1 Linear codes

Suppose we wish to transmit words of length k. In order to correct for errors, we will pad each word by $n-k$ additional bits, and therefore transmit words of length n.

Definition 1.1. A **linear code** C of length n and rank k is a k-dimensional subspace of $\mathbb{F}_2^n$.

Associated to any linear code are two matrices, which we now define.

Definition 1.2. The **code generator matrix** G is an $n \times k$ matrix such that if $v \in \mathbb{F}_2^k$ is an unencoded word, then $Gv \in \mathbb{F}_2^n$ is the corresponding encoded word.

Definition 1.3. A **parity-check matrix** H is an $(n-k) \times n$ matrix with $\ker H = C$.

Definition 1.4. Let $x \in \mathbb{F}_2^n$. The **Hamming weight** of x is the number of non-zero entries; we denote this by $\|x\|$. If $x, y \in \mathbb{F}_2^n$, we define the **Hamming distance** between x and y as $\|x-y\|$.

Let C be a linear code, and let $d > 0$ be the minimal Hamming weight amongst non-zero codewords of C. Observe that this is equal to the minimal Hamming distance between any two distinct codewords of C, since C is a subspace.

Suppose that we send a codeword c, but an error in transmission causes the word $w \in \mathbb{F}_2^n$ to be received instead. If w satisfies $0 < \|w-c\| < d$, then we know that an error has occurred, since the word w cannot be a codeword. Moreover, if we assume that $0 < \|w-c\| < \lfloor d/2 \rfloor$, then we can correct the error by finding the closest word in C to w, which is c.

We now provide perhaps the most well known example of a linear code.

Example 1.5. The **Hamming (7,4) code** is a linear code with matrices

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

Encoding. Suppose we would like to send the word $v = (v_1, v_2, v_3, v_4)$. We calculate the encoded word c given by

$$c = Gv = (p_1, p_2, v_1, p_3, v_2, v_3, v_4),$$

where

$$\begin{aligned} p_1 &= v_1 + v_2 + v_4, \\ p_2 &= v_1 + v_3 + v_4, \\ p_3 &= v_2 + v_3 + v_4. \end{aligned} \tag{$*$}$$

We transmit the word c.

Before we describe the decoding process, we first note the following important property of the code.

Lemma 1.6. If $c \in C$ is non-zero, then $\|c\| \geq 3$. Hence the Hamming encoding can detect up to two errors, and correct up to one error.

Proof. Suppose $c = (p_1, p_2, v_1, p_3, v_2, v_3, v_4) \in C$. If all of the v_i are 0, then each $p_i = 0$ as well so $c = 0$. If exactly one of v_i is 1, then by $(*)$ at least two of the p_i are 1 as well, so $\|c\| \geq 3$. If exactly two of the v_i are 1, then again by $(*)$ at least one of the p_i is 1, so $\|c\| \geq 3$. Lastly, if three or more of the v_i are 1, then clearly $\|c\| \geq 3$. □

Decoding. Suppose we receive the word w. In order to determine if errors have occurred, we compute the vector $s = Hw$.

Case 0, no errors: If $s = 0$ then $w \in C$, so there is no error.

Case 1, one error: If $s \neq 0$, then we assume that one error occurred. Note that this implies that $w = c + e$, where e is a vector with all entries 0, except for a 1 where the error occurred, say in position i. Hence

$$s = Hw = Hc + He = He$$

since $Hc = 0$ as c is a codeword. Since the only non-zero entry of e is in position i, the product He will be equal to the ith column of H. Thus we can determine the location of the error by inspection of Hw.

Case 2, two errors: Note that if two errors occur, we will still be able to detect the presence of an error, since the valid codewords are spaced by Hamming distance at least three. However, if we try to correct the error as in case 1 we will compute an incorrect answer, as the closest valid codeword is no longer correct. If three or more errors occur, it is possible that we will be unable to even detect the presence of an error.

2 BCH codes

BCH codes, named after their inventors Bose, Ray-Chaudhuri and Hocquenghem, are a family of multiple-error correcting codes. In order to describe them, we will require some basic theory of finite fields.

Definition 2.1. For each prime p, the set of integers modulo p forms a field $\mathbb{F}_p$. For each prime p and each $n \in \mathbb{N}$ there is a unique field of order p^n denoted $\mathbb{F}_{p^n}$, which can be realised as the quotient

$$\mathbb{F}_{p^n} = \mathbb{F}_p[x]/(f)$$

where f is an irreducible polynomial of degree n in $\mathbb{F}_p[x]$. Note that $\mathbb{F}_p$ is a subfield of $\mathbb{F}_{p^n}$.

