Coding Theory

RWTH Aachen, SS 2022
RWTH Aachen, SS 2019
Universität Duisburg-Essen, WS 2011
RWTH Aachen, SS 2006

Jürgen Müller

Contents

I

0 Introduction

(0.1) The basic model of communication is that of sending information between communication partners, Alice and Bob say, who communicate through some channel, which might be anything as, for example, a telephone line, radio, an audio compact disc, a keyboard, and so on. This leads to the following questions:

• **Information theory.** What is information? How much information can be sent through a channel per time unit?

• **Coding theory.** The channel might be **noisy**, that is information might be changed randomly when sent through the channel. How is it possible to recover a sufficiently large fraction of the original information from distorted data?

• **Cryptography.** The channel might be **insecure**, that is information which is intended to be kept private to Alice and Bob might be caught by an opponent, Oscar say, or even be changed deliberately by Oscar. How can this be prevented?

(0.2) Alphabets. A finite set $\mathcal{X}$ such that $|\mathcal{X}| \geq 2$ is called an **alphabet**, its elements are called **letters** or **symbols**. A finite sequence $w = [x_1, \ldots, x_n]$ consisting of $n \in \mathbb{N}$ letters $x_i \in \mathcal{X}$ is called a **word** over $\mathcal{X}$ of **length** $l(w) = n$. The empty sequence ϵ is called the **empty word**, and we let $l(\epsilon) := 0$. Let $\mathcal{X}^n$ be the set of all words of length $n \in \mathbb{N}_0$, and let $\mathcal{X}^* := \coprod_{n \in \mathbb{N}_0} \mathcal{X}^n$. For $v, w \in \mathcal{X}^*$ let $vw \in \mathcal{X}^*$ be their **concatenation**. We have $v\epsilon = \epsilon v = v$ and $(uv)w = u(vw)$, for $u, v, w \in \mathcal{X}^*$. Hence $\mathcal{X}^*$ is a monoid, the **free monoid** over $\mathcal{X}$, and the length function $l \colon \mathcal{X}^* \to (\mathbb{N}_0, +) \colon w \mapsto l(w)$ is a monoid homomorphism.

To describe numbers, typically alphabets $\mathbb{Z}_q := \{0, \ldots, q-1\}$ for $q \in \mathbb{N} \setminus \{1\}$ are used, for example $q = 10$. In computer science, the alphabet $\mathbb{Z}_2 = \{0, 1\}$, whose elements are called **binary digits** or **bits**, the alphabet $\mathbb{Z}_8$, whose elements are called **Bytes**, and the **hexadecimal** alphabet $\{0, \ldots, 9, A, B, C, D, E, F\}$ being in bijection with $\mathbb{Z}_{16}$ are used. For interchange of written texts the **Latin** alphabet $\{\mathtt{A}, \ldots, \mathtt{Z}\}$ being in bijection with $\mathbb{Z}_{26}$, and the **American Standard Code for Information Interchange (ASCII)** alphabet being in bijection with $\mathbb{Z}_{128}$ are used.

Then, to transmit information, it has to be **(source) encoded** into words over an alphabet $\mathcal{X}$ suitable for the chosen channel, and words have to be **decoded** again after transmition. Thus, a **code** is just a subset $\emptyset \neq \mathcal{C} \subseteq \mathcal{X}^*$; then $\mathcal{C}$ is interpreted as the set of all words representing sensible information.

1 Parity check codes

Parity check codes are used to detect typing errors. They are not capable of correcting errors, and thus are used whenever the data can easily be reentered

Table 1: Typing errors.

error		frequency
single	$a \longrightarrow b$	79.0%
adjacent transposition	$ab \longrightarrow ba$	10.2%
jump transposition	$abc \longrightarrow cba$	0.8%
twin	$aa \longrightarrow bb$	0.6%
jump twin	$aca \longrightarrow bcb$	0.3%
phonetic	$a0 \longrightarrow 1a$	0.5%
random		8.6%

again. Typical typing errors and their frequencies are given in Table 1; an example of a phonetic error is replacing ‘thirty’ by ‘thirteen’.

(1.1) Example: The ISBN [1968, 2007]. The **International Standard Book Number** is used to identify books, where up to the year 2006 the standard was **ISBN-10**, which from the year 2007 on has been replaced by **ISBN-13**. The ISBN-10 is formed as follows:

The alphabet is $\mathbb{Z}_{11}$, where 10 is replaced by the Roman letter X, and words $[x_1; x_2, \ldots, x_6; x_7, \ldots, x_9; x_{10}] \in \mathbb{Z}_{10}^9 \times \mathbb{Z}_{11}$ have length 10, where X might possibly occur only as a last letter. Here $x_1, \ldots, x_9$ are information symbols, where x_1 is the **group code**, $x_1 \in \{0, 1\}$ referring to English, $x_1 = 2$ referring to French, and $x_1 = 3$ referring to German, $[x_2, \ldots, x_6]$ is the **publisher code**, $[x_7, \ldots, x_9]$ is the **title code**, and x_{10} is a check symbol fulfilling $x_{10} = \sum_{i=1}^{9} ix_i \in \mathbb{Z}_{11}$. Hence a valid ISBN-10 is an element of the $\mathbb{Z}_{11}$-subspace $\{[x_1, \ldots, x_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} ix_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$.

From 2007 on the ISBN-13 is used: After a 3-letter prefix, being a country code 978 or 979 referring to ‘bookland’, the first 9 letters of the ISBN-10 are taken, and then a check symbol is added such that the EAN standard is fulfilled, see (1.2). For example, a valid ISBN-10 is ‘1-58488-508-4’: We have $1 \cdot 1 + 2 \cdot 5 + 3 \cdot 8 + 4 \cdot 4 + 5 \cdot 8 + 6 \cdot 8 + 7 \cdot 5 + 8 \cdot 0 + 9 \cdot 8 = 246 = 4 \in \mathbb{Z}_{11}$. The corresponding ISBN-13 is ‘978-1-58488-508-5’; indeed we have $9 + 3 \cdot 7 + 8 + 3 \cdot 1 + 5 + 3 \cdot 8 + 4 + 3 \cdot 8 + 8 + 3 \cdot 5 + 0 + 3 \cdot 8 = 145 = 5 = -5 \in \mathbb{Z}_{10}$.

(1.2) Example: The EAN [1977]. The **International Article Number (EAN)**, formerly **European Article Number**, is formed as follows: The alphabet is $\mathbb{Z}_{10}$, and words $[x_1, \ldots, x_3; x_4, \ldots, x_7; x_8, \ldots, x_{12}; x_{13}] \in \mathbb{Z}_{10}^{13}$ have length 13. Here $x_1, \ldots, x_{12}$ are information symbols, where $[x_1, \ldots, x_3]$ is the **country code**, $x_1 = 4$ referring to Germany, $[x_4, \ldots, x_7]$ is the **company code**, $[x_8, \ldots, x_{12}]$ is the **article code**, and x_{13} is a check symbol fulfilling $x_{13} = -\sum_{i=1}^{6}(x_{2i-1} + 3x_{2i}) \in \mathbb{Z}_{10}$. Hence a valid EAN is an element of the set

$\{[x_1,\ldots,x_{13}]\in\mathbb{Z}_{10}^{13};\sum_{i=1}^{6}(x_{2i-1}+3x_{2i})+x_{13}=0\in\mathbb{Z}_{10}\}\subseteq\mathbb{Z}_{10}^{13}$.

The bar code printed on goods is formed as follows: Each bar is either black or white, and has width 1, 2, 3 or 4. Each letter is encoded by 4 bars of alternating colors, whose widths add up to 7; see Table 2 where 0 and 1 stand for white and black, respectively. The odd and even type codes for each letter start with white and end with black, where for the even type codes the width patterns are just those for the odd type codes read backwardly. The negative type code for each letter starts with black and ends with white, using the same width pattern as for the odd type code, which hence amounts to just reading the even type code backwardly. In the odd type code for each letter the widths of the black bars add up to an odd number, while in the even type code these sums are even.

An EAN is depicted as follows: There is a prefix 101, then the letters $x_2,\ldots,x_7$ are depicted by odd and even type codes, then there is an infix 01010, then the letters $x_8,\ldots,x_{12}$ are depicted by negative type codes, and finally there is a postfix 101. The choice of the odd and even type codes for $x_2,\ldots,x_7$ is determined by x_1; see Table 2, where $-$ and $+$ stand for odd and even, respectively. Since the even type codes, that is the negative type codes read backwardly, are disjoint from the odd type codes, this allows to read bar codes in either direction and to swap data if necessary, or to read data in two halves.

For example, '4-901780-728619' indeed yields $4+3\cdot 9+0+3\cdot 1+7+3\cdot 8+0+3\cdot 7+2+3\cdot 8+6+3\cdot 1=121=1=-9\in\mathbb{Z}_{10}$, hence this is a valid EAN. The pattern [odd, even, odd, odd, even, even] for $x_1=4$ yields:

101
0001011 0100111 0011001 0111011 0001001 0100111
01010
1000100 1101100 1001000 1010000 1100110 1110100
101

(1.3) Example: The IBAN [2007]. The German general version of the **International Bank Account Number (IBAN)** is formed as follows:

Words $[x_1,x_2;x_3,x_4;x_5,\ldots,x_{12};x_{13},\ldots,x_{22}]\in\mathbb{Z}_{26}^2\times\mathbb{Z}_{10}^{20}$ have length 22, where actually x_1,x_2 are Latin letters, and we identify the Latin alphabet with $\mathbb{Z}_{26}$ by letting `A` $\mapsto 0$, `B` $\mapsto 1$, …, `Z` $\mapsto 25$. Here, $x_1,x_2;x_5,\ldots,x_{12};x_{13},\ldots,x_{22}$ are information symbols, where $[x_1,x_2]$ is the **country code**, for Germany being `DE` $\mapsto[3,4]$, followed by the 8-digit **bank identification number** $[x_5,\ldots,x_{12}]$, and the 10-digit **bank account number** $[x_{13},\ldots,x_{22}]$, where the latter is possibly filled up by leading zeroes; the word $[x_5,\ldots,x_{22}]$ is also called the **Basic Bank Account Number (BBAN)**.

Finally, $[x_3,x_4]$ are check symbols fulfilling the following condition: The concatenation $v:=x_5\cdots x_{22}(x_1+10)(x_2+10)x_3x_4\in\mathbb{Z}_{10}^{24}$ can be considered as a non-negative integer having 24 decimal digits, where $\mathbb{Z}_{26}+10=\{10,\ldots,35\}$. Then v is a valid IBAN if $v\equiv 1 \pmod{97}$.

Table 2: EAN bar code.

letter	odd	code	negative	even	code	odd/even
0	3211	0001101	1110010	1123	0100111	$------$
1	2221	0011001	1100110	1222	0110011	$--+-++$
2	2122	0010011	1101100	2122	0010011	$--++-+$
3	1411	0111101	1000010	1141	0100001	$--+++-$
4	1132	0100011	1011100	2311	0011101	$-+--++$
5	1231	0110001	1001110	1321	0111001	$-++--+$
6	1114	0101111	1010000	4111	0000101	$-+++--$
7	1312	0111011	1000100	2131	0010001	$-+-+-+$
8	1213	0110111	1001000	3121	0001001	$-+-++-$
9	3112	0001011	1110100	2113	0010111	$-++-+-$

Hence allowing for the alphabet $\mathbb{Z}_{97}$, containing the digits $\mathbb{Z}_{10}$ as a subset, the check condition can be rephrased as $(\sum_{i=5}^{22} x_i \cdot 10^{28-i}) + (x_1+10)\cdot 10^4 + (x_2+10)\cdot 10^2 + x_3 \cdot 10 + x_4 = 1 \in \mathbb{Z}_{97}$. Thus check symbols $x_3, x_4 \in \mathbb{Z}_{10}$ can always be found, where $[x_3, x_4] \not\in \{[0,0],[0,1],[9,9]\}$ for uniqueness. Letting $w := [10^{24-i} \in \mathbb{Z}_{97}; i \in \{1,\ldots,18\}] \in \mathbb{Z}_{97}^{18}$, that is

$$w = [56, 25, 51, 73, 17, 89, 38, 62, 45, 53, 15, 50, 5, 49, 34, 81, 76, 27],$$

and $x_1 := 3$ and $x_2 := 4$, entailing $(x_1+10)\cdot 10^4 + (x_2+10)\cdot 10^2 = 62 \in \mathbb{Z}_{97}$, we infer that the valid IBAN can be identified with the set $\{[x_3,\ldots,x_{22}] \in \mathbb{Z}_{10}^{20}; (x_3 \cdot 10 + x_4) + \sum_{i=1}^{18} w_i x_{i+4} = 36 \in \mathbb{Z}_{97}\}$.

For example, given the bank identification number '390 500 00' and the fictious bank account number '0123456789', we get the BBAN '3905 0000 0123 4567 89'. For the latter we get $\sum_{i=1}^{18} w_i x_{i+4} = 65 \in \mathbb{Z}_{97}$, thus the check condition yields $x_3 \cdot 10 + x_4 = 68 \in \mathbb{Z}_{97}$, so that we get the IBAN 'DE68 3905 0000 0123 4567 89'.

(1.4) Parity check codes over $\mathbb{Z}_q$. Let $q \geq 2$ be the **modulus**, let $n \in \mathbb{N}$, and let the **weights** $w := [w_1, \ldots, w_n] \in \mathbb{Z}_q^n$ be fixed. Then $v := [x_1, \ldots, x_n] \in \mathbb{Z}_q^n$ is called **valid** if $vw^{\text{tr}} := \sum_{i=1}^n x_i w_i = 0 \in \mathbb{Z}_q$.

a) We consider single errors: Let $v := [x_1, \ldots, x_n] \in \mathbb{Z}_q^n$ be valid and let $v' := [x_1, \ldots, x_j', \ldots, x_n] \in \mathbb{Z}_q^n$ such that $x_j' \neq x_j$ for some $j \in \{1, \ldots, n\}$. Then we have $v'w^{\text{tr}} = (v'-v)w^{\text{tr}} = (x_j' - x_j)w_j \in \mathbb{Z}_q$, hence $x_j' \neq x_j$ is detected if and only if $x_j' w_j \neq x_j w_j \in \mathbb{Z}_q$. Thus all single errors are detected if and only if all weights $w_j \in \mathbb{Z}_q$ are chosen such that the map $\mu_{w_j} \colon \mathbb{Z}_q \rightarrow \mathbb{Z}_q \colon x \mapsto x w_j$ is injective, or equivalently bijective.

For $y \in \mathbb{Z}_q$ the map $\mu_y \colon \mathbb{Z}_q \rightarrow \mathbb{Z}_q \colon x \mapsto xy$ is injective if and only if $y \in \mathbb{Z}_q^* := \{z \in \mathbb{Z}_q; \gcd(z, q) = 1\}$, the group of units of $\mathbb{Z}_q$: Let $d := \gcd(y, q) \in \mathbb{N}$. If $d > 1$ then we have $0 \neq \frac{q}{d} \in \mathbb{Z}_q$ and $\frac{q}{d} \cdot y = 0 = 0 \cdot y \in \mathbb{Z}_q$, hence μ_y is not injective. Since by the Euclidean Algorithm there are Bézout coefficients $s, t \in \mathbb{Z}$ such that $d = ys + qt \in \mathbb{Z}$, if $d = 1$ then $ys = d = 1 \in \mathbb{Z}_q$, thus from $xy = x'y \in \mathbb{Z}_q$ we get $x = xys = x'ys = x' \in \mathbb{Z}_q$, implying that μ_y is injective.

For example, for the non-prime modulus $q = 10$ used in the EAN we get $\mu_1 = \mathrm{id}_{\mathbb{Z}_{10}}$ and $\mu_3 = (0)(1,3,9,7)(2,6,8,4)(5) \in \mathcal{S}_{\mathbb{Z}_{10}}$, hence the weight tuple $w = [1,3,\ldots,1,3,1] \in (\mathbb{Z}_{10}^*)^{13}$ allows to detect all single errors. For the prime modulus $q = 11$ used in the ISBN-10 we have $\mathbb{Z}_{11}^* = \mathbb{Z}_{11} \setminus \{0\}$, hence again the weight tuple $w = [1,\ldots,10] \in (\mathbb{Z}_{11}^*)^{10}$ allows to detect all single errors. A similar consideration for the IBAN, using the prime modulus $q = 97$, shows that the weight tuple for the BBAN allows to detect all single errors.

b) We consider adjacent transposition errors for $n \geq 2$: Let $v := [x_1, \ldots, x_n] \in \mathbb{Z}_q^n$ be valid and let $v' := [x_1, \ldots, x_{j+1}, x_j, \ldots, x_n] \in \mathbb{Z}_q^n$ such that $x_{j+1} \neq x_j$ for some $j \in \{1, \ldots, n-1\}$. Then we have $v'w^{\mathrm{tr}} = (v' - v)w^{\mathrm{tr}} = (x_j - x_{j+1})(w_{j+1} - w_j) \in \mathbb{Z}_q$. Thus all adjacent transposition errors are detected if and only if the weights fulfill $w_{j+1} - w_j \in \mathbb{Z}_q^*$ for all $j \in \{1, \ldots, n-1\}$.

