TC10 / **Problems 16-30**

S. Xambó

16. In *maximum likelihood decoding* (MLD) of a code C, the received vector y is decoded into the vector $x \in C$ (assuming that it is unique) that maximizes the probability $P(y|x)$ of receiving y when x is sent.

In *minimum error decoding* (MED), y is decoded into the vector $x \in C$ (assuming that it is unique) that maximizes the probability $P(x|y)$ of having sent x when y is received.

Prove that in a symmetric q-ary channel with probability error per symbol p (assuming $p \leq (q-1)/q$, or $p \leq 1/2$ in the binary case):

a) $P(y|x) = (1-p)^{n-s}\left(\frac{p}{q-1}\right)^s$, where $s = hd(y,x)$.

b) Use this formula to deduce that MLD is equivalent to minimum distance decoding (MDD).

c) $P(x|y) = P(y|x)\frac{P(x)}{P(y)}$ and so MED of a received vector y is equivalent to maximizing $P(y|x)P(x) = (1-p)^{n-s}\left(\frac{p}{q-1}\right)^s P(x)$. In particular we see that if the code vectors are equiprobable, then MED coincides with MLD (and with MDD).

17. If we use a code [23,12,7] on a binary symmetric channel with a probability p of error per bit, what is the probability p' of a decoding error in the MDD? Use the expression of p' to find an upper bound for the probability $\bar{p}$ of a bit error after decoding.

18. Let $p_1 < p_2 < \cdots < p_n$ be relatively prime integers. Encode integers u such that $0 \leq u < p_1 \cdots p_k$, for a given positive integer $k \leq n$, as

$$f(u) = (u \bmod p_1, \dots, u \bmod p_n).$$

Show that $|f(u)| \geq n - k + 1$ for any u, with equality holding for some u.

Linear codes

19. The matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & a \\ 1 & 1 & 0 & 0 & 0 & b \\ 1 & 0 & 1 & 0 & 0 & c \\ 0 & 1 & 1 & 1 & 0 & d \end{pmatrix}$$

is the control matrix of a binary code C.

a) List the vectors of C in the case $a = b = c = d = 1$.

b) Prove that it is possible to choose a, b, c, d in such a way that C can correct 1 error and detect 2 errors. Are there values for a, b, c, d such that C corrects 2 errors?

20. Let G_1 and G_2 be generating matrices of linear codes of type $[n_1, k, d_1]$ and $[n_2, k, d_2]$, respectively. Show that the matrices $G_3 = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix}$ and

$G_4 = (G_1|G_2)$ generate liner codes of types $[n_1 + n_2, 2k, d_3]$ and $[n_1 + n_2, k, d_4]$ with $d_3 = \min(d_1, d_2)$ and $d_4 \geq d_1 + d_2$.

21. Let $n = rs$, where r and s are positive. Let C be the binary code of length n formed by the words $x = x_1x_2 \cdots x_n$ such that in the $s \times r$ matrix

$$\begin{pmatrix} x_1 & \cdots & x_r \\ x_{r+1} & \cdots & x_{2r} \\ \vdots & & \vdots \\ x_{(s-1)r+1} & \cdots & x_{sr} \end{pmatrix}$$

the sum of the elements of each column and of each row is 0.

a) Check that C is a linear code and find its dimension and its minimum distance.

b) Propose a decoding scheme that exploits its matrix presentation.

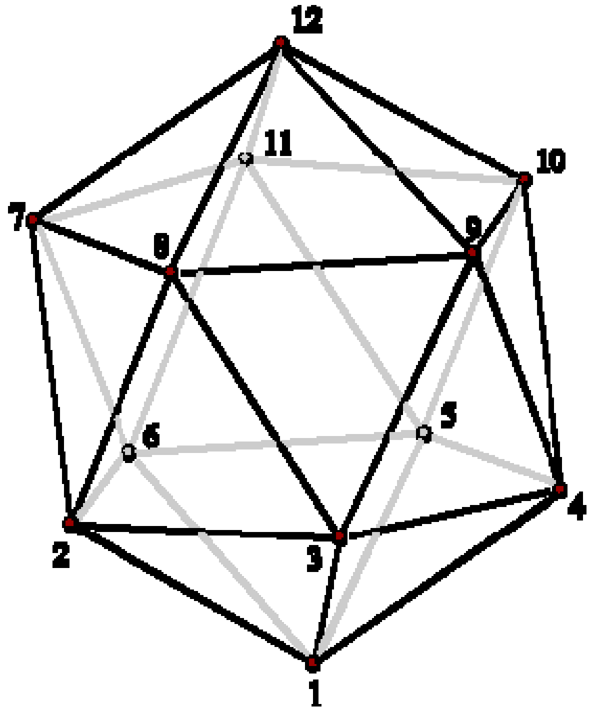
c) Find a generating matrix and a control matrix of the code C in the case $r = 3$ and $s = 4$.

22. Prove that the dual of an MDS linear code is an MDS code.

23 [van Lint, problem 3.8.5] Let C be a $\mathbb{F}_q$-linear code $[n,k]$ and G a generating matrix. Assume that no column of G is identically 0. Prove that the sum of the weights of the vectors of C is $n(q-1)q^{k-1}$.

24. Suppose that H is a check matrix of a linear code C and set $d = d_C$. Let $(x|x') \in C$ and assume $|x'| < d$. Describe a procedure that yields x' in terms of x and H.

[This problem shows that it is possible to recover s erasures in a codeword if $s < d$. This fact is the basis of some applications, including one that makes possible to reconstruct the information stored in an array of n memory discs when some of them fail. One possibility is to code the information by means of a code C of length n and store the successive components $x_1, \ldots, x_n$ of $x \in C$ in n discs $D_1, \ldots, D_n$ (x_i is stored in D_i). If $d = d_C$, then it is possible to recover from the failure of s discs if $s < d$. Schemes of this sort are known as RAIDs, from **R**edundant **A**rrays of **I**nexpensive **D**isks].



