AN INTRODUCTION TO LINEAR AND CYCLIC CODES

MICHAEL CALDERBANK

Abstract. We shall examine a small aspect of coding theory from an algebraic standpoint, examining the algorithms for decoding and encoding from a theoretical perspective. We assume basic knowledge of finite fields, but we shall introduce the theory of linear codes and built up to the machinery for generating cyclic codes and efficiently encoding and decoding them.

Contents

1. Preliminary Finite Field Results 1
2. Minimal Polynomials and Cyclotomic Cosets 3
3. Linear Codes: The Basics 5
4. Encoding and Decoding of Linear Codes 6
5. Cyclic Codes 9
6. Encoding and Decoding of Cyclic Codes 11
7. Conclusion 15
References 15

1. Preliminary Finite Field Results

We will not build up the entire structure of finite fields, but rather we will assume some results without proof, upon which we base the rest of our study.

Definition 1.1. Let F be a field. The *characteristic* of F is the least positive integer p such that $p \cdot 1 = 0$, where 1 is the multiplicative identity of F. If no such p exists, we define the characteristic to be 0.

Theorem 1.2. *A finite field F of characteristic p contains p^n elements for some integer $n \geq 1$.*

Definition 1.3. Let F be a field. The set

$$F[x] := \left\{ \sum_{i=0}^{n} a_i x^i : a_i \in F, n \geq 0 \right\}$$

is called the polynomial ring over F, and it is not hard to verify it is a ring with the usual addition and multiplication.

Theorem 1.4. *Let $f(x)$ be a polynomial over a field F of degree ≥ 1. Then $F[x]/(f(x))$ forms a ring, with addition and multiplication* ($\bmod f(x)$). *Furthermore $F[x]/(f(x))$ is a field if and only if $f(x)$ is irreducible.*

Date: DEADLINE AUGUST 22, 2008.

Lemma 1.5. *For every element β of a finite field F with q elements, we have $\beta^q = \beta$.*

This all builds up to the main characterization of finite fields:

Theorem 1.6. *For any prime p and an integer $n \geq 1$, there exists a unique finite field of p^n elements.*

Henceforth, it makes sense to denote the finite field with q elements by $\mathbb{F}_q$. Let α be a root of $f(x)$, where $f(x)$ is an irreducible polynomial of degree n over a field F. Then the field $F[x]/(f(x))$ can be represented as:

$$F[\alpha] = \{a_0 + a_1\alpha + \cdots + a_{n-1}\alpha^{n-1} : a_i \in F\}$$

This avoids the confusion between an element of $F[x]/(f(x))$ and a polynomial over F. Next, we introduce the concept of a primitive element in our field; it plays in integral part later on when we want to construct codes.

Definition 1.7. An element α in a finite field $\mathbb{F}_q$ is called a *primitive element* (or generator) of $\mathbb{F}_q$ if $\mathbb{F}_q = \{0, \alpha, \alpha^2, \cdots, \alpha^{q-1}\}$.

Definition 1.8. The *order* of a nonzero element $\alpha \in \mathbb{F}_q$, denoted by $\text{ord}(\alpha)$, is the smallest positive integer k such that $\alpha^k = 1$.

Lemma 1.9. *(i) The order ord(α) divides $q-1$ for every $\alpha \in \mathbb{F}_q^*$, where $\mathbb{F}_q^*$ is the multiplicative group.*
(ii) For two nonzero elements $\alpha, \beta \in \mathbb{F}_q^*$, if gcd(ord($\alpha$), ord($\beta$)) = 1, then ord($\alpha\beta$) = ord($\alpha$) $\times$ ord(β).

Proof. (i) Let m be a positive integer satisfying $\alpha^m = 1$. Write $m = a\cdot\text{ord}(\alpha) + b$ for some integers $a \geq 0$ and $0 \leq b < \text{ord}(\alpha)$. Then

$$1 = \alpha^m = \alpha^{a\cdot ord(\alpha)+b} = (\alpha^{ord(\alpha)})^a \cdot \alpha^b = \alpha^b.$$

By the definition of $\text{ord}(\alpha)$, this forces $b = 0$; hence, $\text{ord}(\alpha)$ is a divisor of m. Since $\alpha^{q-1} = 1$, we have that $\text{ord}(\alpha)|(q-1)$.
(ii) Let $r = \text{ord}(\alpha) \times \text{ord}(\beta)$. By construction, $\alpha^r = 1 = \beta^r$, since both $\text{ord}(\alpha)$ and $\text{ord}(\beta)$ are divisors of r. Thus $(\alpha\beta)^r = \alpha^r\beta^r = 1$. Therefore, $\text{ord}(\alpha\beta) \leq \text{ord}(\alpha) \times \text{ord}(\beta)$. Now, let t = $\text{ord}(\alpha\beta)$. We have

$$1 = (\alpha\beta)^{t\cdot ord(\alpha)} = (\alpha^{ord(\alpha)})^t\beta^{t\cdot ord(\alpha)} = \beta^{t\cdot ord(\alpha)}.$$

This implies that $\text{ord}(\beta)$ divides $t\cdot$ $\text{ord}(\alpha)$ by the proof of part (i), and since $\text{ord}(\alpha)$ is relatively prime to $\text{ord}(\beta)$, we have that $\text{ord}(\beta)$ divides t. Similarly, we can show that $\text{ord}(\alpha)$ divides t. Together, these imply that $\text{ord}(\alpha) \times \text{ord}(\beta)$ divides t. Thus, $\text{ord}(\alpha\beta) = t \geq \text{ord}(\alpha) \times \text{ord}(\beta)$. Combined with the inequality above, we have our desired result. □

Theorem 1.10. *(i) A nonzero element of $\mathbb{F}_q$ is a primitive element if and only if its order is $q-1$.*
(ii)Every finite field has at least one primitive element.

Proof. (i) If $\alpha \in \mathbb{F}_q^*$ has order $q-1$, then we must have that all the elements $\alpha, \alpha^2, \ldots, \alpha^{q-1}$ are distinct. This is equivalent to saying that
$\mathbb{F}_q = \{0, \alpha, \alpha^2, \ldots, \alpha^{q-1}\}$
(ii) Let m be the least common multiple of the orders of the elements in $\mathbb{F}_q^*$. If r^k is

a prime power in the factorization of m, then $r^k | \text{ord}(\alpha)$ for some $\alpha \in \mathbb{F}_q^*$ The order of $\alpha^{\frac{ord(\alpha)}{r^k}}$ is r^k. Thus if

$$m = r_1^{k_1} \cdots r_n^{k_n}$$

is the canonical factorization of m for distinct primes $r_1, \ldots, r_n$, then for each $i = 1, \ldots, n$ there exists $\beta_i \in \mathbb{F}_q^*$ with $\text{ord}(\beta_i) = r_i^{k_i}$. Lemma 1.9(ii) implies that there exists $\beta \in \mathbb{F}_q^*$ with $\text{ord}(\beta) = m$. Now $m|(q-1)$ by Lemma 1.9(i), but on the other hand, all the $q-1$ elements of $\mathbb{F}_q^*$ are roots of the polynomial $x^m - 1$ by construction, meaning that $m \geq q-1$. Hence, $\text{ord}(\beta) = m = q-1$, and the result follows from part (i). □

Note that a primitive element is hardly unique. In fact, it is not hard to show that if α is a primitive element of $\mathbb{F}_q$, then α^i is also a primitive element so long as $\gcd(i, q-1) = 1$. While we might not always have easy ways of expressing a primitive root, the fact that it exists in our field is vital for our next section.

