$ denote the images of $f(X)$ and $h(X)$ under the reduction morphism. From (3), we have $(g(X))^p = \tilde{f}(X)\tilde{h}(X)$. Let $\tilde{k}(X) \in \mathbb{F}_p[X]$ be an irreducible factor of $\tilde{f}(X)$. Then, $\tilde{k}(X)$ divides $(g(X))^p$ and thus divides $g(X)$. Let $g(X) = \tilde{k}(X)\tilde{l}(X)$ for some $\tilde{l}(X) \in \mathbb{F}_p$. Substituting this in (2), we have

$$X^n - 1 = \tilde{k}(X)\tilde{m}(X)\tilde{k}(X)\tilde{l}(X)\tilde{d}(X) \tag{4}$$

This means that $\tilde{k}^2(X) \mid X^n - 1$ which would imply that $X^n - 1$ has multiple roots. But, for that, $X^n - 1$ and its derivative nX^{n-1} must have a common root. As $(p, n) = 1$, this is not possible in $\mathbb{F}_p$. Hence, we have reached a contradiction.
Therefore, ζ^p must be a root of $f(X)$. We have checked that for any root ζ of $f(X)$ in $\mathbb{C}$ and any prime number p which does not divide n, the number ζ^p is again a root of $f(X)$. By induction on the number of prime factors of m, it follows that for any integer m with $(m, n) = 1$ the element ζ^m is a root of $f(X)$.
From Definition 1.4.1, we have $\Phi_n(x) = \prod^{n}_{\substack{s=1 \\ (s,n)=1}} (x - \zeta^s)$. So, $\Phi_n(X)|f(X)$. But $f(X)$ was irreducible by assumption. Thus $\Phi_n(X) = f(X)$. □

Theorem 1.4.4. *[Neu13] The finite field $\mathbb{F}_q$ is the $(q-1)^{th}$ cyclotomic field over any of its subfields.*

Proof. In $\mathbb{F}_q$, the $q-1$ non-zero elements are all zeroes of the polynomial $X^{q-1}-1=0$. This polynomial splits in $\mathbb{F}_q$ from Theorem 1.2.2 and cannot not split in any further subfield of $\mathbb{F}_q$. Thus, $\mathbb{F}_q$ is the $(q-1)^{th}$ cyclotomic field of any of its subfields. □

Defintion 1.4.5. The Moebius function μ is the function on $\mathbb{N}$ defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n \text{ is the product of } k \text{ distinct primes} \\ 0 & \text{if } n \text{ is divisible by the square of a prime} \end{cases}$$

Lemma 1.4.6. *For $n \in \mathbb{N}$, the Moebius function satisfies*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases}$$

Proof. The case for $n = 1$ is trivial. For other $n > 1$, we only need to consider divisors d such that $\mu(d)$ is non-zero. Let $p_1, p_2 ... p_k$ be the distinct prime divisors of n. Then, we have

$$\begin{aligned} \sum_{d|n} \mu(d) &= \mu(1) + \sum_{i=1}^{k} \mu(p_i) + \sum_{1 \le i_1 < i_2 \le k} \mu(p_{i_1} p_{i_2}) + ... + \mu(p_1 p_2 ... p_k) \\ &= 1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + ... + \binom{k}{k}(-1)^k \\ &= (1 + (-1))^k = 0 \end{aligned}$$

□

Theorem 1.4.7 (Mobeius Inversion Formula)**.** *[Neu13] Depending on the operation on the group, there exist two version of this formula*
Additive version: Let h and H be two functions from $\mathbb{N}$ into an additively written abelian group G. Then

$$H(n) = \sum_{d|n} h(d) \text{ for all } n \in \mathbb{N}$$

if and only if

$$h(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) H(d) = \sum_{d|n} \mu(d) H\left(\frac{n}{d}\right)$$

for all $n \in \mathbb{N}$

Multiplicative version: let h and H be two functions from $\mathbb{N}$ into a multiplicatively written abelian group G. Then

$$H(n) = \prod_{d|n} h(d) \text{ for all } n \in \mathbb{N}$$

if and only if

$$h(n) = \prod_{d|n} H(d)^{\mu\left(\frac{n}{d}\right)} = \prod_{d|n} H\left(\frac{n}{d}\right)^{\mu(d)}$$

for all $n \in \mathbb{N}$

The Mobeius inversion formula can also be used to find another expression for the cyclotomic polynomial. On substituting $h(n) = \Phi(X)$ and $H(n) = X^n - 1$, in the multiplicative version of the above formula, we get

$$\Phi_n(x) = \prod_{d|n} \left(x^d - 1\right)^{\mu\left(\frac{n}{d}\right)} = \prod_{d|n} \left(x^{\frac{n}{d}} - 1\right)^{\mu(d)}$$

Theorem 1.4.8. *[Neu13] For every finite field* $\mathbb{F}_q$ *and every* $n \in \mathbb{N}$*, the product of all monic irreducible polynomials over* $\mathbb{F}_q$ *whose degree divides* n*, is equal to* $X^{q^n} - X$*. If* $N_q(d)$ *denotes the number of irreducible polynomials of degree* d *over* $\mathbb{F}_q$*, then we find that*

$$N_q(n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d = \frac{1}{n} \sum_{d|n} \mu(d) q^{\frac{n}{d}}$$

Proof. The derivative of $X^{q^n} - X$ is -1, which has no root, hence, $X^{q^n} - X$ has no multiple factors in characteristic p. From Proposition 1.2.4, we know that an irreducible polynomial over $\mathbb{F}_q$ divides $X^{q^n} - X$ if and only if $m \mid n$. Since $X^{q^n} - X$ is monic, all the irreducible divisors must be monic as well. Thus, the product of all the irreducible monic polynomials must be equal to $X^{q^n} - X$. Then on comparing the degrees, we observe

$$q^n = \sum_{d|n} dN_q(d)$$

. Then, using Theorem 1.4.4, substituting q^n as $H(n)$ and $nN_q(d)$ as $h(n)$ in the additive version of the Mobeius inversion formula, we get

$$N_q(n) = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}$$

□

Defintion 1.4.9. If a and n are relatively prime integers, the *order of a modulo n* is the order of the class of a in the multiplicative group $(\mathbb{Z}/n\mathbb{Z})^\times$. In other terms, it is the smallest integer l such that a^l is congruent to 1 modulo n

Proposition 1.4.10. *Let* $\mathbb{F}_q$ *be a finite field with* q *elements and let* n *be a positive integer not divisible by the characteristic of* $\mathbb{F}_q$*. Then the cyclotomic polynomial* Φ_n *splits in* $\mathbb{F}_q[X]$ *into a product of irreducible factors, all of the same degree* d*, where* d *is the order of* q *modulo* n*.*

Proof. Let ζ be the root of Φ_n in a splitting field K of the polynomial Φ_n over $\mathbb{F}_q$. The order of ζ in $K^\times$ is n. Now, let $s = [\mathbb{F}_q[\zeta] : \mathbb{F}_q]$. Then $\mathbb{F}_q[\zeta]/\mathbb{F}_q$ is a Galois extension with a cyclic Galois group of order s, generated by the σ_q. This means that s is the smallest integer such that $\zeta^{q^s} = \zeta$. Then, by the definition of order of q modulo n, we have $s = d$. Since the degree of over $\mathbb{F}_q[\zeta]$ over $\mathbb{F}_q$ is d, and this was for any arbitrary primitive root ζ, we have Φ_n is the product of irreducible factors, all of degree d. □

Proposition 1.4.11. *Φ_n splits completely in $\mathbb{F}_q[X]$ into a product of linear polynomials if and only if $q \equiv 1$ mod n.*

Proof. This follows from Proposition 1.4.10 as a special case of $d = 1$. □

2 Coding Theory

2.1 Introduction

A beautiful application of abstract algebra extends to the theory of coding. The idea is to be able to transmit data reliably over a channel. Messages transferred over a channel are prone to distortions due to 'noise'. Our task is to be able add some symbols in our message, which enable us to detect and correct these errors caused due to the noise.

A simple example would be repetition. Say we need to transfer message 001. We could 'encode' it by just repeating it twice, and send 001001. Due to noise, assume the receiver receives 011001. While the receiver can realise that there has been an error, he cannot identify whether the received message was 011 or 001. If we were to repeat the message thrice, and send 001001001, but the receiver received 001011001, the receiver would assume the correct message was 001. However, the receiver could have also received 011011001, in which case he would 'decode' it as 011, which is incorrect. Repeating the message multiple times is not an efficient method because new types of errors can keep coming in and the message becomes too long to transfer. Thus, the idea is to add the smallest number of extra symbols in order to reduce redundancy, while at the same time be able to detect and correct errors efficiently [LP84].

One commonly used idea is that of a *parity check bit*. Assume our message is $x_1x_2...x_k$. Then encode it by augmenting the message with x_{k+1}, defined to be 1 if the number of 1's are odd in the message, and 0 if the number of 1's are even in the message. What we see here is that this enables us to detect a single error in our message and basically functions in the same manner as that of the single repetition code, while having code length $k+1$ instead of $2k$ [LP84]. Now, we shall define the above concepts more rigorously and see various common codes that have been developed.

We shall assume that elements of a finite field form our symbols/alphabets for coding.

Defintion 2.1.1. A *code* of *length* n on a finite alphabet $\mathbb{F}_q$ is a subset $\mathcal{C}$ of $\mathbb{F}_q^n$. A *word* is an element of $\mathbb{F}_q^n$, while a *codeword* is an element of $\mathcal{C}$.

Defintion 2.1.2. Coding consists of transforming a block of k *message symbols* $a_1a_2...a_k$ into a *codeword* $\mathbf{x} = x_1x_2...x_n$ of length n. Usually, the first k positions are the message symbols, i.e., $x_i = a_i$ for $1 \leq i \leq k$, while the rest $n-k$ symbols are called the *check symbols*.

Defintion 2.1.3. A *linear code* of length n and dimension k over finite field $\mathbb{F}_q$ is a $\mathbb{F}_q$-vector subspace of $\mathbb{F}_q^n$ of dimension k. This is also denoted as a $[n, k]$ code.

Now, we shall define a metric on our space $\mathbb{F}_q^n$.