25 (A construction of the complete binary Golay code). Let $R \in M_{12}(B)$ be the non-incidence matrix of a regular icosahedron (numbering its vertices from 1 to 12, as in the figure, the element R_{ij} is 0 if the vertices i and j are joined by an edge, and 1 otherwise).

a) Calculate R explicitly.

b) Check that $R^2 = I_{12}$, or prove it on the basis of the definition of R.

c) The matrices $G = (I_{12}|R)$ and $H = (R|I_{12})$ satisfy the relation $GH^T = 0$. Since $H = (R|I_{12}) = R(I_{12}|R)$, the code $C = \langle G \rangle = \langle H \rangle$ is self-dual.

d) That C is self-dual implies that all the elements of C have even weight. Prove that in fact the weight of all elements of C is ≥ 4.

e) Using (c), show that if $(x|y) \in C$, $x, y \in B^{12}$, then $(y|x) \in C$, and deduce from this that the weight of every element of C is ≥ 8. Finally note that $C \sim [24,12,8]$.

26. If F is a finite field of q elements, $n = q - 1$ and $\{\alpha_1, \dots, \alpha_n\} = F^*$ (the set of non-zero elements of F), we write $RS_F(k)$ instead of $RS_{\alpha_1,\dots,\alpha_n}(k)$, and we say that it is the Reed–Solomon of dimension k of the field F. In this case the elements h_i and the control matrix H take a particular simple form: prove that

$$h_i = \alpha_i \text{ and } H = V_{1,n-k}(\alpha_1, \dots, \alpha_n),$$

$$\text{where } V_{1,n-k}(\alpha_1, \dots, \alpha_n) = \left(\alpha_i^j\right)_{\substack{1 \le i \le n-k \\ 1 \le j \le n}}.$$

27. Let ρ be a real number such that $0 < \rho < 1$ and t a positive integer. Let F be an arbitrary finite field and $q = |F|$.

a) Show that if the rate of $C = RS_F(k)$ is $> \rho$ and C corrects t errors, then $q \geq 1 + \frac{2t}{1-\rho}$.

What is the minimum q required for a RS code with rate $3/5$ (at least) and which corrects 7 errors? What are the possible parameters for such codes?

b) If we fix q and we need a rate $\geq \rho$, what is the maximum number of errors that we can correct? (Answer: $t \leq \left\lfloor \frac{(1-\rho)(q-1)}{2} \right\rfloor$).

How many errors can we correct if $q = 256$ and the desired rate is $3/4$?

c) If we fix q and t, prove that $\rho \leq 1 - \frac{2t}{q-1}$.

What is the maximum rate that is possible if $q = 256$ and we want to correct at least 10 errors?

28. *Shortened codes*

Given a linear code $C \subset F^n$, let

$$S_n C = \{x' \in F^{n-1} | (x', 0) \in C\}.$$

This is a linear code of length $n-1$ called the *shortening* C by the n-th coordinate (the notion of shortening by the j-th coordinate, $S_j C$, or by a set $J = \{j_1, \dots, j_l\}$ of coordinates, $S_J C$, are defined in a similar way). If $C \sim [n, k, d]$,

(a) Prove that $S_n C \sim [n-1, k', d']$, with $k-1 \leq k' \leq k$ and $d' \geq d$. More generally, $S_J C \sim [n-l, k^*, d^*]$, with $k-l \leq k^* \leq k$ and $d^* \geq d$.

(b) If C is MDS, use (a) and the Singleton inequality to show the codes obtained by shortening C are MDS (or, equivalently, that $k^* = n-l$ and $d^* = d$).

(c) If $C = RS_{\alpha_1,\dots,\alpha_n}(k)$, prove that $S_n C$ is scalarly equivalent to $RS_{\alpha_1,\dots,\alpha_{n-1}}(k-1)$.

(d) In part (a) we have $k' = k - 1$ if and only if not all code-vectors satisfy $x_n = 0$, and $k^* = k - l$ if and only if C is systematic with respect to the positions $j_1, \dots, j_l$.

29. *Decoding of* $C = RS_{\alpha_1,\dots,\alpha_n}(k)$ *by interpolating polynomials*

Let $x = (x_1, \dots, x_n) = \big(f(\alpha_1), \dots, f(\alpha_n)\big) \in C$, $f \in F[X]_k$, be the sent vector. Let $y = x + e$ be the received vector (we say e is the error vector). Let $t = \lfloor (d-1)/2 \rfloor = \lfloor (n-k)/2 \rfloor$ (note that the condition $|e| \le t$ is equivalent to $|e| < d/2$).

(a) Show that there are non-zero polynomials $P(X), Q(X) \in F[X]$ such that $\deg P(X) \le n - t - 1$, $\deg Q(X) \le n - t - k$, and satisfying

$$P(\alpha_i) + y_i Q(\alpha_i) = 0 \text{ for } i = 1, \dots, n.$$

[this condition is equivalent to n homogeneous linear equations in the coefficients of P and Q, and the number of these coefficients is

$$n - t + n - t - k + 1 = n + 1 + n - k - 2t \ge n + 1].$$

(b) Prove that if $|e| \leq t$ then $f(X) = -P(X)/Q(X)$.

(c) Use (a) and (b) to describe and algorithm to decode C.

30. Find a check matrix of $\mathrm{Ham}_7(2)$, the Hamming code over $\mathbb{F}_7$ of codimension 2, and use it to decode the message

3523410610521360.