2. Minimal Polynomials and Cyclotomic Cosets

Definition 2.1. A *minimal polynomial* of an element $\alpha \in \mathbb{F}_{q^m}$ with respect to $\mathbb{F}_q$ is a nonzero monic polynomial $f(x)$ of the least degree in $\mathbb{F}_q[x]$ such that $f(\alpha) = 0$.

Theorem 2.2. *(i) The minimal polynomial of an element of $\mathbb{F}_{q^m}$ with respect to $\mathbb{F}_q$ exists, is unique, and it is also irreducible over $\mathbb{F}_q$.*
(ii) If a monic irreducible polynomial $M(x) \in \mathbb{F}_q[x]$ has $\alpha \in \mathbb{F}_{q^m}$ as a root, then it is the minimal polynomial of α with respect to $\mathbb{F}_q$.

Proof. It's simply a matter of using the division algorithm for polynomials and the definition of minimal polynomial. We omit the details. □

Now for something a little different:

Definition 2.3. Let n be relatively prime to q. The *cyclotomic coset* of q modulo n containing i is defined by:

$$C_i = \{(i \cdot q^j (\mod n)) \in \mathbb{Z}_n : j = 0, 1, \ldots\}.$$

A subset $\{i_1, \ldots, i_t\}$ of $\mathbb{Z}_n$ is called a *complete set of representatives* of cyclotomic cosets of q modulo n if $C_{i_1}, \ldots, C_{i_t}$ are distinct and $\bigcup_{j=1}^{t} C_{i_j} = \mathbb{Z}_n$.

Example 2.4. Consider the cyclotomic cosets of 2 modulo 15:

$$C_0 = \{0\}, \quad C_1 = \{1, 2, 4, 8\}, \quad C_3 = \{3, 6, 9, 12\},$$
$$C_5 = \{5, 10\}, \quad C_7 = \{7, 11, 13, 14\}$$

We have $C_1 = C_2 = C_4 = C_8$, and similarly for other cosets. Also, the set $\{0, 1, 3, 5, 7\}$ is a complete set of representatives of cyclotomic cosets of 2 modulo 15.

The next theorem relates our two previous definitions:

Theorem 2.5. *Let α be a primitive element of $\mathbb{F}_{q^m}$. Then the minimal polynomial of α^i with respect to $\mathbb{F}_q$ is*

$$M^{(i)}(x) := \prod_{j \in C_i} (x - \alpha^j),$$

where C_i is the unique cyclotomic coset of q modulo $q^m - 1$ containing i.

Proof. (i) First of all, it is clear that α^i is a root of $M^{(i)}(x)$ as $i \in C_i$.
(ii) Let $M^{(i)}(x) = a_0 + a_1 x + \cdots + a_r x^r$, where $a_k \in \mathbb{F}_{q^m}$ and $r = |C_i|$. Raising each coefficient to its qth power, we obtain:

$$a_0^q + a_1^q x + \cdots + a_r^q x^r = \prod_{j \in C_i} (x - \alpha^{qj}) = \prod_{j \in C_{qi}} (x - \alpha^j) = \prod_{j \in C_i} (x - \alpha^j) = M^{(i)}(x)$$

Note: We obtained the first product by using the fact that $(a+b)^q = a^q + b^q$ in our finite field (Truly a student's dream expansion), and for the third product, we uesd the fact that $C_i = C_{qi}$.

Hence, $a_k = a_k^q$ for all $0 \leq k \leq r$; which means that a_k are all elements of $\mathbb{F}_q$. This means that $M^{(i)}(x)$ is a polynomial over $\mathbb{F}_q$.
(iii) Since α is a primitive element, we have $\alpha^j \neq \alpha^k$ for two distinct elements j, k of C_i. Hence, $M^{(i)}(x)$ has no multiple roots. Construct $f(x)$ such that $f(x) \in \mathbb{F}_q[x]$ and $f(\alpha^i) = 0$. Let $f(x) = f_0 + f_1 x + \cdots + f_n x^n$ for some $f_k \in \mathbb{F}_q$. Then, for any $j \in C_i$, there exists an integer l such that $j \equiv iq^l \pmod{q^m - 1}$. Hence,

$$\begin{aligned} f(\alpha^j) = f(\alpha^{iq^l}) &= f_0 + f_1 \alpha^{iq^l} + \cdots + f_n \alpha^{niq^l} = f_0^{q^l} + f_1^{q^l} \alpha^{iq^l} + \cdots + f_n^{q^l} \alpha^{niq^l} \\ &= (f_0 + f_1 \alpha^i + \cdots + f_n \alpha^{ni})^{q^l} = f(\alpha^i)^{q^l} = 0. \end{aligned}$$

This implies that $M^{(i)}(x)$ is a divisor of $f(x)$.

Altogether, (i), (ii), (iii) show that $M^{(i)}(x)$ is the minimal polynomial of α^i □

Lastly, we have the final theorem that will be useful when we study cyclic codes:

Theorem 2.6. *Let n be a positive integer with $gcd(q, n) = 1$. Suppose that m is a positive integer satisfying $n|(q^m - 1)$. Let α be a primitive element of $\mathbb{F}_{q^m}$ and let $M^{(j)}(x)$ be the minimal polynomial of α^j with respect to $\mathbb{F}_q$. Let $\{s_1, \ldots, s_t\}$ be a complete set of representatives of cyclotomic cosets of q modulo n. Then the polynomial $x^n - 1$ has the factorization into monic irreducible polynomials over $\mathbb{F}_q$:*

$$x^n - 1 = \prod_{i=1}^{t} M^{(\frac{(q^m-1)s_i}{n})}(x).$$

Proof. Let $r = (q^m - 1)/n$. Then α^r is a primitive nth root of unity, and thus all the roots of $x^n - 1$ are $1, \alpha^r, \alpha^{2r}, \ldots, \alpha^{(n-1)r}$. Thus, by the definition of the minimal polynomial, the polynomials $M^{(ir)}(x)$ are divisors of $x^n - 1$, for all $0 \leq i \leq n-1$. As a result, we have:

$$x^n - 1 = lcm(M^{(0)}(x), M^{(r)}(x), M^{(2r)}(x), \ldots, M^{((n-1)r)}(x)).$$

In order to factorize $x^n - 1$, it suffices to determine all of the distinct polynomials among $M^{(0)}(x), M^{(r)}(x), M^{(2r)}(x), \ldots, M^{((n-1)r)}(x)$. By Theorem 2.5 and the definition of cyclotomic coset, we know that $M^{(ir)}(x) = M^{(jr)}(x)$ if and only if ir and jr are in the same cyclotomic coset of q modulo $q^m - 1 = rn$; This is equivalent to saying that i and j are in the same cyclotomic coset of q modulo n. This implies that $M^{(s_1 r)}(x), M^{(s_2 r)}(x), \ldots, M^{(s_t r)}(x)$ are all the distinct polynomials among $M^{(0)}(x), M^{(r)}(x), M^{(2r)}(x), \ldots, M^{((n-1)r)}(x)$. □

We have the following corollary:

Corollary 2.7. *Let n be a positive integer with $gcd(q, n) = 1$. Then the number of monic irreducible factors of $x^n - 1$ over $\mathbb{F}_q$ is equal to the number of cyclotomic cosets of q modulo n.*

We give an example to clarify all the new language and symbols:

Example 2.8. Consider the polynomial $x^{21} - 1$ over $\mathbb{F}_2$. After a few calculations, we see that $\{0, 1, 3, 5, 7, 9\}$ is a complete set of representatives of cyclotomic cosets of 2 modulo 21. Since 21 is a divisor of $2^6 - 1$, we consider the field $\mathbb{F}_{64}$. Let α be a root of $1 + x + x^6$. We verify that α is indeed a primitive root of $\mathbb{F}_{64}$ (by checking $\alpha^3 \neq 1, \alpha^7 \neq 1, \alpha^9 \neq 1, \alpha^{21} \neq 1$). In our case, we have $q = 2$, $m = 6$, n =21. Thus, $r = \frac{q^m-1}{n} = 3$ in our case. Hence, we list the cyclotomic cosets of 2 modulo 63 containing multiples of 3:

$$C_0 = \{0\}, \quad C_3 = \{3, 6, 12, 24, 48, 33\}, \quad C_9 = \{9, 18, 36\}$$
$$C_{15} = \{15, 30, 60, 57, 51, 39\}, \quad C_{21} = \{21, 42\}, \quad C_{27} = \{27, 54, 45\}.$$

From this, and a few expansions and simplifications later, we obtain:

$$M^{(0)}(x) = 1 + x$$
$$M^{(3)}(x) = \prod_{j \in C_3} (x - \alpha^j) = 1 + x + x^2 + x^4 + x^6$$
$$M^{(9)}(x) = \prod_{j \in C_9} (x - \alpha^j) = 1 + x^2 + x^3$$
$$M^{(15)}(x) = \prod_{j \in C_{15}} (x - \alpha^j) = 1 + x^2 + x^4 + x^5 + x^6$$
$$M^{(21)}(x) = \prod_{j \in C_{21}} (x - \alpha^j) = 1 + x + x^2$$
$$M^{(27)}(x) = \prod_{j \in C_{27}} (x - \alpha^j) = 1 + x + x^3$$

By Theorem 2.6, we obtain the factorization of $x^{21}-1$ over $\mathbb{F}_2$ into monic irreducible polynomials:

$$x^{21} - 1 = M^{(0)}(x) \cdot M^{(3)}(x) \cdot M^{(9)}(x) \cdot M^{(15)}(x) \cdot M^{(21)}(x) \cdot M^{(27)}(x) =$$
$$(1+x)(1+x+x^2+x^4+x^6)(1+x^2+x^3)(1+x^2+x^4+x^5+x^6)(1+x+x^2)(1+x+x^3).$$

3. Linear Codes: The Basics

Definition 3.1. A *linear code* C of length n over $\mathbb{F}_q$ is a subspace of $\mathbb{F}_q^n$. The dimension of the linear code C is the *dimension* of C as a vector space over $\mathbb{F}_q$.

In this manner, we can think of the codewords in C as n-dimensional vectors in $\mathbb{F}_q^n$. These are 'words' using the elements of $\mathbb{F}_q$ as an alphabet. The codes we refer to are not the stuff of secret agents and spies, but they are error-correcting codes. Simply, if we know that part of the message is going to garbled en route (through a noisy channel), then we would like an encoding and decoding system that protects the original message from the inevitable errors. We need a bit more machinery first:

Definition 3.2. Let $\mathbf{x}$ and $\mathbf{y}$ be two elements of $\mathbb{F}_q^n$. (Alternatively, $\mathbf{x}$ and $\mathbf{y}$ are two words of length n over the alphabet $\mathbb{F}_q$). The *(Hamming) distance* from $\mathbf{x}$ to $\mathbf{y}$, denoted by $d(\mathbf{x}, \mathbf{y})$ is defined to be the number of places at which $\mathbf{x}$ and $\mathbf{y}$ differ. If $\mathbf{x} = x_1 \cdots x_n$ and $\mathbf{y} = y_1 \cdots y_n$, then

$$d(\mathbf{x}, \mathbf{y}) = d(x_1, y_1) + \ldots + d(x_n, y_n),$$

where x_i and y_i are elements of $\mathbb{F}_q$ (words of length 1), and

$$d(x_i, y_i) = \{1 \text{ if } x_i \neq y_i \quad 0 \text{ if } x_i = y_i\}.$$

For a code C containing at least two words, the *(minimum) distance* of C, denoted by $d(C)$, is

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

Remark 3.3. It is easy to see that this definition of distance satisfies non-negativity, symmetry, and the triangle inequality, so our code C is living in a metric space.

Definition 3.4. Let $\mathbf{x}$ be a word in $\mathbb{F}_q^n$. The *(Hamming) weight* of $\mathbf{x}$, denoted by $\mathrm{wt}(\mathbf{x})$, is defined to be the number of nonzero coordinates in $\mathbf{x}$; i.e., $\mathrm{wt}(\mathbf{x}) = d(\mathbf{x}, \mathbf{0})$, where $\mathbf{0}$ is the zero word.

Now, suppose that codewords from a code C are being sent over a noisy channel. If we receive $\mathbf{x}$, then we will want to decode that to the nearest possible codeword; this is our *minimum distance decoding rule.* Simply, $\mathbf{x}$ will decode to $\mathbf{c}_x$ if $d(\mathbf{x}, \mathbf{c}_x) = \min_{\mathbf{c} \in C} d(\mathbf{x}, \mathbf{c})$.