Since for the EAN we have $w_{j+1} - w_j \in \{2, 8\} \subseteq \mathbb{Z}_{10} \setminus \mathbb{Z}_{10}^*$, for $j \in \{1, \ldots, 12\}$, adjacent transposition errors are not necessarily detected. Since for the ISBN-10 we have $w_{j+1} - w_j = 1 \in \mathbb{Z}_{11}^*$, for $j \in \{1, \ldots, 9\}$, all adjacent transposition errors are detected; thus in this respect the transition from ISBN-10 to ISBN-13 is not an improvement. Similarly, since for the BBAN the adjacent weights in $\mathbb{Z}_{97}$ are pairwise distinct, all adjacent transposition errors are detected.

(1.5) Parity check codes over arbitrary groups. a) Let G be a finite group, let $n \in \mathbb{N}$, and let $\pi_i \colon G \to G$ for $i \in \{1, \ldots, n\}$ be fixed. Then $[x_1, \ldots, x_n] \in G^n$ is called **valid** if $x_1^{\pi_1} \cdots x_n^{\pi_n} = 1$. For example, letting $G := \mathbb{Z}_q$ and $\pi_i := \mu_{w_i}$, where $w_i \in \mathbb{Z}_q$ for $i \in \{1, \ldots, n\}$, we recover the parity check codes in (1.4); here additionally the π_i are group homomorphisms.

We consider single errors: Let $v := [x_1, \ldots, x_n] \in G^n$ be valid and let $v' := [x_1, \ldots, x_j', \ldots, x_n] \in G^n$ such that $x_j' \neq x_j$ for some $j \in \{1, \ldots, n\}$. Let $y_i := x_i^{\pi_i} \in G$ for $i \in \{1, \ldots, n\}$, and $y_j' := (x_j')^{\pi_j}$. Then v' is valid if and only if $y_1 \cdots y_{j-1} y_j' y_{j+1} \cdots y_n = 1 = y_1 \cdots y_n$, which by multiplying from the left by $y_1^{-1}, \ldots, y_{j-1}^{-1}$, and from the right by $y_n^{-1}, \ldots, y_{j+1}^{-1}$, is equivalent to $(x_j')^{\pi_j} = y_j' = y_j = x_j^{\pi_j}$. Hence we conclude that all single errors are detected if and only if π_j is injective, or equivalently bijective, for all $j \in \{1, \ldots, n\}$.

b) Let π_i be injective for all $i \in \{1, \ldots, n\}$. We consider adjacent transposition errors for $n \geq 2$: Let $v := [x_1, \ldots, x_n] \in G^n$ be valid and let $v' := [x_1, \ldots, x_{j+1}, x_j, \ldots, x_n] \in G^n$ such that $x_{j+1} \neq x_j$ for some $j \in \{1, \ldots, n-1\}$. Let $y_i := x_i^{\pi_i} \in G$ for $i \in \{1, \ldots, n\}$ and $y_j' := x_{j+1}^{\pi_j} \in G$ and $y_{j+1}' := x_j^{\pi_{j+1}} \in G$. Then v' is valid if and only if $y_1 \cdots y_{j-1} y_j' y_{j+1}' y_{j+2} \cdots y_n = 1 = y_1 \cdots y_n$, which by multiplying from the left by $y_1^{-1}, \ldots, y_{j-1}^{-1}$, and from the right by $y_n^{-1}, \ldots, y_{j+2}^{-1}$, is equivalent to $x_{j+1}^{\pi_j} x_j^{\pi_{j+1}} = y_j' y_{j+1}' = y_j y_{j+1} = x_j^{\pi_j} x_{j+1}^{\pi_{j+1}}$. Writing $g := x_j^{\pi_j} \in G$ and $h := x_{j+1}^{\pi_j} \in G$ and letting $\tau_j := \pi_j^{-1} \pi_{j+1}$, we conclude that all adjacent transposition errors are detected if and only if $gh^{\tau_j} \neq hg^{\tau_j}$ for

Table 3: Elements of D_{10}.

x	x	$\overline{x}$		
0	A	1	id	()
1	D	2	α	$(1,2,3,4,5)$
2	G	3	α^2	$(1,3,5,2,4)$
3	K	4	α^3	$(1,4,2,5,3)$
4	L	5	α^4	$(1,5,4,3,2)$
5	N	6	β	$(2,5)(3,4)$
6	S	7	$\alpha\beta$	$(1,5)(2,4)$
7	U	8	$\alpha^2\beta$	$(1,4)(2,3)$
8	Y	9	$\alpha^3\beta$	$(1,3)(4,5)$
9	Z	10	$\alpha^4\beta$	$(1,2)(3,5)$

all $g \neq h \in G$ and $j \in \{1, \ldots, n-1\}$.

(1.6) Example: Serial numbers. Let D_{10} be the dihedral group of order 10, that is the symmetry group of the plane equilateral pentagon; up to isomorphism there are precisely two groups of order 10, the cyclic group $\mathbb{Z}_{10}$ and the non-abelian group D_{10}. Numbering the vertices of the pentagon counterclockwise, the elements of $D_{10} := \langle \alpha, \beta \rangle \leq \mathcal{S}_5$ are as given in Table 3. Using the numbering of the elements given there let $\tau := (1,2,6,9,10,5,3,8)(4,7) \in \mathcal{S}_{D_{10}}$. Then it can be checked that $gh^\tau \neq hg^\tau$ for all $g \neq h \in D_{10}$.

The serial numbers on the former German currency **Deutsche Mark (DM)** are formed as follows: The alphabet is $\mathcal{X} := \{\texttt{0}, \ldots, \texttt{9}, \texttt{A}, \texttt{D}, \texttt{G}, \texttt{K}, \texttt{L}, \texttt{N}, \texttt{S}, \texttt{U}, \texttt{Y}, \texttt{Z}\}$, and words $[x_1, \ldots, x_{10}; x_{11}] \in \mathcal{X}^{11}$ have length 11, where $x_1, \ldots, x_{10}$ are information symbols and x_{11} is a check symbol. Replacing $x_i \in \mathcal{X}$ by $\overline{x}_i \in D_{10}$ as indicated in Table 3, a word is valid if $\overline{x}_1^\tau \cdots \overline{x}_{10}^{\tau^{10}} \overline{x}_{11} = \mathrm{id} \in D_{10}$.

For example, for GG0184220N0 we get elements $[3,3,1,2,9,5,3,3,1,6;1]$, hence $[3^\tau, 3^{\tau^2}, 1^{\tau^3}, 2^{\tau^4}, 9^{\tau^5}, 5^{\tau^6}, 3^{\tau^7}, 3^{\tau^8}, 1^{\tau^9}, 6^{\tau^{10}}; 1] = [8,1,9,5,1,9,5,3,2,10;1]$, and it can be checked that the product of the associated elements equals $\mathrm{id} \in D_{10}$.

(1.7) Complete maps for abelian groups. a) We briefly digress into group theory, inasmuch the above leads to the following definition: Given a finite abelian group G, a bijective map $\sigma \colon G \to G$ is called **complete**, if the map $\tau := (\sigma + \mathrm{id}_G) \colon G \to G \colon g \mapsto g^{\sigma+1} := g^\sigma g$ is bijective again.

It turns out that G has a complete map if either $|G|$ is odd or G has at least two involutions: (It is surprisingly difficult to prove this completely, so that we only give a partial proof, encompassing the accessible pieces; see [Paige, 1947].)

Let G have a unique involution, z say, and assume that both $\sigma \colon G \to G$ and $\tau := \sigma + \mathrm{id}_G$ are bijective, then pairing off the elements of G with their additive inverses yields $\sum_{g\in G} g = z$, and thus $\sum_{g\in G} g = \sum_{g\in G} g^\tau = \sum_{g\in G} g^{\sigma+1} = \sum_{g\in G} g^\sigma + \sum_{g\in G} g = z + z = 0$, a contradiction.

Recalling that G can be written as a direct sum of cyclic groups of prime power order, to prove the existence of a complete map in the remaining cases we may assume that $G = \mathbb{Z}_q$ where q is odd, or $G = \mathbb{Z}_{2^a} \oplus \mathbb{Z}_{2^b}$ or $G = \mathbb{Z}_{2^a} \oplus \mathbb{Z}_{2^b} \oplus \mathbb{Z}_{2^c}$ where $a \geq b \geq c > 0$. If $G = \mathbb{Z}_q$ where q is odd, then both $\sigma = \mathrm{id}_{\mathbb{Z}_q}$ and $\tau = \sigma + \mathrm{id}_{\mathbb{Z}_q} = \mu_2$ are bijective; recall that $2 \in \mathbb{Z}_q^*$. Unfortunately, we are not able to deal with the cases $G = \mathbb{Z}_{2^a} \oplus \mathbb{Z}_{2^b}$ or $G = \mathbb{Z}_{2^a} \oplus \mathbb{Z}_{2^b} \oplus \mathbb{Z}_{2^c}$ here. ♯

Anyway, we have shown that $G = \mathbb{Z}_q$ has a complete map if and only if q is odd.

b) This is related to parity check codes as follows: Given a bijective map $\tau\colon G \to G$, the condition $gh^\tau \neq hg^\tau$, for all $g \neq h \in G$, is equivalent to $g^{\tau-1} \neq h^{\tau-1}$, for all $g \neq h \in G$, that is $\sigma := \tau - \mathrm{id}_G\colon G \to G$ is bijective as well.

Hence, for a parity check code over G with respect to bijections π_j, for $j \in \{1, \ldots, n\}$, which detects all adjacent transposition errors the associated maps $\sigma_j := \pi_j^{-1}\pi_{j+1} - \mathrm{id}_G$, for $j \in \{1, \ldots, n-1\}$, are complete. Conversely, given a complete map $\sigma\colon G \to G$, we may let $\pi_j := (\sigma + \mathrm{id}_G)^j$ for $j \in \{1, \ldots, n\}$. This shows that there is a parity check code over G, for $n \geq 2$, which detects all single errors and all adjacent transposition errors if and only if G has a complete map.

In particular, there is no parity check code over $\mathbb{Z}_{10}$ which detects all single errors and adjacent transposition errors; thus it is not surprising that the EAN does not detect all adjacent transposition errors.

Moreover, if $\pi_i = \mu_{w_i}$ where $w_i \in \mathbb{Z}_q^*$ for $i \in \{1, \ldots, n\}$, then we get $\tau_j = \pi_j^{-1}\pi_{j+1} = \mu_{w_j}^{-1}\mu_{w_{j+1}}$, for $j \in \{1, \ldots, n-1\}$, and $\tau_j - \mathrm{id}_{\mathbb{Z}_q} = \mu_{w_j}^{-1}(\mu_{w_{j+1}} - \mu_{w_j}) = \mu_{w_j}^{-1}\mu_{w_{j+1}-w_j}$ is bijective if and only if $\mu_{w_{j+1}-w_j}$ is bijective, or equivalently $w_{j+1} - w_j \in \mathbb{Z}_q^*$, as we have already seen in (1.4).

Note that for the ISBN-10 we have $\pi_i = \mu_i\colon \mathbb{Z}_{11} \to \mathbb{Z}_{11}$, for $i \in \{1, \ldots, 10\}$, thus $\tau_j = \mu_j^{-1}\mu_{j+1}$ and $\tau_j - \mathrm{id}_{\mathbb{Z}_{11}} = \mu_j^{-1}\mu_{(j+1)-j} = \mu_j^{-1}\mu_1 = \mu_j^{-1}$, for $j \in \{1, \ldots, 9\}$.

(1.8) Complete maps for arbitrary groups. Let G be a finite group. A bijective map $\sigma\colon G \to G$ is called **complete**, if the map $\tau\colon G \to G\colon g \mapsto gg^\sigma$ is bijective again. Note that, since $(g^\tau)^{-1}g = (g^\sigma)^{-1}$ for $g \in G$, we may likewise call σ complete if $G \to G\colon g \mapsto g^\sigma g$ is bijective again. Moreover, by going over to $G \to G\colon g \mapsto g^\sigma(1^\sigma)^{-1}$ we may assume that $1^\sigma = 1 = 1^\tau$.

We are concerned with the question of characterising the groups having complete maps. To this end, we consider the following conditions:
i) G has a complete map.
ii) There is an ordering $\{g_1, \ldots, g_{|G|}\}$ of G such that $g_1 \cdots g_{|G|} = 1$.
iii) We have $\prod_{g\in G} g \in [G, G]$ for some, and hence any ordering, of the factors.
iv) The Sylow 2-subgroups of G are either trivial or non-cyclic.

a) As for the latter three conditions, we proceed to show that **ii)**⇒**iii)** and **iv)**⇔**iii)**. The missing implication **iii)**⇒**ii)** is a special case of results by [Dénes, Hermann, 1982] (which we are not able to present here).

ii)⇒iii): We have $g_1 \cdots g_{|G|} = 1 \in [G,G]$. Since $gh \equiv hg \pmod{[G,G]}$ for $g, h \in G$, being an element of $[G,G]$ is independent of the order in which the product of the elements of G is taken.

iii)⇒iv): Let $\prod_{g \in G} g \in [G,G]$ for some, and hence any, ordering of the elements of G, and assume to the contrary that $\{1\} \neq S \leq G$ is a cyclic Sylow 2-subgroup of G. Then since $\mathrm{Aut}(S)$ is a 2-group we infer that $N_G(S) = C_G(S)$, thus $S \leq Z(N_G(S))$, which by Burnside's p-complement theorem implies that G is 2-nilpotent, that is G has a normal 2-complement $H \trianglelefteq G$, so that $G \cong H \rtimes S$.

Letting $z \in S$ be the unique involution, pairing off the elements of S with their inverses yields $\prod_{s \in S} s = z$. Moreover, since $G/H \cong S$ is abelian, we have $[G,G] \leq H$. Hence we get $\prod_{g \in G} g \equiv \prod_{s \in S} \prod_{h \in H} sh \equiv (\prod_{s \in S} s)^{|H|} \equiv z^{|H|} \equiv z \not\equiv 1 \pmod{H}$. Thus $\prod_{g \in G} g \notin H$, hence $\prod_{g \in G} g \notin [G,G]$, a contradiction.

iv)⇒iii): We have to show that $\prod_{g \in G} g \in [G,G]$ in some, and hence any, ordering of the elements of G. To this end, let $I(G) := \{z \in G \setminus \{1\}; z^2 = 1\}$ be the set of involutions of G. Pairing off the elements of G with their inverses, we have to show that $\prod_{z \in I(G)} z \in [G,G]$. We are done if G has odd order, thus we may assume that G has even order and a non-cyclic Sylow 2-subgroup S.

Next, we observe that for G-conjugate $z, z' \in I(G)$ we have $zz' = z \cdot z^g = z^{-1}g^{-1}zg = [z,g] \in [G,G]$, for some $g \in G$. Now $I(G)$ is a union of G-conjugacy classes, where for a G-conjugacy $C \subseteq I(G)$ of even length we hence have $\prod_{z \in C} z \in [G,G]$. Thus letting $I'(G) := \{z \in I(G); [G\colon C_G(z)] \text{ odd}\}$ be the set of central involutions of G, we have to show that $\prod_{z \in I'(G)} z \in [G,G]$.

Letting $C \subseteq I'(G)$ be a G-conjugacy class, we have $C \cap Z(S) \neq \emptyset$. By Burnside's theorem saying that two normal subsets of S are G-conjugate if and only if they are $N_G(S)$-conjugate, we conclude that $C \cap Z(S) \subseteq I(Z(S))$ is an $N_G(S)$-conjugacy class. Since $S \leq N_G(S)$ centralises $Z(S)$, we conclude that $C \cap Z(S)$ has odd length, so that $C \setminus Z(S)$ has even length, entailing $\prod_{z \in C \setminus Z(S)} z \in [G,G]$. Hence running through all G-conjugacy classes in $I'(G)$, leading to a covering of $I(Z(S))$, we conclude that we have to show that $\prod_{z \in I(Z(S))} z \in [G,G]$.

Since $Z(S) \neq \{1\}$ is abelian, we have $I(Z(S)) \,\dot\cup\, \{1\} \cong \mathbb{Z}_2^d$, for some $d \in \mathbb{N}$, being the largest 2-elementary abelian subgroup of $Z(S)$. Since there are 2^{d-1} vectors in $\mathbb{Z}_2^d$ having entry 0 and 1, respectively, in their i-th component, for $i \in \{1, \ldots, d\}$, the vectors in $\mathbb{Z}_2^d$ have vanishing sum if and only if $d \geq 2$. In other words, in this case we have $\prod_{z \in I(Z(S))} z = 1 \in [G,G]$.