Element	Equivalence class	Binary notation
0	0	0000
1	1	0001
α	x	0010
α^2	x^2	0100
α^3	x^3	1000
α^4	$x+1$	0011
α^5	x^2+x	0110
α^6	x^3+x^2	1100
α^7	x^3+x+1	1011
α^8	x^2+1	0101
α^9	x^3+x	1010
α^{10}	x^2+x+1	0111
α^{11}	x^3+x^2+x	1110
α^{12}	x^3+x^2+x+1	1111
α^{13}	x^3+x^2+1	1101
α^{14}	x^3+1	1001

Table 1: Elements of the field $\mathbb{F}_{16} = \mathbb{F}_2[x]/(x^4+x+1)$.

Definition 2.2. Let $\mathbb{F} \subseteq \mathbb{E}$ be fields, and $\alpha \in \mathbb{E}$. The **minimal polynomial** of α over $\mathbb{F}$, if it exists, is the unique monic polynomial $m \in \mathbb{F}[x]$ of lowest degree such that $m(\alpha) = 0$.

Note that for finite fields, minimal polynomials always exist.

Definition 2.3. If $\mathbb{F}$ is a finite field, then $\mathbb{F}^\times$ is cyclic. We say that $\alpha \in \mathbb{F}$ is a **primitive root** if it generates the group $\mathbb{F}^\times$.

We now describe a BCH code which is capable of correcting two errors, by adding 8 parity-check digits to form encoded words of length 15. The polynomial $m(x) = x^4+x+1$ is irreducible in $\mathbb{F}_2[x]$. To see this, note that it has no linear factors, and is also not divisible by x^2+x+1, which is the only irreducible quadratic polynomial in $\mathbb{F}_2[x]$. Hence $\mathbb{F}_{16} = \mathbb{F}_2[x]/(x^4+x+1)$ is a field with 16 elements. Let α be the equivalence class of x; by definition, the minimal polynomial of α is m. Note that α is primitive in $\mathbb{F}_{16}$.

Lemma 2.4. The minimal polynomial of α^3 over $\mathbb{F}_2$ is $m_3(x) = x^4+x^3+x^2+x+1$.

Proof. The polynomial $m_3(x)$ is irreducible, as it is not divisible by $x, x+1$ or x^2+x+1.

Using the fact that $\alpha^4 = \alpha + 1$, we compute

$$\begin{aligned} m_3(\alpha^3) &= \alpha^{12} + \alpha^9 + \alpha^6 + \alpha^3 + 1 \\ &= (\alpha+1)^3 + \alpha(\alpha+1)^2 + \alpha^2(\alpha+1) + \alpha^3 + 1 \\ &= (\alpha^3 + \alpha^2 + \alpha + 1) + (\alpha^3 + \alpha) + (\alpha^3 + \alpha^2) + \alpha^3 + 1 \\ &= 0. \end{aligned}$$

□

An immediate consequence is that the polynomial in $\mathbb{F}_2[x]$ of smallest degree with both α and α^3 as roots is

$$g(x) = m(x)m_3(x) = x^8 + x^7 + x^6 + x^4 + 1.$$

In fact, this polynomial also has α^2 and α^4 as roots, since for any polynomial in $\mathbb{F}_2[x]$ we have $p(x^2) = p(x)^2$.

The set of codewords is

$$C = \left\{(c_{14}, ..., c_0) \in \mathbb{F}_2^{15} \mid g \text{ divides } c_{14}x^{14} + c_{13}x^{13} + ... + c_1x + c_0\right\}.$$

Encoding. Suppose we wish to encode the 7-bit word

$$v = (v_{14}, v_{13}, ..., v_8).$$

Let $v(x)$ be the polynomial

$$v(x) = v_{14}x^{14} + v_{13}x^{13} + ... + v_8x^8,$$

and perform Euclidean division by $g(x)$, so that $v(x) = q(x)g(x) + r(x)$, and hence $v(x) + r(x) = q(x)g(x)$. Since r has degree at most 7, we can write

$$v(x) + r(x) = v_{14}x^{14} + v_{13}x^{13} + ... + v_8x^8 + r_7x^7 + r_6x^6 + ... + r_1x + r_0.$$

Since g divides $v + r$, we know that

$$c = (v_{14}, ..., v_8, r_7, ..., r_0) \in C.$$

We transmit c.