Defintion 2.1.4. (i) The *Hamming distance* $d(x, y)$ between two vectors $\mathbf{x} = x_1...x_n$ and $\mathbf{y} = y_1...y_n$ in $\mathbb{F}_q^n$ is the number of coordinates at which $\mathbf{x}$ and $\mathbf{y}$ differ.
(ii) The *Hamming weight* $w(\mathbf{x})$ of a vector $\mathbf{x} = x_1...x_n$ in $\mathbf{F}_q^n$ is the number of nonzero coordinates x_i. So $w(\mathbf{x}) = d(x, 0)$.
(iii) The *minimum distance* of a code $d_{min}(\mathcal{C})$ is defined to be

$$d_{min}(\mathcal{C}) = min\ \{\ d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{y}\}$$

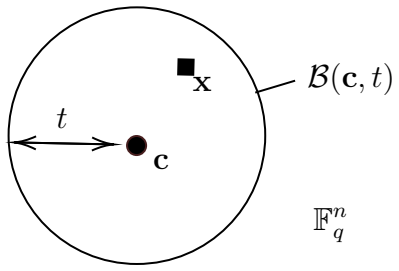
We see that for linear codes, due to linearity, the codeword with smallest non-zero weight has the minimum distance. Given $\mathbf{c_0} \in \mathcal{C}$, we have for $\mathbf{c} \in \mathcal{C}$

$$\min_{\mathbf{c}\in\mathcal{C},\mathbf{c}\neq\mathbf{c_0}} d\left(\mathbf{c}, \mathbf{c_0}\right) = \min_{\mathbf{c}\in\mathcal{C},\mathbf{c}\neq\mathbf{c_0}} d\left(\mathbf{c} - \mathbf{c_0}, \mathbf{0}\right) = \min_{\mathbf{c}\in\mathcal{C},\mathbf{c}\neq\mathbf{0}} d(\mathbf{c}, \mathbf{0})$$

as $\mathbf{c} - \mathbf{c_0}$ must also be a codeword in the linear subspace.

(iv) For a non-negative integer t, we define the *Hamming Ball* $\mathcal{B}(\mathbf{c}, t)$ with centre $\mathbf{c} \in \mathbb{F}_q^n$ and radius t to be the set of elements of distance t from $\mathbf{c}$, i.e.,

$$\mathcal{B}(\mathbf{c}, t) = \{\mathbf{x} \in \mathbb{F}_q^n \mid d(\mathbf{x}, \mathbf{c}) \leq t\}$$



A Hamming Ball of radius t centred at codeword $\mathbf{c}$ consists of word $\mathbf{x}$ such that $d(\mathbf{x}, \mathbf{c}) < t$

Defintion 2.1.5. [Wal20] A *transmission with at most t errors* is a mapping from $f : \mathcal{C} \to \mathbb{F}_q^n$ such that for all $\mathbf{c} \in \mathcal{C}$, $d(f(\mathbf{c}), \mathbf{c}) < t$.

Defintion 2.1.6. Assume our codeword $\mathbf{x} = x_1x_2...x_n$ is received as $f(\mathbf{x}) = y_1y_2...y_n$ due to noise. Then we define our *error word* or *error vector* to be $e(\mathbf{x}) = \mathbf{f}(\mathbf{x}) - \mathbf{x} = y_1 - x_1, y_2 - x_2, ..., y_n - x_n$

In order to detect an error, we need to check if $e(\mathbf{c})$ is zero or not. A code is said to *detect t errors* if for all $\mathbf{c} \in \mathcal{C}$

$$\mathcal{B}(\mathbf{c}, t) \cap \mathcal{C} = \{\mathbf{c}\}$$

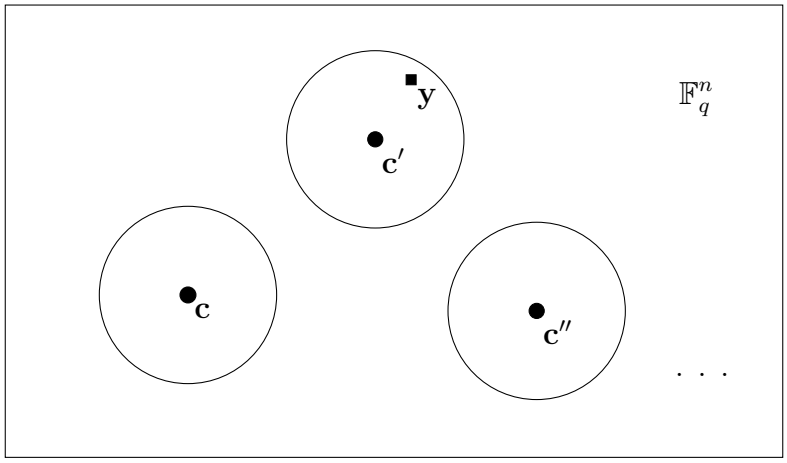
. This would mean a transmission $f(\mathbf{c}) \in \mathcal{C}$ if and only if $e(\mathbf{c}) = 0$.

A code is said to be *t-error correcting* if for all if for all $\mathbf{x} \in \mathbb{F}_q^n$,

$$\#\mathcal{B}(\mathbf{x}, t) \cap \mathcal{C} \leq 1$$

This is to say that the intersection between the number of elements in a ball around $\mathbf{x}$, and $\mathcal{C}$ is less than equal to one. If the above holds, then the transmission $f : \mathcal{C} \to \mathbb{F}_q^n$ is injective, i.e., for every $\mathbf{y} \in f(\mathcal{C})$, we have exactly one $\mathbf{c} \in \mathcal{C}$ such that $f(\mathbf{c}) = \mathbf{y}$ [Wal20].

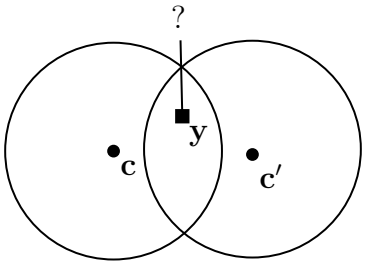
In other words, we can define this balls around different codewords $\mathbf{c}, \mathbf{c}', \mathbf{c}'''...$ as shown in the figure below. If the incoming word $\mathbf{y}$ belongs inside any of the balls, we decode it as the codeword that lies in the centre of the ball.



An incoming word $\mathbf{y}$ belonging to $\mathcal{B}(\mathbf{c}', t)$

Thus, in the above diagram. The incoming word $\mathbf{y}$ would be decoded as the codeword $\mathbf{c}'$.

Our idea is to make the radius t as large as possible, in order to detect as many incoming code words, while at the same time make it small enough to ensure that the balls of two codewords do not intersect.



Hamming ball when $d_{min}(\mathcal{C}) < 2t$

Lemma 2.1.7. *[Wal20] [Hal] A code $\mathcal{C}$ of length n over $\mathbb{F}_q^n$ can detect t errors if and only if $d_{min}(\mathcal{C}) = d \geq t+1$. The code $\mathcal{C}$ can correct t errors if and only if $d \geq 2t+1$.*

Proof. $d \geq t+1$ means that every word in $\mathcal{C}$ is greater than t distance apart. This means that every word of distance $< t$ from any $\mathbf{c} \in \mathcal{C}$, does not belong in $\mathcal{C}$. This means that $\mathcal{B}(\mathbf{c}, t) \cap \mathcal{C} = \{\mathbf{c}\}$ and hence it is an error detecting code.
For t correcting codes, assume $d(\mathcal{C}) \geq 2t+1$. Pick $\mathbf{x} \in \mathbb{F}_q^n$. Let $\mathbf{c_1}, \mathbf{c_2}$ satisfy $d(\mathbf{x}, \mathbf{c_1}) < t$ and $d(\mathbf{x}, \mathbf{c_2}) < t$. The, by triangle inequality, we have $d(\mathbf{c_1}, \mathbf{c_2}) \leq d(\mathbf{x}, \mathbf{c_1}) + d(\mathbf{x}, \mathbf{c_1}) \leq 2t$. But we assumed that $d(\mathcal{C}) \geq 2t+1$. Hence, $\mathbf{c_1} = \mathbf{c_2}$ and we have $\#\mathcal{B}(\mathbf{x}, t) \cap \mathcal{C} \leq 1$.
Conversely, assume $d \leq 2t$. Then we need to show that there exists a word in the intersection of the Hamming balls of two codewords (because it would mean that a ball of radius t around this word would include both the codewords in it). Take $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ such that $d(\mathbf{x}, \mathbf{y}) = d$. If $d < t$, then $\mathbf{x} \in \mathcal{B}(\mathbf{x}, t) \cap \mathcal{B}(\mathbf{y}, t)$ which would mean that $\mathcal{B}(\mathbf{x}, t) \cap \mathcal{C} \geq 2$ and hence it cannot be an t error correcting code. So assume $t < d \leq 2t$. Then $\mathbf{x}$ and $\mathbf{y}$ differ at d places. Let $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n)$. Let $i_1, i_2, ..., i_d < n$ be the places at which $\mathbf{x}$ and $\mathbf{y}$ differ, i.e., $x_{i_j} \neq y_{i_j}$ for $j = 1, 2, ..., d$. Define $\mathbf{z} = (z_1, z_2, ..., z_n)$ where $z_k = y_k$ when $k \notin \{i_1, i_2, ...i_t\}$ and $z_k = x_k$ when $k \in \{i_1, i_2, ...i_t\}$. Then, $d(\mathbf{x}, \mathbf{z}) = d - t < t$ and $d(\mathbf{y}, \mathbf{z}) = t$. Then $\mathbf{z} \in \mathcal{B}(\mathbf{x}, t) \cap \mathcal{B}(\mathbf{y}, t)$. Thus the Hamming Ball around $\mathbf{z}$ of radius t contains both $\mathbf{x}$ and $\mathbf{y}$. Thus it cannot be a t-error correcting code. □

From the above it is now clear that in order to construct a t-error correcting code, we need to choose codewords such that spheres of radius t around them are pairwise disjoint. However, we also know that the volume of the spheres of a fixed radius, packed in a container of a fixed volume must be less than the volume of the container, and thus we have the following bound.