Now it is clear that if our word undergoes $\geq d(C)$ errors in the noisy chanel, then our decoding rule flat out fails, because we could decode to an entirely different codeword. If there is a bound v to the errors our code C can handle, then we say that C is v-error-correcting. We have the following proposition, whose proof we leave as an exercise:

Proposition 3.5. *A code C is v-error-correcting if and only if $d(C) \geq 2v + 1$; equivalently, a code C with distance d is an exactly $\lfloor \frac{d-1}{2} \rfloor$-error-correcting code.*

Before we can delve into the matter of encoding and decoding, we must first introduce the concept of duality; it will be essential when we try to decode:

Definition 3.6. The *dual code* of C is $C^\perp$, the orthogonal complement of the subspace C of $\mathbb{F}_q^n$.

Theorem 3.7. *If we let C be a linear code of length n over $\mathbb{F}_q$, then $C^\perp$ is a linear code and $dim(C) + dim(C^\perp) = n$.*

4. Encoding and Decoding of Linear Codes

Just like all vector spaces, once we know a basis for the code, then we can do all sorts of heavy work:

Definition 4.1. (i) A *generator matrix* for a linear code C is a matrix G whose rows form a basis for C. A generator matrix of the form $(I_k|X)$ is said to be in *standard form.*
(ii) A *parity-check matrix* H for a linear code C is a generator matrix for the dual code $C^\perp$. A parity-check matrix of the form $(Y|I_{n-k})$ is said to be in *standard form.*

Theorem 4.2. *If $G = (I_k|X)$ is the standard form generator matrix of a code C with dimension k and length n and distance d (an $[n, k, d]$ code for short), then a parity-check matrix for C is $H = (-X^T|I_{n-k})$.*

Proof. In order for H to be a generator matrix for the dual code $C^\perp$, if G is a generator matrix for C, then we must have $HG^T = 0$. Our choice of H satisfies this, and by considering the last $n - k$ coordinates, it is clear that the rows of H are linearly independent. Hence, H satisfies all the criteria and we are done. □

Now, set G as the generator matrix of C whose ith row is the vector $\mathbf{r}_i$ in the chosen basis for G. Recall that C has dimension k and length n. Then, given a vector $\mathbf{u} = (u_1, \ldots, u_k) \in \mathbb{F}_q^k$, it is clear that

$$\mathbf{v} = \mathbf{u}G = u_1\mathbf{r}_1 + \cdots + u_k\mathbf{r}_k$$

is a codeword in C. Conversely, any $\mathbf{v} \in C$ can be written uniquely as $\mathbf{v} = \mathbf{u}G$, where $\mathbf{u} = (u_1, \ldots, u_k) \in \mathbb{F}_q^k$. Hence, every word $\mathbf{u} \in \mathbb{F}_q^k$ can be encoded as $\mathbf{v} = \mathbf{u}G$.

Remark 4.3. (i) If a linear code C has a generator matrix G in standard form: $G = (I|X)$, then we have an equally simple form for the parity-check matrix H for C: $H = (-X^T|I)$.

(ii) It is trivial to recover the message $\mathbf{u}$ from the codeword $\mathbf{v} = \mathbf{u}G$, since $\mathbf{v} = \mathbf{u}G = \mathbf{u}(I|X) = (\mathbf{u}, \mathbf{u}X)$. In other words, the first k digits in the codewod $\mathbf{v} = \mathbf{u}G$ give the message $\mathbf{u}$; they are called the *message* digits. The remaining $n - k$ digits are called *check* digits. These check digits are *redundant* and they protect the message against noise.

Now, a code is only of practical significance if there is an efficient decoding scheme. For this, we need some concepts from group theory:

Definition 4.4. Let C be a linear code of length n over $\mathbb{F}_q$, and let $\mathbf{u} \in \mathbb{F}_q^n$ be any vector of length n; we define the *coset* of C determined by $\mathbf{u}$ to be the set$C + \mathbf{u} = \{\mathbf{v} + \mathbf{u} : \mathbf{v} \in C\} = \mathbf{u} + C$.

This coset coincides with the usual notion from group theory, if we consider $\mathbb{F}_q^n$ as a finite abelian group under addition and a linear code C of length n over $\mathbb{F}_q$ as a subgroup of $\mathbb{F}_q^n$. We have the following proposition about cosets, the proof consists of a few lines for each part and is easily supplied by the reader:

Proposition 4.5. *Let C be an $[n, k, d]$ linear code over the finite field $\mathbb{F}_q$. Then,*
(i) every vector of $\mathbb{F}_q^n$ is contained in some coset of C;
(ii) for all $\boldsymbol{u} \in \mathbb{F}_q^n$, $|C + \boldsymbol{u}| = |C| = q^k$;
(iii) for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}_q^n$, $\boldsymbol{u} \in C + \boldsymbol{v}$ implies that $C + \boldsymbol{u} = C + \boldsymbol{v}$;
(iv) two cosests are either identical or they have the same intersection
(v) there are q^{n-k} different cosets of C;
(vi) for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}_q^n$, $\boldsymbol{u} - \boldsymbol{v} \in C$ if and only if $\boldsymbol{u}$ and $\boldsymbol{v}$ are in the same coset.

The last part of the proposition is crucial to decoding. Assuming that the codeword $\mathbf{v}$ is transmitted and the word $\mathbf{w}$ is received, the resulting error pattern $\mathbf{e} = \mathbf{w} - \mathbf{v} \in \mathbf{w} + C$. So by Proposition 4.5(vi), we have that $\mathbf{e}$ and $\mathbf{w}$ are in the same coset. Now, since error patterns of small weight are most likely, we choose a word $\mathbf{e}$ of least weight in the coset $\mathbf{w} + C$ and conclude that $\mathbf{v} = \mathbf{w} - \mathbf{e}$ was the codeword transmitted. Now, this decoding scheme works great for small n, but it gets cumbersome as n gets larger. Instead, we shall make use of syndromes here and for the cyclic codes later.

Definition 4.6. Let C be an $[n,k,d]$-linear code over $\mathbb{F}_q$, and let H be a parity-check matrix for C. For any $\mathbf{w} \in \mathbb{F}_q^n$, the *syndrome* of $\mathbf{w}$ is the word $S_H(\mathbf{w}) = \mathbf{w}H^T \in \mathbb{F}_q^{n-k}$.

