MT361 Error Correcting Codes: Sheet 1

Hand in your answers to questions 1, 2, 3 and 4.
Questions 5 and 6 are optional, unless you are an MSc student, in which case please do at least one of them. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by noon on Thursday 27th January 2011 or handed in at the Thursday lecture.

1. (a) Let C be the binary code with codewords 000, 011, 101 and 110. What are the length and size of C? Explain why C can detect one error. [*Hint: there are an even number of* 1*s in each codeword.*]

(b) Now let C' be the binary code with codewords

$$00000, 01101, 10110, 11011.$$

Find the minimum distance of C'. Hence find the number of errors that C' can (i) detect, (ii) correct.

2. Find all $v \in \mathbf{Z}_2^5$ for which $d(v, 11011) = 3$.

3. Find the information rate and the minimum distance of the binary code

$$0000, 0101, 0011, 0110, 1111, 1010, 1100, 1001.$$

4. (a) Which of the following are valid ISBNs:

1-84628-040-0, 1-84628-400-0, 0-486-68735-X?

(b) Show that if two unequal adjacent digits are interchanged when writing down an ISBN then the result is not a valid ISBN. [**Corrected 23 January adding 'unequal'.**]

5. I have an important decision 'Yes' or 'No' that I wish to communicate to a friend across a crowded room. I can shout to him up to three times. The probability that he mishears on the first shout is p, and on the second and third it is $r > p$.

(a) Why is it reasonable to assume that $r \leq 1/2$?

(b) Explain a three-shout strategy, making it clear how my friend will decode what he hears.

(c) Calculate the probability that the three-shout strategy will successfully communicate the message.

(d) If $p = 1/5$ show that the three-shout strategy is superior if and only if $r < 1/3$.

(e) Show more generally that there is a function $f : [0, 1/2] \to [0, 1/2]$ such that three shouts are superior to a single shout if and only if $r < f(p)$. Sketch the graph of f.

6. Show that a binary code of length 5 and minimum distance 3 has size at most 4.

MT361/461/5461 Error Correcting Codes: Sheet 2

Hand in your answers to questions 1, 2, 3 and 4. [Note 4 is compulsory: corrected 29 January.]

Question 8 is compulsory for MSc students. Questions 5, 6 and 7 are optional for all. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 7th February 2011 or handed in at the Monday lecture.

1. Let C be the binary code with codewords

$$00000, 01101, 10110, 11011.$$

Decode, where possible, (a) 10010, (b) 01111 and (c) 11000 using nearest neighbour decoding.

2. Let $n \in \mathbf{N}$ and let C be the length n repetition code over the q-ary alphabet $\{1, 2, \ldots, q\}$ where $q \geq 2$.

(a) What is the minimum distance of C?

(b) What is the maximum t such that C is t-error detecting?

(c) Show that if $n = 2m + 1$ where $m \in \mathbf{N}$ then C is m-error correcting but not $(m+1)$-error correcting. Prove an analogous result if n is even. [*Please argue directly from Definition 2.16, without using Theorem 3.2.*]

3. Let C be the binary code of length 9 whose codewords are all binary words of the form

$$(u_1, u_2, u_3, u_4, u_1, u_2, u_3, u_4, x)$$

where x is chosen to make the total number of 1s in (u_1, u_2, u_3, u_4, x) even.

(a) Give examples of codewords $u, v \in C$ such that $d(u, v) = 3$ and $d(u, v) = 4$. [*Hint: the word with nine* 0*s is in* C.]

(b) Let

$$u = (u_1, u_2, u_3, u_4, u_1, u_2, u_3, u_4, x)$$
$$v = (v_1, v_2, v_3, v_4, v_1, v_2, v_3, v_4, y)$$

be codewords.

(i) Show that if $d(u_1u_2u_3u_4, v_1v_2v_3v_4) = 1$ then $d(u, v) = 3$.

(ii) Show that if $d(u_1u_2u_3u_4, v_1v_2v_3v_4) \geq 2$ then $d(u, v) \geq 4$.

(iii) Hence find the minimum distance of C.

(c) Deduce that C is 1-error correcting and 2-error detecting. [*You may use Theorem 3.2.*]

4. Suppose that a binary repetition code of length 5 is used to communicate on a channel where each bit in a sent codeword flips (i.e. 0 changes to 1 and 1 changes to 0) independently with probability p. Show that the probability that a codeword is decoded incorrectly is $10p^3 - 15p^4 + 6p^5$. Evaluate this probability if $p = 1/10$.

5. Let $p < 1/2$. A jury consists of three people. Two of them get the verdict wrong with probability p, while the third flips a coin to decide. What is the probability that the majority verdict of the jury is correct? Comment on your answer.

