MATHEMATICAL TRIPOS PART II (2022–2023)
CODING AND CRYPTOGRAPHY
EXAMPLE SHEET 3 OF 4

1 Find generator and parity check matrices for the Hamming $(7, 4)$-code, putting each in the form $(I|B)$ for I an identity matrix of suitable size. Repeat for the parity check extension of this code.

2 The Mariner mission to Mars[1] used the RM$(5, 1)$ code. What was its information rate? What proportion of errors could it correct in a single codeword? How does it compare to the Hamming code of length 31?

3 Show that if C is a linear code, then so are its parity check extension C^+ and puncturing C^-. When is the shortening C' of C a linear code?

(i) Show that extension followed by puncturing does not change a code. Is this true if we replace 'puncturing' by 'shortening'?

(ii) Give an example where shortening reduces the information rate and an example where shortening increases the information rate.

(iii) Show that the minimum distance of C^+ is the least even integer n with $n \geq d(C)$.

(iv) Show that the minimum distance of C^- is $d(C)$ or $d(C) - 1$ and that both cases can occur.

(v) Show that shortening cannot decrease the minimum distance but give examples to show that the minimum distance can stay the same or increase.

4 State the recursive definition of the Reed-Muller codes, using the bar product construction. Use this to compute the rank of RM(d, r). Show that all but two codewords in RM$(d, 1)$ have the same weight.

5 Show that RM$(d, d - 2)$ is the parity check extension of the Hamming $(n, n - d)$ code with $n = 2^d - 1$. (This is useful because we often want codes of length 2^d.)

6 Factor the polynomials $X^3 - 1$ and $X^5 - 1$ into irreducibles in $\mathbb{F}_2[X]$. Hence find all cyclic codes of length 3 or 5 and relate them to codes you have already met.

7 Show directly that the dual code $C^\perp$ of a cyclic code C is cyclic. Explain how the generator polynomials of C and $C^\perp$ are related.

8 Let C be the cyclic code of length $n = 2^d - 1$ defined by a primitive nth root of unity.

(i) Show that if $g(X) \in \mathbb{F}_2[X]$ then $g(X)^2 = g(X^2)$.

(ii) Show that C is a BCH code of design distance 3.

(iii) Deduce that C is (equivalent to) the Hamming $(n, n - d)$-code.

[1]Launched by NASA from Cape Canaveral on 30 May 1971, Mariner 9 was the first spacecraft to orbit another planet, reaching planetary orbit in mid-November and narrowly beating the Soviet probes *Mars 2* and *Mars 3*, which both arrived only weeks later. Once dust storms on the surface had cleared, the orbiter had transmitted 7,329 images, covering 85% of Mars' surface. As of February 2022, Mariner 9's location is unknown; it is either still in orbit, or has already burned up in the Martian atmosphere or crashed into the surface of Mars. The enormous *Valles Marineris* canyon system is named after Mariner 9 in honour of its achievements.

9 Prove the following results.

(i) If K is a field containing $\mathbb{F}_2$, then $(a+b)^2 = a^2 + b^2$ for all $a, b \in K$.

(ii) If $P \in \mathbb{F}_2[X]$ and K is a field containing $\mathbb{F}_2$, then $P(a)^2 = P(a^2)$ for all $a \in K$.

(iii) Let K be a field containing $\mathbb{F}_2$ in which $X^7 - 1$ factorises into linear factors. If β is a root of $X^3 + X + 1$ in K, then β is a primitive root of unity and β^2 is also a root of $X^3 + X + 1$.

(iv) We continue with the notation (iii). The BCH code with $\{\beta, \beta^2\}$ as defining set is Hamming's original (7,4) code.

10 (a) Consider the collection K of polynomials $a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3$ with $a_j \in \mathbb{F}_2$, manipulated subject to the usual rules of polynomial arithmetic and the further condition $1 + \alpha + \alpha^4 = 0$. Show by direct calculation that $K^\times = K \setminus \{0\}$ is a cyclic group under multiplication and deduce that K is a finite field. (Of course, this follows directly from general theory but direct calculation is not uninstructive.)

(b) Let $\alpha \in \mathbb{F}_{16}$ be a root of $X^4 + X + 1$. Let C be the BCH code of length 15 and design distance 5, with defining set the first few powers of α.

(i) Find the minimal polynomial for each element of the defining set, and then compute the generator polynomial of C as the least common multiple of these polynomials.