For simplicity of notation, we will assume that the parity-check matrix H is in standard form, and thus we can drop the suffix H from the syndrome when there is no risk of ambiguity. We have the following proposition, whose proof is similar to the previous proposition

Proposition 4.7. *Let C be an $[n,k,d]$-linear code over $\mathbb{F}_q$ and let H be a parity-check matrix for C. For $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{F}_q^n$, we have*
(i) $S(\boldsymbol{u}+\boldsymbol{v}) = S(\boldsymbol{u}) + S(\boldsymbol{v})$;
(ii) $S(\boldsymbol{u}) = \boldsymbol{0}$ if and only if $\boldsymbol{u}$ is a codeword in C;
(iii) $S(\boldsymbol{u}) = S(\boldsymbol{v})$ if and only if $\boldsymbol{u}$ and $\boldsymbol{v}$ are in the same coset of C.

Remark 4.8. The proposition above asserts that we can identify a coset by its syndrome; conversely, all the words in a given coset yield the same syndrome. Thus, there is a one-to-one correspondence between the cosets and the syndromes. Since the syndromes are in $\mathbb{F}_q^{n-k}$, there are at most q^{n-k} syndromes. By Proposition 4.5(v), we know that there are q^{n-k} cosets, so there must be q^{n-k} corresponding (distinct) syndromes. As a result, *all* the vectors in $\mathbb{F}_q^{n-k}$ appear as syndromes.

Our next step is to construct a syndrome look-up table. The table matches words of least weight in a given coset (coset leaders) with their syndrome. There is more than one way to construct this table, but if we know the distance d of the code C, then we generate all the error patterns $\mathbf{e}$ with $\mathrm{wt}(\mathbf{e}) \leq \lfloor \frac{d-1}{2} \rfloor$. From the definition of distance, and Proposition 4.5, we know that these error patterns $\mathbf{e}$ have to be coset leaders, so then we simply compute the syndrome $S(\mathbf{e})$ for each of these error patterns. Using this table, decoding is very simple:

Decoding procedure using Syndromes:

Step 1: For the received word $\mathbf{w}$, compute the syndrome S($\mathbf{w}$).
Step 2: Find the coset leader $\mathbf{u}$ next to the syndrome S($\mathbf{w}$) =S($\mathbf{u}$) in the syndrome look-up table.
Step 3: Decode $\mathbf{w}$ as $\mathbf{v} = \mathbf{w} - \mathbf{u}$.

We give an example to illustrate this process:

Example 4.9. Let $q = 2$ and let $C = \{0000, 1011, 0101, 1110\}$. First, we construct a parity-check matrix H. This is not hard at all, since the 2nd and 3rd codewords in C constitute a basis for C. Hence

$$G = (I|X) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad H = (-X^T|I) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

Next, we construct our syndrome look-up table for C:

(4.10)

Coset leader $\mathbf{u}$	Syndrome S($\mathbf{u}$)
0000	00
0001	01
0010	10
1000	11

Now, say that we wanted to decode $\mathbf{w} = 1101$. The syndrome is $\mathrm{S}(\mathbf{w}) = \mathbf{w}H^T = 11$. From our syndrome look-up table, we see that the coset leader is 1000. Hence $1101 - 1000 = 0101$ was the most likely codeword sent.

5. Cyclic Codes

Linear codes are nice to study and implement, because they have algebraic structures that ensure easy encoding and decoding. However, we can do more to simplify the implementaion of codes if we require a cyclic shift of a codeword in C to still be a codeword. This requirement smells like a combinatorial structure, but we shall combine the works of the previous section to show that this has an algebraic structure.

Definition 5.1. A subset S of $\mathbb{F}_q^n$ is *cyclic* if $(a_{n-1}, a_0, a_1, \ldots, a_{n-2}) \in S$ whenever $(a_0, a_1, \ldots, a_{n-1}) \in S$. A linear code C is called a *cyclic code* if C is a cyclic set. The word $(u_{n-r}, \ldots, u_{n-1}, u_0, u_1, \ldots, u_{n-r-1})$ is said to be obtained from the word $(u_0, \ldots, u_{n-1}) \in \mathbb{F}_q^n$ by cyclically shifting r positions.

In order to convert the combinatorial structure of cyclic codes into an algebraic one, we consider the following map:

$$\pi : \mathbb{F}_q^n \to \mathbb{F}_q[x]/(x^n - 1), \quad (a_0, a_1, \ldots, a_{n-1}) \mapsto a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}. \tag{5.2}$$

It is clear that this is bijective, and from now on we will sometimes identify $\mathbb{F}_q^n$ with $\mathbb{F}_q[x]/(x^n - 1)$, and a codeword $\mathbf{u} = (u_0, u_1, \ldots, u_{n-1})$ with the polynomial $u(x) = \sum_{i=0}^{n-1} u_i x^i$. Note that by 1.4, we know that $\mathbb{F}_q[x]/(x^n - 1)$ is a ring.

Definition 5.3. Let R be a commutative ring. For simplicity's sake, we shall assume that all rings that we mention are commutative; we will not delve into noncommutative rings. A nonempty subset I of R is called an *ideal* if
(i) both $a + b$ and $a - b$ belong to I, for all a, $b \in I$.
(ii) $r \cdot a \in I$, for all $r \in R$ and $a \in I$.

Definition 5.4. An ideal I of a ring R is called a *principal ideal* if there exists an element $g \in I$ such that $I =< g >$, where

$$< g >:= \{gr : r \in R\}.$$

The element g is called a *generator* of I and I is said to be generated by g. A ring R is called a *principal ideal ring* if every ideal of R is principal.

Example 5.5. In the ring $\mathbb{F}_2[x]/(x^3 - 1)$, the subset

$$I := \{0, 1 + x, x + x^2, 1 + x^2\}$$

is an ideal. In fact, it is a principal ideal, with $I =< 1 + x >$:

$$0 \cdot (1 + x) = 1 + x^3 = 0 = (1 + x + x^2)(1 + x)$$
$$1 \cdot (1 + x) = 1 + x = (x + x^2)(1 + x)$$
$$x \cdot (1 + x) = x + x^2 = (1 + x^2)(1 + x)$$
$$x^2 \cdot (1 + x) = 1 + x^2 = (1 + x)(1 + x)$$

We note that the zero trivial ideal is clearly principal; non-zero ideals are much more interesting:

Theorem 5.6. *Let I be a nonzero ideal in $\mathbb{F}_q[x]/(x^n-1)$ and let $g(x)$ be a nonzero monic polynomial of the least degree in I. Then $g(x)$ is a generator of I. In other words, $\mathbb{F}_q[x]/(x^n-1)$ is a principal ideal ring. Also, $g(x)$ divides x^n-1.*