6. Show that if an ISBN not ending with an X is written down backwards then it is still valid.

7. Suppose that Alice wishes to send a message 'Yes' or 'No' to Bob. She sends 'Yes' with probability 3/4 and 'No' with probability 1/4. They agree to use the length 3 binary repetition code, encoding 'Yes' as 111 and 'No' as 000. When a codeword is sent to Bob, each bit flips independently with probability p.

(a) Find $\mathbf{P}$[001 received | 000 transmitted].

(b) Hence find $\mathbf{P}$[000 transmitted | 001 received].

(c) Show that $\mathbf{P}$[111 transmitted | 001 received] $>$ $\mathbf{P}$[000 transmitted | 001 received] if $p > 1/4$.

(d) Comment on the implications for nearest neighbour decoding.

8. Let C be the Reed–Solomon code with alphabet $\mathbf{F}_5$ defined in Example 4.2(3) of the MSc lecture notes. So $n = 4$ and the codewords are

$$u(f) = (f(0), f(1), f(2), f(3))$$

for $f(x) = ax + b$, with $a, b \in \mathbf{F}_5$.

(a) Find $u(f)$ if $f(x) = 3x + 1$.

(b) Suppose that the word 1312 is received. Show that 1312 is not in C. Find a codeword $u \in C$ such that $d(u, 1312) = 1$.

9. Let $n \in \mathbf{N}$. Show that if $2e \geq n$ then no code of length n can be e-error correcting. [*As in Question 2, please argue directly from Definition 2.16 without using Theorem 3.2*]

MT361/461/5461 Error Correcting Codes: Sheet 3

Hand in your answers to questions 1, 2, 3, 4.

Questions 5 and 6 are compulsory for MSc students. Questions 7 is optional, but highly recommended. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 14th February 2011 or handed in at the Monday lecture.

1. Let C be a code of length n over an alphabet A. Suppose that C has minimum distance $2t+1$ where $t \in \mathbf{N}$. Let v be a word over A of length n. Show that there is at most one codeword in C within distance t of v.

2. Alice wants to send Bob a single message 'Yes' or 'No' using a binary symmetric channel with cross-over probability p. They consider two possible schemes.

(1) Using the binary repetition code of length 3 as a 1-error correcting code, encoding 'Yes' as 111 and 'No' as 000.

(2) Using the binary repetition code of length 2 as a 1-error detecting code, encoding 'Yes' as 11 and 'No' as 00.

Suppose that Alice sends Bob the message 'Yes'.

(a) Show that using scheme (1) Bob decodes Alice's message as 'No' if and only if he receives 100, 010, 001 or 000. Show that this event has probability $3p^2 - 2p^3$.

(b) Now suppose that Alice and Bob use scheme (2).

(i) Show that Bob decodes Alice's message as 'No' with probability p^2.

(ii) Find the probability that Bob detects that an error has occurred.

(iii) Suppose that whenever Bob detects an error he asks for retransmission, repeating until he receives one of the codewords 00 or 11. Find the probability that Bob eventually decodes Alice's message as 'No'. [*Hint: you will need to sum a geometric series.*]

(iv) Let b be the average number of bits that are sent using (2). Show that $b = 2\big(1 - 2p(1-p)\big) + 2p(1-p)(2+b)$, and hence find b.

(c) Compare the relative merits of schemes (1) and (2) when the channel has cross-over probabilities 1/10 and 1/4.

3. Let C be a binary (n, M, d)-code. For each codeword $u = u_1u_2 \dots u_n \in C$ we define $\bar{u} = u_1u_2 \dots u_nu_1u_2 \dots u_n$. Let $D = \{\bar{u} : u \in C\}$.

(a) Find D in the special case when $C = \{000, 011, 101, 110\}$.

(b) Find the parameters (i.e. length, size, minimum distance) of D in terms of n, M and d.

4. The object of this question is to show that $A_2(8,5) = 4$. Suppose that C is a binary code of length 8 with minimum distance 5.

(a) Explain why we may assume that C contains 00000000.

(b) Show that the weight of every non-zero codeword in C is 5 or more. (The *weight* of a binary codeword is the number of 1s it contains.)

(c) Show that C contains at most one codeword of weight 6 or more.

(d) Show that any two codewords in C of weight 5 are both equal to 1 in exactly two positions.

(e) Hence show that C has size ≤ 4 and give an example where C has the largest possible size. Deduce that $A_2(8,5) = 4$.

5. (MSc / MSci) Construct a $(4, 25, 3)$ Reed–Solomon code over $\mathbf{F}_5$. [**Corrected 16 to 25 on 10 February. If you prefer, construct a** $(4, 16, 3)$ **Reed–Solomon code over** $\mathbf{F}_4$ **using the tables for** $\mathbf{F}_4$ **given in the first lecture.**]

6. (MSc / MSci) Let C be the Reed–Solomon code $RS_{7,5,3}$ where polynomials are evaluated at $0, 1, 2, 3, 4 \in \mathbf{F}_7$. Find a polynomial $f(x) \in \mathbf{F}_7[x]$ of degree < 3 such that $f(0) = 0$, $f(1) = 0$, $f(2) = 1$. Find the corresponding codeword $u(f)$. Hence find the minimum distance of C without using Theorem 4.4.