Decoding. Suppose we receive a word w with associated polynomial $w(x)$. We can write $w(x) = c(x) + e(x)$ for some polynomial e of degree at most 15. We assume that at most two errors occurred during transmission, and hence at most two coefficients in $e(x)$ are non-zero.

To begin decoding, we first calculate $s_i = w(\alpha^i)$ for $1 \leq i \leq 4$. Note that since $c(\alpha^i) = 0$, we have $s_i = e(\alpha^i)$. Form the matrix

$$S = \begin{bmatrix} s_1 & s_2 \\ s_2 & s_3 \end{bmatrix}.$$

We will see that the rank of S is equal to the number of errors which occurred.

Case 0, no errors: If no errors occurred, then each $s_i = 0$. So $\alpha, \alpha^2, \alpha^3$ and α^4 are all roots of w, and hence $w \in C$.

Case 1, one error: If one error occurred, then $e(x) = x^a$ for some $0 \leq a \leq 14$. Hence

$$S = \begin{bmatrix} \alpha^a & \alpha^{2a} \\ \alpha^{2a} & \alpha^{3a} \end{bmatrix}$$

which has rank 1. Moreover, we can determine where the error occurred by comparing to Table 1; since $s_1 = \alpha^a$, we know that the error occurred in the bit corresponding to the coefficient of x^a in $w(x)$.

Case 2, two errors: If two errors occurred, then $e(x) = x^a + x^b$ and hence

$$S = \begin{bmatrix} \alpha^a + \alpha^b & \alpha^{2a} + \alpha^{2b} \\ \alpha^{2a} + \alpha^{2b} & \alpha^{3a} + \alpha^{3b} \end{bmatrix}$$

Hence $\det S = \alpha^{a+3b} + \alpha^{3a+b}$. Note that this determinint cannot be zero, since this would require $a + 3b \equiv 3a + b$ modulo 15 and thus $a \equiv b$. So $\operatorname{rank} S = 2$.

To determine the locations of the errors, we solve the linear system

$$S \begin{bmatrix} \sigma_0 \\ \sigma_1 \end{bmatrix} = \begin{bmatrix} s_3 \\ s_4 \end{bmatrix}. \qquad (*)$$

The reason for this is that it will turn out that the roots of $x^2 + \sigma_1 x + \sigma_0$ are precisely α^a and α^b, so by consulting Table 1 again we can find both errors. We will prove this in the next section.

Example 2.5. Suppose we wish to transmit the word $v = (1, 1, 0, 1, 0, 1, 0)$. We have

$$\underbrace{x^{14} + x^{13} + x^{11} + x^9}_{v(x)} = \underbrace{(x^6 + x^4 + x)}_{q(x)}\underbrace{(x^8 + x^7 + x^6 + x^4 + 1)}_{g(x)} + \underbrace{(x^7 + x^6 + x^5 + x^4 + x)}_{r(x)},$$

and thus we transmit the word

$$c = (1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0).$$

Suppose that we instead receive the word

$$w = (1, 1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0).$$

The error polynomial is $e(x) = x^{11} + x^4$, but we don't yet know this. Using Table 1, we compute

$$\begin{aligned}
s_1 &= w(\alpha) = \alpha^{14} + \alpha^{13} + \alpha^9 + \alpha^7 + \alpha^6 + \alpha^5 + \alpha = \alpha^{13} \\
s_2 &= w(\alpha^2) = \alpha^{28} + \alpha^{26} + \alpha^{18} + \alpha^{14} + \alpha^{12} + \alpha^{10} + \alpha^2 = \alpha^{11} \\
s_3 &= w(\alpha^3) = \alpha^{42} + \alpha^{39} + \alpha^{27} + \alpha^{21} + \alpha^{18} + \alpha^{15} + \alpha^3 = \alpha^{10} \\
s_4 &= w(\alpha^4) = \alpha^{56} + \alpha^{52} + \alpha^{36} + \alpha^{28} + \alpha^{24} + \alpha^{20} + \alpha^4 = \alpha^7.
\end{aligned}$$

Hence

$$S = \begin{bmatrix} s_1 & s_2 \\ s_2 & s_3 \end{bmatrix} = \begin{bmatrix} \alpha^{13} & \alpha^{11} \\ \alpha^{11} & \alpha^{10} \end{bmatrix},$$

which has rank 2. The inverse of S is

$$S^{-1} = \frac{1}{\alpha^{23} + \alpha^{22}} \begin{bmatrix} \alpha^{10} & \alpha^{11} \\ \alpha^{11} & \alpha^{13} \end{bmatrix} = \alpha^{-11} \begin{bmatrix} \alpha^{10} & \alpha^{11} \\ \alpha^{11} & \alpha^{13} \end{bmatrix} = \begin{bmatrix} \alpha^{14} & 1 \\ 1 & \alpha^2 \end{bmatrix}.$$