Proof. For any polynomial $f(x)$ of I, we have $f(x) = s(x)g(x) + r(x)$ for some polynomials $s(x), r(x) \in \mathbb{F}_q[x]$ with $\deg(r(x)) < \deg(g(x))$. This forces $r(x) = 0$, since $r(x) = f(x) - s(x)g(x) \in I$ and $g(x)$ has the lowest degree. Hence $I =< g(x) >$. The second part of the theorem follows if we substitute $x^n - 1$ for $f(x)$. (Note: we can do this, since $x^n - 1$ is the zero-element of $\mathbb{F}_q[x]/(x^n-1)$. □

Now, we are ready for the main theorem connecting cyclic codes and ideals, it will form the backbone of our further work:

Theorem 5.7. *Let π be the linear map defined in Equation 5.2. Then a nonempty subset C of $\mathbb{F}_q^n$ is a cyclic code if and only if $\pi(C)$ is an ideal of $\mathbb{F}_q[x]/(x^n-1)$.*

Proof. Suppose that $\pi(C)$ is an ideal of $\mathbb{F}_q[x]/(x^n-1)$. Then for any $\alpha, \beta \in \mathbb{F}_q \subset \mathbb{F}_q[x]/(x^n-1)$ and $\mathbf{a}, \mathbf{b} \in C$, we have $\alpha\pi(\mathbf{a}) + \beta\pi(\mathbf{b})$ is an element of $\pi(C)$; i.e., $\pi(\alpha\mathbf{a} + \beta\mathbf{b}) \in \pi(C)$, hence $\alpha\mathbf{a} + \beta\mathbf{b}$ is a codeword of C. This shows that C is a linear code.

Now let $\mathbf{c} = (c_0, c_1, \ldots, c_{n-1})$ be a codeword of C. The polynomial

$$\pi(\mathbf{c}) = c_0 + c_1x + \cdots + c_{n-2}x^{n-2} + c_{n-1}x^{n-1}$$

is an element of $\pi(C)$. Since $\pi(C)$ is an ideal, the element

$$\begin{aligned} x\pi(\mathbf{c}) &= c_0x + c_1x^2 + \cdots + c_{n-2}x^{n-1} + c_{n-1}x^n \\ &= c_{n-1} + c_0x + c_1x^2 + \cdots + c_{n-2}x^{n-1} + c_{n-1}(x^n - 1) \\ &= c_{n-1} + c_0x + c_1x^2 + \cdots + c_{n-2}x^{n-1} \end{aligned}$$

is in $\pi(C)$; so $(c_{n-1}, c_0, c_1, \ldots, c_{n-2})$ is a codeword of C. This means that C is cyclic.

Conversely, suppose that C is a cyclic code. Then, we know that Definition 5.3(i) is satisfied automatically. For any polynomial

$$f(x) = f_0 + f_1x + \cdots + f_{n-2}x^{n-2} + f_{n-1}x^{n-1} = \pi(f_0, f_1, \ldots, f_{n-1})$$

of $\pi(C)$ with $(f_0, f_1, \ldots, f_{n-1}) \in C$, the polynomial

$$xf(x) = f_{n-1} + f_0x + f_1x^2 + \cdots + f_{n-2}x^{n-1}$$

is also an element of $\pi(C)$ since C is cyclic. Thus $x^2f(x) = x(xf(x))$ is an element of $\pi(C)$. By induction, we know that $x^if(x)$ belongs to $\pi(C)$ for all $i \geq 0$. Since $\pi(C)$ is a linear code and π a linear transformation, $\pi(C)$ is a linear space over $\mathbb{F}_q$. Hence, for any $g(x) = g_0 + g_1x + \cdots + g_{n-1}x^{n-1} \in \mathbb{F}_q[x]/(x^n-1)$, the polynomial

$$g(x)f(x) = \sum_{i=0}^{n-1} g_i(x^if(x))$$

is an element of $\pi(C)$. Therefore, $\pi(C)$ is an ideal of $\mathbb{F}_q[x]/(x^n-1)$ since Definition 5.3(ii) is also satisfied. □

It is clear that we must have a *unique* monic polynomial $g(x)$ of the least degree in every nonzero ideal I, and by Theorem 5.7, we have that $g(x)$ is a generator of I. Thus, the following definition makes sense:

Definition 5.8. The unique monic polynomial of the least degree in I of $\mathbb{F}_q[x]/(x^n-1)$ is called the *generator polynomial* of I. For a cyclic code C, the generator polynomial of $\pi(C)$ is also called the *generator polynomial* of C.

It is easy to see that each monic divisor of x^n-1 is the generator polynomial of some cyclic code C in $\mathbb{F}_q^n$. From Theorem 5.7 we have the following corollary:

Corollary 5.9. *There is a one-to-one correspondence between the cyclic codes in $\mathbb{F}_q^n$ and the monic divisors of $x^n-1 \in \mathbb{F}_q[x]$.*

Remark 5.10. We can even specify the dimension of the cyclic code easily. If we let $g(x)$ be the generator polynomial of an ideal of $\mathbb{F}_q[x]/(x^n-1)$, then the corresponding cyclic code has dimension k if the degree of $g(x)$ is $n-k$.

Now, using Theorem 2.6, we factorize x^n-1 over $\mathbb{F}_q$, and we can generate cyclic codes from each of the monic divisors of x^n-1. Now we see why we went to all the trouble to define cyclotomic cosets: we get cyclic codes for free!

Example 5.11. Based on the factorization: $x^7-1 = (1+x)(1+x^2+x^3)(1+x+x^3) \in \mathbb{F}_2[x]$, we know that there are only two binary $[7,3]$ cyclic codes:

$$\begin{aligned} <(1+x)(1+x^2+x^3)> &= \{0000000, 1110100, 0111010, 0011101, \\ &\quad 1001110, 0100111, 1010011, 1101001\} \end{aligned}$$

and

$$\begin{aligned} <(1+x)(1+x+x^3)> &= \{0000000, 1011100, 0101110, 0010111, \\ &\quad 1001011, 1100101, 1110010, 0111001\} \end{aligned}$$

6. Encoding and Decoding of Cyclic Codes

We have seen that a cyclic code is determined by its generator polynomial, so such a code should also have generator matrices determined by this polynomial.

Theorem 6.1. *Let $g(x) = g_0 + g_1x + \cdots + g_{n-k}x^{n-k}$ be the generator polynomial of a cyclic code C in $\mathbb{F}_q^n$ with deg(g(x)) $= n-k$. Then the matrix*

$$G = \begin{pmatrix} g(x) \\ xg(x) \\ \cdot \\ \cdot \\ \cdot \\ x^{k-1}g(x) \end{pmatrix} = \begin{pmatrix} g_0 & g_1 & \cdot & \cdot & \cdot & g_{n-k} & 0 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & g_0 & g_1 & \cdot & \cdot & \cdot & g_{n-k} & 0 & 0 & \cdot & \cdot & 0 \\ \cdot & & & & & & & & & & & \cdot \\ \cdot & & & & & & & & & & & \cdot \\ \cdot & & & & & & & & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & g_0 & g_1 & \cdot & \cdot & \cdot & \cdot & g_{n-k} \end{pmatrix}$$

is a generator matrix of C.