(ii) If possible, determine the error positions of the following received words

(a) $r(X) = 1 + X^6 + X^7 + X^8$

(b) $r(X) = 1 + X + X^4 + X^5 + X^6 + X^9$

(c) $r(X) = 1 + X + X^2$

(d) $r(X) = 1 + X + X^7$.

(Your answer to (a) may help with the computations.)

11 Let C be a linear code of length n with A_j codewords of weight j. The *weight enumerator polynomial* is

$$W_C(x, y) = \sum_{j=0}^{n} A_j x^j y^{n-j}.$$

(i) We transmit a codeword through a BSC with error probability p. Give a formula, in terms of the weight enumerator polynomial, for the probability that the word received is a codeword.

(ii) Show that $W_C(x, y) = W_C(y, x)$ if and only if $W_C(1, 0) = 1$.

(iii) Show that the weight enumerator polynomial for RM$(d, 1)$ is

$$y^{2^d} + (2^{d+1} - 2)x^{2^{d-1}} y^{2^{d-1}} + x^{2^d}.$$

12 Show that if $2^k \sum_{i=0}^{d-2} \binom{n-1}{i} < 2^n$ then $A(n, d) \geq 2^k$. Compare with the GSV bound in the case $n = 10$ and $d = 3$. (Hint: Construct a parity check matrix for a linear code by choosing one column at a time.)

Further Problems

13 Let $C \leqslant \mathbb{F}_2^n$ be a linear code of dimension k.
(i) Show that $\sum_{x \in C}(-1)^{x.y} = 2^k$ if $y \in C^\perp$ and that this sum is 0 if $y \notin C^\perp$.
(ii) For $t \in \mathbb{R}$. show that

$$\sum_{y \in \mathbb{F}_2^n} t^{w(y)}(-1)^{x.y} = (1-t)^{w(x)}(1+t)^{n-w(x)}$$

.

(iii) By using (i) and (ii) to evaluate

$$\sum_{x \in C} \left(\sum_{y \in \mathbb{F}_2^n} (-1)^{x.y} (\frac{s}{t})^{w(y)} \right)$$

in two different ways, obtain the MacWilliams identity[2]

$$W_{C^\perp}(s,t) = 2^{-\dim C} W_C(t-s, t+s).$$

14 Show that $\mathrm{RM}(d,r)$ has dual code $\mathrm{RM}(d, d-r-1)$. [Hint: first show that every codeword in $\mathrm{RM}(d, d-1)$ has even weight.]

15 Note that $V(3,23)$ is a power of 2. We will construct a perfect 3-error correcting $(23,12)$-code (called the *binary Golay code*[3]), starting from the factorisation

$$X^{23} - 1 = (X-1)f_1(X)f_2(X)$$

in $\mathbb{F}_2[X]$ where $f_1(X) = 1 + X + X^5 + X^6 + X^7 + X^9 + X^{11}$ and $f_2(X) = X^{11}f_1(1/X)$. (So $f_2(X)$ is obtained from $f_1(X)$ by reversing the sequence of coefficients.)

(i) Show that if $g(X) \in \mathbb{F}_2[X]$ and β is a root of g in some field extension of $\mathbb{F}_2$) then β^2 is also a root of g.

(ii) Make a list of the powers of 2 mod 23. Deduce that the cyclic code C with generator polynomial $f_1(X)$ has minimum distance at least 5.

(iii) Show that $C^\perp$ is a subcode of C. Deduce that the parity check extension of C is a self-dual linear code.

(iv) Show that any self-dual linear code, generated by vectors of weight a multiple of 4, has minimum distance a multiple of 4.

(v) Deduce that C is a perfect 3-error correcting code.

SM, Lent Term 2023
Comments on and corrections to this sheet may be emailed to sm137@cam.ac.uk

[2]This amazing result is named for Jessie MacWilliams, a Cambridge alumna who moved to the US after Cambridge and, amongst many achievements, in 1977 co-authored (with Neil Sloane) an encyclopaedic book about the theory of error-correcting codes.

[3]The Voyager 1 & 2 spacecraft transmitted colour pictures of Jupiter and Saturn in 1979 and 1980. Colour transmission requires three times the amount of data than Mariner 9 needed, so a Golay (24,12,8) code was used. It is only 3-error correcting, but its transmission rate is much higher. Voyager 2 went on to Uranus and Neptune and the code was switched to a so-called Reed-Solomon code for its higher error correcting capabilities.