

## Lecture 6: Hamming codes. Reed-Muller codes.

Guest lecturer: Grigory Kabatiansky

`g.kabatiansky@skoltech.ru`

Course instructor: Alexey Frolov

`al.frolov@skoltech.ru`

Teaching Assistant: Stanislav Kruglik

`stanislav.kruglik@skolkovotech.ru`

February 10, 2017

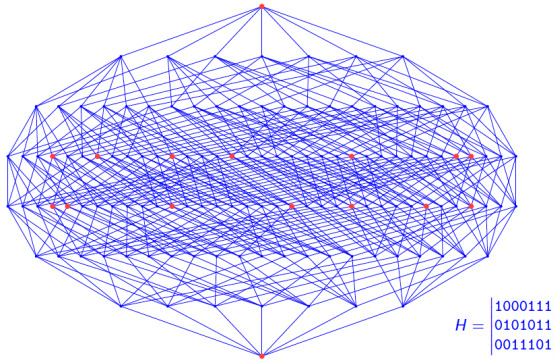
- 1 Hamming code
- 2 Reed-Muller codes
- 3 Code distance

Hamming code is determined by its parity check matrix (PCM) which consists of all non-zero column vectors. Any two columns of such code are linearly independent and there exist 3 linearly dependent columns

Parameters:

$$n = 2^m - 1, k = 2^m - m - 1, d = 3.$$

# Hamming code



# Hamming code

$$H_2 = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad H_3 = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix},$$

$$H_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

# Ulam's problem

Choose arbitrary number from 0 to  $2^{20} - 1$  and try to determine it with minimal number of “yes-no” questions.

What is the minimal number of questions? The answer is 20 questions. Let us ask a questions of form “does the number belong to a subset”. This number of questions is obtained if we take rows of Hamming code PCM as characteristic vectors of such subsets.

And what if the number of questions if some of the answers may be wrong? – 25 – 26 questions

# Syndrome calculation

$$\begin{aligned}
 \begin{pmatrix} S_1 \\ \vdots \\ S_m \end{pmatrix} &= H_m \begin{pmatrix} y_1 \\ \vdots \\ y_{2^m-1} \end{pmatrix} = \begin{array}{c|c|c}
 0 & 0 & 0 & \dots & 0 & 1 & 1 & 1 & 1 & \dots & 1 \\
 \hline
 & H_{m-1} & & & & 0 & & & & & H_{m-1} \\
 & & & & & 0 & & & & & \\
 & & & & & \vdots & & & & & \\
 & & & & & 0 & & & & & 
 \end{array} \begin{pmatrix} y_1 \\ \vdots \\ y_{2^m-1} \end{pmatrix} = \\
 &= \begin{pmatrix} y_{2^m-1} \oplus y_{2^{m-1}+1} \oplus \dots \oplus y_{2^{m-1}} \\ H_{m-1} \begin{pmatrix} y_1 \\ \vdots \\ y_{2^{m-1}-1} \end{pmatrix} \oplus H_{m-1} \begin{pmatrix} y_{2^{m-1}+1} \\ \vdots \\ y_{2^m-1} \end{pmatrix} \end{pmatrix} = \\
 &= \begin{pmatrix} y_{2^m-1} \oplus y_{2^{m-1}+1} \oplus \dots \oplus y_{2^{m-1}} \\ H_{m-1} \begin{pmatrix} y_1 \oplus y_{2^{m-1}+1} \\ \dots \\ y_{2^{m-1}-1} \oplus y_{2^m-1} \end{pmatrix} \end{pmatrix}
 \end{aligned}$$

Hamming code can correct 1 error. The syndrome is a column of the PCM, which corresponds to error.

# Hamming bound and Hamming code

$$n = 2^m - 1$$

$$d = 3$$

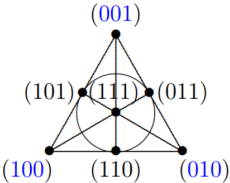
$$|\mathcal{C}| = 2^{n-m} = \frac{2^n}{2^m} = \frac{2^n}{n+1}$$

Hamming code lies on the Hamming bound. Such codes are called perfect codes. Another example of perfect code is Golay codes with parameters  $n = 23$ ,  $k = 12$ ,  $d = 7$



# Hamming code and Fano plane

The Hamming Code and the projective plane of order the  $PG(2, 2)$  (Fano plane) are closely related.