Theorem 2.1.8 (Hamming Bound)**.** *[LP84] The parameters q, n, t, M of a t-error correcting code of length n over $\mathbb{F}_q$ with M codewords satisfy the inequality*

$$M\left(1 + (q-1)\binom{n}{1} + ... + (q-1)^t\binom{n}{t}\right) \leq q^n$$

Proof. The total 'volume' of a vector space over $\mathbb{F}_q$ of length n is the total number of words, which is q^n.
Let v denote the number of words in each ball. As the ball is centred around a codeword of radius t, each word within the ball can deviate at at most t coordinates from the centre. We can see that the number of words differing at 0 coordinates form the codeword is 1. The number of words differing at 1 coordinate from the codeword is $(q-1)$. Similarly, the number of words different at t positions from the codeword is $\binom{n}{t}(q-1)^t$. Therefore, the number of possible words in $\mathcal{B}(\mathbf{c}, t)$ is

$$v = \sum_{k=0}^{t} \binom{n}{k}(q-1)^k$$

Since M is the total number of codewords, we have $M.v \leq q^n$. Thus, the above inequality holds. □

Defintion 2.1.9. A t-error correcting code is called a *perfect code* if the above upper bound is an equality.

Perfect codes imply that the Hamming balls around the codewords occupy the entire Hamming space without overlap, where each sphere has radius t, and the codewords have minimum distance $2t + 1$.

2.2 Linear Codes

We shall now focus on the specific class of codes called linear codes.

Defintion 2.2.1 (Parity Check Matrix)**.** Let H be an $(n-k) \times n$ matrix of rank $n-k$ with elements in $\mathbb{F}_q$. The set of all n-dimensional vectors $\mathbf{x}$ satisfying $H\mathbf{x}^T = 0$ over $\mathbf{F}_q$ is called a *linear (block) code* $\mathcal{C}$ over $\mathbb{F}_q$ of block length n.

If the matrix $\mathbf{H}$ is arranged in the form $[\mathbf{A}, I_{n-k}]$, where I_{n-k} denotes the identity matrix of rank $n-k$, then the first k symbols of the codeword $\mathbf{x}$ are the message symbols and the following $(n-k)$ symbols are the check symbols mention in Definition 2.1.2. Then $\mathcal{C}$ is called a *systemic code* and $\mathbf{H}$ is said to be in *standard form.*

Two linear codes are *equivalent* if one can be obtained from the other by a series of operations of the following two types: 1) an arbitrary permutation of the coordinate positions, and 2) in any coordinate position, multiplication by any non-zero scalar.

Example 2.2.2 (Repetition Code)**.** Let our message consist of only one symbol $a_1 \in \mathbb{F}_2$, and let the rest of the $n-1$ symbols be the check symbols where a_1 is repeated $n-1$ times. This is called the $(n, 1)$ repetition code. The parity check matrix would be

$$\begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & \\ 1 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}$$

The only codewords here are 00...000 and 11...111 .

Example 2.2.3. The parity check matrix for the $(n, n-1)$ parity check code mentioned in subsection 2.1 is

$$\begin{bmatrix} 1 & 1 & ... & 1 \end{bmatrix}$$

Defintion 2.2.4 (Generator Matrix)**.** A *generator matrix* $\mathbf{G} = [I_k, -\mathbf{A}^T]$ of the $[n, k]$ linear code $\mathcal{C}$ over $\mathbb{F}_q$ is a $k \times n$ matrix. Clearly, $\mathbf{GH^T} = \mathbf{0}$. [6]

Example 2.2.5. i) (Repetition code) The generator matrix for the repetition code is

$$\begin{bmatrix} 1 & 1 & 1 & ... & 1 \end{bmatrix}$$

ii) (Parity check code) The generator matrix for the parity check code mentioned earlier is

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 1 \\ & \vdots & & \ddots & & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 \end{bmatrix}$$

iii) The generator matrix for the Hamming code, which we will come across later, is given as below

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

Proposition 2.2.6. *The rows of* $\mathbf{G}$ *form a basis for* $\mathcal{C}$*.*

Proof. The rows of $\mathbf{G}$ are linearly independent by the definition. Now let $\mathbf{r}$ be a row in $\mathbf{G}$. Then. $\mathbf{rH^T} = 0$, which implies $\mathbf{Hr^T} = 0$. Then from Definition 2.2.1, $\mathbf{r} \in \mathcal{C}$. Since the dimension of the nullspace of $\mathbf{H}$ would be the same as the dimension of $\mathcal{C}$, we have that $n - Rk(\mathbf{H}) = n - n + k = k = Rk(\mathbf{G})$, the rows of $\mathbf{G}$ form the basis. □

An interesting fact to note here is that $\mathcal{C}$ is a vector space of dimension k in $\mathbb{F}_q^n$. Its orthogonal complement is a vector space of dimension $n-k$ and also forms a linear $(n, n-k)$ code.

Defintion 2.2.7. Define the dot product of two vectors $\mathbf{u} = u_1u_2...u_n$ and $\mathbf{v} = v_1v_2...v_n$ in $\mathbb{F}_q^n$ as $\mathbf{u}.\mathbf{v} = u_1v_1 + u_2v_2 + ... + u_nv_n$. Two vectors are said to be orthogonal if $\mathbf{u}.\mathbf{v} = 0$. Let $\mathcal{C}$ be linear (n, k) code. Then, its *dual (orthogonal) code* $\mathcal{C}^\perp$ of $\mathcal{C}$ is defined to be

$$\mathcal{C}^\perp = \{\mathbf{u} \mid \mathbf{u}.\mathbf{v} = 0 \ \forall \ \mathbf{v} \in \mathcal{C}\}$$

Then one can see that $\mathbf{GH^T} = \mathbf{0}$ and $\mathbf{HG^T} = \mathbf{0}$ and that the generator matrix and parity check matrix for $\mathcal{C}$, form the parity check matrix and generator matrix for $\mathcal{C}^\perp$ respectively.

Example 2.2.8. Let $\mathcal{C}$ be a binary $(5, 3)$ code with generator matrix

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

Then we can convert G to standard form using the following elementary row operations.

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

$$\begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

The parity check matrix for $\mathcal{C}$ would be

$$H = [A^T I_k] = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

The elements of the dual code would just be linear combinations of the parity check matrix H. Thus,

$$\mathcal{C}^{\perp} = \{(00000), (10010), (11101), (01111)\}$$

Now, we will go on define some further properties of these linear codes.

Proposition 2.2.9 (Singleton Bound)**.** *Let* $\mathbf{H}$ *be a parity check matrix for a* (n, k, d) *linear code. Let* $mld(\mathbf{H})$ *denote the minimum number of linearly dependent columns in* $\mathbf{H}$*. Then* $mld(\mathbf{H}) = d$ *and* $d \leq n - k + 1$*.*

Proof. First we need to show that $d = mld(\mathbf{H})$. From Definition 2.1.4, the minimum distance is the same as the minimum weight of the codewords. First, let us assume that any set of $d-1$ columns are linearly independent. Let $\mathbf{c} = (c_1 c_2 ... c_n) \in \mathcal{C}$ and $\bar{h}_1, \bar{h}_2, ... \bar{h}_n$ denote the columns of $\mathbf{H}$. As $\mathbf{Hc^T} = 0$, we have $c_1\bar{h}_1 + c_2\bar{h}_2 + ... + c_n\bar{h}_n = 0 \implies \sum_{i=1}^{n} c_i\bar{h}_i = 0$. The weight of $\mathbf{c}$ is the number of non-zero $c_i's$. If $w(\mathbf{c}) < d$, then at most $d-1$ $c_i's$ are non-zero. Let $k \leq d-1$ be the number of non-zero c_i's. Then w.l.o.g, we have $c_1\bar{h}_1 + ... + c_k\bar{h}_k + 0.\bar{h}_{k+1} + ... + 0.\bar{h}_n = 0$. But since $d - 1$ vectors are linearly independent and and $k \leq d - 1$, it would mean we have a non trivial linear combination of less than d columns of $\mathbb{H}$ which sums to 0, i.e., $c_1\bar{h}_1 + ... + c_k\bar{h}_k = 0$. This contradicts our assumption that there are $d-1$ linearly independent columns. Hence the minimum weight must be greater than d.
Now conversely, assume the minimum weight of the code to be greater than d. Let $t < d$ columns be linearly dependent. Then there exists scalars $\lambda_i's$ such that $\sum_{i=1}^{t} \lambda_i\bar{h}_i = 0$, where not all λ_i are zero. Now construct a vector which has λ_i in the i^{th} position $1 \leq i \leq t$, and zero elsewhere. By construction, this is a codeword $\mathbf{c} \in \mathcal{C}$, because $\mathbf{Hc^T} = 0$ as $\sum_{i=10}^{t} \lambda_i\bar{h}_i + \sum_{j=t+1}^{n} 0.\bar{h}_i = 0$. But then, weight of $\mathbf{c}$ is t, which is less than d and thus contradicts our initial assumption. Thus we have the minimum number of linearly dependent columns to be equal to the minimum distance of a code.
Since $mld(\mathbf{H})$ is the minimum number of dependent columns of $\mathbf{H}$, then we must have $mld(\mathbf{H}) \leq Rk(\mathbf{H}) + 1$. As $Rk(\mathbf{H}) = n - k$, we have $d \leq n - k + 1$. □

To prove the existence of an (n, k, d) linear code, we just need to show that there exists a $(n - k) \times k$ matrix $\mathbf{H}$, such that $mld(\mathbf{H}) = d$. The following theorem shows that such a code can *exist*, but does not provide an efficient way to find it.

Theorem 2.2.10 (Gilbert Varshamov Bound)**.** *[LP84] If*

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

then we can construct a linear (n,k) code with minimum distance d.