Hence we may assume that $d = 1$, that is $Z(S)$ has a unique involution, z say, and we have to show that $z \in [G,G]$. Assume that S is abelian, then writing $S = Z(S)$ as a direct product of cyclic groups, we conclude that S is cyclic, a contradiction. Hence S is non-abelian, thus $[S,S] \trianglelefteq S$ is a non-trivial normal subgroup. Thus we have $[S,S] \cap Z(S) \neq \{1\}$, entailing $z \in [S,S] \leq [G,G]$. ♯

b) As for the existence of complete maps, the following is straightforward:

i)⇒ii): Letting $\sigma\colon G \to G$ be a complete map, we consider the cycles of the bijection $G \to G\colon g \mapsto (g^\sigma)^{-1}$. Picking $1 \neq g_1 \in G$, for $i \geq 1$ we successively let $g_{i+1} := (g_i^\sigma)^{-1} \in G$, until we get $g_{s+1} = (g_s^\sigma)^{-1} = g_1$; since $g_1 g_1^\sigma = g_1^\tau \neq 1$ we have $s \geq 2$. Then we get $g_1^\tau \cdots g_s^\tau = g_1 g_1^\sigma \cdot g_2 g_2^\sigma \cdots g_s g_s^\sigma = g_1 \cdot g_1^\sigma g_2 \cdots g_{s-1}^\sigma g_s \cdot g_s^\sigma = g_1 g_s^\sigma = 1$. Hence proceeding like this for all the cycles of the above bijection, we get an ordering $\{g_1, \ldots, g_{|G|}\}$ of the elements of G such that $g_1 \cdots g_{|G|} = 1$.

i)⇒iii): Letting $\sigma\colon G \to G$ be a complete map, we get $\prod_{g \in G} g \equiv \prod_{g \in G} g g^\sigma \equiv \prod_{g \in G} g \cdot \prod_{g \in G} g^\sigma \equiv (\prod_{g \in G} g)^2 \pmod{[G,G]}$, thus $\prod_{g \in G} g \equiv 1 \pmod{[G,G]}$. ♯

Actually, [Paige, 1951] has conjectured that **ii)⇒i)**, and [Hall, Paige, 1955] have conjectured that **iv)⇒i)**, only indicating that ii) implies iii), and that iii) implies iv). The implication **iv)⇒i)** has an involved proof which has been completed only recently (where we are only able to present a very rough sketch):

iv)⇒i): If $|G|$ is odd, then $\sigma := \mathrm{id}_G$ is complete, since the map $G \to G\colon g \mapsto g g^\sigma = g^2$ is a bijection again. Hence we may assume that $|G|$ is even. Recall that, by Burnside's p-complement theorem, any non-abelian simple group has non-cyclic Sylow 2-subgroups.

Firstly [Hall, Paige, 1955] showed that the alternating groups have complete maps. Next [Dalla-Volta, Gavioli, 2001] showed that a minimal counterexample is almost simple or has a center of even order. Then [Wilcox, 2009] showed that a minimal counterexample is actually simple, and that simple groups of Lie type, excluding the Tits group, have complete maps. This reduced the problem, by the classification of finite simple groups, to the sporadic simple groups. Now [Evans, 2009] showed that the Tits group and the sporadic simple groups, excluding the Janko group J_4, have complete maps. Finally [Bray, 2018] showed that J_4 has complete maps. ♯

c) For comparison, we return to the case of abelian groups: Let G be abelian, let $I(G) \,\dot{\cup}\, \{1\} \cong \mathbb{Z}_2^d$ for some $d \in \mathbb{N}_0$, and let $z := \prod_{g \in I(G)} g \in G$. Recalling that G can be written as a direct product of cyclic groups of prime power order, we conclude that $|G|$ is odd if $d = 0$, that G has a non-trivial cyclic Sylow 2-subgroup if $d = 1$, and that G has a non-cyclic Sylow 2-subgroup if $d \geq 2$.

Thus we have $z = 1$ if and only if $d = 0$ or $d \geq 2$. Hence, pairing off the elements of G with their inverses yields $\prod_{g \in G} g = \prod_{g \in I(G)} g = z$, showing **iv)⇒ii)**. Moreover, G has precisely $2^d - 1$ involutions, by (1.7) showing **iv)⇒i)** (up to the unproven pieces there). ♯

2 Information theory

(2.1) Information. Let $\mathcal{X}$ be an alphabet, and let $\mu\colon \mathcal{P}(\mathcal{X}) \to \mathbb{R}_{\geq 0}$ be a probability distribution, that is **i)** $\mu(\mathcal{X}) = 1$, and **ii)** $\mu(\mathcal{A} \cup \mathcal{B}) = \mu(\mathcal{A}) + \mu(\mathcal{B})$ for all $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$ such that $\mathcal{A} \cap \mathcal{B} = \emptyset$.

a) To model the information content of a symbol $x \in \mathcal{X}$, we use the frequency of its occurrence, which is given by μ. Then the information content should be the smaller the more often it occurs. Moreover, for independent events their information contents should add up, while the associated probabilities multiply. Hence letting $\mathcal{S} := \{a \in \mathbb{R}; 0 < a \leq 1\}$, this motivates to let an **information measure** be a strongly decreasing continuous map $\iota\colon \mathcal{S} \to \mathbb{R}$ such that $\iota(ab) = \iota(a) + \iota(b)$ for all $a, b \in \mathcal{S}$. Then the **information content** of a possible elementary event $x \in \mathcal{X}$, that is $\mu(x) > 0$, by abusing notation is given as $\iota(x) := \iota(\mu(x))$.

We show that information measures are unique up to normalization: Given an information measure ι, we consider the continuous map $\eta\colon \mathbb{R}_{\leq 0} \to \mathbb{R}\colon a \mapsto \iota(\exp(a))$, which hence fulfills $\eta(a+b) = \iota(\exp(a+b)) = \iota(\exp(a)\exp(b)) = \iota(\exp(a)) + \iota(\exp(b)) = \eta(a) + \eta(b)$ for all $a, b \in \mathbb{R}_{\leq 0}$. Letting $\alpha := -\eta(-1) \in \mathbb{R}$, we get $\eta(-n) = -\alpha n$ for all $n \in \mathbb{N}_0$, and from that $\eta(-\frac{n}{m}) = -\alpha \cdot \frac{n}{m}$ for all $n \in \mathbb{N}_0$ and $m \in \mathbb{N}$, hence η being continuous we infer $\eta(a) = \alpha a$ for all $a \in \mathbb{R}_{\leq 0}$. Thus from $\iota(\exp(a)) = \eta(a) = \alpha a = \alpha \ln(\exp(a))$, for all $a \in \mathbb{R}_{\leq 0}$, we get $\iota(a) = \alpha \ln(a)$, for all $a \in \mathcal{S}$. Since ι is strongly decreasing we have $\alpha < 0$. Conversely, for any $\alpha < 0$ the map $\mathcal{S} \to \mathbb{R}_{\geq 0}\colon a \mapsto \alpha \ln(a)$ indeed is an information measure. Hence it remains to normalize:

The information content of a binary digit from the alphabet $\mathbb{Z}_2$, carrying the uniform distribution, is set to be 1, hence $1 = \iota(\frac{1}{2}) = \alpha \ln(\frac{1}{2})$, that is $\alpha = -\frac{1}{\ln(2)}$. Thus henceforth we let $\iota(a) = -\frac{\ln(a)}{\ln(2)} = -\log_2(a) = \log_2(\frac{1}{a})$, for all $a \in \mathcal{S}$.

b) The **average information content** or **entropy** of $\mathcal{X} = \{x_1, \ldots, x_q\}$, letting $p_i := \mu_{\mathcal{X}}(x_i) \in \mathbb{R}_{\geq 0}$ for $i \in \{1, \ldots, q\}$, and $\mathcal{I} := \{i \in \{1, \ldots, q\}, p_i > 0\}$, is the expected value $H(\mathcal{X}) = H(\mu) := -\sum_{i \in \mathcal{I}} p_i \log_2(p_i) \in \mathbb{R}_{\geq 0}$ of the information content. Since we have $\lim_{a \to 0^+}(a \log_2(a)) = \lim_{a \to \infty}(\frac{-\log_2(a)}{a}) = 0$, the function $\mathcal{S} \to \mathbb{R}_{\geq 0}\colon a \mapsto -a \log_2(a)$ can be continuously extended to $\mathcal{S} \,\dot{\cup}\, \{0\}$. Thus we may let $H(\mathcal{X}) = -\sum_{i=1}^{q} p_i \log_2(p_i)$, saying that impossible elementary events do not contribute to the average information content.

We have $H(\mathcal{X}) = 0$ if and only if all summands in the defining sum are zero, or equivalently we have either $p_i = 0$ or $\log_2(p_i) = 0$, for all $i \in \{1, \ldots, q\}$, the latter case being equivalent to $p_i = 1$; since $\sum_{i=1}^{q} p_i = 1$ this in turn is equivalent to $p_i = 1$ for a unique $i \in \{1, \ldots, q\}$, and $p_j = 0$ for $j \neq i$, that is μ is concentrated in x_i for some $i \in \{1, \ldots, q\}$.

Moreover, we always have $H(\mathcal{X}) \leq \log_2(|\mathcal{X}|)$, with equality if and only if $\mathcal{X}$ carries the uniform distribution: We make use of the **Jensen inequality** for (strictly) concave functions, which applied to the logarithm function says that for $\lambda_1, \ldots, \lambda_q \in \mathbb{R}_{>0}$ such that $\sum_{i=1}^{q} \lambda_i = 1$, and $\alpha_1, \ldots, \alpha_q \in \mathbb{R}_{>0}$ we have $\sum_{i=1}^{q} \lambda_i \log_2(\alpha_i) \leq \log_2(\sum_{i=1}^{q} \lambda_i \alpha_i)$, where (by strictness) equality occurs if and only if $\alpha_1 = \cdots = \alpha_q$. Thus we get $H(\mathcal{X}) = -\sum_{i \in \mathcal{I}} p_i \log_2(p_i) = \sum_{i \in \mathcal{I}} p_i \log_2(\frac{1}{p_i}) \leq \log_2(\sum_{i \in \mathcal{I}} p_i \cdot \frac{1}{p_i}) = \log_2(\sum_{i \in \mathcal{I}} 1) \leq \log_2(q) = \log_2(|\mathcal{X}|)$, with equality if and only if $p_1 = \cdots = p_q = \frac{1}{q}$. ♯

For example, we consider $\mathcal{X} := \mathbb{Z}_2 = \{0, 1\}$ with elementary probabilities

$\mu_{\mathcal{X}}(0) = p$ and $\mu_{\mathcal{X}}(1) = 1-p$ for some $0 \leq p \leq 1$. Then we get average information content $H(\mu) = H(p) := -p\log_2(p) - (1-p)\log_2(1-p)$. Differentiating yields $\frac{\partial}{\partial p}H(p) = -\log_2(p) + \log_2(1-p) = \log_2(\frac{1-p}{p})$, for all $0 < p < 1$. Since $H(0) = H(1) = 0$ and $H(p) > 0$, for all $0 < p < 1$, we infer that $H(p)$ has a unique maximum for $p = 1-p = \frac{1}{2}$, where $H(\frac{1}{2}) = 1$. Thus indeed the average information content of the symbols in $\mathbb{Z}_2$ is maximized if and only if $\mathbb{Z}_2$ carries the uniform distribution.

The relevance of these notions is elucidated by the **First Main Theorem** of information theory, **Shannon's Theorem on source coding**:

(2.2) Theorem: Shannon [1948]. Let $h\colon \mathcal{X} \to (\mathbb{Z}_2)^* \setminus \{\epsilon\}$ be an injective and **prefix-free** encoding, that is for all $v \in \operatorname{im}(h)$ and $w \in (\mathbb{Z}_2)^* \setminus \{\epsilon\}$ we have $vw \notin \operatorname{im}(h)$. Then for the average length of the code words in $h(\mathcal{X})$ we have

$$\sum_{i=1}^{q} \mu_{\mathcal{X}}(x_i) \cdot l(h(x_i)) \geq H(\mathcal{X}).$$

Proof. i) We first show the **Kraft-McMillan inequality** [1949, 1956]: Letting $l_i := l(h(x_i)) \in \mathbb{N}$, for $i \in \{1,\ldots,q\}$, we have $\sum_{i=1}^{q} 2^{-l_i} \leq 1$.

In order to see this, we may assume that $l_1 \leq \cdots \leq l_q$. Then, for $i \in \{1,\ldots,q\}$, there are $2^{l_q - l_i}$ words in $(\mathbb{Z}_2)^{l_q}$ having $h(x_i) \in (\mathbb{Z}_2)^{l_i}$ as their prefix. Since h is prefix-free we conclude that the latter sets of words are pairwise disjoint. Thus, since there are 2^{l_q} words in $(\mathbb{Z}_2)^{l_q}$, we get $\sum_{i=1}^{q} 2^{l_q - l_i} \leq 2^{l_q}$, thus $\sum_{i=1}^{q} 2^{-l_i} \leq 1$.

ii) Let again $p_i := \mu_{\mathcal{X}}(x_i) \in \mathbb{R}_{\geq 0}$ for $i \in \{1,\ldots,q\}$. Then we show the **Gibbs inequality**: Letting $\alpha_1,\ldots,\alpha_q \in \mathbb{R}_{>0}$ such that $\sum_{i=1}^{q} \alpha_i = 1$, then $H(\mathcal{X}) \leq -\sum_{i=1}^{q} p_i \log_2(\alpha_i)$, with equality if and only if $\alpha_i = p_i$ for all $i \in \{1,\ldots,q\}$. This is seen as follows:

Letting $\mathcal{I} := \{i \in \{1,\ldots,q\}, p_i > 0\}$ again, applying the Jensen inequality we get $\sum_{i\in\mathcal{I}} p_i(\log_2(\alpha_i) - \log_2(p_i)) = \sum_{i\in\mathcal{I}} p_i \log_2(\frac{\alpha_i}{p_i}) \leq \log_2(\sum_{i\in\mathcal{I}} p_i \cdot \frac{\alpha_i}{p_i}) = \log_2(\sum_{i\in\mathcal{I}} \alpha_i) \leq \log_2(1) = 0$, implying $\sum_{i=1}^{q} p_i \log_2(p_i) \geq \sum_{i=1}^{q} p_i \log_2(\alpha_i)$. Moreover, we have equality if and only if $\mathcal{I} = \{1,\ldots,q\}$ and $\frac{\alpha_1}{p_1} = \cdots = \frac{\alpha_q}{p_q} =: \rho$, in which case we have $1 = \sum_{i=1}^{q} \alpha_i = \rho \cdot \sum_{i=1}^{q} p_i = \rho$.

iii) Finally let $\alpha_i := \frac{2^{-l_i}}{\alpha} \in \mathbb{R}_{>0}$ for $i \in \{1,\ldots,q\}$, where $\alpha := \sum_{i=1}^{q} 2^{-l_i} \in \mathbb{R}_{>0}$, so that $\sum_{i=1}^{q} \alpha_i = 1$. By the Kraft-McMillan inequality we have $\alpha \leq 1$, thus $\log_2(\alpha) \leq 0$. Hence the Gibbs inequality yields $H(\mathcal{X}) \leq -\sum_{i=1}^{q} p_i \log_2(\alpha_i) = -\sum_{i=1}^{q} p_i(\log_2(2^{-l_i}) - \log_2(\alpha)) = \log_2(\alpha) + \sum_{i=1}^{q} p_i l_i \leq \sum_{i=1}^{q} p_i l_i$. ♯

The inequality in Shannon's Theorem can be interpreted as follows: We consider the set $(\mathbb{Z}_2)^* \setminus \{\epsilon\}$ of possible code words. For the set $\mathbb{Z}_2^n$ of words of length $n \in \mathbb{N}$, carrying the uniform distribution, which is obtained from the uniform distribution on $\mathbb{Z}_2$ by choosing the symbols in the words independently, we get $\mu(w) = \frac{1}{2^n}$, for all $w \in \mathbb{Z}_2^n$, thus $\iota(w) = -\log_2(\frac{1}{2^n}) = \log_2(2^n) = n = l(w)$; thus summing over $\mathbb{Z}_2^n$ also yields $H(\mathbb{Z}_2^n) = -2^n \cdot \frac{1}{2^n} \cdot \log_2(\frac{1}{2^n}) = \log_2(2^n) = n$.

Hence the left hand side of the inequality is the average information content of the genuine code words $h(\mathcal{X})$, with respect to the uniform distribution on $\mathbb{Z}_2^n$ for all $n \in \mathbb{N}$, and Shannon's Theorem says that this cannot possibly be strictly smaller that the average information content of the original alphabet $\mathcal{X}$.

