

1. Let C be a binary $[n, k]$ -code. Fix $1 \leq i \leq n$. Let C_i be the set of all codewords in C whose i -th digit is zero. So

$$C_i = \{x_1 \dots x_n \in C : x_i = 0\}$$

- (a) Prove that C_i is linear.
- (b) Prove that C_i is a subgroup of C of index at most two.
- (c) Deduce that exactly one of the following holds:
 - The i -th digit in every codeword in C is zero.
 - The i -th digit in exactly half of the codewords is zero.

Proof : (a) Let $\mathbf{x}, \mathbf{y} \in C_i$ and $a, b \in \{0, 1\}$. Looking at the i -position, we see that

$$(a\mathbf{x} + b\mathbf{y})_i = ax_i + by_i = a \cdot 0 + b \cdot 0 = 0$$

Hence $a\mathbf{x} + b\mathbf{y} \in C_i$. So C_i is linear.

(b) Recall that the index of C_i in C (notation : $[C : C_i]$) is the number of cosets of C_i in C .

If $C_i = C$ then C_i is the only coset of C_i in C and $[C : C_i] = 1$.

So we may assume that $C_i \neq C$. Fix $\mathbf{y} \in C \setminus C_i$. So \mathbf{y} has a one in the i -th position. We claim that $\mathbf{y} + C_i$ is the set of all codewords whose i -th digit is one. Pick $\mathbf{x} \in \mathbf{y} + C_i$. Then $\mathbf{x} = \mathbf{y} + \mathbf{c}$ for some $\mathbf{c} \in C_i$. Hence $\mathbf{x} \in C$ since $\mathbf{y}, \mathbf{c} \in C$ and $x_i = (\mathbf{y} + \mathbf{c})_i = y_i + c_i = 1 + 0 = 1$. Pick $\mathbf{u} \in C$ with a one in the i -th position. Put $\mathbf{v} = \mathbf{u} - \mathbf{y}$. Then $\mathbf{v} \in C$ since $\mathbf{u}, \mathbf{y} \in C$ and $v_i = (\mathbf{u} - \mathbf{y})_i = u_i - y_i = 1 - 1 = 0$. So $\mathbf{v} \in C_i$ and $\mathbf{u} = \mathbf{y} + (\mathbf{u} - \mathbf{y}) = \mathbf{y} + \mathbf{v} \in \mathbf{y} + C_i$, which proves our claim.

Since the code C is binary, we have that every codeword has either a zero or a one in the i -th position. Hence $C = C_i \cup (\mathbf{y} + C_i)$. So C is the disjoint union of two cosets of C_i . Since the cosets of C_i in C form a partition of C , we must have that C_i has exactly two cosets in C . So $[C : C_i] = 2$

(c) From (b), it follows that $[C : C_i] \in \{1, 2\}$.

If $[C : C_i] = 1$ then $C_i = C$ and so every codeword has a zero in the i -th position.

So we may assume that $[C : C_i] = 2$. Since every coset of C_i in C has the same number of elements (namely $|C_i|$) and C_i has two cosets in C , we get that $|C_i| = \frac{1}{2}|C|$. So exactly half of the codewords have a zero in the i -th position. Since the code is binary, we get that exactly half of the codewords have a one in the i -th position. \square

2. Let C be a binary $[n, k, d]$ -code.

- (a) Prove that the sum of the weights of all the codewords in C is at most $n2^{k-1}$.
- (b) Prove that $d \leq \frac{n2^{k-1}}{2^k - 1}$.
- (c) Suppose that $2d > n$. Prove the *Plotkin Bound* :

$$|C| \leq \frac{2d}{2d - n}$$

Proof : (a) Let $SW(C)$ denote the sum of the weights of all the codewords. So

$$SW(C) = \sum_{\mathbf{x} \in C} w(\mathbf{x})$$

We write all the codewords in a matrix. Then $SW(C)$ is the number of ones in this matrix. Since C is a binary $[n, k, d]$ -code, we end up with a $2^k \times n$ -matrix. We add up all the ones in the first column, then all the ones in the second column, etc. Finally, we add up all these sums.