Hence the solution to the linear system $(*)$ is

$$\begin{bmatrix} \sigma_0 \\ \sigma_1 \end{bmatrix} = S^{-1} \begin{bmatrix} s_3 \\ s_4 \end{bmatrix} = \begin{bmatrix} \alpha^{14} & 1 \\ 1 & \alpha^2 \end{bmatrix} \begin{bmatrix} \alpha^{10} \\ \alpha^7 \end{bmatrix} = \begin{bmatrix} \alpha^{24} + \alpha^7 \\ \alpha^{10} + \alpha^9 \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha^{13} \end{bmatrix}.$$

Consider the polynomial $p(x) = x^2 + \sigma_1 x + \sigma_0 = x^2 + \alpha^{13} x + 1$. Note that

$$p(\alpha^4) = \alpha^8 + \alpha^{17} + 1 = 0, \text{ and}$$
$$p(\alpha^{11}) = \alpha^{22} + \alpha^{24} + 1 = 0.$$

Hence we conclude the errors occurred in bits which correspond to the coefficients of x^4 and x^{11}, which is correct.

3 General BCH codes

Let $m \geq 3$ and $t < 2^{m-1}$ be positive integers. In this section we will prove that there exists a binary BCH code C with

- block length $n = 2^m - 1$,
- minimum distance $d \geq 2t + 1$,
- number of parity-check digits $n - k \leq mt$.

Hence C can correct at most t errors.

Let $\mathbb{F} = \mathbb{F}_{2^m}$, and let $\alpha \in \mathbb{F}$ be a primitive element. Let $g \in \mathbb{F}_2[x]$ be the unique polynomial of lowest degree with $\alpha, \alpha^2, ..., \alpha^{2t}$ as roots. We can compute g by taking the least common multiple of the minimal polynomials of $\alpha, \alpha^3, ..., \alpha^{2t-1}$. We define the set of codewords C as

$$C = \left\{ (c_{n-1}, ..., c_0) \in \mathbb{F}_2^n \mid g \text{ divides } c_{n-1}x^{n-1} + ... + c_1 x + c_0 \right\}.$$

Lemma 3.1. Let

$$H = \begin{bmatrix} \alpha^{n-1} & \alpha^{n-2} & \dots & \alpha & 1 \\ \alpha^{2(n-1)} & \alpha^{2(n-2)} & \dots & \alpha^2 & 1 \\ \alpha^{3(n-1)} & \alpha^{3(n-2)} & \dots & \alpha^3 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha^{2t(n-1)} & \alpha^{2t(n-2)} & \dots & \alpha^{2t} & 1 \end{bmatrix}.$$

Then $\ker H = C$.

Proof. Let $v \in \mathbb{F}_2^n$. Note that $Hv = 0$ if and only if

$$(\alpha^{(n-1)i}, \alpha^{(n-2)i}, ..., \alpha^i, 1) \cdot (v_{n-1}, v_{n-2}, ..., v_1, v_0) = 0.$$

for each $1 \leq i \leq 2t$. This occurs if and only if α^i is a root of $v_{n-1}x^{n-1} + ... + v_1x + v_0$ for each $1 \leq i \leq 2t$, thus $v \in C$. □

Remark 3.2. Note that H is not strictly speaking a parity-check matrix as defined previously, since its entries are elements of $\mathbb{F}$, not $\mathbb{F}_2$. However, if each entry of H is replaced by its corresponding binary m-tuple over $\mathbb{F}_2$, written in column form, we obtain a binary parity-check matrix for C.

Lemma 3.3. The code C has minimum distance at least $2t+1$, and hence can correct t errors.