The lower bound given is best possible, where we may assume that $H(\mathcal{X}) > 0$: For the **Huffman encoding**, see Exercises (14.9) and (14.10), the average length of code words is bounded above by $H(\mathcal{X}) + 1$, so that the lower bound is attained up to a factor of $1 + \frac{1}{H(\mathcal{X})}$. Now, replacing $\mathcal{X}$ by $\mathcal{X}^n$, where the symbols in words are chosen independently, so that we have $H(\mathcal{X}^n) = n \cdot H(\mathcal{X})$ as will be shown in (2.4) below, we get a factor of $1 + \frac{1}{n \cdot H(\mathcal{X})} \to 1$, for $n \to \infty$.

(2.3) Noisy channels. a) The standard model for this is the **binary channel**: The data consists of elements of $\mathcal{X} := \mathbb{Z}_2$, sent with probability distribution $\mu_{\mathcal{X}}$, and being distorted by the channel, so that the received elements of $\mathcal{Y} := \mathbb{Z}_2$ carry the probability distribution $\mu_{\mathcal{Y}}$, where the noise is described by the conditional distribution $\mu_{\mathcal{Y}|\mathcal{X}}$, thus $\mu_{\mathcal{Y}}(j) = \sum_{i \in \mathcal{X}} \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}|i}(j)$, for $j \in \mathcal{Y}$.

The **symmetric** binary channel with error probability $0 \leq p \leq \frac{1}{2}$ is given by $\mu_{\mathcal{Y}|i}(i+1) = p$ and $\mu_{\mathcal{Y}|i}(i) = 1-p$, for $i \in \mathcal{X}$. In other words, $\mu_{\mathcal{Y}|\mathcal{X}}$ is given by the transition matrix $M(p) := \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$, entailing $\mu_{\mathcal{Y}} = [\mu_{\mathcal{Y}}(0), \mu_{\mathcal{Y}}(1)] = [\mu_{\mathcal{X}}(0), \mu_{\mathcal{X}}(1)] \cdot M(p) = [\mu_{\mathcal{X}}(0)(1-p) + \mu_{\mathcal{X}}(1)p, \mu_{\mathcal{X}}(0)p + \mu_{\mathcal{X}}(1)(1-p)]$.

In particular, the **quiet** binary channel is given by $p = 0$, that is $\mu_{\mathcal{Y}|i}(j) = \delta_{ij}$, hence we have $\mu_{\mathcal{Y}} = \mu_{\mathcal{X}}$; and the **completely noisy** binary channel is given by $p = \frac{1}{2}$, that is $\mu_{\mathcal{Y}|i}(j) = \frac{1}{2}$ for $i \in \mathcal{X}$ and $j \in \mathcal{Y}$, hence $\mu_{\mathcal{Y}}$ is the uniform distribution, independently of $\mu_{\mathcal{X}}$.

If $\mathcal{X}$ carries the uniform distribution, then we have $\mu_{\mathcal{Y}} = [\frac{1}{2}, \frac{1}{2}] \cdot M(p) = [\frac{1}{2}, \frac{1}{2}]$, saying that $\mathcal{Y}$ carries the uniform distribution as well, independently of p. Conversely, if $0 \leq p < \frac{1}{2}$, then we have $M(p)^{-1} = \frac{1}{1-2p} \cdot \begin{bmatrix} 1-p & -p \\ -p & 1-p \end{bmatrix} \in \mathrm{GL}_2(\mathbb{R})$, hence if $\mathcal{Y}$ carries the uniform distribution, then $\mu_{\mathcal{X}} = [\frac{1}{2}, \frac{1}{2}] \cdot M(p)^{-1} = [\frac{1}{2}, \frac{1}{2}]$, that is $\mathcal{X}$ necessarily carries the uniform distribution as well.

b) For decoding purposes we provide the transition matrix $\widetilde{M}(p)$ describing the conditional probability $\mu_{\mathcal{X}|\mathcal{Y}}$ as well: Bayes's Theorem says that $\mu_{\mathcal{X}|j}(i)\mu_{\mathcal{Y}}(j) = \mu_{\mathcal{X}\times\mathcal{Y}}(i,j) = \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}|i}(j)$, for $i \in \mathcal{X}$ and $j \in \mathcal{Y}$. Hence we have $\widetilde{M}(p)^{\mathrm{tr}} \cdot \mathrm{diag}[\mu_{\mathcal{Y}}(j); j \in \mathcal{Y}] = \mathrm{diag}[\mu_{\mathcal{X}}(i); i \in \mathcal{X}] \cdot M(p)$. Assuming that $\mu_{\mathcal{Y}}(j) \neq 0$ for all $j \in \mathcal{Y}$, this yields

$$\widetilde{M}(p) = \begin{bmatrix} \frac{\mu_{\mathcal{X}}(0)(1-p)}{\mu_{\mathcal{X}}(0)(1-p)+\mu_{\mathcal{X}}(1)p} & \frac{\mu_{\mathcal{X}}(1)p}{\mu_{\mathcal{X}}(0)(1-p)+\mu_{\mathcal{X}}(1)p} \\ \frac{\mu_{\mathcal{X}}(0)p}{\mu_{\mathcal{X}}(0)p+\mu_{\mathcal{X}}(1)(1-p)} & \frac{\mu_{\mathcal{X}}(1)(1-p)}{\mu_{\mathcal{X}}(0)p+\mu_{\mathcal{X}}(1)(1-p)} \end{bmatrix}.$$

In particular, if $\mathcal{X}$ carries the uniform distribution, thus $\mathcal{Y}$ carrying the uniform distribution as well, then we have $\widetilde{M}(p) = M(p)$.

If $\mu_{\mathcal{Y}}(j) = 0$ for some $j \in \mathcal{Y}$, then we have $p = 0$ and $\mu_{\mathcal{X}}(j) = \mu_{\mathcal{Y}}(j) = 0$, entailing that $\mu_{\mathcal{X}|k}(j) = 0$ for all $k \in \mathcal{Y}$. Hence, if $\mu_{\mathcal{Y}}(0) = 0$ or $\mu_{\mathcal{Y}}(1) = 0$, then $\widetilde{M}(0) = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ and $\widetilde{M}(0) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, respectively, thus $\widetilde{M}(0) = \lim_{p \to 0^+} \widetilde{M}(p)$.

(2.4) Capacity. a) We still consider a noisy channel working over an alphabet $\mathcal{X} = \mathcal{Y}$, with associated probability distributions $\mu_{\mathcal{X}}$ and $\mu_{\mathcal{Y}}$, respectively.

Given an elementary event $j \in \mathcal{Y}$, the conditional distribution $\mu_{\mathcal{X}|j}$ describes the probability distribution on the sent symbols in $\mathcal{X}$ provided j is received. Then for $\mu_{\mathcal{X}|j}$ we get $H(\mathcal{X}|j) = -\sum_{i \in \mathcal{X}} \mu_{\mathcal{X}|j}(i) \log_2(\mu_{\mathcal{X}|j}(i))$, describing the average information content of $\mathcal{X}$ which upon seeing $j \in \mathcal{Y}$ is lost due to noise. Hence the **average conditional information content** or **conditional entropy** $H(\mathcal{X}|\mathcal{Y}) = \sum_{j \in \mathcal{Y}} \mu_{\mathcal{Y}}(j) H(\mathcal{X}|j) = -\sum_{j \in \mathcal{Y}} \sum_{i \in \mathcal{X}} \mu_{\mathcal{Y}}(j) \mu_{\mathcal{X}|j}(i) \log_2(\mu_{\mathcal{X}|j}(i))$ is the average information content of $\mathcal{X}$ which is lost by transport through the channel. Thus the **capacity** $C(\mathcal{X}|\mathcal{Y}) := H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y})$ is the average information content being transported through the channel.

b) We proceed to show that **i)** $H(\mathcal{X} \times \mathcal{Y}) \leq H(\mathcal{X}) + H(\mathcal{Y})$, with equality if and only if $\mathcal{X}$ and $\mathcal{Y}$ are independent, and **ii)** $H(\mathcal{X} \times \mathcal{Y}) = H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y})$.

Thus we have $C(\mathcal{X}|\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X} \times \mathcal{Y}) \geq 0$ indeed, with equality if and only if $\mathcal{X}$ and $\mathcal{Y}$ are independent. Moreover, we have $C(\mathcal{X}|\mathcal{Y}) = C(\mathcal{Y}|\mathcal{X})$, saying that the capacity of the channel is independent of the direction of information transport. It remains to show (i) and (ii):

i) We may assume that $\mu_{\mathcal{X} \times \mathcal{Y})}(i,j) \neq 0$, and thus $\mu_{\mathcal{X}}(i) \neq 0 \neq \mu_{\mathcal{Y}}(j)$, for all $i \in \mathcal{X}$ and $j \in \mathcal{Y}$. We have $H(\mathcal{X} \times \mathcal{Y}) = -\sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X} \times \mathcal{Y}}(i,j) \log_2(\mu_{\mathcal{X} \times \mathcal{Y}}(i,j))$ and $H(\mathcal{X}) + H(\mathcal{Y}) = -\sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X} \times \mathcal{Y}}(i,j)(\log_2(\mu_{\mathcal{X}}(i)) + \log_2(\mu_{\mathcal{Y}}(j)))$, from which we get

$$H(\mathcal{X} \times \mathcal{Y}) - H(\mathcal{X}) - H(\mathcal{Y}) = \sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X} \times \mathcal{Y}}(i,j) \log_2(\frac{\mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j)}{\mu_{\mathcal{X} \times \mathcal{Y}}(i,j)}),$$

which by Jensen's inequality entails

$$H(\mathcal{X} \times \mathcal{Y}) - H(\mathcal{X}) - H(\mathcal{Y}) \leq \log_2(\sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X} \times \mathcal{Y}}(i,j) \cdot \frac{\mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j)}{\mu_{\mathcal{X} \times \mathcal{Y}}(i,j)}),$$

where the double sum on the right hand side equals $\sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j) = (\sum_{i \in \mathcal{X}} \mu_{\mathcal{X}}(i)) \cdot (\sum_{j \in \mathcal{Y}} \mu_{\mathcal{Y}}(j)) = 1$, thus $H(\mathcal{X} \times \mathcal{Y}) - H(\mathcal{X}) - H(\mathcal{Y}) \leq 0$. Moreover, we have equality if and only if there is $m \in \mathbb{R}$ such that $\frac{\mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j)}{\mu_{\mathcal{X} \times \mathcal{Y}}(i,j)}) = m$ for all $i \in \mathcal{X}$ and $j \in \mathcal{Y}$, in which case we get $1 = \sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j) = m \cdot \sum_{i \in \mathcal{X}} \sum_{j \in \mathcal{Y}} \mu_{\mathcal{X} \times \mathcal{Y}}(i,j) = 1$, saying that $\mu_{\mathcal{X} \times \mathcal{Y}}(i,j) = \mu_{\mathcal{X}}(i)\mu_{\mathcal{Y}}(j)$.

ii) Using Bayes's Theorem, saying that $\mu_{\mathcal{X}|j}(i)\mu_{\mathcal{Y}}(j) = \mu_{\mathcal{X}\times\mathcal{Y}}(i,j)$, we get

$$\begin{array}{rcl} H(\mathcal{X}\times\mathcal{Y}) & = & -\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{X}\times\mathcal{Y}}(i,j)\log_2(\mu_{\mathcal{X}\times\mathcal{Y}}(i,j)) \\ & = & -\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{Y}}(j)\mu_{\mathcal{X}|j}(i)(\log_2(\mu_{\mathcal{Y}}(j)) + \log_2(\mu_{\mathcal{X}|j}(i))) \\ & = & -\sum_{i\in\mathcal{X}}\sum_{j\in\mathcal{Y}} \mu_{\mathcal{Y}}(j)\mu_{\mathcal{X}|j}(i)\log_2(\mu_{\mathcal{X}|j}(i)) \\ & & -\sum_{i\in\mathcal{X}} \mu_{\mathcal{X}|j}(i)\cdot\sum_{j\in\mathcal{Y}} \mu_{\mathcal{Y}}(j)\log_2(\mu_{\mathcal{Y}}(j) \\ & = & H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y}). \end{array}$$

c) For the symmetric binary channel with error probability $0 \leq p \leq \frac{1}{2}$, working over the alphabet $\mathcal{X} = \mathbb{Z}_2 = \mathcal{Y}$, using the transition matrix $M(p)$ describing $\mu_{\mathcal{Y}|\mathcal{X}}$, we obtain $H(\mathcal{Y}|\mathcal{X}) = -p\log_2(p) - (1-p)\log_2(1-p)$, thus the capacity is $C(p) = H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X}) = H(\mathcal{Y}) + p\log_2(p) + (1-p)\log_2(1-p)$. In particular, for the quiet channel we have $\mu_{\mathcal{Y}} = \mu_{\mathcal{X}}$ and $C(0) = H(\mathcal{Y}) = H(\mathcal{X})$, saying that the average information content of $\mathcal{X}$ is completely transported through the channel; and for the completely noisy channel the alphabet $\mathcal{Y}$ carries the uniform distribution in any case, hence we have $H(\mathcal{Y}) = 1$ and $C(\frac{1}{2}) = 1 + \log_2(\frac{1}{2}) = 0$, saying that no information is transported through the channel.

In general, we have $0 \leq H(\mathcal{Y}) \leq 1$, where the maximum $H(\mathcal{Y}) = 1$ is attained precisely for the uniform distribution on $\mathcal{Y}$. Hence the maximum capacity is given as $C_{\max}(p) = 1 + p\log_2(p) + (1-p)\log_2(1-p)$. Moreover, if $0 \leq p < \frac{1}{2}$ this is equivalent to $\mathcal{X}$ carrying the uniform distribution, in other words if and only if $H(\mathcal{X}) = 1$, saying that the maximum capacity of the channel is attained if and only if $\mathcal{X}$ has maximum average information content.

(2.5) Redundancy. The idea of error correcting **channel coding**, to be used for noisy channels, is to add redundancy. This is measured as follows:

Letting $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, any subset $\emptyset \neq \mathcal{C} \subseteq \mathcal{X}^n$ is called **block code** of **length** $n \in \mathbb{N}$ and **order** $m := |\mathcal{C}| \in \mathbb{N}$. Assuming the uniform distribution on $\mathcal{X}^n$, the relative information content of a word in $\mathcal{C}$, as compared to viewing it as an element of $\mathcal{X}^n$, is given as the **information rate** $\rho(\mathcal{C}) = \rho_{\mathcal{X}^n}(\mathcal{C}) := \frac{\iota(|\mathcal{C}|)}{\iota(|\mathcal{X}^n|)} = \frac{-\log_2(|\mathcal{C}|)}{-\log_2(|\mathcal{X}^n|)} = \frac{\log_q(|\mathcal{C}|)}{\log_q(|\mathcal{X}^n|)} = \frac{\log_q(m)}{n}$. We have $0 \leq \rho(\mathcal{C}) \leq 1$, where $\rho(\mathcal{C}) = 1$ if and only if $\mathcal{C} = \mathcal{X}^n$, that is no redundancy is added. Thus the larger $\rho(\mathcal{C})$ the better $\mathcal{C}$ is, in terms of information transport.

For example, in parity check codes, words over $\mathcal{X}$ consisting of $k \in \mathbb{N}_0$ **information symbols** are encoded into words of length $n \in \mathbb{N}$ by adding $n-k$ **check symbols**; thus allowing for $\mathcal{X}^k$ to start with we obtain $\rho(\mathcal{X}^k) = \frac{\log_q(|\mathcal{X}^k|)}{n} = \frac{k}{n}$. Similarly, if $\mathcal{X} = \mathbb{F}_q$ is the field with q elements, and $\mathcal{C} \leq \mathbb{F}_q^n$ is an $\mathbb{F}_q$-subspace, then we have $\rho(\mathcal{C}) = \frac{\log_q(|\mathcal{C}|)}{n} = \frac{\dim_{\mathbb{F}_q}(\mathcal{C})}{n}$.

(2.6) Maximum likelihood decoding. a) If words are sent through a noisy channel, they are susceptible to random errors, where we assume that errors occurring in distinct positions in a word are independent of each other, that is we have a **channel without memory**. Hence if a word is received, the question arises how to decode it again: We again consider the binary symmetric channel with error probability $0 \leq p < \frac{1}{2}$, working over the alphabet $\mathcal{X} = \mathbb{Z}_2 = \mathcal{Y}$. We assume that $\mathcal{X}$ carries the uniform distribution, or equivalently that $\mathcal{Y}$ carries the uniform distribution, so that the transition matrix describing $\mu_{\mathcal{X}|\mathcal{Y}}$ is $M(p)$.