Let $1 \leq i \leq n$. It follows from Ex. #1(c) that at most half of the codewords have a one in the i -th position. Hence the number of ones in the i -th column of the matrix is at most half of $|C|$, which is 2^{k-1} . Since there are n columns in the matrix, we get that

$$SW(C) \leq n \cdot 2^{k-1}$$

(b) Recall that the minimum distance of a linear code equals its minimum weight. By definition of minimum weight, we have that $w(\mathbf{x}) \geq w(C)$ for all $\mathbf{x} \in C$ with $\mathbf{x} \neq \mathbf{0}$. Since $|C| = 2^k$, we get

$$SW(C) = \sum_{\mathbf{x} \in C} w(\mathbf{x}) = \sum_{\mathbf{0} \neq \mathbf{x} \in C} w(\mathbf{x}) \geq \sum_{\mathbf{0} \neq \mathbf{x} \in C} w(C) = (|C| - 1)w(C) = (2^k - 1)d$$

Combining this with (a), we get

$$(2^k - 1)d \leq SW(C) \leq n \cdot 2^{k-1}$$

Hence

$$d \leq \frac{n \cdot 2^{k-1}}{2^k - 1}$$

(c) From (b), it follows that

$$2d \leq 2 \frac{n \cdot 2^{k-1}}{2^k - 1} = \frac{n \cdot 2^k}{2^k - 1} = \frac{n|C|}{|C| - 1}$$

Since C is a binary $[n, k, d]$ -code and so $|C| = 2^k$. Hence

$$2d(|C| - 1) \leq n|C|$$

and so

$$(2d - n)|C| \leq 2d$$

Since $2d - n > 0$, we get that

$$|C| \leq \frac{2d}{2d - n} \quad \square$$

3. The codewords in the first order Reed-Muller code $R(1, 3)$ are of the form

$$x_1x_2 \dots x_7x_8 = [a_3 \ a_2 \ a_1 \ a_0] G(1, 3)$$

(a) Find the parity check sums for a_1 , a_2 and a_3 .

(b) Decode 10000011.

(c) Decode 10101010.

Solution : (a) We have that

$$x_1x_2 \dots x_7x_8 = [a_3 \ a_2 \ a_1 \ a_0] \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

So

$$\begin{cases} x_1 = a_0 \\ x_2 = a_1 + a_0 \\ x_3 = a_2 + a_0 \\ x_4 = a_2 + a_1 + a_0 \\ x_5 = a_3 + a_0 \\ x_6 = a_3 + a_1 + a_0 \\ x_7 = a_3 + a_2 + a_0 \\ x_8 = a_3 + a_2 + a_1 + a_0 \end{cases}$$

Hence we get the following parity check sums :

$$\begin{aligned} a_1 &= x_1 + x_2 = x_3 + x_4 = x_5 + x_6 = x_7 + x_8 \\ a_2 &= x_1 + x_3 = x_2 + x_4 = x_5 + x_7 = x_6 + x_8 \\ a_3 &= x_1 + x_5 = x_2 + x_6 = x_3 + x_7 = x_4 + x_8 \end{aligned}$$

(b) We calculate the parity check sums for a_1 , a_2 and a_3 :

$$\begin{aligned} a_1 & \{1, 0, 0, 0\} \quad \text{so } a_1 = 0 \\ a_2 & \{1, 0, 1, 1\} \quad \text{so } a_2 = 1 \\ a_3 & \{1, 0, 1, 1\} \quad \text{so } a_3 = 1 \end{aligned}$$

We easily get that

$$[a_3 \ a_2 \ a_1 \ 0] G(1,3) = [1 \ 1 \ 0 \ 0] \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} = 00111100$$

Since $10000011 - 00111100 = 10111111$, we get that $a_0 = 1$. So we decode 10000011 as

$$00111100 + 11111111 = 11000011$$

(c) We proceed as in (b). The parity check sums for a_1 , a_2 and a_3 are :

$$\begin{aligned} a_1 & \{1, 1, 1, 1\} \quad \text{so } a_1 = 1 \\ a_2 & \{0, 0, 0, 0\} \quad \text{so } a_2 = 0 \\ a_3 & \{0, 0, 0, 0\} \quad \text{so } a_3 = 0 \end{aligned}$$

We easily get that

$$[a_3 \ a_2 \ a_1 \ 0] G(1,3) = [0 \ 0 \ 1 \ 0] \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} = 01010101$$

Since $10101010 - 01010101 = 11111111$, we get that $a_0 = 1$. So we decode 10101010 as

$$01010101 + 11111111 = 10101010$$

Note that this implies that the received word is a codeword. □

4. Consider the binary matrix

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

Show that G is a generator matrix for the second order Reed-Muller code $R(2, 3)$.

