

1. (a) The codewords of C are $\{00000, 10011, 01101, 11110\}$.
- (b) The code C has length 5, dimension 2, distance 3, and rate $2/5$.
- (c) The dimension of C^\perp is $5 - 2 = 3$.
- (d) By Algorithm 2.13 a generating matrix and a parity check matrix for C are

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix} \quad \mathbf{H} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The columns of \mathbf{H} are a basis for the dual code C^\perp so

$$C^\perp = \{00000, 01100, 10010, 11001, 11110, 10101, 01011, 00111\}.$$

The distance of C^\perp is 2.

- (e) The distance of C is 3 and the distance of C^\perp is 2. If $a \in C$ and $b \in C^\perp$ then the distance of the code D would be $\min\{2 \times 3, 2\} = 2$. If $b \in C$ and $a \in C^\perp$ then the distance of the code D would be $\min\{2 \times 2, 3\} = 3$. Thus it would be better (more error detection/correction capabilities) to form the code D by taking $a \in C^\perp$ and $b \in C$.

2. (a) The incidence matrix for the $(7, 7, 3, 3, 1)$ block design is

$$\begin{matrix} & & & & & & & \text{blocks} \\ & & & & & & & \left(\begin{array}{ccccccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{array} \right) \\ \text{elements} & & & & & & & \end{matrix}.$$

- (b) The weight of a codeword is the number of ones in the corresponding row of the incidence matrix. This is the number of blocks in which the element occurs. By definition, this is a constant r for each element, so every codeword has constant weight r .
- (c) There are r ones in each row. By definition, every pair of elements occurs in exactly λ of the blocks. Hence, given any pair of rows (Row 1 and Row 2), each with r ones in them, there must be $(r - \lambda)$ places in which Row 1 has a one and Row 2 has a zero, and there must be $(r - \lambda)$ places in which Row 1 has a zero and Row 2 has a one. Hence they differ in $2(r - \lambda)$ places. This is true for every pair of rows, so the distance of the code is $2(r - \lambda)$.
- (d) Since the distance of the code is $2(r - \lambda)$ (which is even), this code can detect $2(r - \lambda) - 1$ errors and can correct

$$\frac{2(r - \lambda) - 2}{2} = r - \lambda - 1 \text{ errors.}$$

2. (e) Codes defined in this way are not linear. Adding two codewords of the code defined by the incidence matrix in part (a) gives a word with weight 4 but the codewords all have weight 3. In general, adding two codewords would give a word of weight $2(r - \lambda)$, whereas codewords must have weight r .
3. (a) There is no row of zeros in \mathbf{H}_C and no pair of rows of \mathbf{H}_C sum to zero, so $\delta_C \geq 3$. Rows 1, 2 and 3 are linearly dependent, so C has distance 3. It can detect 2 errors and correct 1 error.
- (b) $(1\ 0\ 1\ 0\ 0\ 1)\mathbf{H}_C = (1\ 1\ 0) \neq (0\ 0\ 0)$, so $101001 \notin C$. Since 110 is the fifth row of \mathbf{H}_C , the most likely codeword is 101011.
- (c) Since $\delta_C = 3$, the words of least weight have weight 3. Moreover, in each such word the three ones occur in positions corresponding to three rows of \mathbf{H} that sum to 000. Thus the words of least weight are:

$$111000, \quad 100110, \quad 010011 \quad \text{and} \quad 001101.$$

- (d) Since \mathbf{H}_C is a parity check matrix for C , we know that \mathbf{H}_C^T is a generating matrix for C^\perp . So

$$\mathbf{G}_{C^\perp} = \mathbf{H}_C^T = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

We now apply Algorithm 2.13 to \mathbf{G}_{C^\perp} . The RREF form of \mathbf{G}_{C^\perp} is

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

We rearrange columns to get

$$\mathbf{G}'_{C^\perp} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Thus

$$\mathbf{H}'_{C^\perp} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and so} \quad \mathbf{H}_{C^\perp} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

\mathbf{H}_{C^\perp} is a parity check matrix for the dual code C^\perp .

4. (a) If $\delta = 3$, then $t = 1$ and the Hamming bound gives

$$|C| \leq \frac{2^n}{\binom{n}{0} + \binom{n}{1}} = \frac{2^n}{1+n}.$$

Let $\lfloor z \rfloor$ denote the largest integer less than or equal to z . Then since we have $n = 16$,

$$|C| \leq \left\lfloor \frac{2^{16}}{1+16} \right\rfloor = \left\lfloor \frac{2^{16}}{17} \right\rfloor = 3855.$$

Since C is a linear code, we need $|C|$ to be a power of 2. Since $2^{12} = 4096$ and $2^{11} = 2048$, an upper bound on the number of codewords in such a linear code is given by $|C| \leq 2^{11}$.