7. Alice knows a polynomial f with coefficients in the natural numbers, of unknown degree. Bob can pick any number $x \in \mathbf{Z}$ and ask Alice to tell him $f(x)$. After hearing Alice's answer, Bob may then pick $y \in \mathbf{Z}$ and ask for $f(y)$, and so on. Find a strategy for Bob that will determine f in as few questions as possible.

8. You and two of your friends are on your way to a party. At the party, a white or black hat will be put on each person's head. You can see your friends' hats, but not your own.

When the host says 'Go' you may stay silent, or say either 'White' or 'Black'. Everyone who speaks, must speak at the same time: you cannot wait to see what your friends do. Then

- if at least one person speaks, and everyone who speaks says the colour of their own hat, you all get some cake;
- if anyone gets it wrong, or everyone stays silent, there is no cake.

Find a good strategy. [*Hint: if you guess at random, and your two friends stay silent, the chance of success is* 1/2. *So you should aim to beat this.*] Now think about the variant game played with four (or more) people.

MT361/461/5461 Error Correcting Codes: Sheet 4

Hand in your answers to questions 1, 2, 3, 4.

Questions 5 is compulsory for MSc students. Questions 6 to 10 are optional. Question 6 is in the style of past exam questions. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 21st February 2011 or handed in at the Monday lecture.

1. Consider the codes

$$C_1 = \{0000, 1100, 1010, 1001\}$$
$$C_2 = \{1101, 0001, 0100, 0111\}$$
$$C_3 = \{0000, 1100, 0110, 0011\}$$
$$C_4 = \{0000, 1100, 1010, 0110\}$$

(a) Show that C_1 is equivalent to C_2.

(b) Show that C_2 is not equivalent to C_3. [*Hint: Example 4.9 is relevant.*] Is C_1 equivalent to C_3?

(c) (Optional.) Is C_4 equivalent to one of C_1, C_2 or C_3?

2. Let X and Y be the mutually orthogonal Latin squares (MOLs) shown below.

$$\begin{array}{ccc} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{array} \qquad \begin{array}{ccc} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{array}$$

(a) Use X and Y to construct a $(4, 9, 3)$-code over the alphabet $\{0, 1, 2\}$ containing the codeword 0000.

(b) Decode the received words 1201, 2222 and 2110 using nearest neighbour decoding.

3. Use the construction in Lemma 5.5 to find a pair of MOLs of order 5.

4. Find the pair of MOLs of order 4 corresponding to the $(4, 16, 3)$-code whose codewords are listed below.

0000	1111	2222	3333	0123	1032	2301	3210
0231	1320	2013	3102	0312	1203	2130	3021

5. (**MSc, MSci**) Consider the Reed–Solomon code $RS_{5,4,2}$ over $\mathbf{F}_5$ where polynomials are evaluated at $a_1 = 0$, $a_2 = 1$, $a_3 = 2$, $a_4 = 3$. Suppose that you receive the words (i) 2413, (ii) 1033, (iii) 1032. In each case solve the Key Equation for $Q(x)$ and $E(x)$ and decode (where possible) the received word.

6. What does it mean to say that a binary code is an (n, M, d)-code?

For which of the following parameters either give a binary code with these parameters, or show that no such code can exist:

(i) $(6, 2, 6)$; (ii) $(6, 4, 4)$; (iii) $(6, 3, 5)$; (iv) $(6, 32, 2)$.

7. Let $q \geq 2$ and let $A = \{0, 1, \ldots, q-1\}$. Suppose that C is a q-ary code of length $n \geq 2$ and minimum distance $n - 1$.

(a) Show that if $u = u_1 u_2 \ldots u_n$ and $v = v_1 v_2 \ldots v_n$ are codewords in C then

$$d(u_1 u_2, v_1 v_2) \geq 1.$$

(b) By putting codewords into pigeonholes according to their first two symbols, show that $|C| \leq q^2$.

(c) Deduce that $A_2(n, n-1) \leq q^2$. (This is a special case of the Singleton bound. The case $n = 4$ shows that codes constructed from mutually orthogonal Latin squares are as large as possible.)

8. Let $d \in \mathbf{N}$ be odd and let $n \geq d$. Suppose that C is a binary code of length n and minimum distance d. Define a new code C^+ of length $n + 1$ by adding a final bit to each codeword in C to make the total number of 1s even. (For example, if $C = \{00, 01, 10, 11\}$ then $C^+ = \{000, 011, 101, 110\}$.)