Proof : It follows from the definition of the Reed-Muller codes that

$$R(2, 3) = R(2, 2) \otimes R(1, 2)$$

We know that $G(1, 2) = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ is a generator matrix for $R(1, 2)$. Note that $R(2, 2)$ is the set of all binary words of length four. We claim that

$$G(2, 2) := \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

is a generator matrix for $R(2, 2)$. The only thing we have to show is that the rank of $G(2, 2)$ is four, but this is obvious since $G(2, 2)$ is an upper uni-triangular matrix.

Hence it follows from Theorem 4.2 that $G' := \begin{bmatrix} 0_{3 \times 4} & G(1, 2) \\ G(2, 2) & G(2, 2) \end{bmatrix}$ is a generator matrix for $R(2, 3)$. So

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

Thus the rows of G' form a basis for $R(2, 3)$. Since the rows of G are the same as the rows of G' , we get that the rows of G form a basis for $R(2, 3)$. Hence G is a generator matrix for $R(2, 3)$. \square

5. Explain the probabilities on page 30 in the notes (I updated the notes. The correct percentages are 26.49%, 0.6929%, 0.0139% and 99.99%. First write your answers in terms of a general symbol-error-probability p . Then evaluate your answers (you will need a computer to find 99.99%).

Solution : Let p be the probability that one digit is received incorrectly.

- **No error correction**

A pixel (or grayscale) is now a message with six digits : $a_1a_2a_3a_4a_5a_6$. If an error occurs during transmission, we will not be able to recover. The probability that one digit is received correctly is $1 - p$. So the probability that we receive a pixel correctly is $(1 - p)^6$. Hence the probability that a pixel is received incorrectly is

$$1 - (1 - p)^6 \quad ; \quad \text{for } p = 5\% : 26.49\%$$

- **Repeating every digit five times**

A pixel had six digits $a_1a_2a_3a_4a_5a_6$. We send each digit of this pixel five times. So we are using the repetition code $\{00000, 11111\}$ for each digit of the pixel. We will decode a digit of the pixel correctly if at most two errors occur during transmission. The probability of at most two errors occurring during the transmission of say $a_1a_1a_1a_1a_1$ is

$$(1 - p)^5 + \binom{5}{1}p(1 - p)^4 + \binom{5}{2}p^2(1 - p)^3 \quad \text{or} \quad \sum_{i=0}^2 \binom{5}{i}p^i(1 - p)^{5-i}$$

So the probability we decode correctly all six digits of the pixel is

$$\left(\sum_{i=0}^2 \binom{5}{i}p^i(1 - p)^{5-i} \right)^6$$

Hence the probability that a pixel is decoded incorrectly now drops to

$$1 - \left(\sum_{i=0}^2 \binom{5}{i} p^i (1-p)^{5-i} \right)^6 \quad ; \quad \text{for } p = 5\% : 0.6929\%$$

- **Using $R(1, 5)$**

A pixel is now a binary word of length 32. Since $R(1, 5)$ is seven-error correcting, we will decode a pixel correctly if at most seven errors occur during transmission. The probability that at most seven errors occur during transmission is

$$\sum_{i=0}^7 \binom{32}{i} p^i (1-p)^{32-i}$$

So the probability that a pixel is decoded incorrectly drops even further down to

$$P := 1 - \sum_{i=0}^7 \binom{32}{i} p^i (1-p)^{32-i} \quad ; \quad \text{for } p = 5\% : 0.0139\%$$

- **Whole picture using $R(1, 5)$**

Since a picture is 832×700 , we get that a picture contains $M := 832 \cdot 700 = 582,400$ pixels. We just showed that the probability that a pixel is decoded incorrectly is $P = 1 - \sum_{i=0}^7 \binom{32}{i} p^i (1-p)^{32-i}$. Hence the probability that (after decoding) the whole picture contains at most 116 incorrect pixels is

$$\sum_{i=0}^{116} \binom{M}{i} P^i (1-P)^{M-i} \quad ; \quad \text{for } p = 5\% : 99.9899\%$$