- (b) Now if $\delta = 3$ and $n = 16$, then Corollary 3.6 says that a linear code C with distance 3 and length n exists with

$$|C| \geq \frac{2^{15}}{\binom{15}{0} + \binom{15}{1}} = \frac{2^{15}}{1+15} = \frac{2^{15}}{2^4} = 2^{11}.$$

Thus there exists a linear code C with distance 3 and length 16 having 2^{11} codewords, and by the Hamming bound, no linear code of distance 3 and length 16 can have more codewords.

5. (a) Consider $r = 0$. The code $RM(0, m)$ consists of the zero word and the all ones word, each of length 2^m . Thus its dimension is $1 = \sum_{i=0}^0 \binom{m}{i} = \binom{m}{0}$.

Consider $r = m$. The code $RM(m, m)$ is K^{2^m} , all the binary words of length 2^m . Thus its dimension is $2^m = \sum_{i=0}^m \binom{m}{i}$.

We now proceed by mathematical induction.

We must prove that the dimension of $RM(r, m)$, where $0 < r < m$, is $\sum_{i=0}^r \binom{m}{i}$.

By inspection of the code on page 51 of the lecture notes, the dimension of $RM(1, 2)$ is 3 and $3 = 1 + 2 = \sum_{i=0}^1 \binom{2}{i} = \sum_{i=0}^r \binom{m}{i}$. Thus the statement is true for $r = 1$ and $m = 2$.

Suppose that for $m' < m$ and for all $0 < r' < m'$, the dimension of $RM(r', m')$ is $\sum_{i=0}^{r'} \binom{m'}{i}$.

Consider $RM(r, m)$. Since $RM(r, m)$ is obtained using the $(a \mid a + b)$ construction where $A = RM(r, m-1)$ and $B = RM(r-1, m-1)$, Theorem 3.23 and our induction hypothesis tell us that the dimension of $RM(r, m)$ can be calculated as follows.

$$\begin{aligned} \dim RM(r, m) &= \dim RM(r, m-1) + \dim RM(r-1, m-1) \\ &= \sum_{i=0}^r \binom{m-1}{i} + \sum_{i=0}^{r-1} \binom{m-1}{i} \\ &= \binom{m-1}{0} + \sum_{i=1}^r \binom{m-1}{i} + \sum_{i=1}^r \binom{m-1}{i-1} \\ &= 1 + \sum_{i=1}^r \left(\binom{m-1}{i} + \binom{m-1}{i-1} \right) \\ &= \binom{m}{0} + \sum_{i=1}^r \binom{m}{i} \\ &= \sum_{i=0}^r \binom{m}{i} \end{aligned}$$

Thus the dimension of $RM(r, m)$ is $\sum_{i=0}^r \binom{m}{i}$.

5. (b) The number of cosets of a code of length n and dimension k is 2^{n-k} . From Theorem 3.37 we know that $RM(1, m)$ has length $n = 2^m$ and dimension $k = \sum_{i=0}^1 \binom{m}{i} = 1 + m$. Thus the number of cosets of $RM(1, m)$ is $2^{2^m - m - 1}$.
- (c) Since $RM(1, m)$ has distance 2^{m-1} , it is $\frac{2^{m-1}}{2} - 1 = (2^{m-2} - 1)$ -error correcting. Thus every error pattern of length 2^m and weight less than or equal to $2^{m-2} - 1$ must be a unique coset leader. There are

$$1 + \binom{2^m}{1} + \binom{2^m}{2} + \cdots + \binom{2^m}{2^{m-2} - 1}$$

such words. There may be other unique coset leaders, so this gives us a lower bound on the number of cosets which have unique coset leaders. Thus $RM(1, m)$ has at least

$$1 + \binom{2^m}{1} + \cdots + \binom{2^m}{2^{m-2} - 1} = \sum_{i=0}^{2^{m-2}-1} \binom{2^m}{i} \text{ cosets which have unique coset leaders.}$$

- (d) We apply Algorithm 3.43 and start with $\bar{w} = (1, 1, -1, 1, -1, -1, 1, 1)$.

$$\begin{aligned} w_1 = \bar{w}\mathbf{L}_3^1 &= (2, 0, 0, -2, -2, 0, 2, 0) \\ w_2 = w_1\mathbf{L}_3^2 &= (2, -2, 2, 2, 0, 0, -4, 0) \\ w_3 = w_2\mathbf{L}_3^3 &= (2, -2, -2, 2, 2, -2, 6, 2) \end{aligned}$$

The largest component of w_3 has magnitude 6, and occurs in position 6. Since

$$v(6) = 011$$

and it was positive, the most likely message word was 1011 (with corresponding code-word 11000011).