(a) Show that the distance between any two codewords in C^+ is even. Hence show that C^+ has minimum distance $d + 1$.

(b) Deduce that $A_2(n + 1, d + 1) \geq A_2(n, d)$.

9. Let $n, d \in \mathbf{N}$ where $n \geq d$. Let C be a code of length $n+1$ and minimum distance $d + 1$.

(a) Suppose that two codewords at distance $d+1$ in C differ in position i. Show that the code $C^\star$ obtained by removing position i from all codewords in C has length n and minimum distance d.

(b) Deduce that $A_2(n + 1, d + 1) \leq A_2(n, d)$. Use the previous question to show that if d is odd then $A_2(n, d) = A_2(n + 1, d + 1)$.

10. (For people who know some basic group theory.) Let G be a finite group of order n with group operation $\circ$. Suppose that $G = \{g_1, g_2, \ldots, g_n\}$. Show that the matrix X defined by $X_{ij} = g_i \circ g_j$ is a Latin square over the alphabet G.

MT361/461/5461 Error Correcting Codes: Sheet 5

Hand in your answers to questions 1, 2, 3.

Questions 5 is compulsory for MSc students. Questions 4, 6 and 7 are optional. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 28th February 2011 or handed in at the Monday lecture.

1. Suppose that C is a $(4, q^2, 3)$-code over the alphabet $A = \{0, 1, \ldots, q-1\}$.

(a) Show that if $u = (u_1, u_2, u_3, u_4)$ and $u' = (u'_1, u'_2, u'_3, u'_4) \in C$ are distinct codewords then $(u_1, u_2) \neq (u'_1, u'_2)$.

(b) Deduce that for all $i, j \in A$ there is a unique codeword, say (i, j, X_{ij}, Y_{ij}), whose first two positions are (i, j). [*Hint: C has size q^2.*]

(c) By (b) we know that

$$C = \{(i, j, X_{ij}, Y_{ij}) : i, j \in A\}$$

(i) Prove that the rows of the matrix X have distinct entries. [*Hint: suppose row i has a repeated entry, so $X_{ij} = X_{ij'}$ where $j \neq j'$. What does this imply about the codewords whose first two positions are (i, j) and (i, j')?*]

(ii) Prove that the columns of X have distinct entries.

(iii) Deduce that X is a Latin square.

(iv) Prove that X and Y are MOLs.

2. (a) Let H be a Hadamard matrix of order n. Show that the $2n \times 2n$ matrix

$$K = \begin{pmatrix} H & H \\ H & -H \end{pmatrix}$$

is a Hadamard matrix of order $2n$.

(b) Starting from the 2×2 Hadamard matrix

$$\begin{pmatrix} + & + \\ + & - \end{pmatrix}$$

use (a) to construct Hadamard matrices with orders 4 and 8. Hence write down the codewords in (i) a binary $(4, 8, 2)$ code and (ii) binary $(8, 16, 4)$-code. [*For* (ii)*, please write down 8 codewords, and then explain how the other* 8 *are obtained.*]

(c) Use nearest neighbour decoding to decode (where possible) the received words 01010111, 10011110 and 11000000 using the code in (b)(ii).

(d) Use (b)(ii) to show that there is a binary $(7, 8, 4)$-code. Deduce from the Plotkin bound that $A_2(7, 4) = 8$.

3. Let C be the ternary code with codewords $000, 111, 222, 012, 021, 120, 102, 201, 210$. Let $C^\star$ be the code obtained from C by puncturing it in its final position. Write down the codewords in $C^\star$ and find the length, size, and minimum distance of $C^\star$.

4. Suppose that C is a binary code of length n and minimum distance d.

(a) Use the Plotkin bound to show that if $d > 3n/4$ then C has at most 2 codewords.

(b) Use Lemma 4.11 and equivalences of codes to prove the stronger result that if $d > 2n/3$ then C has at most 2 codewords.

5. The purpose of this question is to give a geometric proof of the Plotkin bound. Let C be a binary (n, M, d)-code where $2d > n$. For each codeword $u \in C$ define an associated vector $x(u) \in \mathbf{R}^n$ by

$$x(u)_i = \begin{cases} 1 & \text{if } u_i = 0, \\ -1 & \text{if } u_i = 1. \end{cases}$$

Let $x \cdot y$ be the usual dot product of vectors $x, y \in \mathbf{R}^n$.

(a) Show that $x(u) \cdot x(u) = n$ for all $u \in C$.

(b) Let $u, u' \in C$ be distinct codewords. Show that $x(u) \cdot x(u') = n - 2d(u, u')$ and deduce that $x(u) \cdot x(u') \leq -(2d - n)$.

(c) Let $z = \sum_{u \in C} x(u)$. By considering $z \cdot z$ prove the Plotkin bound.

(d) (Optional, but maybe instructive.) Find z if C is a binary $(5, 4, 3)$-code.