[6]

Proof. For a linear (n, k) code, we need to construct a $(n-k) \times k$ parity check matrix $\mathbf{H}$. For this, let the first column be any $(n-k)$ tuple over $\mathbb{F}_q^n$. There are totally q^{n-k} such tuples. The second column must be $(n-k)$ tuple that is not a scalar multiple of the first. Suppose we have chosen $j-1$ columns such that any $d-1$ columns are linearly independent. The j^{th} column must be chosen from column vectors of length $n-k$ such that it is not the 0 vector, it is not any of the previous $j-1$ columns or their scalar multiples, not the sum of any two columns or their scalar multiples... . We can stop this process at $d-2$ vectors and any of their scalar multiples. Thus there are

$$\sum_{i=0}^{d-2} \binom{j-1}{i} (q-1)^i$$

possible choices for us to choose from, such that we have $d-1$ linearly independent vectors. Thus, if the inequality of the theorem holds, it will be possible to choose a j^{th} column which is linearly independent of any $d-2$ of the first $j-1$ columns. This construction can be carried out in such a way that $\mathbf{H}$ has rank $n-k$. The resulting code has minimum distance d as there are $d-1$ linearly independent columns. □

Note, the Gilbert Varshamov bound only proves existence, and failure to attain the bound does not imply non-existence.

Example 2.2.11. Can a $(5, 2, 3)$ code exist? On applying the Gilbert Varshamov bound, we have

$$2^{5-2} > 1 + \binom{5-1}{1} \ldots \binom{5-1}{3-2}$$
$$\implies 8 > 1 + 4 \implies 8 > 5$$

Therefore, yes. Such a code can exist.

Example 2.2.12. Can a (5,3,3) binary code exist? We have $n = 5, k = 3$ and $d = 3$. If we apply the Gilbert Varshamov bound, we have

$$2^{5-3} > 1 + \binom{5-1}{1} \ldots \binom{5-1}{3-2}$$
$$\implies 4 > 1 + 4 \implies 4 > 5$$

This is not possible and hence we *cannot draw any conclusion* from this.
On the other hand, let us apply the Hamming Bound to this. Since we are dealing with linear codes, the number of codewords of dimension 3 must be 2^3. Thus, we have

$$2^5 \geq 2^3 \left(1 + \binom{5}{1}\right)$$
$$\implies 32 \geq 48$$

which is false. Hence we may conclude that there cannot exist such a code.

Now given the number of codewords and length of the codeword, the minimum distance must satisfy the following inequality.

Theorem 2.2.13 (Plotkin Bound)**.** *[LP84] If there is a linear code of length n with M codewords and minimum distance d over $\mathbb{F}_q$, then*

$$d \leq \frac{nM(q-1)}{(M-1)q}$$

(6)

Proof. Let $\mathcal{C}$ be our linear (n,k) code over $\mathbb{F}_q$. Let $1 \leq i \leq n$ be such that $\mathcal{C}$ contains a codeword with nonzero i^{th} component. Define D to be the subspace of $\mathcal{C}$ containing all codewords with i^{th} component zero. Then we have that in $\mathcal{C}/D$, there are q elements, which correspond to the q choices for the i^{th} coordinate of the codeword. We know that $|\mathcal{C}| = q^k = M$, so we have that $M/|D| = |\mathcal{C}/D|$, i.e., $|D| = q^{k-1}$. Now, the sum of the weights of all the codewords in $\mathcal{C}$ can be at most $nq^{k-1}(q-1)$. Since the total number of codewords of nonzero weight is $q^k - 1$, we get $nq^{k-1}(q-1) \geq$ sum of all weights $\geq d\left(q^k - 1\right)$, so

$$d \leq \frac{nq^{k-1}(q-1)}{q^k - 1} \leq \frac{nq^k(q-1)}{(q^k-1)\,q} = \frac{nM(q-1)}{(M-1)q}$$

□

2.3 Hamming Codes

The Hamming code is a linear error-correcting code which can detect up to two simultaneous bit errors, and correct single-bit errors; thus, reliable communication is possible when the Hamming distance between the transmitted and received bit patterns is less than or equal to one. Below, we shall see how to construct the binary Hamming code.

Pick any positive integer m. Then we construct $m \times (2^m - 1)$ matrix $\mathbf{H}$ by making the columns to be binary representations of $1, 2, ..., 2^m - 1$. Now $\{\mathbf{x} \mid \mathbf{H}\mathbf{x}^\mathbf{T} = 0\}$ is a linear code of dimension $2^m - 1 - m$. Call this $(2^m - 1, 2^m - 1 - m)$ linear code the binary Hamming code of length $2^m - 1$ and dimension $2^m - 1 - m$ [Fie04].

Defintion 2.3.1. A binary code C_m of length $n = 2^m - 1$ where $m \geq 2$, with an $m \times (2^m - 1)$ parity-check matrix $\mathbf{H}$ whose columns consist of all nonzero binary vectors of length m is called a binary Hamming code.

Example 2.3.2. The binary $(7, 4)$ Hamming code (i.e., $m = 3$) has the parity check matrix

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

The code $\mathcal{C}$ consists of the following vectors

$$\{(0,0,0,0,0,0,0), (1,0,0,0,0,1,1), (0,1,0,0,1,0,1), (0,0,0,1,1,1,1),$$

$$(0,0,1,1,0,0,1),(0,0,1,0,0,1,1),(0,0,1,0,1,1,0),(0,1,1,0,0,1,1),$$
$$(0,1,1,0,1,1,1),(0,1,0,1,1,1,1),(1,0,0,0,0,1,1),(1,0,1,1,0,1,0),$$
$$(1,1,1,0,0,0,0),(1,0,0,1,1,0,0),(1,1,0,0,1,1,0),(1,1,1,1,1,1,1)\}$$

The standard generator matrix of the Hamming code is then

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

Following this, we can also define the generalised Hamming code over q elements.

Defintion 2.3.3. A *generalized Hamming code* over $\mathbb{F}_q$ can be defined by an $m \times (q^m - 1)/(q-1)$ parity-check matrix $\mathbf{H}$ such that no two columns of $\mathbf{H}$ are multiples of each other. This gives us a $(\frac{q^m-1}{q-1}, \frac{q^m-1}{q-1} - m)$ code over $\mathbb{F}_q$.

Proposition 2.3.4. *The minimum distance between binary Hamming codes is 3.*

Proof. This is true because from Proposition 2.2.9, we know that $d = mld(\mathbf{H})$. If $d = 1$, it would mean that we have one linearly dependent vector, which implies that $\mathbf{0}$ must be in our check matrix, which is not the case. In our construction, only any two but not any three columns are linearly independent, thus we have that $mld(\mathbf{H}) = 3$. □

What is special about the Hamming code is the fact that it is a *perfect 1- error correcting code* and the minimum distance between binary Hamming codes is 3.

For $t = 1$, the Hamming Bound gives us

$$M(1 + n(q-1)) \leq q^n$$

We need to show that this in fact an equality. From the construction of a Hamming code, we see that $n = \frac{q^m-1}{q-1}$. As dimension of the code would then be $n - m$, we have $M = q^{n-m}$. Substituting in the equation, we have

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

Defintion 2.3.5. The dual of a $(\frac{q^m-1}{q-1}, \frac{q^m-1}{q-1} - m, 3)$ Hamming code, is called a simplex code which is a linear $(\frac{q^m-1}{q-1}, m, q^{m-1})$ code.

Proposition 2.3.6. *The minimum distance of a binary simplex code is* 2^{m-1} *and all non-zero codewords have weight* 2^{m-1}*.*

Proof. The parity check matrix for a binary Hamming code C_m consists of all the binary non-zero $m-tuples$, forming a matrix of dimension $m \times 2^m - 1$. The rows of this matrix would form the generator matrix for the binary simplex code. The Hamming code would be a $(n, n-m)$ code where $n = 2^m - 1$, while the simplex code would be a (n, m) code. Call the generator matrix of the simplex code Δ. The number of $0's$ and $1's$ appear equal number of times in the $m \times 2^m$ matrix, and thus the weight of each row in Δ is equal, i.e., it is $2^m/2 = 2^{m-1}$.

Now, we must show that any linear combination of any of the row vectors must also have weight 2^{m-1}. Take any two row vectors of Δ. Then in order to have weight 2^{m-1}, they must disagree at 2^{m-1} coordinates, which means they agree at $2^{m-1} - 1$ coordinates. We shall break this down into three cases.

1. $\mathbf{x_1}, \mathbf{x_2}$ agree at $\geq 2^{m-1}$ coordinates.

WLOG, pick the first two *row* vectors from Δ, call them $\mathbf{x_1}$ and $\mathbf{x_2}$ and assume they agree at least at the first $\geq 2^{m-1}$ coordinates.
We are now left with $m-2$ *rows*. Let us consider the *columns* in this $m-2 \times 2^m - 1$ matrix. Pick an arbitrary *column* vector $\mathbf{y}$. There are 2^{m-2} possibilities for $\mathbf{y}$. Then, $\mathbf{y}$ cannot occur ≥ 3 times in the first 2^{m-1} columns because the only possibilities are

$$\begin{bmatrix} 1 \\ 1 \\ \mathbf{y} \end{bmatrix} \; or \; \begin{bmatrix} 0 \\ 0 \\ \mathbf{y} \end{bmatrix}$$

(because we assumed that the first two *rows* in Δ agreed on at least the first 2^{m-1} coordinates). If we had another repetition of either of the above vectors, it would imply that Δ did not have distinct non-zero $m-tuples$.
We cannot have $\mathbf{y}$ occur only once, because then we would not be able to fill in all the columns in the first 2^{m-1} positions as $2^{m-2} < 2^{m-1}$. Thus, we must have each $\mathbf{y}$ occur exactly twice. If each $\mathbf{y}$ occurs twice, then the zero vector must also occur twice. As the columns in Δ are distinct, we must have the two possibilities to be

$$\begin{bmatrix} 1 \\ 1 \\ \mathbf{0} \end{bmatrix} \; or \; \begin{bmatrix} 0 \\ 0 \\ \mathbf{0} \end{bmatrix}$$

However, Δ does not consist of any zero column vector and hence we cannot have $\mathbf{x_1}$ and $\mathbf{x_2}$ agree at $\geq 2^{m-1}$ coordinates.

2. $\mathbf{x_1}, \mathbf{x_2}$ agree at $< 2^{m-1} - 1$ which means they disagree at $> 2^{m-1}$ coordinates.