6. The weight of w is 7 so we form the word w^* by adding a zero to the end of w . The syndrome of $w^* = 010\ 010\ 001\ 000\ 001\ 001\ 101\ 000$ is

$$s = w^*\mathbf{H} = 010\ 010\ 001\ 000 + 101\ 101\ 110\ 110 = 111\ 111\ 111\ 110.$$

Since this is the last row of \mathbf{B} , w is a codeword of C_{23} so we conclude that the most likely transmitted word is

$$v = w = 010\ 010\ 001\ 000\ 001\ 001\ 101\ 00.$$

7. Note that since we require 64 codewords, we need dimension $k = 6$. Thus we need $n \geq 6$.

We have a channel with reliability $p = 0.9995$ and we would like a linear code C such that $\Theta_p(C) > 0.99995$. If we transmitted codewords of length n that had no error correction capability, then $\Theta_p(C) = (0.9995)^n$. But $(0.9995)^n < 0.99995$ for all positive n , so we require some error correction. Thus we need $\delta \geq 3$. If $\delta = 3$ then we can correct all error patterns of weight 1 (and maybe some others) so

$$\Theta_p(C) \geq (0.9995)^n + n(0.9995)^{n-1}(0.0005).$$

Using a calculator (or Maple or Matlab),

$$(0.9995)^n + n(0.9995)^{n-1}(0.0005) \geq 0.99995 \quad \text{for } 6 \leq n \leq 20.$$

Thus if there exists a linear code of length n , dimension $k = 6$ and distance $\delta = 3$ for some n between 6 and 20, then it would be a suitable code. Since we are looking for the most efficient code, we want to find the smallest n for which such a code exists.

The Hamming bound (with $\delta = 3$ so $t = 1$) says that

$$|C| \leq \frac{2^n}{1+n}.$$

With $n = 6, 7, 8, 9, 10$ the Hamming bound gives

$$|C| \leq 9, \quad |C| \leq 16, \quad |C| \leq 28, \quad |C| \leq 51 \quad \text{and} \quad |C| \leq 93.$$

Thus the smallest value of n that may be possible is $n = 10$.

By Theorem 3.5, there exists a linear code with length $n = 10$, dimension $k = 6$ and distance at least $\delta = 3$ if

$$\binom{9}{0} + \binom{9}{1} < 2^{10-6}, \quad \text{that is, } 10 < 16.$$

Thus there exists a $(10, 6, 3)$ linear code and it would satisfy the requirements of PixPix.

This is the most efficient code that can be used by PixPix since they require $k = 6$, and $n = 10$ is the smallest n that will give any error correction, and if you tried for more error correction then you would have to increase the length of the codewords.

8. Since $1 + x^5 = (1 + x)(1 + x + x^2 + x^3 + x^4)$, we obtain 2 proper linear cyclic codes of length 5.

Generator	Dimension of code	Code
$1 + x$	4	{00000, 11000, 01100, 00110, 00011, 10001, 10100, 01010, 00101, 10010, 01001, 11110, 01111, 10111, 11011, 11101}
$1 + x + x^2 + x^3 + x^4$	1	{00000, 11111}

9. (a) Using the parity check matrix \mathbf{H} from the lecture notes, the syndrome of w is

$$s = w\mathbf{H} = 11110010 \quad \text{so } s(x) = 1 + x + x^2 + x^3 + x^6.$$

Now apply the Lemma from the lecture notes to calculate s_i for $i = 1, 2, \dots$ until a syndrome is found that contains a burst error pattern with burst length at most 4.

$$\begin{aligned} s_1(x) &= xs(x) &= x + x^2 + x^3 + x^4 + x^7 && \text{so } s_1 = 01111001 \\ s_2(x) &= xs_1(x) &= x^2 + x^3 + x^4 + x^5 + (1 + x^4 + x^6 + x^7) && \\ & &= 1 + x^2 + x^3 + x^5 + x^6 + x^7 && \text{so } s_2 = 10110111 \\ s_3(x) &= xs_2(x) &= x + x^3 + x^4 + x^6 + x^7 + (1 + x^4 + x^6 + x^7) && \\ & &= 1 + x + x^3 && \text{so } s_3 = 11010000 \end{aligned}$$

Thus s_3 is a syndrome with burst length at most 4. Hence $e_3 = 000\ 000\ 011\ 010\ 000$ and $e = 000\ 011\ 010\ 000\ 000$. Thus the most likely transmitted codeword is

$$v = w + e = 100\ 011\ 111\ 101\ 110.$$

(b) Using the cyclic structure of C and by using coset leaders that contain cyclic burst errors of shortest length, all error patterns that contain a cyclic burst error of length at most 4 are in distinct cosets, and hence can be corrected. Thus an error pattern of weight 4, in which the four ones occur together, can be corrected.