6. Let C be an (n, M, d)-code where $n \geq 2d$. By putting the codewords in C into pigeonholes according to their final $n - 2d$ positions and applying Corollary 7.7 to the codes of length $2d$ given by removing the final $n - 2d$ positions from the codewords in each pigeonhole, prove that

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

Compare this bound with the Singleton bound for binary codes.

7. Let p be a prime such that $p \equiv 3 \bmod 4$ and let $\mathbf{F}_p$ be the final field with p elements. We say that $x \in \mathbf{F}_p$ is a *square* if there exists $y \in \mathbf{F}_p$ such that $x = y^2$.

Let Q be the $p \times p$ matrix Q whose rows and columns are indexed by $i, j \in \{0, 1, \ldots, p-1\}$ and where

$$Q_{ij} = \begin{cases} -1 & \text{if } i - j \text{ is a square in } \mathbf{F}_p \\ 1 & \text{otherwise.} \end{cases}$$

It is known that the matrix H obtained from Q by adding an extra row and an extra column consisting entirely of 1s is a Hadamard matrix of order $p + 1$. Use this construction to find a Hadamard matrix of order 12.

MT361/461/5461 Error Correcting Codes: Sheet 6

Hand in your answers to questions 1, 2 and 3.

All other questions are optional. Question 5 is in the style of past exam questions. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 7th March 2011 or handed in at the Monday lecture.

1. Let C be the binary code with codewords 00000, 11100, 00111 and 11011.

(a) Show that C is linear.

(b) Find the Hamming balls of radius 1 about 11100 and 11011 and check that they are disjoint.

(c) How many binary words of length 5 are not within distance 1 of any codeword in C?

2. Let $u = u_1u_2 \dots u_n$ and $u' = u'_1u'_2 \dots u'_n$ be binary words of length n. We write $(u \mid u')$ for the word $u_1u_2 \dots u_nu'_1u'_2 \dots u'_n$ of length $2n$.

Let C and C' be linear binary codes of length n. Let D be the code of length $2n$ consisting of all words of the form $(u \mid u + u')$ where $u \in C$ and $u' \in C'$

(a) Show that D is linear, by checking the axioms for linearity.

(b) Show that $|D| = |C||C'|$.

(c) Suppose that C has minimum distance $d \leq n/2$. Show that if C' is the binary repetition code of length n then the minimum distance of D is $2d$. [*Hint: find the weights of codewords in D.*]

(d) (Optional.) Show that whenever d is a power of 2 there exists a binary linear $(2d, 4d, d)$-code. (The code you get is a Hadamard code.)

3. A binary code of length n and minimum distance $2e+1$ is said to be perfect if the Hamming balls of radius e about codewords partition the set $\{0,1\}^n$. (In other words, the balls are disjoint, and every binary word of length n is in some ball.)

(a) Show that if n is odd then the binary repetition code of length n is perfect.

(b) Show that if C is a perfect 1-error correcting binary code of length n then $|C| = 2^n/(n+1)$ and $n = 2^r - 1$ for some $r \in \mathbf{N}$. Express $|C|$ in terms of r.

(c) Why is 'perfect' not defined for codes of even minimum distance?

4. Show that a Hadamard matrix of order ≥ 4 has order divisible by 4. [*Hint: if r, r' and r'' are three rows of H then, by reordering columns, we may assume that r and r' are equal in their first $n/2$ positions and unequal in their final $n/2$ positions. What restrictions does this put on r''?*]

5. (a) Let u be a binary word of length n and let $e \in \mathbf{N}$. Define the *Hamming ball of radius e about u*.

(b) Let C be a binary code of length n. Suppose that C is e-error correcting.

(i) Let $u, u' \in C$ be distinct codewords. Show that the Hamming balls of radius e about u and u' are disjoint.

(ii) Prove that

$$|C| \leq \frac{2^n}{\sum_{k=0}^{e} \binom{n}{e}}.$$

6. Let $n \in \mathbf{N}$ and let $1 \leq e \leq n$.

(a) Show that if $0 < \lambda < 1/2$ then $f(k) = (1-\lambda)^{n-k}\lambda^k$ is a decreasing function of k.

(b) Use the identity

$$1 = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}$$

and (a) to show that

$$1 \geq \big(1 - e/n\big)^{n-e} \big(e/n\big)^e \sum_{k=0}^{e} \binom{n}{k}.$$

(c) Hence show that

$$\sum_{k=0}^{e} \binom{n}{k} \leq 2^{nH(e/n)}$$

where H is Shannon's entropy function, defined by

$$H(\lambda) = -\lambda \log \lambda - (1-\lambda)\log(1-\lambda).$$

7. Show that in a linear binary code either all the codewords have even weight, or half have odd weight and half have even weight.