WLOG assume that the first two *row* vectors $\mathbf{x_1}$ and $\mathbf{x_2}$ disagree at the first $2^{m-1} + 1$ coordinates. Pick an arbitrary *column* vector $\mathbf{y}$. Since $\mathbf{x_1}$ and $\mathbf{x_2}$ disagree at the first $2^{m-1}+1$ positions, then using the same logic as in 1., $\mathbf{y}$ cannot occur ≥ 3 times as we have the two

possibilities to be

$$\begin{bmatrix}1\\0\\\mathbf{y}\end{bmatrix} \text{ or } \begin{bmatrix}0\\1\\\mathbf{y}\end{bmatrix}$$

Similarly, we cannot have $\mathbf{y}$ occur only once, as $2^{m-2} \times 1 < 2^{m-1}$. If $\mathbf{y}$ occurred exactly twice, then we would have $2^{m-2} \times 2 = 2^{m-1} < 2^{m-1} + 1$. Therefore, $\mathbf{x_1}$ and $\mathbf{x_2}$ cannot agree at $< 2^{m-1} - 1$ positions.

3. The only case we are left with is they agree at *exactly* $2^{m-1} - 1$ positions. This must be true, for consider the non zero possibilities of $\mathbf{y}$. There are $2^{m-2} - 1$ possibilities. Each can occur twice. Then we have $2^{m-1} - 2$ columns filled. This leaves us with one column in the $2^{m-1} - 1$ positions at which $\mathbf{x_1}$ and $\mathbf{x_2}$ agree. This would be occupied by

$$\begin{bmatrix}1\\1\\\mathbf{0}\end{bmatrix}.$$

Thus, the weight of all codewords in a simplex code is 2^{m-1}. □

This proof can also be extended for any q, where the minimum distance of any q-ary code is q^{m-1}, and codewords have equal weight [Hal].

We define the *weight distribution* to be the number of codewords of given weight in a code. If we look at the codewords mentioned in Example 2.3.2, we find that there are seven codewords of weight 3 and 4 each, and one codeword of weight 0 and 7 each. If A_i denotes the number of words of weight i, we have $A_0 = A_7 = 0$ and $A_3 = A_4 = 7$.

Defintion 2.3.7. If A_i denotes the number of words of weight i, then the polynomial

$$A(X, Y) = \sum_{i=0}^{n} A_i X^i Y^{n-i}$$

in $\mathbb{C}[X, Y]$ is called the *weight enumerator* of a code.

From the above discussion, the weight enumerator of a binary $(7, 4)$ Hamming code is $X^7 + 7X^4Y^3 + 7X^3Y^4 + Y^7$.

We will now proceed to define MacWilliam's identity, which gives an elegant way to relate the weight distribution of a code with the weight distribution of its dual code. However, before that we need to prove some lemmas.

Defintion 2.3.8. A homomorphism χ from an additive group $(\mathbb{F}_q, +)$ to a multiplicative group $(\mathbb{C}, .)$, is called a *character* of $(\mathbb{F}_q)$. The character $\chi_0 : \mathbb{F}_q \to 1$ is the *trivial character*.

Lemma 2.3.9. *[LP84] (i)* $\sum_{a \in \mathbb{F}_q} \chi(a) = 0$ *for a nontrivial character of* $\mathbb{F}_q$.
(ii) For any $a, b \in \mathbb{F}_q$, $a \neq -b$, *we have* $\sum \chi(a)\chi(b) = 0$ *where the summation runs through all characters of* $\mathbb{F}_q$.

Proof. i) Let χ be a non trivial character. Then there exists $b \in \mathbb{F}_q$ such that $\chi(b) \neq 1$ and $b \neq 0$. Now,

$$\begin{aligned}\sum_{a\in\mathbb{F}_q} \chi(a) &= \sum_{a\in\mathbb{F}_q} \chi(a+b) \\ &= \sum_{a\in\mathbb{F}_q} \chi(a)\chi(b) \\ &= \chi(b) \sum_{a\in\mathbb{F}_q} \chi(a)\end{aligned}$$

$$\implies \Big(1 - \sum_{a\in\mathbb{F}_q} \chi(b)\Big) \sum_{a\in\mathbb{F}_q} \chi(a) = 0$$

But $\chi(b) \neq 1$ and hence we have $\sum_{a\in\mathbb{F}_q} \chi(a) = 0$.

ii) Since $a + b$ is not 0, we know there exists a non-trivial χ_1 such that $\chi_(a + b) = \chi_1(a)\chi_1(b) \neq 1$. Then we have,

$$\chi_1(a)\chi_1(b) \sum_{\chi} \chi(a)\chi(b) = \sum_{\chi} \chi(a)\chi(b)$$

Then, using the same procedure as in i), we have $\sum_\chi \chi(a)\chi(b) = 0$. □

Defintion 2.3.10. Let χ be a non trivial additive character of $\mathbb{F}_q$ and let $\mathbf{v}.\mathbf{u}$ denote the dot product for $\mathbf{v}, \mathbf{u} \in \mathbb{F}_q^n$. Then, for a fixed $\mathbf{v}$, define

$$\begin{aligned}\chi_{\mathbf{v}} : \mathbb{F}_q^n &\to \mathbb{C} \\ \mathbf{u} &\mapsto \chi(\mathbf{u}.\mathbf{v})\end{aligned}$$

where $\mathbf{u} \in \mathbb{F}_q^n$. If V is a vector space over $\mathbb{C}$ and f is a mapping from $\mathbb{F}_q^n$ to V, then we define

$$\begin{aligned}g_f : \mathbb{F}_q^n &\to V \\ \mathbf{u} &\mapsto \sum_{v\in\mathbb{F}_q^n} \chi_{\mathbf{v}}(\mathbf{u}) f(\mathbf{u})\end{aligned}$$

Lemma 2.3.11. *Let W be a subspace of $\mathbb{F}_q^n, W^\perp$ its orthogonal complement, $f : \mathbb{F}_q^n \to V$ a mapping from $\mathbb{F}_q^n$ into a vector space V over $\mathbb{C}$ and χ a nontrivial additive character of $\mathbb{F}_q$. Then*

$$\sum_{\mathbf{v}\in W} g_f(\mathbf{u}) = |W| \sum_{\mathbf{v}\in W^\perp} f(\mathbf{v})$$

Proof. We have

$$\sum_{\mathbf{u}\in W} g_f(\mathbf{u}) = \sum_{\mathbf{u}\in W} \sum_{\mathbf{v}\in\mathbb{F}_q^n} \chi_{\mathbf{v}}(\mathbf{u}) f(\mathbf{v}) = \sum_{\mathbf{v}\in\mathbb{F}_q^n} \sum_{\mathbf{u}\in\mathbf{W}} \chi(\mathbf{v}\cdot\mathbf{u}) f(\mathbf{v})$$

$$= |W| \sum_{\mathbf{v} \in W^\perp} f(\mathbf{v}) + \sum_{\mathbf{v} \notin W^\perp} \sum_{a \in \mathbf{F}_q} \sum_{\substack{\mathbf{u} \in \mathbf{W} \\ \mathbf{u}.\mathbf{v} = a}} \chi(a) f(\mathbf{v})$$

Here, we get the first term $|W| \sum_{\mathbf{v} \in W^\perp} f(\mathbf{v})$ because if $\mathbf{u} \in \mathbf{W}$ and $\mathbf{v} \in W^\perp$, then $\chi(\mathbf{u}.\mathbf{v}) = \chi(0) = 1$, and since we are summing over $\mathbf{u}$, we get the cardinality of W, which would be q^k. Also, the map $\mathbf{u} \to (\mathbf{u}.\mathbf{v})$ is a surjective map from $W \to \mathbb{F}_q$ since $\mathbf{v} \in W$. So the number of vectors $\mathbf{u} \in W$ such that $\mathbf{v}.\mathbf{u} = a$, where $a \in \mathbb{F}_q$, is a constant $|W|/q$. Since $\sum_{a \in \mathbf{F}_q} \chi(a) = 0$ by i) in Lemma 2.3.9, we have

$$\sum_{\mathbf{u} \in W} g_f(\mathbf{u}) = |W| \sum_{\mathbf{v} \in W^\perp} f(\mathbf{v}) + \frac{|W|}{q} \sum_{\mathbf{v} \notin W^\perp} f(\mathbf{v}) \sum_{a \in \mathbf{F}_q} \chi(a) = |W| \sum_{\mathbf{v} \in W^\perp} f(\mathbf{v})$$

□

Now let V be the space of polynomials in two indeterminates X, Y over $\mathbb{C}$, and f be defined as $f(\mathbf{v}) = X^{w(\mathbf{v})} Y^{n - w(\mathbf{v})}$, where $w(\mathbf{v})$ denotes the weight of $\mathbf{v} \in \mathbb{F}_q^n$.

Theorem 2.3.12 (MacWilliam's Identity)**.** *[LP84] Let $\mathcal{C}$ be a linear (n, k) code over $\mathbb{F}_q$ and $\mathcal{C}^\perp$ its dual code. If $A(X, Y)$ is the weight enumerator of $\mathcal{C}$ and $A^\perp(X, Y)$ is the weight enumerator of $A^\perp$, then*

$$A^\perp(X, Y) = \frac{1}{q^k} A(Y - X, Y + (q-1)X)$$

Proof. Let $f : \mathbb{F}_q^n \to \mathbb{C}[X, Y]$ be as defined above. Then

$$A^\perp(X, Y) = \sum_{v \in W^\perp} f(\mathbf{v})$$

Let g_f be defined as in Definition 2.3.10, and for $v \in \mathbb{F}_q$, define

$$|v| := \begin{cases} 1 & \text{if } v \neq 0 \\ 0 & \text{if } v = 0 \end{cases}$$

For $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{F}_q^n$ we have

$$g_f(\mathbf{u}) = \sum_{\mathbf{v} \in \mathbb{F}_q^n} \chi(\mathbf{v} \cdot \mathbf{u}) X^{w(\mathbf{v})} Y^{n - w(\mathbf{v})}$$

$$= \sum_{v_1, \ldots, v_n \in \mathbb{F}_q} \chi\left(u_1 v_1 + \ldots + u_n v_n\right) X^{|v_1| + \ldots + |v_n|} Y^{(1 - |v_1|) + \ldots + (1 - |v_n|)}$$

$$= \sum_{v_1, \ldots, v_n \in \mathbb{F}_q} \prod_{i=1}^{n} \left(\chi\left(u_i v_i\right) X^{|v_i|} Y^{1 - |v_i|}\right)$$

$$= \prod_{i=1}^{n} \sum_{v \in \mathbb{F}_q} \left(\chi\left(u_i v\right) X^{|v|} Y^{1 - |v|}\right)$$