If C were 3-error correcting, then all error patterns of weight three would have to occur in distinct cosets. There are $2^{n-k} = 2^{15-7} = 2^8 = 256$ cosets of C . There are $\binom{15}{3} = 455$ error patterns of weight three. Clearly, these error patterns cannot occur in distinct cosets, and so there will be some error patterns of weight three that cannot be corrected.

10. The code $C = RM(2, 5)$ has length, dimension and distance as follows $n_C = 32$, $k_C = 16$ and $\delta_C = 8$. The Extended Golay code $D = C_{24}$ has length, dimension and distance as follows $n_D = 24$, $k_D = 12$ and $\delta_D = 8$. Let codewords of C be c_1, c_2, \dots and codewords of D be d_1, d_2, \dots . Thus the cross-interleaving process works as follows. Take the first 12 message words of length 16 and encode them to obtain 12 codewords of C :

$$\begin{aligned} c_1 &= c_{1,1} c_{1,2} c_{1,3} \dots c_{1,32} \\ c_2 &= c_{2,1} c_{2,2} c_{2,3} \dots c_{2,32} \\ &\vdots \\ c_{12} &= c_{12,1} c_{12,2} c_{12,3} \dots c_{12,32} \end{aligned}$$

Then each message word of length 12 $c_{1,i} c_{2,i} \dots c_{12,i}$ is encoded to a codeword of D and these are interleaved to depth 3.

10. The interleaving of the first 3 codewords of D is depicted below.

$$\begin{aligned}d_1 &= d_{1,1} d_{1,2} d_{1,3} \dots d_{1,24} \\d_2 &= d_{2,1} d_{2,2} d_{2,3} \dots d_{2,24} \\d_3 &= d_{3,1} d_{3,2} d_{3,3} \dots d_{3,24}\end{aligned}$$

The digits are then transmitted as $d_{1,1} d_{2,1} d_{3,1} d_{1,2} d_{2,2} d_{3,2} \dots$

Since $\delta_D = 8$, the code D can detect up to 7 errors. Since consecutive digits are spaced 3 bits apart, we can thus detect a burst of up to $3 \times 7 = 21$ errors in the transmitted stream of digits, provided that each codeword is affected by at most one such burst.

Case 1: Suppose that the 21 errors occurred within the $24 \times 3 = 72$ digits of one block of three codewords from code D .

If 21 errors occurred in the transmitted stream of digits (and we assume that these are the only errors affecting the given codewords) then the corresponding three codewords of D (say d_1, d_2, d_3) would have all their digits flagged. The corresponding message words would have their digits flagged as well so in undoing the cross-interleaving we would have the following

$$\begin{aligned}c_1^* &= * * * c_{1,4} c_{1,5} c_{1,6} c_{1,7} \dots c_{1,32} \\c_2^* &= * * * c_{2,4} c_{2,5} c_{2,6} c_{2,7} \dots c_{2,32} \\&\vdots \\c_{12}^* &= * * * c_{12,4} c_{12,5} c_{12,6} c_{12,7} \dots c_{12,32}\end{aligned}$$

Since $\delta_C = 8$ and there are only 3 flagged digits in any received word, we can identify the corresponding codeword and correct the errors.

Case 2: Suppose that the 21 errors occurred across two blocks of three codewords from code D (say between digits 63 and 84).

If 21 errors occurred in the transmitted stream of digits affecting two blocks of three codewords (and we assume that these are the only errors affecting the given codewords) then the corresponding six codewords of D (say $d_1, d_2, d_3, d_4, d_5, d_6$) would have all their digits flagged. The corresponding message words would have their digits flagged as well so in undoing the cross-interleaving we would have the following

$$\begin{aligned}c_1^* &= * * * * * * c_{1,7} c_{1,8} c_{1,9} c_{1,10} c_{1,11} \dots c_{1,32} \\c_2^* &= * * * * * * c_{2,7} c_{2,8} c_{2,9} c_{2,10} c_{2,11} \dots c_{2,32} \\&\vdots \\c_{12}^* &= * * * * * * c_{12,7} c_{12,8} c_{12,9} c_{12,10} c_{12,11} \dots c_{12,32}\end{aligned}$$

Since $\delta_C = 8$ and there are 6 flagged digits in any received word, we are able to identify a unique corresponding codeword and correct the errors.