Proof. It is sufficient to show that $g(x), xg(x), \ldots, x^{k-1}g(x)$ form a basis of C. Well, it is clear that they are linear independent over $\mathbb{F}_q$, and since we know $\dim(C) = k$, we are done. □

From our study of linear codes, we know that we can find a parity-check matrix of a cylic code C by manipulating the generator matrix of C (See 4.2). However, since the dual-code (3.6) $C^\perp$ is also a cyclic code, we should be able to leapfrog the step of finding a generator matrix and go straight from the generator polynomial of C to a generator polynomial of $C^\perp$.

Definition 6.2. Let $h(x) = \sum_{i=0}^{k} a_i x^i$ be a polynomial of degree $k (a_k \neq 0)$ over $\mathbb{F}_q$. We define the reciprocal polynomial $h_R(x)$ of $h(x)$ by

$$h_R(x) := x^k h(1/x) = \sum_{i=0}^{k} a_{k-i} x^i.$$

Theorem 6.3. *Let $g(x)$ be the generator polynomial of a q-ary $[n,k]$-cyclic code C. Put $h(x) = (x^n - 1)/g(x)$. Then $h_0^{-1} h_R(x)$ is the generator polynomial of $C^\perp$, where h_0 is the constant term of $h(x)$.*

Proof. Let $g(x) = \sum_{i=0}^{n-1} g_i x^i$ and let $h(x) = \sum_{i=0}^{n-1} h_i x^i$. Then

$$h_R(x) = \frac{1}{x^{n-k-1}} \sum_{i=0}^{n-1} h_{n-i-1} x^i,$$

where $k = \deg(h(x))$. Now, consider the product

$$\begin{aligned} 0 &\equiv g(x)h(x) \\ &\equiv (g_0h_0 + g_1h_{n-1} + \cdots + g_{n-1}h_1) + (g_0h_1 + g_1h_0 + \cdots + g_{n-1}h_2)x + \\ &\quad (g_0h_2 + g_1h_1 + \cdots + g_{n-1}h_3)x^2 + \cdots + \\ &\quad (g_0h_{n-1} + g_1h_{n-2} + \cdots + g_{n-1}h_0)x^{n-1} (\mod x^n - 1) \end{aligned}$$

By construction, each coefficient of each power of x in the last line above must be zero. Looking at these coefficients, we obtain $\mathbf{g}_i \cdot (h_{n-1}, h_{n-2}, \ldots, h_1, h_0) = 0$, for all $i = 0, 1, \ldots, n-1$, where $\mathbf{g}_i$ is the vector obtained from $(g_0, g_1, \ldots, g_{n-1})$ by cyclically shifting i positions. Therefore, $(h_{n-1}, h_{n-2}, \ldots, h_1, h_0)$ is a codeword of $C^\perp$ since $\{\mathbf{g}_0, \mathbf{g}_1, \ldots, \mathbf{g}_{n-1}\}$ generates C by Theorem 6.1.

By cyclically shifting the vector $(h_{n-1}, h_{n-2}, \ldots, h_1, h_0)$ by $k+1$ positions, we obtain the vector corresponding to $h_R(x)$. This implies that $h_R(x)$ is a codeword since $C^\perp$ is cyclic.

Since $\deg(h_R(x) = \deg(h(x)) = k$, the set $\{h_R(x), xh_R(x), \ldots, x^{n-k-1}h_R(x)\}$ is a basis of $C^\perp$. Hence, $C^\perp$ is generated by $h_R(x)$. Thus, the monic polynomial $h_0^{-1}h_R(x)$ is the generator polynomial of $C^\perp$. □

A quick corollary allows to construct the parity-check matrix for C directly from the generator polynomial $g(x)$:

Corollary 6.4. *Let C be a $[n,k,d]$-cyclic code with generator polynomial $g(x)$. Put $h(x) = (x^n - 1)/g(x)$. Let $h(x) = h_0 + h_1 x + \cdots + h_k x^k$. Then the matrix*

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

is a parity-check matrix of C

Although the parity-check matrix is not in standard-form, a few row manipulations will yield a parity-check matrix of the form $H = (I_{n-k}|A)$. When we decode, we shall assume that the parity-check matrix is in standard form.

Theorem 6.5. *Let $H = (I_{n-k}|A)$ be a parity-check matrix of a cyclic code C over $\mathbb{F}_q$. Let $g(x)$ be the generator polynomial of C. Then the syndrome of a vector $\boldsymbol{w} \in \mathbb{F}_q^n$ is equal to $w(x)(\bmod g(x))$, where $w(x)$ is the corresponding polynomial of $\boldsymbol{w}$.*

Proof. For each column vector of A, we can associate a polynomial of degree at most $n-k-1$ and write A as

$$A = (a_0(x), a_1(x), \ldots, a_{k-1}(x)).$$

By duality, we know that $G = (-A^T|I_k)$ is a generator matrix for C. Therefore, $x^{n-k+i} - a_i(x)$ is a codeword of C. Put $x^{n-k+i} - a_i(x) = q_i(x)g(x)$ for some $q_i(x) \in \mathbb{F}_q[x]/(x^n - 1)$; so

$$a_i(x) = x^{n-k+i} - q_i(x)g(x)$$

Suppose $w(x) = w_0 + w_1 x + \cdots + w_{n-1}x^{n-1}$. For the syndrome $\mathbf{s} = \mathbf{w}H^T$ of $\mathbf{w}$, the corresponding polynomial $s(x)$ is

$$\begin{aligned} s(x) &= w_0 + w_1 x + \cdots + w_{n-k-1}x^{n-k-1} + w_{n-k}a_0(x) + \cdots + w_{n-1}a_{k-1}(x) \\ &= \sum_{i=0}^{n-k-1} w_i x^i + \sum_{j=0}^{k-1} w_{n-k+j}(x^{n-k+j} - q_j(x)g(x)) \\ &= \sum_{i=0}^{n-1} w_i x^i - \left(\sum_{j=0}^{k-1} w_{n-k+j} q_j(x) \right) g(x) \\ &\equiv w(x)(\bmod g(x)). \end{aligned}$$

As the polynomial $s(x)$ has degree at most $n-k-1$, we are done. □

Remark 6.6. Theorem 6.5 shows that $w(x) - s(x)$ is a codeword, where $s(x)$ is the syndrome of $w(x)$. Now, if $\mathrm{wt}(s(x)) \leq \lfloor \frac{d(C)-1}{2} \rfloor$, then we can safely decode $w(x)$ to $w(x) - s(x)$, since there is no ambiguity. Unfortunately, this is not always the case, and so we will need a bit more machinery.