8. Ten people are told that they will be lined up so that the person at the back can see the 9 people in front of him, and the person directly in front of him can see the 8 people in front of her, and so on. A hat, which may be either black or white, will be put on each person.

Starting from the back of the line, each person is asked to shout the colour of the hat they are wearing. Anyone who gets the colour of their hat wrong is shot (with a water pistol if you prefer, but everyone in front of a person will be in no doubt as to what happens). Of course a person cannot see their own hat.

Advise the ten on a good strategy.

MT361/461/5461 Error Correcting Codes: Sheet 7

Hand in your answers to questions 1 and 2.

All other questions are optional. Question 1 revises previous material: the square code will be used as an example in later lectures. Question 4 is in the style of past exam questions. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 14th March 2011 or handed in at the Monday lecture.

1. The square code is the binary code S of length 8 with codewords

$$\{(u_1, u_2, u_3, u_4, u_1 + u_2, u_3 + u_4, u_1 + u_3, u_2 + u_4) : u_1, u_2, u_3, u_4 \in \mathbf{Z}_2\}.$$

The name comes from the representation of the codewords as four message bits, surrounded by four check bits.

u_1	u_2	$u_1 + u_2$
u_3	u_4	$u_3 + u_4$
$u_1 + u_3$	$u_2 + u_4$	

You will find this representation helpful when decoding received words.

(a) Show that if exactly one error occurs when a codeword is transmitted then it is possible to find the position of the error, and hence correct it. Illustrate your answer by decoding the received words 00100110, 01001100 and 01101110.

(b) By (a) we know that S is 1-error correcting. Prove this in another way using Lemma 9.6 and Corollary 3.3.

(c) Find, by a method of your choice, the maximum t such that (i) S is t-error detecting; (ii) S is t-error correcting.

(d) Explain how S can be used to communicate a number between 0 and 15 down the binary symmetric channel. Illustrate your answer by encoding 5 and decoding the received word in a case where a single error occurs in the channel. (See Question 6 for an application.)

2. (a) Let C be the linear binary $(5, 4, 3)$-code with codewords 00000, 11100, 00111 and 11011. Construct a standard array for C taking 10001 as one of the coset leaders. Use your standard array to decode the received words (i) 01011, (ii) 11101, (iii) 00011.

(b) Let D be linear code with codewords $\{u, u + 10001 : u \in C\}$. Construct a standard array for D with four rows and use it decode the three words in (a). [*You should find that one word is decoded differently.*]

(c) What are the parameters (i.e. length, size, minimum distance) of D?

3. Let C be a linear binary code of length n and let $1 \leq i \leq n$. Show that either all codewords in C have 0 in their i-th position, or half of the codewords have 0 in their i-th position and half have 1. [*Hint: if u is a codeword with $u_i = 1$, consider the map $C \to C$ defined by $v \mapsto v + u$.*]

4. (a) Define the *weight* of a binary codeword.

(b) Prove that the minimum distance of a linear binary code C is equal to the minimum weight of its non-zero codewords.

(c) Let C be a linear binary code of length n with odd minimum distance d. Define the *parity check extension* of C and show that it has minimum distance $d + 1$.

5. Suppose that C is a binary code of length n and minimum distance at least δn where $0 < \delta < 1$. Show that if $\delta = 1/2$ then $|C|$ can be arbitrarily large as n tends to infinity, but if $\delta > 1/2$ then $|C|$ is bounded as n tends to infinity.

6. Alice knows a number between 0 and 15. Bob must give Alice a list of 8 questions about the number. Alice will then answer all of them, but may, if she wishes, lie in one answer. Use the square code to find a good strategy for Bob.

MT361/461/5461 Error Correcting Codes: Sheet 8

Hand in your answers to questions 1 and 2.

All other questions are optional. Question 7 is based on the final part of Question 5 on the 2010 exam paper. I will be happy to discuss any of the questions in office hours.

To be returned to McCrea 240 by 4pm on Monday 21st March 2011 or handed in at the Monday lecture.

1. For each of the two linear binary codes below determine their parameters $[n, m, d]$ and find generator and parity check matrices for them:

(i) $C = \{0000, 1010, 0101, 1111\}$,

(ii) $D = \{0000, 1010, 1100, 0110, 0001, 1011, 1101, 0111\}$.

Find a linear code D' equivalent to D and a generator matrix for D' in standard form.

2. Let S be the square code consisting of all codewords of the form

$$\{(u_1, u_2, u_3, u_4, u_1 + u_2, u_3 + u_4, u_1 + u_3, u_2 + u_4) : u_1, u_2, u_3, u_4 \in \mathbf{Z}_2\}$$

You may assume that S is linear.

(a) Find a generator matrix for S in standard form. Write down the corresponding parity check matrix.

(b) Let S' be the parity check extension of S. Find a generator matrix and a parity check matrix for S'.

