

This Coding Theory Assignment is compulsory, and contributes 7% towards your final assessment. It should be submitted by 14:00 (the start of the lecture) on Monday 1 June 2009. You should hand in your assignment to Barbara either at the lecture or earlier. In the absence of exceptional circumstances (such as documented illness), assignments submitted after 14:00 but within 1 hour of that time will receive at most half marks, and assignments more than 1 hour late will not be marked. Make sure your name and student number are written on your submission!

1. **(6 marks)** Let C be the linear code of length 5 defined by adding three bits x_3 , x_4 and x_5 to each of the binary sequences x_1x_2 of length 2, subject to the conditions:

$$x_3 = x_2 \quad x_4 = x_1 \quad x_5 = x_1 + x_2.$$

- List the codewords of C .
 - State the length, dimension, distance and rate of C .
 - State the dimension of the dual code C^\perp .
 - List the codewords of C^\perp and state its distance.
 - Let D be the code formed by the $(a \mid a + b)$ construction using the two codes C and C^\perp . In this construction, which option would give better error detection/correction capabilities: $a \in C$ and $b \in C^\perp$ or $a \in C^\perp$ and $b \in C$? Explain your answer.
2. **(10 marks)** A (v, b, r, k, λ) block design is an arrangement of v objects into b sets (called *blocks*) each of size k , such that each of the v objects occurs in r blocks, and every pair chosen from the v objects occurs in precisely λ blocks.

For example, a $(7, 7, 3, 3, 1)$ block design is an arrangement of $v = 7$ items into $b = 7$ blocks each of size $k = 3$ such that every item occurs $r = 3$ times and every pair of items occurs in $\lambda = 1$ of the blocks. If the v elements are given by $V = \{1, 2, 3, 4, 5, 6, 7\}$, then a possible $(7, 7, 3, 3, 1)$ block design is

$$D = \{\{1, 4, 5\}, \{2, 3, 5\}, \{3, 4, 7\}, \{2, 4, 6\}, \{1, 2, 7\}, \{5, 6, 7\}, \{1, 3, 6\}\}.$$

Given any block design with $V = \{v_1, v_2, \dots, v_v\}$ and blocks $B = \{b_1, b_2, \dots, b_b\}$, its *incidence matrix* is a $v \times b$ array A (v rows and b columns) in which the entry in row i and column j is 1 iff v_i occurs in block b_j , and is 0 otherwise.

- Find the incidence matrix for the $(7, 7, 3, 3, 1)$ block design given above.
- Given any (v, b, r, k, λ) block design, we can create a code by taking each row of the incidence matrix as a codeword. Explain why such a code has constant weight (that is, explain why every codeword has the same weight).
- What is the distance of the code defined by the incidence matrix of a (v, b, r, k, λ) block design? Show that this distance is equal to the distance between every pair of distinct codewords in the code.
- How many errors does such a code detect? How many errors does it correct?
- Are codes defined in this way linear? Explain your answer briefly.

(continues overleaf)

3. (6 marks) The following matrix \mathbf{H}_C is a parity check matrix for a linear code C .

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

- (a) Determine how many errors C can detect and how many errors C can correct.
 (b) Determine the most likely transmitted codeword corresponding to the received word 101001. Explain your answer.
 (c) Use \mathbf{H}_C to list all the words of least weight in C .
 (d) Find a parity check matrix \mathbf{H}_{C^\perp} for the dual code C^\perp .

4. (5 marks) Consider a linear code with distance $\delta = 3$ and length $n = 16$.

- (a) Use the Hamming bound to show that an upper bound on the number of codewords in such a linear code is 2^{11} .
 (b) Prove that such a linear code with 2^{11} codewords exists.

5. (10 marks)

- (a) Prove (by mathematical induction or otherwise) that the r th order Reed-Muller code $RM(r, m)$ has dimension $k = \sum_{i=0}^r \binom{m}{i}$. This is property 2 of Theorem 3.37. (You can use the following identities $\binom{m}{i} = \binom{m-1}{i} + \binom{m-1}{i-1}$ and $\sum_{i=0}^m \binom{m}{i} = 2^m$.)

- (b) Determine the number of cosets of the first order Reed-Muller code $RM(1, m)$.

- (c) Show that $\sum_{i=0}^{2^{m-2}-1} \binom{2^m}{i}$ is a lower bound on the number of cosets which have a unique coset leader in an SDA for $RM(1, m)$.

The above calculations show that as m gets large, creating and storing an SDA for the Reed-Muller code $RM(1, m)$ is infeasible. Luckily there are other decoding methods that can be used, such as Algorithm 3.43 from the lecture notes.

- (d) Suppose that a message has been encoded using the code $RM(1, 3)$. Use Algorithm 3.43 from the lecture notes to determine the most likely intended message word(s) if we receive the word $w = 11010011$.

6. (4 marks) The Golay code C_{23} is used to transmit messages. Words have been encoded using the generating matrix $\mathbf{G} = (\mathbf{I}_{12} \ \hat{\mathbf{B}})$ where $\hat{\mathbf{B}}$ is the 12×11 matrix described in Section 3 of your lecture notes. Determine the most likely transmitted codeword for the following received word.

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

(continues next page)

7. (7 marks) The company PixPix creates photographs using pixels of 64 different colours. They wish to transmit these photographs across a binary symmetric channel that has reliability 0.9995 and is subject to randomly scattered noise. The received information will be decoded using IMLD. The company would like the probability of a pixel in the received image being assigned an incorrect colour to be less than 0.00005. Determine the parameters (n, k, δ) of the most efficient linear code that could be used by PixPix. Explain your working.

8. (3 marks) Determine all the proper linear cyclic codes of length 5. For each code, state its generator, its dimension and list its codewords.

9. (5 marks) Let C be the 4 cyclic burst error correcting linear cyclic code of length 15 with generator

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

A parity check matrix \mathbf{H} for C is given in Example 4.39 of your lecture notes.

(a) Use \mathbf{H} and the decoding method given in Section 4 of your lecture notes to decode the received word

$$w = 100\ 000\ 101\ 101\ 110.$$

(b) Explain why C can correct some error patterns of weight 4 but is not a 3-error correcting code.

10. (7 marks) Let C be the Reed-Muller code $RM(2, 5)$ and let $D = C_{24}$ be the Extended Golay code. Suppose that C is cross-interleaved with D to form codewords. Determine how long a burst of errors the decoding algorithm described in the lecture notes will correct if the resulting codewords are interleaved to depth 3 before they are transmitted.