Let $\emptyset \neq \mathcal{C} \subseteq \mathcal{X}^n$ be a block code of length $n \in \mathbb{N}$. If the word $v \in \mathcal{Y}^n$ is received, then it is decoded to some $c \in \mathcal{C}$ which has maximum probability to be the word sent, that is $\mu_{\mathcal{X}^n|v}(c) = \max\{\mu_{\mathcal{X}^n|v}(w) \in \mathbb{R}; w \in \mathcal{C}\}$, hence this is called **maximum likelihood (ML) decoding**. This is turned into combinatorics:

For $x = [x_1, \ldots, x_n] \in \mathcal{X}^n$ and $y = [y_1, \ldots, y_n] \in \mathcal{X}^n$ we let $d(x, y) := |\{i \in \{1, \ldots, n\}; x_i \neq y_i\}| \in \{0, \ldots, n\}$ be their **Hamming distance**. Since distinct positions are considered to be independent, for all $w \in \mathcal{X}^n$ we have $\mu_{\mathcal{X}^n|v}(w) = p^{d(v,w)}(1-p)^{n-d(v,w)} = (1-p)^n(\frac{p}{1-p})^{d(v,w)}$. If $0 < p < \frac{1}{2}$, then from $0 < \frac{p}{1-p} < 1$ we infer that the function $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{>0} \colon a \mapsto (\frac{p}{1-p})^a$ is strictly decreasing; if $p = 0$, then we have $\mu_{\mathcal{X}^n|v}(w) = \delta_{v,w}$. Thus we choose $c \in \mathcal{C}$ having minimum distance to $v \in \mathcal{Y}^n$, that is $d(v, c) = \min\{d(v, w) \in \mathbb{N}_0; w \in \mathcal{C}\}$, being called **nearest neighbor decoding**. In practice, although **complete decoding** is desired, if this does not determine c uniquely, we revert to **partial decoding** by **unique nearest neighbor decoding**, and mark v as an **erasure**.

b) If $c \in \mathcal{C}$ is sent, let $0 \leq \gamma_c \leq 1$ be the probability that c is not recovered, and the expected value $\gamma(\mathcal{C}) := \frac{1}{|\mathcal{C}|} \cdot \sum_{c \in \mathcal{C}} \gamma_c$ is called the **average error probability** of $\mathcal{C}$; thus the smaller $\gamma(\mathcal{C})$ the better $\mathcal{C}$ is, as far as erroneous decoding is concerned. The question whether there are good codes, in the sense of having a large information rate and a small average error probability at the same time, is generally answered by the **Second Main Theorem** of information theory, **Shannon's Theorem on channel coding**, which we proceed to prove. But note that the proof is a pure existence proof giving no clue at all how to actually find good codes. We need a lemma first:

(2.7) Lemma: Chernoff inequality. Let $\mathcal{X}$ be an alphabet with probability distribution μ, and let $X_1, \ldots, X_n$, for $n \in \mathbb{N}$, be independent random variables with values in $\{0, 1\}$, such that $\mu(X_i = 1) = p$, for $0 \leq p \leq 1$. Then $X := \sum_{i=1}^n X_i$ is binomially distributed, that is we have $\mu(X = d) = \binom{n}{d} \cdot p^d(1-p)^{n-d}$, for $d \in \{0, \ldots, n\}$, and for $0 \leq \epsilon \leq 1$ we have $\mu(X \geq (1+\epsilon)pn) \leq e^{-\frac{1}{2}\epsilon^2 pn}$.

Proof. The first assertion is a matter of counting. Next, for $t > 0$ we have the special case $t \cdot \mu(X \geq t) = \sum_{x \in \mathcal{X}} \mu(x) t \delta_{X(x) \geq t} \leq \sum_{x \in \mathcal{X}} \mu(x) X(x) \delta_{X(x) \geq t} \leq \sum_{x \in \mathcal{X}} \mu(x) X(x) = E(X)$ of the **Markov inequality**; note that here we only use that X has non-negative values. Now, for $t \in \mathbb{R}$ and $i \in \{1, \ldots, n\}$ we have

$$E(\exp(tX_i)) = \mu(X_i = 0) \cdot \exp(0) + \mu(X_i = 1) \cdot \exp(t) = 1 + p(\exp(t) - 1),$$

hence we obtain $E(\exp(tX)) = E(\exp(t \cdot \sum_{i=1}^n X_i)) = E(\prod_{i=1}^n \exp(tX_i)) = \prod_{i=1}^n E(\exp(tX_i)) = (1 + p(\exp(t) - 1))^n$, which using the convexity of the exponential function yields $E(\exp(tX)) \leq \exp((\exp(t)-1)pn)$. Thus we get

$$\begin{array}{rcl} \mu(X \geq (1+\epsilon)pn) & = & \mu(\exp(tX) \geq \exp(t(1+\epsilon)pn)) \\ & \leq & \exp(-t(1+\epsilon)pn) \cdot E(\exp(tX)) \\ & = & \exp(-t(1+\epsilon)pn + (\exp(t)-1)pn). \end{array}$$

Letting $t := \ln(1+\epsilon)$ this yields $\mu(X \geq (1+\epsilon)pn) \leq \exp(pn(\epsilon-(1+\epsilon)\ln(1+\epsilon)))$. Finally, the Taylor expansion $(1+\epsilon)\ln(1+\epsilon) = \epsilon + \frac{1}{2}\epsilon^2 - \frac{1}{6}\epsilon^3 + \cdots < \epsilon + \frac{1}{2}\epsilon^2$, for $0 \leq \epsilon \leq 1$, entails $\mu(X \geq (1+\epsilon)pn) \leq \exp(-\frac{1}{2}\epsilon^2 pn)$ as asserted. $\sharp$

(2.8) Theorem: Shannon [1948]. We still consider the symmetric binary channel over the alphabet $\mathcal{X} = \mathbb{Z}_2 = \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ carry the uniform distribution, with error probability $0 \leq p < \frac{1}{2}$. Then for any $0 < \rho < 1 + p\log_2(p) + (1-p)\log_2(1-p) = C_{\max}(p)$ and $\epsilon > 0$ there is a code $\emptyset \neq \mathcal{C} \subseteq \mathcal{X}^n$, for some $n \in \mathbb{N}$, such that $\rho(\mathcal{C}) \geq \rho$ and $\gamma(\mathcal{C}) < \epsilon$.

Proof. If $p = 0$, then we may take $\mathcal{C} = \mathcal{X}$, fulfilling $\rho(\mathcal{X}) = 1$ and $\gamma(\mathcal{X}) = 0$, hence we may assume that $p > 0$. For $n \in \mathbb{N}$ let $m := 2^{\lceil \rho n \rceil} \in \mathbb{N}$ and $\Gamma_n := \{\mathcal{C} \subseteq \mathcal{X}^n; |\mathcal{C}| = m\}$; note that $|\Gamma_n| = \binom{2^n}{m} > 0$ for $n \gg 0$. Hence for $\mathcal{C} \in \Gamma_n$ we have $\rho(\mathcal{C}) = \frac{\log_2(m)}{n} = \frac{\lceil \rho n \rceil}{n} \geq \rho$. Now let $\mathcal{C}_0 \in \Gamma_n$ such that $\gamma(\mathcal{C}_0)$ is minimal. Hence $\gamma(\mathcal{C}_0)$ is bounded above by the expected value $E_{\Gamma_n}(\gamma(\mathcal{C}))$, subject to codes $\mathcal{C} \in \Gamma_n$ being chosen according to the uniform distribution on Γ_n.

If some word in $\mathcal{X}^n$ is sent, then the probability that the received word contains precisely $d \in \{0, \ldots, n\}$ errors is given by the binomial distribution $\beta(d) = \binom{n}{d} \cdot p^d(1-p)^{n-d}$. Thus the Chernoff inequality yields $\mu(\beta \geq (1+a)np) \leq \exp(-\frac{1}{2}a^2np)$, for any $0 \leq a \leq 1$. For $n \in \mathbb{N}$ let $0 < a_n < 1$ such that $a_n \to 0$ and $na_n^2 \to \infty$, for $n \to \infty$; for example we may let $a_n \sim \frac{1}{\ln(n)}$. Letting $\delta = \delta_n := \lfloor (1+a_n)np \rfloor$ yields $\mu(\beta > \delta) \leq \exp(-\frac{1}{2}a_n^2 np) < \epsilon$, for $n \gg 0$.

Moreover, we have $\frac{\delta}{n} \to p < \frac{1}{2}$, for $n \to \infty$; hence we have $\delta < \frac{n}{2} - 1$, for $n \gg 0$. The equation $n^n = (d + (n-d))^n \geq \binom{n}{d} d^d (n-d)^{n-d}$ yields $\binom{n}{d} \leq \frac{n^n}{d^d(n-d)^{n-d}}$, for $d \in \{0, \ldots, n\}$. Letting $\mathcal{B}_\delta(v) := \{w \in \mathcal{X}^n; d(v,w) \leq \delta\}$ be the **sphere** or **ball** with **radius** δ around $v \in \mathcal{X}^n$, the unimodularity of binomial coefficients yields, for $n \gg 0$,

$$b = b_\delta := |\mathcal{B}_\delta(v)| = \sum_{d=0}^{\delta} \binom{n}{d} < \frac{n}{2} \cdot \binom{n}{\delta} \leq \frac{n^{n+1}}{2\delta^\delta (n-\delta)^{n-\delta}} = \frac{n}{2(\frac{\delta}{n})^\delta (1-\frac{\delta}{n})^{n-\delta}}.$$

If $\mathcal{C} \in \Gamma_n$, then we decode by unique nearest neighbor decoding with respect to δ, that is, given $v \in \mathcal{Y}^n$, if $\mathcal{C} \cap \mathcal{B}_\delta(v) = \{c\}$ then we decode v to c, otherwise we mark v as an erasure. For $c \in \mathcal{C}$ let $\chi_c \colon \mathcal{Y}^n \to \{0,1\}$ be defined by $\chi_c(v) := 1$ if $d(v,c) \leq \delta$, and $\chi_c(v) := 0$ if $d(v,c) > \delta$. Let $\varphi_c \colon \mathcal{Y}^n \to \mathbb{N}_0$ be defined by

$$\varphi_c(v) := 1 - \chi_c(v) + \sum_{c' \in \mathcal{C} \setminus \{c\}} \chi_{c'}(v) = \left\{ \begin{array}{cl} |\mathcal{C} \cap \mathcal{B}_\delta(v)| + 1, & \text{if } d(v,c) > \delta, \\ |(\mathcal{C} \setminus \{c\}) \cap \mathcal{B}_\delta(v)|, & \text{if } d(v,c) \leq \delta; \end{array} \right.$$

thus in particular we have $\varphi_c(v) = 0$ if and only if $\mathcal{C} \cap \mathcal{B}_\delta(v) = \{c\}$.

Hence for $c \in \mathcal{C}$ we have $\gamma_c \leq \sum_{v \in \mathcal{Y}^n} \mu_{\mathcal{Y}^n|c}(v)\varphi_c(v)$, where moreover we have $\sum_{v \in \mathcal{Y}^n} \mu_{\mathcal{Y}^n|c}(v)(1 - \chi_c(v)) = \sum_{v \in \mathcal{Y}^n \setminus \mathcal{B}_\delta(c)} \mu_{\mathcal{Y}^n|c}(v) = \mu(\beta > \delta) < \epsilon$, for $n \gg 0$. Thus we get

$$\gamma(\mathcal{C}) = \frac{1}{m} \cdot \sum_{c \in \mathcal{C}} \gamma_c < \epsilon + \frac{1}{m} \cdot \sum_{v \in \mathcal{Y}^n} \sum_{[c,c'] \in \mathcal{C}^2,\, c \neq c'} \mu_{\mathcal{Y}^n|c}(v)\chi_{c'}(v).$$

Hence averaging over all $\binom{2^n}{m}$ many m-subsets of Γ_n, distinct code words being chosen uniformly and independently, since any 2-subset of $\mathcal{X}^n$ is contained in precisely $\binom{2^n-2}{m-2}$ of its m-subsets, we get

$$E_{\Gamma_n}(\gamma(\mathcal{C})) < \epsilon + \frac{1}{m} \cdot \frac{m(m-1)}{2^n(2^n-1)} \cdot \sum_{v \in \mathcal{Y}^n} \sum_{[c,c'] \in (\mathcal{X}^n)^2,\, c \neq c'} \mu_{\mathcal{Y}^n|c}(v)\chi_{c'}(v).$$

For all $v \in \mathcal{Y}^n$ fixed we have $\sum_{c \in \mathcal{X}^n} \chi_c(v) = b$ as well as $\sum_{c \in \mathcal{X}^n} \mu_{\mathcal{Y}^n|c}(v) = \sum_{c \in \mathcal{X}^n} p^{d(v,c)}(1-p)^{n-d(v,c)} = \sum_{d=0}^{n} \binom{n}{d} p^d (1-p)^{n-d} = 1$. Thus we get

$$\gamma(\mathcal{C}_0) \leq E_{\Gamma_n}(\gamma(\mathcal{C})) < \epsilon + \frac{(m-1)b}{2^n - 1} < \epsilon + \frac{mb}{2^n} < \epsilon + \frac{2^{\lceil \rho n \rceil - n - 1} \cdot n}{(\frac{\delta}{n})^\delta (1 - \frac{\delta}{n})^{n-\delta}}.$$

This entails $\frac{\log_2(\gamma(\mathcal{C}_0) - \epsilon)}{n} < \frac{\lceil \rho n \rceil - n - 1 + \log_2(n)}{n} - \frac{\delta}{n}\log_2(\frac{\delta}{n}) - (1 - \frac{\delta}{n})\log_2(1 - \frac{\delta}{n}) \to \rho - 1 - p\log_2(p) - (1-p)\log_2(1-p) < 0$, for $n \to \infty$. Hence there is $\alpha > 0$ such that $\gamma(\mathcal{C}_0) - \epsilon < 2^{-n\alpha}$, for $n \gg 0$. $\sharp$

II

3 Block codes

(3.1) Hamming distance. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, and let $n \in \mathbb{N}_0$. Letting $v = [x_1, \ldots, x_n] \in \mathcal{X}^n$ and $w = [y_1, \ldots, y_n] \in \mathcal{X}^n$, then $d(v, w) := |\{i \in \{1, \ldots, n\}; x_i \neq y_i\}| \in \{0, \ldots, n\}$ is called their **Hamming distance**; recall that we have already used this for $q = 2$ in (2.6).

The Hamming distance defines a discrete metric on $\mathcal{X}^n$: We have positive definiteness $d(v, w) \in \mathbb{R}_{\geq 0}$, where $d(v, w) = 0$ if and only if $v = w$; we have symmetry $d(v, w) = d(w, v)$; and the triangle inequality holds: Letting $u := [z_1, \ldots, z_n] \in \mathcal{X}^n$, from $\{i \in \{1, \ldots, n\}; x_i \neq z_i\} = \{i \in \{1, \ldots, n\}; y_i = x_i \neq z_i\} \,\dot\cup\, \{i \in \{1, \ldots, n\}; y_i \neq x_i \neq z_i\} \subseteq \{i \in \{1, \ldots, n\}; y_i \neq z_i\} \cup \{i \in \{1, \ldots, n\}; x_i \neq y_i\}$ we get $d(v, u) \leq d(v, w) + d(w, u)$.

An **isometry** of $\mathcal{X}^n$ is a map $\varphi \colon \mathcal{X}^n \to \mathcal{X}^n$ such that $d(v, w) = d(v^\varphi, w^\varphi)$, for all $v, w \in \mathcal{X}^n$; it follows from positive definiteness that any isometry is injective, hence is bijective. Thus the set $I(\mathcal{X}^n)$ of all isometries of $\mathcal{X}^n$ forms a group, called the **isometry group** of $\mathcal{X}^n$. We determine $I(\mathcal{X}^n)$:

Given permutations $\pi_i \in \mathcal{S}_q$, for all $i \in \{1, \ldots, n\}$, this yields an isometry $[\pi_1, \ldots, \pi_n]$ acting component-wise, and any permutation in $\mathcal{S}_n$ induces an isometry by permuting the components. Hence $G := I(\mathcal{X}^n)$ contains a subgroup isomorphic to the semidirect product $\mathcal{S}_{\mathcal{X}}^n \rtimes \mathcal{S}_n$; thus G acts transitively on $\mathcal{X}^n$.

Let $0 \in \mathcal{X}$ be a fixed element, and let $H := \mathrm{Stab}_G([0, \ldots, 0])$; hence we have $[G \colon H] = q^n$. For $v = [x_1, \ldots, x_n] \in \mathcal{X}^n$ let $\mathrm{supp}(v) := \{i \in \{1, \ldots, n\}; x_i \neq 0\}$ be the **support** of v. Hence we conclude that H acts transitively on the various sets $\{v \in \mathcal{X}^n; |\mathrm{supp}(v)| = s\}$, for $s \in \{0, \ldots, n\}$. Now let $h \in H$.