(c) Prove that S' is 3-error detecting and 1-error correcting. [*You may use any general results proved in the course.*]

(d) Give examples to show that S' is not 4-error detecting or 2-error correcting.

3. Show that if C is a linear binary code of length n then the code $C^\star$ obtained by puncturing C in its final position is also linear.

4. Let C be a linear binary code of length n. For $i \in \{1, 2, \ldots, n\}$ let $e(i)$ be the word with 1 in position i and 0 in all other positions.

(a) Show that C is one-error correcting if and only if the cosets

$$C,\ C + e(1),\ \ldots,\ C + e(n)$$

are pairwise disjoint.

(b) Suppose that H is a parity check matrix for C. Show that if v, $v' \in \mathbf{Z}_2^n$ then $C + v = C + v'$ if and only if $vH = v'H$. Hence show that C is one-error correcting if and only if the rows of H are distinct and non-zero.

5. (MSc, MSci) The definitions of generator and parity check matrices extend in an obvious way to linear codes over general finite fields. Find generator and parity check matrices for the Reed–Solomon code $RS_{5,4,2}$ in which polynomials are evaluated at $0, 1, 2, 3 \in \mathbf{F}_5$.

6. (MSc, MSci) Find all the 3-ary cyclic codes of length 5, giving both a generator polynomial and a generator matrix for each. You may assume that

$$x^5 - 1 = (x - 1)(x^4 + x^3 + x^2 + x + 1).$$

is the factorization of $x^5 - 1$ into irreducible polynomials in $\mathbf{F}_3[x]$.

7. Consider the Hadamard code $C = \{0000, 1111, 1010, 0101, 1100, 0011, 1001, 0110\}$ of length 4.

(i) Prove that this is a linear code and give a basis for it.

(ii) Write down a generator matrix G and calculate the corresponding parity check matrix H for C.

(iii) Construct a standard array for C. List the possible coset leaders in each coset.

8. The Grand Vizier has 27 coins. He knows that one coin is counterfeit, and is lighter than the other 26. Using a balance that can weigh any subset of the coins against any other subset, show how to find the counterfeit coin in three weighings.

9. (a) By generalizing the square code S, define a linear binary $[n^2 + 2n, n^2, 3]$-code for each $n \in \mathbf{N}$. Show that the rate of these codes tends to 1 as $n \to \infty$.

(b) Define a *cube code* by analogy with the square code. By extending these codes, find a family of two-error correcting linear binary codes whose rate tends to 1 as their length tends to infinity.

10. Let G be an $m \times n$ matrix with entries in $\mathbf{Z}_2$. We may regard G as a linear map $\mathbf{Z}_2^n \to \mathbf{Z}_2^m$ acting on column vectors of length n.

(a) Show that $\operatorname{im} G = \{Gw : w \in \mathbf{Z}_2^n\}$ is a subspace of $\mathbf{Z}_2^m$.

(b) Suppose that $v(1), \ldots, v(k)$ is a basis for $\ker G = \{v \in \mathbf{Z}_2^n : Gv = 0\}$. Show that if this basis is extended to a basis $v(1), \ldots, v(k), w(1), \ldots, w(n-k)$ for $\mathbf{Z}_2^n$ then

$$Gw(1), \ldots, Gw(n - k)$$

is a basis for $\operatorname{im} G$.

(c) Deduce that $\dim \ker G + \dim \operatorname{im} G = n$. (This is the rank-nullity theorem.)

(d) Use (c) to show that if C is a linear binary code then $\dim C + \dim C^\perp = n$.

MT361/461/5461 Error Correcting Codes: Sheet 9

Questions 1 and 2 cover the material in the final week of lectures.

Question 8 is based on Question 5 on the 2008 exam paper. Questions 9 and 10 are similar to the MSc / MSci questions on Reed–Solomon and cyclic codes on the exam paper.

I will be happy to answer questions by email over the vacation. My email address is `mark.wildon@rhul.ac.uk`. If you would like to go through any of these questions, please see me in the second week of next term.

1. Let

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

Let C be the linear binary code of length 7 defined by $C = \{u \in \mathbf{Z}_2^7 : uH = 0\}$.

(a) Show that $(u_1, u_2, 1, u_4, 0, 0, 0) \in C$ for a unique choice of $u_1, u_2, u_3 \in \mathbf{Z}_2$.

(b) Show more generally that if $u_3, u_5, u_6, u_7 \in \mathbf{Z}_2$ are given then there exist unique $u_1, u_2, u_4 \in \mathbf{Z}_2$ such that $(u_1, u_2, u_3, u_4, u_5, u_6, u_7) \in C$.

(c) Prove that

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

is a generator matrix for C.

2. Let C be the Hamming $[7, 4, 3]$-code as defined in question 1.

(a) Explain how C can be used to transmit numbers between 0 and 15 using a binary channel. Illustrate your answer by encoding the number 8.