Proof. Suppose for a contradiction that there exists a non-zero $c \in C$ such that $\|c\| \leq 2t$. Let $c_{e_1}, c_{e_2}, ..., c_{e_q}$ be the non-zero components of c. Then

$$0 = Hc = \begin{bmatrix} \alpha^{e_1} & \alpha^{e_2} & \dots & \alpha^{e_q} \\ \alpha^{2e_1} & \alpha^{2e_2} & \dots & \alpha^{2e_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{2te_1} & \alpha^{2te_2} & \dots & \alpha^{2te_q} \end{bmatrix} \begin{bmatrix} c_{e_1} \\ c_{e_2} \\ \vdots \\ c_{e_q} \end{bmatrix} = \begin{bmatrix} \alpha^{e_1} & \alpha^{e_2} & \dots & \alpha^{e_q} \\ \alpha^{2e_1} & \alpha^{2e_2} & \dots & \alpha^{2e_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{2te_1} & \alpha^{2te_2} & \dots & \alpha^{2te_q} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Since $q = \|c\| \leq 2t$, we can remove the bottom $2t - q$ rows from the matrix to obtain the equation

$$0 = \begin{bmatrix} \alpha^{e_1} & \alpha^{e_2} & \dots & \alpha^{e_q} \\ \alpha^{2e_1} & \alpha^{2e_2} & \dots & \alpha^{2e_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{qe_1} & \alpha^{qe_2} & \dots & \alpha^{qe_q} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

This is only possible if

$$0 = \det \begin{bmatrix} \alpha^{e_1} & \alpha^{e_2} & \dots & \alpha^{e_q} \\ \alpha^{2e_1} & \alpha^{2e_2} & \dots & \alpha^{2e_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{qe_1} & \alpha^{qe_2} & \dots & \alpha^{qe_q} \end{bmatrix} = \alpha^{e_1+...+e_q} \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha^{e_1} & \alpha^{e_2} & \dots & \alpha^{e_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(q-1)e_1} & \alpha^{(q-1)e_2} & \dots & \alpha^{(q-1)e_q} \end{bmatrix}$$

But this is impossible, as it is a Vandermonde determinant which is non-zero since each α^{e_i} is distinct. □

Encoding. Suppose we wish to send the message

$$v = (v_{n-1}, v_{n-2}, ..., v_{n-k}).$$

To do this, form the polynomial

$$v(x) = v_{n-1}x^{n-1} + v_{n-2}x^{n-2} + ... + v_{n-k}x^{n-k}$$

and perform Euclidean division by $g(x)$

$$v(x) = q(x)g(x) + r(x)$$

where

$$r(x) = r_{n-k-1}x^{n-k-1} + ... + r_1x + r_0$$

is a polynomial of degree at most $n-k-1$. By construction $g(x)$ divides $v(x)+r(x)$, and hence

$$c = (v_{n-1}, ..., v_{n-k}, r_{n-k-1}, ..., r_0) \in C.$$

We transmit c.

Decoding. Suppose we receive the word w. Let $w(x)$ be the associated polynomial, and write $w(x) = c(x) + e(x)$, where $e(x)$ is the error. If the errors occur at positions $e_1, ..., e_q$ (where $q \leq t$) then we have

$$e(x) = x^{e_1} + ... + x^{e_q}.$$

To decode, we begin by calculating $s_i = w(\alpha^i)$ for $1 \leq i \leq 2t$. Note that

$$\begin{aligned} s_i &= c(\alpha^i) + e(\alpha^i) \\ &= e(\alpha^i) && \text{since } c \in C \\ &= \alpha^{ie_1} + ... + \alpha^{ie_q}. \end{aligned}$$

In order to find the error locations, we need to solve these equations for $e_1, ..., e_q$.

Proposition 3.4. Let

$$S = \begin{bmatrix} s_1 & s_2 & \dots & s_t \\ s_2 & s_3 & \dots & s_{t+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_t & s_{t+1} & \dots & s_{2t-1} \end{bmatrix}.$$

Then the number of errors q is equal to the rank of S, and moreover, the $q \times q$ principal minor of S is invertible.

Proof. For $1 \leq i \leq q$, let

$$u_i = (\alpha^{e_i}, \alpha^{2e_i}, ..., \alpha^{te_i}).$$

We claim that $u_1, ..., u_q$ span the row space of S. To see this, note that for any $1 \leq i \leq q$, we have

$$\begin{aligned}\sum_{j=1}^{n} \alpha^{(i-1)e_j} u_j &= \sum_{j=1}^{n} \alpha^{(i-1)e_j} (\alpha^{e_j}, \alpha^{2e_j}, ..., \alpha^{te_j}) \\ &= \left(\sum_{j=1}^{n} \alpha^{ie_j}, \sum_{j=1}^{n} \alpha^{(i+1)e_j}, ..., \sum_{j=1}^{n} \alpha^{(i+t-1)e_j} \right) \\ &= (s_i, s_{i+1}, ..., s_{i+t-1}).\end{aligned}$$

This is precisely the ith row of S. So each row of S can be written as a linear combination of the u_i and thus $\operatorname{rank} S \leq q$.