If $u_i = 0$ we have $\chi(u_i v) = \chi(0) = 1$, and then that term in the product is $(q-1)X + Y$. For $u_i \neq 0$ the factor is $Y + X\sum_{v\in\mathbb{F}_q^*}\chi(v) = Y - X$. Therefore

$$g_f(\mathbf{u}) = (Y-X)^{w(\mathbf{u})}(Y+(q-1)X)^{n-w(\mathbf{u})}$$

Lemma 2.3.11 implies

$$|C|A^{\perp}(X,Y) = |C|\sum_{v\in C^{\perp}} f(\mathbf{v}) = \sum_{u\in C} g_f(\mathbf{u}) = A(Y-X, Y+(q-1)X)$$

Since $|\mathcal{C}| = q^k$, we have the above identity. □

Let us see an application of this identity. In order to find the weight enumerator of the binary Hamming code, we will use the weight enumerator of the simplex code, which is easier to find because of Proposition 2.3.6. Let the binary Hamming code be a $(n, n-m)$ code, where $n = 2^m - 1$ while the simplex code would then be a (n, m) code. As the weight enumerator is given by $A(X,Y) = \sum_{i=0}^{n} A_i X^i Y^{n-i}$ for the simplex code, we have $A_0 = 1$ and $A_{2^{(m-1)}} = 2^m - 1 = n$ (from Proposition 2.3.6), and the weight enumerator must be

$$Y^n + nX^{2^{(m-1)}}Y^{2^{(m-1)}-1}$$

Using MacWilliams identity, we have

$$A^{\perp}(X,Y) = \frac{1}{q^k}A(Y-X, Y+(q-1)X)$$

Hence, the weight enumerator of the binary Hamming code must be

$$\frac{1}{2^{n-m}}(Y-X)^n + n(Y-X)^{2^{m-1}}(Y+X)^{2^{(m-1)}-1}$$

2.4 Cyclic Codes

Cyclic codes are linear codes that satisfy certain properties and have a rich algebraic structure. The idea we use here is the fact that it is extremely useful to associate codewords with polynomials.

Defintion 2.4.1. A *cyclic shift* T is a map $T : \mathbb{F}_q^n \to \mathbb{F}_q^n$ that maps

$$(c_0, c_1, ..., c_{n-1}) \mapsto (c_{n-1}, c_0, c_1, ..., c_{n-2})$$

Defintion 2.4.2. A linear code of length n is called a *cyclic code* if it is invariant under a cyclic shift, i.e.,

if $(c_0, c_1, ..., c_{n-1}) \in \mathcal{C}$, then $(c_{n-1}, c_0, c_1, ..., c_{n-2}) \in \mathcal{C}$.[7]

Let $\mathbf{c} = (c_1, c_2, ..., c_{n-1})$ be a codeword in $\mathcal{C}$, then we associate a polynomial called the *code polynomial* with it, defined as $c(X) = c_0 + c_1X + ... + c_{n-1}X^{n-1}$. This would mean that the shifted codeword $T(\mathbf{c}) = (c_{n-1}, c_0, ..., c_{n-2})$ would have the associated code polynomial $\tilde{c}(X) = c_{n-1} + c_0X + ... + c_{n-2}X^{n-1}$.

We see that $\tilde{c}(X) = Xc(X) - c_{n-1}(X^n - 1)$. This means that $\tilde{c}(X)$ is the remainder when $Xc(X)$ is divided by $X^n - 1$. Thus we have $\tilde{c}(X) = Xc(X)\ (mod(X^n - 1))$. Define the isomorphism $\psi : \mathbb{F}_q^n \to \mathbb{F}_q[X]/(X^n - 1)$ that maps

$$(c_0, c_1, ..., c_{n-1}) \mapsto c_0 + c_1X + ... + c_{n-1}X^{n-1}$$

Theorem 2.4.3. *[Hal] [Wal20] A linear code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is cyclic if and only if $\mathcal{C}$ is a principal ideal in $\mathbb{F}_q[X]/(X^n-1)$, generated by $g(X) \in \mathcal{C}$. This $g(X)$ is a monic divisor of X^n-1.*

Proof. Assume $\mathcal{C}$ is cyclic. We need to show $\psi(\mathcal{C})$ is an ideal.

Let $\mathbf{c} = (c_0, ...c_{n-1}) \in \mathcal{C}$. Then $T(\mathbf{c}) = (c_{n-1}, c_0, ..., c_{n-2}) \in \mathcal{C}$ because $\mathcal{C}$ is cyclic. Now, consider $Xc(X)$. We have

$$Xc(X) = c_{n-1} + c_0X + ... + c_{n-2}X^{n-1} \pmod{X^n-1}$$

As $(c_{n-1}, c_0, ..., c_{n-2}) \in \mathcal{C}$, we have $Xc(X) \in \psi(\mathcal{C})$. Similarly, we have that for any i, $X^ic(X) \in \psi(\mathcal{C})$. By linearity for any $a_i \in \mathbb{F}_q$, we have $a_iXc(X) \in \psi(\mathcal{C})$, and thus $\psi(\mathcal{C})$ is an ideal.

Conversely, assume $\psi(\mathcal{C})$ is an ideal. If $c(X) \in \psi(\mathcal{C})$, then $Xc(X) \in \psi(\mathcal{C})$. This means that $(c_{n-1}, c_0, ..., c_{n-2}) \in \mathcal{C}$ and thus $\mathcal{C}$ is cyclic.

In order to show that $\psi(\mathcal{C}$ is a principal ideal, we consider $\mathcal{I}$ to be an ideal in $\mathbb{F}_q[X]/(X^n-1)$. Let $g(X)$ be a monic polynomial of least degree in $\mathcal{I}$. Then for any $h(X) \in \mathcal{I}$ by division algorithm, we have $h(X) = g(X)q(X) + r(X)$, for some $q(X), r(X) \in \mathbb{F}_q[X]$ and $deg(r(X)) < deg(g(X))$. Then, we have $r(X) = h(X) - g(X)q(X)$ which means $r(X) \in \mathcal{I}$. But $g(X)$ is supposed to have least degree in $\mathcal{I}$, hence $r(X) = 0$. Thus $\mathcal{I}$ is a principal ideal generated by $g(X)$.

Now, let $X^n-1 = g(X)h(X)+s(X)$ for some $h(X), s(X) \in \mathbb{F}_q$ with $deg(s(X)) < deg(g(X))$. Since X^n-1 is in the kernel of $\mathbb{F}_q[X]/(X^n-1)$, we have $-s(X) = h(X)g(X)$ which implies that $s(X) \in \mathcal{I}$. Then using the same argument as above, $s(X) = 0$ and thus $g(X) \mid X^n-1$. □

Defintion 2.4.4. The polynomial $g(X)$ is called the *generator polynomial* of the cyclic code $\mathcal{C}$. The polynomial $h(X)$ is called the *check polynomial* of the cyclic code $\mathcal{C}$.

If the generator polynomial is $g(X) = g_0 + g_1X + ... + g_mX^m$, with $m < n$ then for a linear (n, k) code, where $k = n - m$, the generator matrix is defined to be [6]

$$\mathbf{G} = \begin{bmatrix} g_0 & g_1 & \cdots & g_m & 0 & \cdots & 0 \\ 0 & g_0 & \cdots & g_{m-1} & g_m & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & g_0 & g_1 & \cdots & g_m \end{bmatrix} = \begin{bmatrix} g \\ xg \\ \vdots \\ x^{k-1}g \end{bmatrix}$$

Example 2.4.5. Let us keep $n = 7$ and $q = 2$. Over $\mathbb{F}_2$, $X^7 - 1$ factorises into

$$X^7 - 1 = (X + 1)(X^3 + X + 1)(X^3 + X^2 + 1)$$

The generator polynomials are as follows:

i) $1 = 1$

The matrix generated by this polynomial is 7×7 identity matrix. The code is the entire space $\mathbb{F}_2^7$.

ii) $X + 1 = X + 1$

The matrix generated by this polynomial is

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

The code generated by this matrix is equivalent to the $(7, 6)$ parity check matrix.

iii) $X^3 + X + 1 = X^3 + X + 1$
The matrix generated by this polynomial is

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

We see that the code generated by this matrix is a $(7, 4)$ linear code. This code is equivalent to the $(7, 4)$ Hamming code with minimum distance 3.

iv) $X^3 + X^2 + 1 = X^3 + X^2 + 1$
The matrix generated by this polynomial is

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

We see that the code generated by this matrix is a $(7, 4)$ linear code. This code is also equivalent to the $(7, 4)$ Hamming code.

v) $(X + 1)(X^3 + X + 1) = X^4 + X^3 + X^2 + 1$
The matrix generated by this polynomial is

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

This code is a linear $(7, 3)$ code, which is equivalent to the $(7, 3)$ simplex code.

vi) $(X + 1)(X^3 + X^2 + 1) = X^4 + X^2 + X + 1$
The matrix generated by this polynomial is

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

This will also be a linear $(7,3)$ code, which is equivalent to the $(7,3)$ simplex code.

vii) $(X^3+X+1)(X^3+X^2+1) = X^6+X^5+X^4+X^3+X^2+X+1$
The matrix generated by this polynomial is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

This corresponds to the $(7,1)$ repetition code.

viii) $(X+1)(X^3+X^2+1)(X^3+X+1) = X^7-1$ The 0 code is generated by this polynomial. [7]

To illustrate the above claims, let us consider iii). We have

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

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

This corresponds to the same generator matrix mentioned in Example 2.3.2. We can do the same with iv),v),vi) and vii).