(b) Use syndrome decoding to decode the received words 1011010, 0011011, 1100111 as numbers between 0 and 15.

3. Show that the Hamming $[7, 4, 3]$-code is perfect, in the sense defined in Question 3 of Sheet 6.

4. (a) Construct a linear ternary one-error correcting code of length 12 with a 12×3 parity check matrix.

(b) You have 12 pennies, one of which *might* be counterfeit, and of a different weight to the others. Using three weighings on a balance find out whether there is a counterfeit penny, and if so, determine whether it is heavy or light.

5. Use the Singleton bound to show that if C is a linear binary $[n, m, d]$-code then $m \leq n - d + 1$.

6. Let C be the binary code with generator matrix

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

(a) Write down a parity check matrix for C in standard form.

(b) Construct a syndrome look-up table for C giving coset leaders for all 16 possible syndromes.

(c) Decode the received words $0110110, 1111010$ and 1110111, using the table from (b).

7. Suppose that C is a linear binary code of length n with parity check matrix H. Let $d(C)$ be the minimum distance of C. For each $i \in \{1, 2, \ldots, n\}$, let $e(i)$ be the word defined by

$$e(i)_j = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that if there exist codewords u, $u' \in C$ such that $u = u' + e(i)$ then the i-th row of H has all entries equal to zero.

(b) Show that if there exist codewords u, $u' \in C$ such that $u + e(i) = u' + e(j)$ then the i-th row and the j-th row of H are equal.

(c) Hence show that if the rows of H are distinct and non-zero then $d(C) \geq 3$.

(d) Show conversely that if H has a zero row or two equal rows then $d(C) \leq 2$.

[*Hint: in (a) and (b) use Theorem 11.19, that if $u \in C$ then $uH = 0$.*]

8. (a) Let C be a binary code of length n and dimension m with generator matrix G.

(i) Define the *dual code* $C^{\perp}$.

(ii) What does it mean to say that H is a *parity check matrix* for C? How many rows and columns does a parity check matrix for C have?

(iii) What is the *syndrome* of a received word $v \in \mathbf{Z}_2^n$ with respect to H?

(b) Let $C = \{0000, 1111\}$ be the binary repetition code of length 4.

(i) Write down a generating matrix for C in standard form and hence find a parity check matrix for C.

(ii) Construct a table of coset leaders and their syndromes.

(iii) Perform syndrome decoding on the word 0011, documenting all your steps. How could you modify your table so that 0011 is decoded as the other codeword in C?

9. (**MSc, MSci**) Let $n, k \in \mathbf{N}$ and let p be a prime such that $p \geq n \geq k$. Let $a_1, \ldots, a_n \in \mathbf{F}_p$ be distinct elements.

(a) Define the *Reed–Solomon code* $RS_{p,n,k}$ associated to these parameters.

(b) Prove that Reed–Solomon codes are linear.

(c) Show that the minimum distance of the Reed–Solomon code $RS_{p,n,k}$ is at least $n - k + 1$. (Any general results you use to show that two polynomials are equal should be clearly stated, but need not be proved.)

(d) Suppose that $n - k = 2e$ where $e \in \mathbf{N}$. Show that if the codeword u is transmitted and the word v is received then, provided $d(u, v) \leq e$, there exist polynomials $E(x)$ of degree $\leq e$ and $Q(x)$ of degree $\leq k + e - 1$ such that the Key Equation

$$Q(a_i) = v_i E(a_i)$$

holds for all i. Explain briefly how a solution to the Key Equation can be used to decode v.

10. (**MSc, MSci**)

(a) What does it mean to say that a binary code is *cyclic*?

(b) Explain how codewords of length n in a binary cyclic code can be represented by polynomials of degree $< n$. If $f(x) \in \mathbf{F}[x]/(x^n - 1)$ represents $(u_0, u_1, \ldots, u_{n-1})$, show that $(u_{n-1}, u_0, u_1, \ldots, u_{n-2})$ is represented by $xf(x)$.

(c) What is meant by a *generator polynomial* for a cyclic code?

(d) Let C be the binary cyclic code of length 7 with generator polynomial $g(x) = x^3 + x + 1$.

(i) Write down a generator matrix for C.

(ii) Let C^- be the binary code consisting of all codewords in C whose weight is even. Show that C^- is cyclic and find a generator polynomial for C^-.

You may find it helpful to note that $x^7 + 1 = (x + 1)(x^3 + x + 1)(x^3 + x^2 + 1)$ where the factors are irreducible.

11. (**MSc, MSci**). Let $n \in \mathbf{N}$ and suppose that r divides n. Show that $g(x) = x^r + 1 \in \mathbf{F}_2[x]$ divides $x^n + 1$. Find an explicit basis for the the cyclic binary code of length n with generator polynomial $g(x)$ and determine its dimension and minimum distance.