Let first $v \neq w \in \mathcal{X}^n$ such that $|\mathrm{supp}(v)| = |\mathrm{supp}(w)| = 1$. If $\mathrm{supp}(v) = \mathrm{supp}(w)$ then from $d(v^h, w^h) = d(v, w) = 1$ we conclude that $\mathrm{supp}(v^h) = \mathrm{supp}(w^h)$, while if $\mathrm{supp}(v) \neq \mathrm{supp}(w)$ then from $d(v^h, w^h) = d(v, w) = 2$ we conclude that $\mathrm{supp}(v^h) \neq \mathrm{supp}(w^h)$. Hence h induces a permutation of the components of $\mathcal{X}^n$, and permutations within any of the components. In other words the action of h on $\{v \in \mathcal{X}^n; |\mathrm{supp}(v)| = 1\}$ is induced by an element of $\mathcal{S}_{\mathcal{X} \setminus \{0\}}^n \rtimes \mathcal{S}_n$.

Now let $v = [x_1, \ldots, x_n] \in \mathcal{X}^n$ such that $\mathrm{supp}(v) = \{i_1, \ldots, i_s\}$, for some $s \in \{1, \ldots, n\}$, and let $w_j := [0, \ldots, 0, x_{i_j}, 0, \ldots, 0]$, for $j \in \{1, \ldots, s\}$. Since $d(v^h, w_j^h) = d(v, w_j) = s - 1$, we infer that the non-zero components of v^h are determined by the non-zero components of the various w_j^h. Thus the action of h on $\mathcal{X}^n$ is induced by the element of $\mathcal{S}_{\mathcal{X} \setminus \{0\}}^n \rtimes \mathcal{S}_n$ describing its action on $\{v \in \mathcal{X}^n; |\mathrm{supp}(v)| = 1\}$, entailing that H is isomorphic to a subgroup of $\mathcal{S}_{\mathcal{X} \setminus \{0\}}^n \rtimes \mathcal{S}_n$. Hence from $H \leq \mathcal{S}_{\mathcal{X} \setminus \{0\}}^n \rtimes \mathcal{S}_n \leq \mathcal{S}_{\mathcal{X}}^n \rtimes \mathcal{S}_n \leq G$ and $[\mathcal{S}_{\mathcal{X}} \colon \mathcal{S}_{\mathcal{X} \setminus \{0\}}] = q$ we conclude that $G \cong \mathcal{S}_{\mathcal{X}}^n \rtimes \mathcal{S}_n$ and $H \cong \mathcal{S}_{\mathcal{X} \setminus \{0\}}^n \rtimes \mathcal{S}_n$. ♯

b) Let $\mathcal{X} = \mathbb{F}_q$ be the field with q elements, let $0_n := [0, \ldots, 0] \in \mathbb{F}_q^n$ and let $1_n := [1, \ldots, 1] \in \mathbb{F}_q^n$. For $v = [x_1, \ldots, x_n] \in \mathbb{F}_q^n$ let $\mathrm{wt}(v) := d(v, 0_n) \in \{0, \ldots, n\}$ be the **Hamming weight** of v, and let $\mathrm{supp}(v) := \{i \in \{1, \ldots, n\}; x_i \neq 0\}$ be the **support** of v; hence we have $\mathrm{wt}(v) = |\mathrm{supp}(v)|$.

An $\mathbb{F}_q$-linear isometry of $\mathbb{F}_q^n$ is called a **linear isometry**, the group $I_n(\mathbb{F}_q) \leq \mathrm{GL}_n(\mathbb{F}_q)$ of all linear isometries is called the **linear isometry group** of $\mathbb{F}_q^n$. We determine $I_n(\mathbb{F}_q)$:

We have $d(v + u, w + u) = d(v, w)$, for all $u, v, w \in \mathbb{F}_q^n$, thus we have $d(v, w) = d(v - w, 0_n) = \mathrm{wt}(v - w)$. Since $I_n(\mathbb{F}_q)$ fixes $0_n \in \mathbb{F}_q^n$, we infer $\mathrm{wt}(v^g) = \mathrm{wt}(v)$ for all $v \in \mathbb{F}_q^n$ and $g \in I_n(\mathbb{F}_q)$. Hence for the i-th unit vector $e_i = [0, \ldots, 0, 1, 0, \ldots, 0] \in \mathbb{F}_q^n$, where $i \in \{1, \ldots, n\}$, we have $e_i^g = x_i e_{i^\pi}$, where $\pi \in \mathcal{S}_n$ and $x_i \in \mathbb{F}_q^* := \mathbb{F}_q \setminus \{0\}$. Thus g is described by a monomial matrix $\mathrm{diag}[x_1, \ldots, x_n] \cdot P_\pi \in \mathrm{GL}_n(\mathbb{F}_q)$, where $P_\pi \in \mathrm{GL}_n(\mathbb{F}_q)$ is the permutation matrix associated with $\pi \in \mathcal{S}_n$.

Conversely, any invertible diagonal matrix and any permutation matrix, and thus any monomial matrix, gives rise to a linear isometry. Thus $I_n(\mathbb{F}_q) \leq \mathrm{GL}_n(\mathbb{F}_q)$ is the subgroup of monomial matrices, hence $I_n(\mathbb{F}_q) \cong (\mathbb{F}_q^*)^n \rtimes \mathcal{S}_n$; in particular, it acts transitively on $\{v \in \mathbb{F}_q^n; \mathrm{wt}(v) = i\}$, for all $i \in \{0, \ldots, n\}$. ♯

(3.2) Minimum distance. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, and let $\emptyset \neq \mathcal{C} \subseteq \mathcal{X}^n$ be a **block code** of **length** $n \in \mathbb{N}$ and **order** $m := |\mathcal{C}| \in \mathbb{N}$; note that we do not distinguish between information and check symbols. If $m = 1$ then $\mathcal{C}$ is called **trivial**. Codes $\mathcal{C}, \mathcal{C}' \subseteq \mathcal{X}^n$ are called **equivalent**, if there is an isometry $g \in I(\mathcal{X}^n)$ such that $\mathcal{C}^g = \mathcal{C}'$; the **automorphism group** $\mathrm{Aut}(\mathcal{C})$ of $\mathcal{C}$ is the group of all isometries $g \in I(\mathcal{X}^n)$ such that $\mathcal{C}^g = \mathcal{C}$.

If $\mathcal{C}$ is non-trivial, then $d(\mathcal{C}) := \min\{d(v,w) \in \mathbb{N}; v \neq w \in \mathcal{C}\} \in \{1, \ldots, n\}$ is called the **minimum distance** of $\mathcal{C}$; if $\mathcal{C}$ is trivial we let $d(\mathcal{C}) := \infty$. If $d(\mathcal{C}) = d$ then $\mathcal{C}$ is called an (n,m,d)**-code** over $\mathcal{X}$. We have $d(\mathcal{X}^n) = 1$, and equivalent codes have the same minimum distance.

For any non-trivial (n,m,d)-code $\mathcal{C}$ we have the **Singleton bound** [1964] $\log_q(m) \leq n - d + 1$: We consider the map $\alpha\colon \mathcal{X}^n \to \mathcal{X}^{n-d+1}\colon [x_1, \ldots, x_n] \mapsto [x_1, \ldots, x_{n-d+1}]$; since for any $v \neq w \in \mathcal{C}$ we have $d(v,w) \geq d$, we infer that the restriction $\alpha|_{\mathcal{C}}\colon \mathcal{C} \to \mathcal{X}^{n-d+1}$ is injective, thus $m = |\mathcal{C}| \leq q^{n-d+1}$. Note that in the above argument we can choose any $n-d+1$ components instead. If we have equality $d-1 = n - \log_q(m)$, then $\mathcal{C}$ is called a **maximum distance separable (MDS)** code; in particular, $\mathcal{X}^n$ is the only MDS code such that $d = 1$.

b) Let $\mathcal{X} = \mathbb{F}_q$ be the field with q elements, and let $\mathcal{C} \leq \mathbb{F}_q^n$ be a **linear code**; specifically, if $q = 2$ or $q = 3$ then $\mathcal{C}$ is called **binary** and **ternary**, respectively. Let $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \in \mathbb{N}_0$ be the **dimension** of $\mathcal{C}$, and let $d(\mathcal{C}) = d$ be its minimum distance, then $\mathcal{C}$ is called an $[n,k,d]$**-code** over $\mathbb{F}_q$; in particular $\mathcal{C}$ is an (n, q^k, d)-code, and the Singleton bound for $k \geq 1$ reads $d - 1 \leq n - k$.

Moreover, if $\mathcal{C}$ is non-trivial then $\mathrm{wt}(\mathcal{C}) := \min\{\mathrm{wt}(v) \in \mathbb{N}; 0_n \neq v \in \mathcal{C}\} \in \{1, \ldots, n\}$ is called the **minimum weight** of $\mathcal{C}$; if $\mathcal{C}$ is trivial we let $\mathrm{wt}(\mathcal{C}) := \infty$. Then we have $\mathrm{wt}(\mathcal{C}) = d$, that is the minimum distance and the minimum weight of $\mathcal{C}$ coincide: We may assume that $\mathcal{C}$ is non-trivial, that is $k \geq 1$; since $\mathrm{wt}(v) = d(v, 0_n) \geq d$ for all $0_n \neq v \in \mathcal{C}$, we have $\mathrm{wt}(\mathcal{C}) \geq d$; conversely, for all $v \neq w \in \mathcal{C}$ we have $0_n \neq v - w \in \mathcal{C}$ and $d(v,w) = d(v-w, 0_n) = \mathrm{wt}(v-w) \geq \mathrm{wt}(\mathcal{C})$, hence we have $d \geq \mathrm{wt}(\mathcal{C})$ as well.

Codes $\mathcal{C}, \mathcal{C}' \leq \mathbb{F}_q^n$ are called **linearly equivalent**, if there is an $\mathbb{F}_q$-linear isometry $g \in I_n(\mathbb{F}_q)$ such that $\mathcal{C}^g = \mathcal{C}'$; the **linear automorphism group** $\mathrm{Aut}_{\mathbb{F}_q}(\mathcal{C})$ of $\mathcal{C}$ is the group of all $\mathbb{F}_q$-linear isometries $g \in I_n(\mathbb{F}_q)$ such that $\mathcal{C}^g = \mathcal{C}$. Linearly equivalent codes have the same dimension and the same minimum weight.

(3.3) Error correction. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$. For $n \in \mathbb{N}$ and $r \in \mathbb{N}_0$ let $\mathcal{B}_r(v) := \{w \in \mathcal{X}^n; d(v,w) \leq r\}$ be the **sphere** or **ball** with **radius** r around $v \in \mathcal{X}^n$. Hence independently of $v \in \mathcal{X}^n$ we have $|\mathcal{B}_r(v)| = \sum_{d=0}^{\min\{r,n\}} |\{w \in \mathcal{X}^n; d(v,w) = d\}| = \sum_{d=0}^{\min\{r,n\}} \binom{n}{d} \cdot (q-1)^d \in \mathbb{N}$; recall that we have already used this for $q = 2$ in (2.8).

Let $\mathcal{C}$ be an (n,m,d)-code over $\mathcal{X}$. Then $\mathcal{C}$ is called (e,f)**-error detecting**, for some $e \in \{0, \ldots, n\}$ and $f \in \{e, \ldots, n\}$, if $\mathcal{B}_f(v) \cap \mathcal{B}_e(w) = \emptyset$ for all $v \neq w \in \mathcal{C}$. In particular, if $f = e + 1$, that is $\mathcal{B}_f(v) \cap \mathcal{B}_{f-1}(w) = \emptyset$ for all $v \neq w \in \mathcal{C}$, then $\mathcal{C}$ is called f**-error detecting**; and if $e = f$, that is $\mathcal{B}_e(v) \cap \mathcal{B}_e(w) = \emptyset$ for all $v \neq w \in \mathcal{C}$, then $\mathcal{C}$ is called **e-error correcting**. Thus, if $\mathcal{C}$ is (e,f)-error detecting then it is e-error correcting, and in this case it is e-error detecting.

Let $\mathcal{C}$ be (e,f)-error detecting. Then for $u \in \mathcal{X}^n$ we have $u \in \mathcal{B}_f(v)$, for some $v \in \mathcal{C}$, if and only if u is obtained from v by at most f errors. In this case, u has distance at least $e+1$ from any $v \neq w \in \mathcal{C}$; thus it is detected whether an error has occurred. If at most $\min\{f, e+1\}$ errors have occurred, the number of errors is detected, but if $f = e+1$ and f errors have occurred then u is not uniquely nearest neighbor decodable, while if at most e errors have occurred, then u is correctly decoded by unique nearest neighbor decoding.

Moreover, if $\mathcal{C}$ is e-error correcting then since any element of $\mathcal{X}^n$ is contained in at most one sphere $\mathcal{B}_e(v)$, where $v \in \mathcal{C}$, we have the **Hamming bound** [1960] or **sphere packing bound** $m \cdot \sum_{i=0}^{e} \binom{n}{i} \cdot (q-1)^i \leq q^n = |\mathcal{X}^n|$.

b) If $\mathcal{C}$ is trivial, then $\mathcal{C}$ is n-error correcting; hence let $\mathcal{C}$ be non-trivial. Then $d \in \mathbb{N}$ is related to the error correction and detection properties of $\mathcal{C}$ as follows:

The code $\mathcal{C}$ is (e,f)-error detecting if and only if $e+f \leq d-1$: Let $e+f$ be as large as possible, that is $e+f = d-1$. Assume that there are $v \neq w \in \mathcal{C}$ such that $u \in \mathcal{B}_f(v) \cap \mathcal{B}_e(w) \neq \emptyset$, then we have $d(v,w) \leq d(v,u) + d(u,w) \leq f+e = d-1$, a contradiction; hence $\mathcal{C}$ is (e,f)-error detecting. Conversely, let $v \neq w \in \mathcal{C}$ such that $d(v,w) = d$, then $\mathcal{B}_{f+1}(v) \cap \mathcal{B}_e(w) \neq \emptyset$ and $\mathcal{B}_f(v) \cap \mathcal{B}_{e+1}(w) \neq \emptyset$, thus $\mathcal{C}$ is neither $(e, f+1)$-error detecting, nor $(e+1, f)$-error detecting if $e < f$.

In particular, $\mathcal{C}$ is f-error detecting if and only if $2f \leq d$, and $\mathcal{C}$ is e-error correcting if and only if $2e+1 \leq d$. In other words, $\mathcal{C}$ always is $\lfloor \frac{d-1}{2} \rfloor$-error correcting, and if d is even then $\mathcal{C}$ moreover is $\frac{d}{2}$-error detecting; note that $\lfloor \frac{d-1}{2} \rfloor = \frac{d-1}{2}$ if d is odd, but $\lfloor \frac{d-1}{2} \rfloor = \frac{d}{2} - 1$ if d is even.

Example. Let $\mathcal{C} := \{v \in \mathbb{Z}_q^n; vw^{\text{tr}} = 0 \in \mathbb{Z}_q\}$ be a parity check code over $\mathbb{Z}_q$, for $q \geq 2$ and $n \geq 2$ and weights $w \in (\mathbb{Z}_q^*)^n$. Since $w_n \in \mathbb{Z}_q^*$, for all $[x_1, \ldots, x_{n-1}] \in \mathbb{Z}_q^{n-1}$ there is a unique $x_n \in \mathbb{Z}_q$ such that $v := [x_1, \ldots, x_{n-1}, x_n] \in \mathcal{C}$. Hence the information rate of $\mathcal{C}$ is $\rho(\mathcal{C}) = \frac{\log_q(|\mathcal{C}|)}{n} = \frac{\log_q(q^{n-1})}{n} = \frac{n-1}{n} = 1 - \frac{1}{n}$.

Moreover, for $x_{n-1} \neq x'_{n-1} \in \mathbb{Z}_q$ there is $x'_n \in \mathbb{Z}_q$ such that $[x_1, \ldots, x'_{n-1}, x'_n] \in \mathcal{C}$ as well, implying that $d(\mathcal{C}) \leq 2$. Since $w_i \in \mathbb{Z}_q^*$, for all $i \in \{1, \ldots, n\}$ we infer that we have $[x_1, \ldots, x_{i-1}, x'_i, x_{i+1}, \ldots, x_n] \notin \mathcal{C}$, whenever $x_i \neq x'_i \in \mathbb{Z}_q$. This says that $\mathcal{B}_1(v) \cap \mathcal{C} = \{v\}$, entailing that $\mathcal{C}$ has minimum distance $d(\mathcal{C}) = 2$, and thus is 0-error correcting and 1-error detecting.