Proposition 2.4.6. *If $\mathcal{C}$ is the cyclic code of length n with check polynomial $h(X)$, then*

$$\mathcal{C} = \{c(X) \in \mathbb{F}_q[X]/(X^n-1) \mid c(X)h(x) \pmod{X^n+1}\}$$

Proof. If $c(X) \in \mathcal{C}$, then by Theorem 2.4.3, $c(X) = g(X)q(X)$. Then we have $c(X)h(X) = q(X)g(X)h(X) = q(X)(X^n-1) \pmod{X^n-1} = 0 \pmod{X^n-1}$.
Conversely, let $c(X)$ be any polynomial in $\mathbb{F}_q[X]/(X^n-1)$ such that
$c(X)h(X) = p(X) \pmod{X^n-1}$. Then,

$$\begin{aligned} c(X)h(X) &= p(X)(X^n-1) \\ &= p(X)g(X)h(X) \end{aligned}$$

and so,

$$(c(X) - p(X)g(X))h(X) = 0 \pmod{X^n-1}$$

As $h(X)g(X) = X^n-1$, $h(X) \neq 0$ and thus we have $c(X) = p(X)g(X)$ and is thus contained in $\mathcal{C}$. □

Proposition 2.4.7. *[Hal] The dual code of a cyclic code is also cyclic.*

Proof. Let a cyclic code $\mathcal{C}$ of length n have generator polynomial $g(x)$ of degree r and check polynomial $h(x)$ of degree $k = n - r = dim(\mathcal{C})$. As $h(x)$ is a divisor of $X^n - 1$, it is the generator polynomial of some other cyclic code $\mathcal{D}$ of length n and dimension $n - k = n - (n - r) = r$. We have

$$\mathcal{C} = \{q(X)g(X) \mid q(X) \in \mathbb{F}_q[X]_k\}$$

and

$$\mathcal{D} = \{p(X)h(X) \mid p(X) \in \mathbb{F}_q[X]_r\}$$

where $\mathbb{F}_q[X]_k$ and $\mathbb{F}_q[X]_r$ denote the ring of polynomials over $\mathbb{F}_q$ with degree less than k and r respectively. Let $c(X) = q(X)g(X) \in C$, so that $\deg(q(X)) \leq k - 1$ and let $d(X) = p(X)h(X) \in D$, so that $deg(p(X)) \leq r - 1$.

$$\begin{aligned} c(X)d(X) &= q(X)g(X)p(X)h(X) \\ &= q(X)p(X)\left(X^n - 1\right) \\ &= s(X)\left(X^n - 1\right) \\ &= s(X)X^n - s(X) \end{aligned}$$

where $s(X) = q(X)p(X)$ and

$$deg(s(X)) \leq (k-1) + (r-1) = r + k - 2 = n - 2 < n - 1$$

Therefore the coefficient of X^{n-1} in $c(X)d(X)$ is 0. If $c(X) = \sum_{i=0}^{n-1} c_i X^i$ and $d(X) = \sum_{j=0}^{n-1} d_j X^j$, then in general the coefficient of X^m in $c(X)d(X)$ is $\sum_{i+j=m} c_i d_j$. But for $m = n = 1$, our coefficient of X^{n-1} is zero, and thus we have

$$\begin{aligned} 0 &= \sum_{i+j=n-1} c_i d_j \\ &= \sum_{i=0}^{n-1} c_i d_{n-1-i} \\ &= c_0 d_{n-1} + c_1 d_{n-2} + \cdots + c_i d_{n-i} + \cdots + c_{n-1} d_0 \\ &= \mathbf{c} \cdot \mathbf{d}^* \end{aligned}$$

where $\mathbf{c} = (c_0, c_1, ..., c_{n-1})$ and $\mathbf{d}^* = (d_{n-1}, d_{n-2}, ..., d_0)$. This means that the dot product of each codeword of $\mathcal{C}$ with the reverse of the codeword $\mathbf{d}$ in$\mathcal{D}$ is zero. So we have $\mathcal{D}^{[-1]} \subset \mathcal{C}^{\perp}$, where $\mathcal{D}^{[-1]}$ denotes the reverse code of $\mathcal{D}$[2] [Hal].

On the other hand, we know that $dim(\mathcal{C}^{\perp}) = n - dim(\mathcal{C}) = n - n + r = r$. But the $dim(\mathcal{D}) = dim(\mathcal{D}^{\perp}) = r$. Hence, $\mathcal{C}^{\perp}$ is cyclic. □

As we saw in Example 2.4.5, the dual code of the Hamming code, i.e, the simplex code, is also cyclic.

[2]The reverse code of a code $\mathcal{C}$ is got by reversing each codeword, and is still cyclic. Refer to J.I Hall, *"Notes on Coding Theory", Chapter 8: Cyclic Codes.*

2.5 Generalised Reed Solomon Codes

Take a finite field $\mathbb{F}_q$ and choose $\mathbf{v} = (v_1, v_2, ..., v_n)$ and $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)$ in $\mathbb{F}_q^n$ and $\alpha, v_i \in \mathbb{F}_q$, where each α_i is distinct and $v_i \neq 0$. For $0 \leq k \leq n$ we define *Generalised Reed Solomon codes* to be as follows

$$\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v}) = \{(v_1 f(\alpha_1)), (v_2 f(\alpha_2)), ..., (v_n f(\alpha_n)) \mid f(X) \in \mathbb{F}_q[X]_k\}$$

where $\mathbb{F}_q[X]_k$ denotes the polynomial in $\mathbb{F}_q[X]$ with degree less than k.

If $f(X)$ is a polynomial, then we denote $\mathbf{f}$ to be the code generated by that polynomial. This basically means that we use the evaluation homomorphism mentioned in Proposition 1.1.3, on $f(X)$ at $\boldsymbol{\alpha}$ and then scale it by $\mathbf{v}$. So we have

$$ev_{(\boldsymbol{\alpha},\mathbf{v})} f(X) = \{(v_1 f(\alpha_1)), (v_2 f(\alpha_2)), ..., (v_n f(\alpha_n))$$

Theorem 2.5.1 (7)**.** *The* $\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})$ *is* (n, k) *linear code with length* $n < |\mathbb{F}_q|$*, and* $d_{min} = n - k + 1$ *when* $k \neq 0$*.*

Proof. From the definition, as each α_i is distinct, we must have $n < q$.

For $a \in \mathbb{F}_q$ and $f(X), g(X) \in \mathbb{F}_q[X]_k$, we have $af(X) + g(X)$ is also in $\mathbb{F}_q[X]_k$. And so we have

$$ev_{(\boldsymbol{\alpha},\mathbf{v})}(af(X) + g(X)) = a(ev_{(\boldsymbol{\alpha},\mathbf{v})}(f(X))) + ev_{(\boldsymbol{\alpha},\mathbf{v})}(g(X)) = a\mathbf{f} + \mathbf{g}$$

and thus $\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})$ is a (n, k) linear code over $\mathbb{F}_q$.

Let $f(X), g(X) \in \mathbb{F}_q[X]_k$ and set $h(X) = f(X) - g(X) \neq 0$. Then $h(X)$ must also belong in $\mathbb{F}_q[X]_k$. Then $\mathbf{h} = \mathbf{f} - \mathbf{g}$ and $w(\mathbf{h}) = d(\mathbf{f}, \mathbf{g})$. But weight of $\mathbf{h}$ would be n minus the number of zeroes in $\mathbf{h}$. As all v_i's are non-zero, this implies that the weight must be n minus the number of roots of $h(X)$ in $\{\alpha_1, ..., \alpha_n\}$. As the degree of $h(X)$ can be at most $k - 1$, which means we can have at most $k - 1$ roots to this polynomial, we have that the weight of $\mathbf{h}$ must be at least $n - (k - 1) = n - k + 1$. So $d_{min} \geq n - k + 1$. However from the Singleton bound mentioned in Proposition 2.2.9, we must have an equality. Thus $d_{min} = n - k + 1$.

Since each distinct polynomial $f(X), g(X) \in \mathbf{F}_q[X]_k$ would generate a different codeword, we have the dimension of the code to be k with q^k codewords. □

It is obvious now that any $f_1(X), ..., f_k(X) \in \mathbb{F}_q[X]_k$ would form the basis of a code. Usually, we take the set of monomials $\{1, X, ..., X^{k-1}\}$. Our codewords would then be

$$ev_{(\boldsymbol{\alpha},\mathbf{v})}(X^i) = \{v_1 \alpha_1^i, v_2 \alpha_2^1, ..., v_n \alpha_n^i\}$$

Thus, our canonical generator matrix would then be

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_n \\ v_1\alpha_1 & v_2\alpha_2 & \cdots & v_n\alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ v_1\alpha_1^{k-1} & v_2\alpha_2^{k-1} & \cdots & v_n\alpha_n^{k-1} \end{bmatrix}$$

From polynomial interpolation, we know that any polynomial of degree less than k is distinctly determined by its values at k (or more) points. In the above situation, any codeword with as many as k entries 0, corresponds to a polynomial of degree less than k, which is 0 at at least k points, and hence must be the zero polynomial itself.

We will see that given an n-tuple $\mathbf{f}$, we can construct a unique a polynomial for it. For this we would require Lagrange's interpolation.

Lemma 2.5.2 (Lagrange's interpolation)**.** *Let $f(X)$ be a polynomial of degree d in $F[x]$, where F is any field. Assume that, for distinct $\alpha_1, ..., \alpha_n$ of F with $d < n$, we have $f(\alpha_i) = \beta_i$. Then*

$$f(X) = \sum_{i=1}^{n} \beta_i \left(\prod_{j \neq i} \frac{x - \alpha_j}{\alpha_i - \alpha_j} \right)$$

Proof. Call the polynomial on the right hand side as $g(X)$. Then $g(X)$ is clearly a polynomial of degree at most $n-1$ as for every i, we will have $n-1$ $j's$. Also, clearly, $g(\alpha_i) = \beta_i$. Now, $f(X) - g(X)$ must have degree at most $n-1$ (as $d < n$). A polynomial of degree $n-1$ can have at most $n-1$ roots, but here, α_i for $i = 1, 2, ..., n$ are all roots of $f(X) - g(X)$. Hence the polynomial $f(X) - g(X)$ must be the zero polynomial and thus $f(X) = g(X)$. □

Now, we will define some terms. Define

$$L(X) = \prod_{i=1}^{n} (X - \alpha_i)$$

and

$$L_i(X) = L(X)/(X - \alpha_i) = \prod_{i \neq j} (X - \alpha_j)$$

Then the polynomials $L(X)$ and $L_i(X)$ are of degree n and $n-1$ respectively. Given $\mathbf{f}$, the we know that our i^{th} coordinate is basically $v_i f(\alpha_i)$. Thus, using Lagrange's Interpolation, we have

$$f(X) = \sum_{i=1}^{n} \frac{L_i(X)}{L_i(\alpha_i)} f(\alpha_i)$$

Now, we will see that the dual code of a $\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})$ is in fact, also a *GRS* code.