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}, G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \leftrightarrow$$


# Extended binary Hamming code

The extended binary Hamming code is the code obtained from binary hamming code by adding a check bit. In the PCM we add additional row and column. PCM is presented below.

$$\bar{H} = \left[ \begin{array}{ccc|c} & & & 0 \\ & H & & \vdots \\ & & & 0 \\ \hline 1 & \cdots & 1 & 1 \end{array} \right]$$

Because we add addition parity check bit our code distance is increased by one, so the parameters of code are the following:  $n = 2^m, k = 2^m - m - 1, d = 4$ .

# Non-binary Hamming code

PCM  $\mathbf{H}$  of non-binary Hamming code has the property that its columns are made up of precisely one nonzero vector from each vector subspace of dimension 1 of  $\mathbb{F}_q^m$ .

$$H_m^q = \left[ \begin{array}{cccccc} 0 & 0 & 0 & & 0 & 1 \\ 0 & 0 & 0 & & 1 & * \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & & * & * \\ 0 & 0 & 1 & & * & * \\ 0 & 1 & * & & * & * \\ 1 & * & * & & * & * \end{array} \right] \left. \vphantom{\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ 1 \end{array}} \right\} m$$

The parameters of the resulting code are as follows.  $n = \frac{q^m - 1}{q - 1}$ ,  $k = n - m$  and  $|\mathcal{C}| = \frac{q^n}{1 + n(q - 1)}$ . The code is also perfect.

# Open problem

Are there perfect codes with  $d = 3$  if  $q \neq p^m$ , where  $p$  is a prime number?

# Dual code to binary Hamming code

The dual of a code  $\mathcal{C}$  is denoted by  $\mathcal{C}^\perp$ . For a linear code  $\mathcal{C}$  with parity check matrix  $\mathbf{H}$ ,  $\mathcal{C}^\perp$  is the linear code with generator matrix  $\mathbf{H}$ .

The dual of the Hamming code is a linear code with parameters  $[2^m - 1, m]$  with a generator matrix whose rows are all the nonzero  $m$ -bit vectors. This is the Simplex code. If we include the zero column, we obtain the Hadamard code  $[2^m, m]$ . We note that the Hadamard code is the most redundant linear code in which no two codeword symbols are equal in every codeword. Its distance is equal to  $n/2$

# Outline

- 1 Hamming code
- 2 Reed-Muller codes
- 3 Code distance

Let us consider a boolean function (Zhegalkin polynomial) from  $m$  variables with degree no more than  $s$ . The coefficients correspond to information bits. A value of this polynomial in all points of  $m$ -dimensional Boolean cube is a codeword (evaluation code). This code is called a  $RM(m, s)$  code. The parameters of this code are the following:  $n = 2^m$ ,  $k = \sum_{i=0}^s \binom{m}{i}$ .

Let us introduce some notation

- $U$  – linear  $[n, k_U, d_U]$  code;
- $V$  – linear  $[n, k_V, d_V]$  code;

Plotkin construction

$$\mathcal{C} = U \triangle V = (U, U + V) = \{(u, u + v) : u \in U, v \in V\}.$$

Properties: linear,  $n(\mathcal{C}) = 2n$ ,  $k(\mathcal{C}) = k_U + k_V$ .



## Problem (1 point)

Prove, that

$$d(\mathcal{C}) = \min\{d_V, 2d_U\}.$$

## Lemma

$$RM(s, m) = RM(m - 1, s) \triangle RM(m - 1, s - 1).$$

## Proof.

$$f(x_1, x_2, \dots, x_m) = Q(x_1, x_2, \dots, x_{m-1}) + x_m P(x_1, x_2, \dots, x_{m-1}),$$

where  $\deg Q \leq s$  and  $\deg P \leq s - 1$ . □

## Lemma

$$d(RM(m, s)) = 2^{m-s}.$$

## Proof.

By induction. Base

$$d(RM(m, 1)) = 2^{m-1}.$$

Thus,

$$\begin{aligned} d(RM(m, s)) &= \min\{d(RM(m-1, s)), 2d(RM(m-1, s-1))\} \\ &= 2^{m-s}. \end{aligned}$$



Thank you for your attention!