(3.4) Covering radius. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, and let $\mathcal{C}$ be an (n,m,d)-code over $\mathcal{X}$. The minimum $c \in \{0, \ldots, n\}$ such that $\mathcal{X}^n = \bigcup_{v \in \mathcal{C}} \mathcal{B}_c(v)$ is called the **covering radius** $c(\mathcal{C})$ of $\mathcal{C}$. Hence we have $c(\mathcal{C}) = 0$ if and only if $\mathcal{C} = \mathcal{X}^n$, and if $\mathcal{C}$ is trivial then $c(\mathcal{C}) = n$. Letting $\mathcal{C}$ be non-trivial, the covering radius is related to the minimum distance as follows:

If d is odd, letting $e := \frac{d-1}{2}$ we have $\mathcal{B}_e(v) \cap \mathcal{B}_e(w) = \emptyset$ for all $v \neq w \in \mathcal{C}$, hence since $\mathcal{B}_e(v) \setminus \mathcal{B}_{e-1}(v) \neq \emptyset$, we conclude that $c(\mathcal{C}) \geq e$. If d is even, letting $f := \frac{d}{2}$

we have $\mathcal{B}_f(v) \cap \mathcal{B}_{f-1}(w) = \emptyset$ for all $v \neq w \in \mathcal{C}$, hence since $\mathcal{B}_f(v) \setminus \mathcal{B}_{f-1}(v) \neq \emptyset$, we conclude that $c(\mathcal{C}) \geq f$.

b) If $c(\mathcal{C}) = e := \lfloor \frac{d-1}{2} \rfloor$, that is $\mathcal{X}^n = \coprod_{v \in \mathcal{C}} \mathcal{B}_e(v)$, then $\mathcal{C}$ is called **perfect**. In this case we have unique nearest neighbor decoding for any element of $\mathcal{X}^n$, but $\mathcal{C}$ incorrectly decodes any word containing at least $e+1$ errors. In other words, $\mathcal{C}$ is perfect if and only if the Hamming bound is an equality, that is we have $m \cdot \sum_{i=0}^{e} \binom{n}{i} \cdot (q-1)^i = q^n$. Hence $\mathcal{C}$ is perfect with $d=1$ if and only if $\mathcal{C} = \mathcal{X}^n$.

As for the existence of perfect codes in general, if d is even then $c(\mathcal{C}) \geq f = \frac{d}{2}$ and $\mathcal{B}_f(v) \cap \mathcal{B}_f(w) \neq \emptyset$, for all $v \neq w \in \mathcal{C}$ such that $d(v,w) = d$, imply that there are no perfect codes in this case. The picture changes if d is odd, but still perfect codes are rare; fulfilling the Hamming bound is not sufficient.

If d is odd and $c(\mathcal{C}) = e+1 = \frac{d+1}{2}$, or d is even and $c(\mathcal{C}) = f = \frac{d}{2}$, the code $\mathcal{C}$ is called **quasi-perfect**; in this case there are elements of $\mathcal{X}^n$ which do not allow for unique nearest neighbor decoding, and $\mathcal{C}$ incorrectly decodes any word containing at least $e+2$ respectively $f+1$ errors, and possibly some words containing $e+1$ respectively f errors.

Example. The **repetition code** $\mathcal{C} := \{[x, \ldots, x] \in \mathcal{X}^n; x \in \mathcal{X}\}$, for $n \in \mathbb{N}$, has information rate is $\rho(\mathcal{C}) = \frac{\log_q(|\mathcal{C}|)}{n} = \frac{1}{n}$. Its minimum distance is $d(\mathcal{C}) = n$, hence $\mathcal{C}$ is $\lfloor \frac{n-1}{2} \rfloor$-error correcting, and if n is even it is $\frac{n}{2}$-error detecting.

In particular, if $\mathcal{X} = \mathbb{F}_2$ then we have $\mathcal{C} = \{0_n, 1_n\} \leq \mathbb{F}_2^n$. Since for all $v \in \mathbb{F}_2^n$ we have $d(v, 0_n) \leq \lfloor \frac{n}{2} \rfloor$ or $d(v, 1_n) \leq \lfloor \frac{n}{2} \rfloor$, we get $c(\mathcal{C}) = \frac{n-1}{2}$ if n is odd, that is $\mathcal{C}$ perfect, and $c(\mathcal{C}) = \frac{n}{2}$ if n is even, that is $\mathcal{C}$ quasi-perfect.

Example. Let $\mathcal{C} := \langle [1,0,0,0,1,1], [0,1,0,1,0,1], [0,0,1,1,1,0] \rangle_{\mathbb{F}_2} \leq \mathbb{F}_2^6$, hence $\mathcal{X} = \mathbb{F}_2$, and the elements of $\mathcal{C}$ consist of the rows of the following matrix:

$$\left[\begin{array}{cccccc}
. & . & . & . & . & . \\
\hline
1 & . & . & . & 1 & 1 \\
. & 1 & . & 1 & . & 1 \\
. & . & 1 & 1 & 1 & . \\
\hline
1 & 1 & . & 1 & 1 & . \\
1 & . & 1 & 1 & . & 1 \\
. & 1 & 1 & . & 1 & 1 \\
\hline
1 & 1 & 1 & . & . & .
\end{array}\right] \in \mathbb{F}_2^{8 \times 6}.$$

(Actually, $\mathcal{C}$ is obtained from the Hamming $[7,4,3]$-code, see (4.2) and (6.2), by shortening with respect to the 4-th component, see (4.5).)

Thus we have the minimum distance $d = d(\mathcal{C}) = \mathrm{wt}(\mathcal{C}) = 3$, that is $\mathcal{C}$ is a $[6,3,3]$-code, which hence is 1-error correcting. Using $q=2$ and $e = \frac{d-1}{2} = 1$, the Hamming bound yields $2^3 \cdot (\binom{6}{0} + \binom{6}{1}) = 56 < 64 = 2^6$, thus $\mathcal{C}$ is not perfect. Hence the covering radius is $c(\mathcal{C}) \geq e+1 = 2$, and we show that actually $c(\mathcal{C}) = 2 = \frac{d+1}{2}$, saying that $\mathcal{C}$ is quasi-perfect:

Let $u \in \mathbb{F}_2^6$. Since the Hamming distance is translation invariant, by adding a suitable element of $\mathcal{C}$ we may assume that $u = [0,0,0,*,*,*]$. Since the permutation $(1,2,3)(4,5,6) \in \mathcal{S}_6$ induces a linear isometry leaving $\mathcal{C}$ invariant, we may assume that $u \in \{0_6, [0,0,0,1,0,0], [0,0,0,1,1,0], [0,0,0,1,1,1]\}$. Now we observe that $u \in \mathcal{B}_2(v)$ for some $v \in \mathcal{C}$. ♯

(3.5) Theorem: Tietäväinen, van Lint [1973]. Let $\mathcal{C} \subseteq \mathbb{F}_q^n$ be a perfect $(n, m, 2e+1)$-code, where $e \in \mathbb{N}$; in particular we have $\mathcal{C} \neq \mathbb{F}_q^n$ and $n \geq 3$.

a) If $e \geq 2$, then $\mathcal{C}$ is equivalent to a linear code, and linearly equivalent to
i) the binary repetition $[n,1,n]$-code $\{0_n, 1_n\}$, where $n \geq 5$ is odd;
ii) the **binary Golay** $[23,12,7]$**-code** $\mathcal{G}_{23}$, see (13.1);
iii) the **ternary Golay** $[11,6,5]$**-code** $\mathcal{G}_{11}$, see (13.2).

b) If $e = 1$, then $n = \frac{q^k-1}{q-1}$ for some $k \geq 2$, and $m = q^{n-k}$. If $\mathcal{C}$ is linear, then it is linearly equivalent to the **Hamming** $[n, n-k, 3]$**-code** $\mathcal{H}_k$, see (6.1). ♯

In particular, for $q = 2$ and $k = 2$ we recover the binary repetition $[3,1,3]$-code $\{0_3, 1_3\}$. Actually, there are non-linear codes having the parameters of Hamming codes, and their classification still is an open problem.

4 Linear codes

(4.1) Generator matrices. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a linear code of length $n \in \mathbb{N}$ over $\mathbb{F}_q$, and let $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \in \{0, \ldots, n\}$. A matrix $G \in \mathbb{F}_q^{k \times n}$ whose rows form an $\mathbb{F}_q$-basis of $\mathcal{C}$ is called a **generator matrix** of $\mathcal{C}$; hence we have $\mathcal{C} = \mathrm{im}(G) = \{vG \in \mathbb{F}_q^n; v \in \mathbb{F}_q^k\}$; in particular, for $k = 0$ we have $G \in \mathbb{F}_q^{0 \times n}$, and for $k = n$ we may choose the identity matrix $G = E_n \in \mathbb{F}_q^{n \times n}$. Then $v \in \mathbb{F}_q^k$ is encoded into $vG \in \mathbb{F}_q^n$, and conversely $w \in \mathcal{C} = \mathrm{im}(G) \leq \mathbb{F}_q^n$ is decoded by solving the system of $\mathbb{F}_q$-linear equations $[X_1, \ldots, X_k] \cdot G = w \in \mathbb{F}_q^n$, which since $\mathrm{rk}_{\mathbb{F}_q}(G) = k$ has a unique solution.

Since $\mathrm{rk}_{\mathbb{F}_q}(G) = k$, by Gaussian row elimination and possibly column permutation G can be transformed into **standard** form $[E_k \mid A] \in \mathbb{F}_q^{k \times n}$, where $A \in \mathbb{F}_q^{k \times (n-k)}$. Row operations leave the row space $\mathcal{C}$ of G invariant; and column permutations amount to permuting the positions of symbols, thus transform $\mathcal{C}$ into a linearly equivalent code. Hence in this case $[x_1, \ldots, x_k] \in \mathbb{F}_q^k$ is encoded into $[x_1, \ldots, x_k; y_1, \ldots, y_{n-k}] \in \mathbb{F}_q^n$, where $[y_1, \ldots, y_{n-k}] = [x_1, \ldots, x_k] \cdot A \in \mathbb{F}_q^{n-k}$. Thus the first k symbols can be considered as information symbols, and the last $n-k$ symbols as check symbols; since information and check symbols can be distinguished like this $\mathcal{C}$ is called **separable**. Moreover, the projection $\mathcal{C} \to \mathbb{F}_q^k \colon [z_1, \ldots, z_n] \mapsto [z_1, \ldots, z_k]$ onto the first k positions is a bijection; hence $\mathcal{C}$ is called **systematic** on the information symbols.

(4.2) Check matrices. a) Let $\mathbb{F}_q$ be the field with q elements. For $n \in \mathbb{N}$ let $\langle \cdot, \cdot \rangle \colon \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{F}_q \colon [[x_1, \ldots, x_n], [y_1, \ldots, y_n]] \mapsto x \cdot y^{\mathrm{tr}} = \sum_{i=1}^n x_i y_i$ be the **standard $\mathbb{F}_q$-bilinear form** on $\mathbb{F}_q^n$. This form is symmetric and non-degenerate, and we have $\langle vM, w\rangle = \langle v, wM^{\mathrm{tr}}\rangle$ for all $v, w \in \mathbb{F}_q^n$ and $M \in \mathbb{F}_q^{n\times n}$.

For a code $\mathcal{C} \leq \mathbb{F}_q^n$, the orthogonal space $\mathcal{C}^\perp := \{v \in \mathbb{F}_q^n; \langle v, w\rangle = 0 \in \mathbb{F}_q \text{ for all } w \in \mathcal{C}\} \leq \mathbb{F}_q^n$ with respect to the standard $\mathbb{F}_q$-bilinear form is called the associated **dual code**. Letting $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \in \{0, \ldots, n\}$, we have $\dim_{\mathbb{F}_q}(\mathcal{C}^\perp) = n - k$. Moreover, we have $\mathcal{C} \leq (\mathcal{C}^\perp)^\perp$, and from $\dim_{\mathbb{F}_q}((\mathcal{C}^\perp)^\perp) = n - (n-k) = k = \dim_{\mathbb{F}_q}(\mathcal{C})$ we get $(\mathcal{C}^\perp)^\perp = \mathcal{C}$. If $\mathcal{C} \leq \mathcal{C}^\perp$ then $\mathcal{C}$ is called **weakly self-dual**, and if $\mathcal{C} = \mathcal{C}^\perp$ then $\mathcal{C}$ is called **self-dual**; in the latter case we have $n - k = \dim_{\mathbb{F}_q}(\mathcal{C}^\perp) = \dim_{\mathbb{F}_q}(\mathcal{C}) = k$, thus $n = 2k$ is even.

b) If $G \in \mathbb{F}_q^{k\times n}$ is a generator matrix of $\mathcal{C}$, then we have $\mathcal{C}^\perp = \{v \in \mathbb{F}_q^n; Gv^{\mathrm{tr}} = 0 \in \mathbb{F}_q^{k\times 1}\} = \{v \in \mathbb{F}_q^n; vG^{\mathrm{tr}} = 0 \in \mathbb{F}_q^k\}$. Hence if $H \in \mathbb{F}_q^{(n-k)\times n}$ is a generator matrix of $\mathcal{C}^\perp$, then $\mathcal{C} = (\mathcal{C}^\perp)^\perp = \{v \in \mathbb{F}_q^n; vH^{\mathrm{tr}} = 0 \in \mathbb{F}_q^{n-k}\} = \ker(H^{\mathrm{tr}}) \leq \mathbb{F}_q^n$. Thus H is called a **check matrix** of $\mathcal{C}$, and instead of using a generator matrix the code $\mathcal{C}$ can also be defined by a check matrix. In particular, for $k = n$ we have $H \in \mathbb{F}_q^{0\times n}$, and for $k = 0$ we may choose the identity matrix $H = E_n \in \mathbb{F}_q^{n\times n}$.

If $G = [E_k \mid A] \in \mathbb{F}_q^{k\times n}$ is in standard form, then $H = [-A^{\mathrm{tr}} \mid E_{n-k}] \in \mathbb{F}_q^{(n-k)\times n}$ is a generator matrix of $\mathcal{C}^\perp$, also being called a **standard** check matrix for $\mathcal{C}$: We have $\mathrm{rk}_{\mathbb{F}_q}(H) = n - k$, and $HG^{\mathrm{tr}} = [-A^{\mathrm{tr}} \mid E_{n-k}] \cdot \left[\frac{E_k}{A^{\mathrm{tr}}}\right] = -A^{\mathrm{tr}} \cdot E_k + E_{n-k} \cdot A^{\mathrm{tr}} = 0 \in \mathbb{F}_q^{(n-k)\times k}$.

Duality interferes with linear equivalence nicely inasmuch a code $\mathcal{C}' \leq \mathbb{F}_q^n$ is linearly equivalent to $\mathcal{C}$ if and only if its dual $(\mathcal{C}')^\perp$ is linearly equivalent to $\mathcal{C}^\perp$: It suffices to show one direction. If $\mathcal{C}'$ is linearly equivalent to $\mathcal{C}$, then there is $M \in I_n(\mathbb{F}_q)$ such that $\mathcal{C}' = \mathcal{C}\cdot M$. Then we have $\mathcal{C}'\cdot(HM^{-\mathrm{tr}})^{\mathrm{tr}} = \mathcal{C}\cdot MM^{-1}H^{\mathrm{tr}} = \{0\}$. Thus $HM^{-\mathrm{tr}} \in \mathbb{F}_q^{(n-k)\times n}$, having rank $\mathrm{rk}_{\mathbb{F}_q}(HM^{-\mathrm{tr}}) = n - k$, is a check matrix for $\mathcal{C}'$, entailing that $(\mathcal{C}')^\perp$, having $HM^{-\mathrm{tr}}$ as a generator matrix, is linearly equivalent to $\mathcal{C}^\perp$, which has H as a generator matrix.