Lemma 6.7. *Let C be a q-ary $[n,k]$-cyclic code with generator polynomial $g(x)$. Let $s(x) = \sum_{i=0}^{n-k-1} s_i x^i$ be the syndrome of $w(x)$. Then the syndrome of the cyclic shift $xw(x)$ is equal to $xs(x) - s_{n-k-1}g(x)$.*

Proof. By Theorem 6.5, it is sufficient to show that $xs(x) - s_{n-k-1}g(x)$ is the remainder of $xw(x)$ divided by $g(x)$. Let $w(x) = q(x)g(x) + s(x)$. Then

$$xw(x) = xq(x)g(x) + xs(x) = (xq(x) + s_{n-k-1})g(x) + (xs(x) - s_{n-k-1}g(x)).$$

Since $\deg(xs(x) - s_{n-k-1}g(x)) < n-k = \deg(g(x))$. □

Now, we can generalize this to see that the syndrome of the cyclic shift $x^i w(x)$ of a word $w(x)$ can be computed through the syndrome of the cyclic shift $x^{i-1}w(x)$. Thus, the syndromes of $w(x), xw(x), x^2w(x), \ldots$ can be computed inductively.

Before we begin our theorem for syndrome decoding of cyclic codes, we have one more definition related to error patterns:

Definition 6.8. A *cyclic run* of 0 of length l of an n-tuple is a succession of l cyclically consecutive zero components.

Note that: $\mathbf{e} = (0, 0, 1, 2, 0, 0, 0, 1, 0, 0)$ has a cyclic run of 0 of length 4.

Decoding algorithm for Cyclic Codes

Let C be a q-ary $[n, k, d]$-cyclic code with generator polynomial $g(x)$. Let $w(x)$ be a received word with error pattern $e(x)$, where $\mathrm{wt}(e(x)) \leq \lfloor \frac{d-1}{2} \rfloor$ and $e(x)$ has a cyclic run of 0 of length at least k. The goal is to determine $e(x)$.

Step 1: Compute the syndromes of $x^i w(x)$, for $i = 0, 1, 2, \ldots$, and denote by $s_i(x)$ the syndrome $(x^i w(x) (\bmod\ g(x))$.

Step 2: Find m such that $\mathrm{wt}(s_m(x)) \leq \lfloor \frac{d-1}{2} \rfloor$.

Step 3: Compute the remainder $e(x)$ of $x^{n-m} s_m(x)$ divided by $x^n - 1$. Decode $w(x)$ to $w(x) - e(x)$.

Proof. First of all, we show that such an m in Step 2 exists. By assumption, we know that there exists an error pattern $e(x)$ such that $e(x)$ has a cyclic run of 0 of length at least k. Thus, there exists an integer $m \geq 0$ such that the cyclic shift of the error pattern $e(x)$ through m positions has all its non-zero coefficients within the first $n-k$ positions. Note that this cyclic shift of the error pattern $e(x)$ through m positions is in fact $(x^m w(x) (\bmod\ x^n - 1)) = s_m(x)$. Since the weight is invariant under a cyclic shift, we have our m that works.

The word $t(x) := (x^{n-m} s(x) (\bmod\ x^n - 1)$ is a cyclic shift of $(\mathbf{s}_m, \mathbf{0})$ through $n - m$ positions, where $\mathbf{s}_m$ is the vector of $\mathbb{F}_q^{n-k}$ corresponding to the polynomial $s_m(x)$ It is clear that the weight of $t(x)$ is the same as the weight of $s_m(x)$. Hence $\mathrm{wt}(t(x)) \leq \lfloor \frac{d-1}{2} \rfloor$. As

$$\begin{aligned} x^m(w(x) - t(x)) &\equiv x^m(w(x) - x^{n-m} s_m(x)) \\ &\equiv x^m w(x) - x^n s_m(x) \\ &\equiv s_m(x) - x^n s_m(x) \\ &\equiv (1 - x^n) s_m(x) \equiv 0(\ \bmod\ g(x)) \end{aligned}$$

and x^m is co-prime to $g(x)$, we have that $w(x) - t(x)$ is divisible by $g(x)$; in other words, $w(x) - t(x)$ is a codeword. As $t(x)$ and $e(x)$ are in the same coset, because of their weights we must have that $e(x) = t(x) = (x^{n-m} s_m(x) (\bmod\ x^n - 1)$. □

Example 6.9. Consider the binary $[15, 7]$-cyclic code generated by $g(x) = 1 + x^4 + x^6 + x^7 + x^8$. We can check from the parity-check matrices that the minimum distance is 5. An error pattern with weight at most 2 must have a cyclic run of 0 of length at least 7. Thus, we can correct such an error pattern using what we have proven above. Now, consider the received word:

$$w(x) = 110011101100010 = 1 + x + x^4 + x^5 + x^6 + x^8 + x^9 + x^13.$$

We can compute the syndromes $s_i(x)$ of $x^i w(x)$ until $\mathrm{wt}(s_i(x)) \leq 2$:

(6.10)

i	$s_m(x)$
0	$1 + x^2 + x^5 + x^7$
1	$1 + x + x^3 + x^4 + x^7$
2	$1 + x + x^2 + x^5 + x^6 + x^7$
3	$1 + x + x^2 + x^3 + x^4$
4	$x + x^2 + x^3 + x^4 + x^5$
5	$x^2 + x^3 + x^4 + x^5 + x^6$
6	$x^3 + x^4 + x^5 + x^6 + x^7$
7	$1 + x^5$

Then we decode $w(x) = 110011101100010$ to $w(x) - x^8 s_7(x) = w(x) - x^8 - x^{13} = 1+x+x^4+x^5+x^6+x^9 = 110011100100000$. Note that we only used the parity-check matrix to check for distance; it was not specifically used in the decoding procedure.

7. Conclusion

After establishing some well-known results of finite fields, we proceeded to define cyclotomic cosets so that we could obtain a formula for factoring $x^n - 1$ over a finite field. Then, we proceeded to lay the background for linear codes and coding theory in general. While these two sections seem to be unrelated, when we defined cyclic codes later, we were able to see that each of the monic factors of $x^n - 1$ generated a cyclic code that we could use to correct errors. In fact, the encoding and decoding process is simplified due to the pleasing algebraic structures of cyclic codes. The area of cyclic codes is very rich, and while we only looked at the general codes to correct random errors, there is a whole other set of codes if we have some more information about the error pattern.

References

[1] S. Ling and C. Xing Coding Theory: A First Course. Cambridge University Press. 2004.
[2] J. Bierbrauer Introduction to Coding Theory Chapman and Hall/CRC Press. 2005