To show that $\operatorname{rank} S = q$, it now suffices to prove the second part of the claim. Let U_q be the $q \times q$ principal minor of S:

$$U_q = \begin{bmatrix} s_1 & \dots & s_q \\ \vdots & \ddots & \vdots \\ s_q & \dots & s_{2q-1} \end{bmatrix}.$$

We claim that $U_q = ADA^T$, where

$$D = \begin{bmatrix} \alpha^{e_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha^{e_q} \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \alpha^{e_1} & \alpha^{e_2} & \dots & \alpha^{e_q} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(q-1)e_1} & \alpha^{(q-1)e_2} & \dots & \alpha^{(q-1)e_q} \end{bmatrix}.$$

For a matrix B, let $B_{i,j}$ denote the entry in row i and column j. We compute:

$$\begin{aligned}(ADA^T)_{i,j} &= \sum_{k=1}^{q} \sum_{l=1}^{q} A_{ik} D_{kl} A^T_{lj} \\ &= \sum_{k=1}^{q} A_{ik} D_{kk} A_{jk} \\ &= \sum_{k=1}^{q} \alpha^{(i-1)e_k} \alpha^{e_k} \alpha^{(j-1)e_k} \\ &= \sum_{k=1}^{q} \alpha^{(i+j-1)e_k} \\ &= s_{i+j-1} \\ &= (U_q)_{i,j}.\end{aligned}$$

So $U_q = ADA^T$. Since A is a Vandermonde matrix, it is invertible, and hence so is U_q. $\square$

The above proposition allows us to determine the number of errors. In fact, with a little more work we can also use it to determine the locations of the errors as well.

Proposition 3.5. Consider the unique solution $(\sigma_0, ..., \sigma_{q-1})$ to the matrix equation

$$U_q \begin{bmatrix} \sigma_0 \\ \vdots \\ \sigma_{q-1} \end{bmatrix} = \begin{bmatrix} s_{q+1} \\ \vdots \\ s_{2q} \end{bmatrix}$$

Then the roots of $x^q + \sigma_{q-1}x^{q-1} + ... + \sigma_1 x + \sigma_0$ in $\mathbb{F}$ are $\alpha^{e_1}, ..., \alpha^{e_q}$.

Proof. Let

$$\sigma(x) = (x - \alpha^{e_1}) \dots (x - \alpha^{e_q}) = x^q + \sigma_{q-1}x^{q-1} + ... + \sigma_1 x + \sigma_0.$$

We need to show that the coefficients σ_i satisfy the given matrix equation. Note that for $1 \leq i \leq q$ we have

$$0 = \sigma(\alpha^{e_i}) = \alpha^{qe_i} + \sigma_{q-1}\alpha^{(q-1)e_i} + ... + \sigma_1\alpha^{e_i} + \sigma_0.$$

and hence

$$\alpha^{qe_i} = \sigma_{q-1}\alpha^{(q-1)e_i} + ... + \sigma_1\alpha^{e_i} + \sigma_0. \qquad (*)$$

We compute:

$$\begin{aligned}
\sum_{j=0}^{q-1} s_{i+j}\sigma_j &= \sum_{j=0}^{q-1}\sum_{k=0}^{q} \alpha^{(i+j)e_k}\sigma_j \\
&= \sum_{k=0}^{q} \alpha^{ie_k} \sum_{j=0}^{q-1} \alpha^{je_k}\sigma_j \\
&= \sum_{k=0}^{q} \alpha^{ie_k}\alpha^{qe_k} && \text{by } (*) \\
&= \sum_{k=0}^{q} \alpha^{(i+q)e_k} \\
&= s_{i+q}.
\end{aligned}$$

Hence

$$\begin{bmatrix} s_1 & \dots & s_q \\ \vdots & \ddots & \vdots \\ s_q & \dots & s_{2q-1} \end{bmatrix} \begin{bmatrix} \sigma_0 \\ \vdots \\ \sigma_{q-1} \end{bmatrix} = \begin{bmatrix} s_{q+1} \\ \vdots \\ s_{2q} \end{bmatrix}$$

as claimed. □

4 References

1. L.N. Childs, *A Concrete Introduction to Higher Algebra*, Springer, New York, 2009.

2. Y.S. Han, *BCH Codes* talk slides, National Taipei University, Taiwan.