Theorem 2.5.3. *[Hal] Given $\mathbf{u} = (u_1, ..., u_n)$ and defining $\mathbf{v} = (v_1, ..., v_n)$ by the relation $u_i^{-1} = v_i \prod_{i \neq j} (\alpha_j - \alpha_i)$, we have*

$$\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})^{\perp} = \mathbf{GRS}_{n,n-k}(\boldsymbol{\alpha}, \mathbf{u})$$

Proof. To prove this, we essentially just need to prove that the product between every $\mathbf{f}$ in $\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})$ and $\mathbf{g}$ in $\mathrm{GRS}_{n,n-k}(\boldsymbol{\alpha}, \mathbf{u})$ is 0. By the definition, we have $u_i = v_i^{-1} L_i(\alpha_i)^{-1}$. Let $\mathbf{f} = ev_{\boldsymbol{\alpha},\mathbf{v}}(f(X))$ and $\mathbf{g} = ev_{\boldsymbol{\alpha},\mathbf{u}}(g(X))$. As $\mathbf{f}$ is in $\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})$, $f(X)$ is a polynomial of degree less than k, while $\mathbf{g}$ is in $\mathrm{GRS}_{n,n-k}(\boldsymbol{\alpha}, \mathbf{u})$ and so $g(X)$ is a polynomial of degree less than $n-k$. Then $f(X) \cdot g(X)$ must have degree less than equal to $n-2$ at most. From Lagrange's interpolation,we have

$$f(X)g(X) = \sum_{i=1}^{n} \frac{L_i(X)}{L_i(\alpha_i)} f(\alpha_i)\, g(\alpha_i)$$

Equating the coefficient of X^{n-1} from the two sides gives:

$$\begin{aligned}
0 &= \sum_{i=1}^{n} \frac{1}{L_i(\alpha_i)} f(\alpha_i)\, g(\alpha_i) \\
&= \sum_{i=1}^{n} \frac{v_i \cdot v_i^{-1}}{L_i(\alpha_i)} f(\alpha_i) g(\alpha_i) \\
&= \sum_{i=1}^{n} (v_i f(\alpha_i)) \left(\frac{v_i^{-1}}{L_i(\alpha_i)} g(\alpha_i) \right) \\
&= \sum_{i=1}^{n} (v_i f(\alpha_i)) (u_i g(\alpha_i)) \\
&= \mathbf{f} \cdot \mathbf{g}
\end{aligned}$$

Thus, we have $\mathbf{f} \cdot \mathbf{g} = 0$ and $\mathrm{GRS}_{n,k}(\boldsymbol{\alpha}, \mathbf{v})^{\perp} = \mathrm{GRS}_{n,n-k}(\boldsymbol{\alpha}, \mathbf{u})$. □

Lemma 2.5.4. *Given $\alpha_i \in \mathbb{F}_q$, we have $L_i(\alpha_i) = L'(\alpha_i)$ where $L'(X)$ is the formal derivative of $L(X)$.*

Proof. We know that $L_i(X) = L(X)/(X - \alpha_i)$ which implies $(X - \alpha_i)L_i(X) = L(X)$. Taking derivatives on both sides, by product rule, we have

$$L_i'(X)(X - \alpha_i) + L_i(X) = L'(X)$$

At $X = \alpha_i$, we have

$$L_i'(\alpha_i)(\alpha_i - \alpha_i) + L_i(X) = L'(X)$$

and so $L_i(\alpha_i) = L'(\alpha_i)$ as required □

Lemma 2.5.5. *If $\{\alpha_1, ..., \alpha_n\}$ consists of the n^{th} roots of unity in $\mathbb{F}_q$, then $L_i(\alpha_i) = n\alpha_i^{-1}$. If $n = q - 1$, then $L_i(\alpha_i) = n\alpha_i^{-1}$.*

Proof. The proof follows directly from Lemma 2.5.4. As the α_i's are n^{th} roots of unity, we have $L(X) = X^n - 1$. This implies $L'(X) = nX^{n-1}$. At $X = \alpha_i$, we have $L'(X) = n\alpha^{n-1} = n\alpha^n\alpha^{-1} = n\alpha^{-1}$. If $n = q - 1$, as $\mathbb{F}_q$ has characteristic p, the result follows. □

Let us now consider a special case of Generalised Reed Solomon codes that give rise to cyclic codes. First, lets define the following.

Let $a \in \mathbb{Z}$. For ζ, a primitive n^{th} of unity in $\mathbb{F}_q$, where $n \mid q - 1$, we define

$$\begin{aligned}
\boldsymbol{\zeta}^{(a)} &= \left((\zeta^0)^a, (\zeta^1)^a, ..., (\zeta^n)^a\right) \\
&= \left((\zeta^a)^0, (\zeta^a)^1, ..., (\zeta^a)^n\right)
\end{aligned}$$

Then, clearly, we have $\boldsymbol{\zeta}^{(0)} = (1, 1, ..., 1)$ and $\boldsymbol{\zeta}^{(1)} = \boldsymbol{\zeta}$.

Another observation to be made is that under the cyclic shift T, we see

$$\begin{aligned} T(\boldsymbol{\zeta}^{(a)}) &= \left(\left(\zeta^{n-1}\right)^a, \left(\zeta^0\right)^a, \ldots, \left(\zeta^j\right)^a, \ldots, \left(\zeta^{n-2}\right)^a\right) \\ &= \zeta^{-a}\left(\left(\zeta^0\right)^a, \left(\zeta^1\right)^a, \ldots, \left(\zeta^j\right)^a, \ldots, \left(\zeta^{n-1}\right)^a\right) \\ &= \zeta^{-a}\zeta^{(a)} \end{aligned}$$

Thus, a cyclic shift on $\zeta^{(a)}$ is always a scalar multiple of $\zeta^{(a)}$.

Theorem 2.5.6. $\mathrm{GRS}_{n,k}(\boldsymbol{\zeta}, \boldsymbol{\zeta}^a)$ *is cyclic.*

Proof. For $0 \leq i \leq k-1$ and $0 \leq j \leq n-1$, the (i,j)-entry of the canonical generator matrix is

$$v_j\zeta_j^i = (\zeta^j)^a(\zeta^j)^i \ = \zeta^{ja}\zeta^{ji} = (\zeta^j)^{i+a}$$

This means that our canonical generator matrix has its k rows to be $\boldsymbol{\zeta}^{(a+i)}$ for $i = 0, 1, ..., k-1$. As noticed above, this is invariant under the cyclic shift operator T, and is thus a cyclic code. □

Theorem 2.5.7. *[Hal] If* $\zeta^n = 1$ *where* $n = q-1$*, and* $\boldsymbol{\zeta} = (\zeta^0, ..., \zeta^{n-1})$*, then*

$$\mathrm{GRS}_{n,k}\left(\boldsymbol{\zeta}, \boldsymbol{\zeta}^{(a)}\right)^{\perp} = \mathrm{GRS}_{n,n-k}\left(\boldsymbol{\zeta}, \boldsymbol{\zeta}^{(1-a)}\right)$$

Proof. From Theorem 2.5.3, we know

$$\mathrm{GRS}_{n,k}\left(\boldsymbol{\zeta}, \boldsymbol{\zeta}^{(a)}\right)^{\perp} = \mathrm{GRS}_{n,n-k}(\boldsymbol{\zeta}, \mathbf{u})$$

where, for $0 \leq j \leq n-1$ and $\mathbf{v} = \boldsymbol{\zeta}^{(a)}$, and by definition, we have $u_j = v_j^{-1}L_j\left(\zeta^j\right)^{-1}$. From Lemma 2.5.5, $L_j\left(\zeta^j\right) = \left(\zeta^j\right)^{-1}$. Thus

$$\begin{aligned} u_j &= \left(\left(\zeta^j\right)^a\right)^{-1}\left(\left(\zeta^j\right)^{-1}\right)^{-1} \\ &= \zeta^{-ja}\zeta^j \\ &= \left(\zeta^j\right)^{1-a} \end{aligned}$$

Therefore $\mathbf{u} = \boldsymbol{\zeta}^{(1-a)}$, and so

$$\mathrm{GRS}_{n,k}\left(\boldsymbol{\zeta}, \boldsymbol{\zeta}^{(\zeta)}\right)^{\perp} = \mathrm{GRS}_{n,n-k}\left(\boldsymbol{\zeta}, \boldsymbol{\zeta}^{(1-\zeta)}\right)$$

□

When $n = q-1$, such codes are called *Primitive Reed Solomon Codes.*

References

[LP84] Rudolf Lidl and Guter Pilz. *Applied Abstract Algebra, Undergraduate Texts in Mathematics*. Springer-Verlag, 1984.

[Fie04] James Fiedler. "Hamming Codes". 2004. URL: https://orion.math.iastate.edu/linglong/Math690F04/HammingCodes.pdf.

[Con13] Keith Conrad. "Finite Fields". Connecticut University, Mansfield, 2013.

[LP13] I S Luthar and I B S Passi. *Algebra, Volume IV: Field Theory*. Narosa, 2013.

[Neu13] Max Neunhoffer. "Finite Fields". 2013. URL: http://www.math.rwth-aachen.de/~Max.Neunhoeffer/Teaching/ff/ffchap3.pdf.

[Wal20] Michel Waldschmidt. "An introduction to the theory of Finite Fields". 2020. URL: https://webusers.imj-prg.fr/~michel.waldschmidt/articles/pdf/FiniteFields.pdf.

[Hal] J.I. Hall. "Notes on Coding Theory". Dept. of Mathematics, Michigan State University, East Lansing, MI 48824 USA. URL: https://users.math.msu.edu/users/halljo/classes/codenotes/coding-notes.html.