Example. Let the binary **Hamming code** $\mathcal{H} \leq \mathbb{F}_2^7$ be given by the following generator matrix $G \in \mathbb{F}_2^{4\times 7}$, or equivalently its standard form G',

$$G := \begin{bmatrix} . & . & . & 1 & 1 & 1 & 1 \\ . & 1 & 1 & 1 & 1 & . & . \\ 1 & 1 & . & . & 1 & 1 & . \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad G' := \left[\begin{array}{cccc|ccc} 1 & . & . & . & . & 1 & 1 \\ . & 1 & . & . & 1 & . & 1 \\ . & . & 1 & . & 1 & 1 & . \\ . & . & . & 1 & 1 & 1 & 1 \end{array}\right].$$

A check matrix $H \in \mathbb{F}_2^{3\times 7}$, or equivalently its standard form H', is given as

$$H := \begin{bmatrix} . & . & . & 1 & 1 & 1 & 1 \\ . & 1 & 1 & . & . & 1 & 1 \\ 1 & . & 1 & . & 1 & . & 1 \end{bmatrix} \quad \text{and} \quad H' := \left[\begin{array}{cccc|ccc} . & 1 & 1 & 1 & 1 & . & . \\ 1 & . & 1 & 1 & . & 1 & . \\ 1 & 1 & . & 1 & . & . & 1 \end{array}\right].$$

We have $k = \dim_{\mathbb{F}_2}(\mathcal{H}) = 4$, that is $m = |\mathcal{H}| = 2^4 = 16$. From inspecting the elements of $\mathcal{H}$, as given by the rows of the matrices below, or from (4.3) below, we get $d = d(\mathcal{H}) = 3$, thus $\mathcal{H}$ is a $[7, 4, 3]$-code. Moreover, we have $m \cdot \sum_{i=0}^{\frac{d-1}{2}} \binom{7}{i} = 16 \cdot (1 + 7) = 128 = 2^7 = |\mathbb{F}_2^7|$, thus by the Hamming bound we conclude that $\mathcal{H}$ is perfect; see also (6.2).

$$\begin{bmatrix} . & . & . & . & . & . & . \\ . & . & . & 1 & 1 & 1 & 1 \\ . & . & 1 & . & 1 & 1 & . \\ . & . & 1 & 1 & . & . & 1 \\ . & 1 & . & . & 1 & . & 1 \\ . & 1 & . & 1 & . & 1 & . \\ . & 1 & 1 & . & . & 1 & 1 \\ . & 1 & 1 & 1 & 1 & . & . \end{bmatrix} \qquad \begin{bmatrix} 1 & . & . & . & . & 1 & 1 \\ 1 & . & . & 1 & 1 & . & . \\ 1 & . & 1 & . & 1 & . & 1 \\ 1 & . & 1 & 1 & . & 1 & . \\ 1 & 1 & . & . & 1 & 1 & . \\ 1 & 1 & . & 1 & . & . & 1 \\ 1 & 1 & 1 & . & . & . & . \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(4.3) Theorem. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a non-trivial $[n, k, d]$-code, let $\mathbb{F}_q \subseteq F$ be a field extension, and let $H \in F^{(n-k)\times n}$ be a **generalized check matrix** of $\mathcal{C}$, that is $\mathcal{C} = \ker(H^{\mathrm{tr}}) \cap \mathbb{F}_q^n$ is a **subfield subcode** of $\ker(H^{\mathrm{tr}}) \leq F^n$; note that we do not require that H has full F-rank. Then any $(d-1)$-subset of columns of H is $\mathbb{F}_q$-linearly independent, and there is some $\mathbb{F}_q$-linearly dependent d-subset of columns of H.

In particular, for $F = \mathbb{F}_q$ the matrix H is just a check matrix of $\mathcal{C}$, inasmuch the condition $\mathrm{rk}_{\mathbb{F}_q}(H) = n - \dim_{\mathbb{F}_q}(\ker(H^{\mathrm{tr}})) = n - k$ is fulfilled automatically; thus we recover the Singleton bound $d - 1 \leq \mathrm{rk}_{\mathbb{F}_q}(H) = n - k$ for linear codes.

Proof. Let $d' \in \mathbb{N}$ such that any $(d'-1)$-subset of columns of H is $\mathbb{F}_q$-linearly independent, while there is some $\mathbb{F}_q$-linearly dependent d'-subset of columns of H. Let $d' \geq 2$, and let $0 \neq v \in \mathbb{F}_q^n$ such that $\mathrm{wt}(v) \leq d'-1$. Hence $vH^{\mathrm{tr}} \in F^{n-k}$ is a non-trivial $\mathbb{F}_q$-linear combination of at most $d'-1$ rows of H^{tr}. Since the latter are $\mathbb{F}_q$-linearly independent we have $vH^{\mathrm{tr}} \neq 0$, hence $v \notin \mathcal{C}$, implying that $d \geq d'$. Conversely, for $d' \geq 1$, picking an $\mathbb{F}_q$-linearly dependent d'-subset of columns of H, there is $0 \neq v \in \mathbb{F}_q^n$ such that $\mathrm{wt}(v) \leq d'$ and $vH^{\mathrm{tr}} = 0 \in F^{n-k}$, thus we have $v \in \mathcal{C}$, implying that $d \leq d'$. ♯

(4.4) Syndrome decoding. Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $n \in \mathbb{N}$, be given by a check matrix $H \in \mathbb{F}_q^{(n-k)\times n}$, where $k := \dim_{\mathbb{F}_q}(\mathcal{C})$. Then for $v \in \mathbb{F}_q^n$ let $vH^{\mathrm{tr}} \in \mathbb{F}_q^{n-k}$ be the **syndrome** of v with respect to H.

We consider the quotient $\mathbb{F}_q$-vector space $\mathbb{F}_q^n/\mathcal{C}$, whose elements $\overline{v} := v + \mathcal{C} = \{v + w \in \mathbb{F}_q^n; w \in \mathcal{C}\} \in \mathbb{F}_q^n/\mathcal{C}$, for $v \in \mathbb{F}_q^n$, are called the **cosets** with respect to $\mathcal{C}$. Since $\mathcal{C} = \ker(H^{\mathrm{tr}}) \leq \mathbb{F}_q^n$ and $\mathrm{rk}_{\mathbb{F}_q}(H) = n - k$, by the homomorphism theorem we have $\mathbb{F}_q^n/\mathcal{C} = \mathbb{F}_q^n/\ker(H^{\mathrm{tr}}) \cong \mathrm{im}(H^{\mathrm{tr}}) = \mathbb{F}_q^{n-k}$ as $\mathbb{F}_q$-vector spaces. Thus the syndromes are in natural bijection with the cosets with respect to $\mathcal{C}$, in particular there are $|\mathbb{F}_q^n/\mathcal{C}| = |\mathbb{F}_q^{n-k}| = q^{n-k}$ syndromes.

To decode a possibly erroneous vector $v \in \mathbb{F}_q^n$, we proceed as follows: Let $w \in \mathcal{C}$ be the codeword sent, and let $u := v - w \in \mathbb{F}_q^n$ be the associated **error**

vector, thus we have $d(v,w) = \mathrm{wt}(u)$. Moreover, for the syndromes we have $vH^{\mathrm{tr}} = (w+u)H^{\mathrm{tr}} = uH^{\mathrm{tr}} \in \mathbb{F}_q^{n-k}$, that is we have $\overline{v} = \overline{u} \in \mathbb{F}_q^n/\mathcal{C}$. Hence v is uniquely nearest neighbor decodable if and only if the coset $v + \mathcal{C} \subseteq \mathbb{F}_q^n$ possesses a unique element $u \in \mathbb{F}_q^n$ of minimum weight, and in this case $v \in \mathbb{F}_q^n$ is decoded to $v - u \in \mathbb{F}_q^n$.

The elements of the coset $v+\mathcal{C} \subseteq \mathbb{F}_q^n$ having minimum weight are called the **coset leaders** of $\overline{v} \in \mathbb{F}_q^n/\mathcal{C}$. In general coset leaders are not unique. Still, the zero vector $0_n \in \mathbb{F}_q^n$ always is the unique coset leader of the coset $\overline{0}_n \in \mathbb{F}_q^n/\mathcal{C}$; and if $\mathcal{C}$ is e-error correcting, for some $e \in \{0,\ldots,n\}$, and the coset $v+\mathcal{C} \subseteq \mathbb{F}_q^n$ possesses an element $u \in \mathbb{F}_q^n$ of weight $\mathrm{wt}(u) \leq e$, then u is the unique coset leader of $\overline{v} \in \mathbb{F}_q^n/\mathcal{C}$. In practice, coset leaders for all syndromes in $\mathrm{im}(H^{\mathrm{tr}}) = \mathbb{F}_q^{n-k}$ are computed once and stored into a table, so that then coset leaders are found by computing syndromes and subsequent table lookup. Finding coset leaders algorithmically, being called the **decoding problem for linear codes**, is an NP-hard problem; hence the aim is to find codes having fast decoding algorithms.

Example: Parity check codes. a) Let $\mathcal{C} = \{v \in \mathbb{F}_q^n; vw^{\mathrm{tr}} = 0 \in \mathbb{F}_q\} \leq \mathbb{F}_q^n$ be defined by the check matrix $0 \neq H := [w] = [w_1,\ldots,w_n] \in \mathbb{F}_q^n$, where up to linear equivalence we may assume that $w_n = 1$, that is H is in standard form; hence we have $k = \dim_{\mathbb{F}_q}(\mathcal{C}) = n-1$. Thus the standard generator matrix is given as $G := [E_{n-1} \mid y^{\mathrm{tr}}] \in \mathbb{F}_q^{(n-1)\times n}$, where $y := -[w_1,\ldots,w_{n-1}] \in \mathbb{F}_q^{n-1}$. Hence $[x_1,\ldots,x_{n-1}] \in \mathbb{F}_q^{n-1}$ is encoded into $[x_1,\ldots,x_{n-1}; -\sum_{i=1}^{n-1} x_i w_i] \in \mathbb{F}_q^n$, and for $v = [x_1,\ldots,x_n] \in \mathbb{F}_q^n$ we get the syndrome $vH^{\mathrm{tr}} = \sum_{i=1}^n x_i w_i \in \mathbb{F}_q$. From (4.3) we get $d(\mathcal{C}) \in \{1,2\}$, where for $n \geq 2$ we have $d(\mathcal{C}) = 2$ if and only if $w_i \neq 0$ for all $i \in \{1,\ldots,n\}$.

b) In particular, for $q = 2$ and $w = 1_n \in \mathbb{F}_2^n$, any $[x_1,\ldots,x_{n-1}] \in \mathbb{F}_2^{n-1}$ is encoded into $[x_1,\ldots,x_{n-1}; \sum_{i=1}^{n-1} x_i] \in \mathbb{F}_2^n$, which indeed amounts to adding a **parity check** symbol; and $v = [x_1,\ldots,x_n] \in \mathbb{F}_2^n$ has syndrome $vH^{\mathrm{tr}} = \sum_{i=1}^n x_i \in \mathbb{F}_2$, hence we have $v \in \mathcal{C}$ if and only if $\mathrm{wt}(v) \in \mathbb{N}_0$ is even, thus $\mathcal{C}$ is called the binary **even-weight code** of length n. For $n \geq 2$ we have $d(\mathcal{C}) = 2$, thus $\mathcal{C}$ is a binary $[n, n-1, 2]$-code. Indeed, it is the unique one: Any such code has a check matrix in $\mathbb{F}_2^n$ without zero entries, thus being equal to 1_n.

Since any $v \in \mathbb{F}_2^n$ has distance at most 1 from an element of $\mathcal{C}$, the covering radius equals $c(\mathcal{C}) = 1 = \frac{d(\mathcal{C})}{2}$, thus $\mathcal{C}$ is quasi-perfect. We have $\mathbb{F}_2^n = (0+\mathcal{C}) \,\dot\cup\, (v+\mathcal{C})$, where $v \in \mathbb{F}_2^n$ is any vector such that $\mathrm{wt}(v)$ is odd, corresponding to the syndromes $0 \in \mathbb{F}_2$ and $1 \in \mathbb{F}_2$, respectively; thus any vector of weight 1 is a coset leader of the coset $\mathbb{F}_2^n \setminus \mathcal{C}$, which is not uniquely nearest neighbor decodable.

Example: Repetition codes. a) Let $\mathcal{C}$ be given by the standard generator matrix $G := [1_n] \in \mathbb{F}_q^n$, hence we have $k = \dim_{\mathbb{F}_2}(\mathcal{C}) = 1$ and $d(\mathcal{C}) = n$. Thus the standard check matrix is $H := [-1_{n-1}^{\mathrm{tr}} \mid E_{n-1}] \in \mathbb{F}_q^{(n-1)\times n}$. Then $x \in \mathbb{F}_q$ is encoded into $[x,\ldots,x] \in \mathbb{F}_q^n$, and for $v = [x_1,\ldots,x_n] \in \mathbb{F}_q^n$ we get the syndrome $vH^{\mathrm{tr}} = [x_2 - x_1, \ldots, x_n - x_1] \in \mathbb{F}_q^{n-1}$.

b) In particular, for $q = 2$ we have $\mathcal{C} = \{0_n, 1_n\} \leq \mathbb{F}_2^n$; recall that $\mathcal{C}$ is perfect if n is odd, and quasi-perfect if n is even. The repetition code is the unique binary $[n, 1, n]$-code; it is weakly self-dual if and only if $\langle 1_n, 1_n\rangle = 0$, which holds if and only if n is even. The generator matrix $G = [1_n] \in \mathbb{F}_2^n$ is a check matrix of the even-weight code of length n, hence the latter coincides with $\mathcal{C}^\perp$.

Any $v = [x_1, \ldots, x_n] \in \mathbb{F}_2^n$ has syndrome $[x_2 + x_1, \ldots, x_n + x_1] \in \mathbb{F}_2^{n-1}$; and the coset affording syndrome $w \in \mathbb{F}_2^{n-1}$ equals $[0 \mid w] + \mathcal{C} \subseteq \mathbb{F}_2^n$, where $\operatorname{wt}([0 \mid w]) = \operatorname{wt}(w)$ and $\operatorname{wt}([0 \mid w] + 1_n) = n - \operatorname{wt}(w)$. Thus, given v, computing its syndrome and finding coset leaders yields the following decoding algorithm:

For n odd, coset leaders are uniquely given as $[0 \mid w]$ if $\operatorname{wt}(w) \leq \frac{n-1}{2} =: e$, and $[0 \mid w] + 1_n$ if $\operatorname{wt}(w) \geq \frac{n+1}{2} = e + 1$; in both cases the coset leaders have weight at most e. Thus if $\operatorname{wt}(v) \leq e$ and $x_1 = 0$, then v has syndrome $[x_2, \ldots, x_n]$, and is decoded to $v + [0 \mid x_2, \ldots, x_n] = 0_n$; if $x_1 = 1$, then v has syndrome $[x_2 + 1, \ldots, x_n + 1]$, and is decoded to $v + ([0 \mid x_2 + 1, \ldots, x_n + 1] + 1_n) = 0_n$; if $\operatorname{wt}(v) \geq e + 1$ and $x_1 = 0$, then v is decoded to $v + ([0 \mid x_2, \ldots, x_n] + 1_n) = 1_n$; if $x_1 = 1$, then v is decoded to $v + [0 \mid x_2 + 1, \ldots, x_n + 1] = 1_n$.

For n even, coset leaders are uniquely given as $[0 \mid w]$ if $\operatorname{wt}(w) \leq \frac{n}{2} - 1 =: e$, and $[0 \mid w] + 1_n$ if $\operatorname{wt}(w) \geq \frac{n}{2} + 1 = e + 2$, where in both cases the coset leaders have weight at most e, but for $\operatorname{wt}(w) = \frac{n}{2} = e + 1$ we have $\operatorname{wt}([0 \mid w]) = \operatorname{wt}([0 \mid w] + 1_n) = \operatorname{wt}(w)$, in which case coset leaders are not unique. Thus if $\operatorname{wt}(v) \leq e$ and $x_1 = 0$, then v is decoded to $v + [0 \mid x_2, \ldots, x_n] = 0_n$; if $x_1 = 1$, then v is decoded to $v + ([0 \mid x_2 + 1, \ldots, x_n + 1] + 1_n) = 0_n$; if $\operatorname{wt}(v) \geq e + 2$ and $x_1 = 0$, then v is decoded to $v + ([0 \mid x_2, \ldots, x_n] + 1_n) = 1_n$; if $x_1 = 1$, then v is decoded to $v + [0 \mid x_2 + 1, \ldots, x_n + 1] = 1_n$; but if $\operatorname{wt}(v) = e + 1$ then v is not uniquely nearest neighbor decodable.

(4.5) Modifying codes. Let $\mathcal{C}$ be an $[n, k, d]$-code over $\mathbb{F}_q$, with generator matrix $G \in \mathbb{F}_q{}^{k\times n}$ and check matrix $H \in \mathbb{F}_q^{(n-k)\times n}$.

a) i) Puncturing by deleting the n-th component, for $n \geq 2$, yields the code $\mathcal{C}^\bullet := \{[x_1, \ldots, x_{n-1}] \in \mathbb{F}_q^{n-1}; [x_1, \ldots, x_n] \in \mathcal{C}\} \leq \mathbb{F}_q^{n-1}$. Using the $\mathbb{F}_q$-linear map $\mathbb{F}_q^n \to \mathbb{F}_q^{n-1} \colon [x_1, \ldots, x_n] \mapsto [x_1, \ldots, x_{n-1}]$, having kernel $\langle [0, \ldots, 0, 1]\rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^n$, shows that $k - 1 \leq k^\bullet := \dim_{\mathbb{F}_q}(\mathcal{C}^\bullet) \leq k$. If $d \geq 2$, or $x_n = 0$ for all $[x_1, \ldots, x_n] \in \mathcal{C}$, then $k^\bullet = k$, amounting to deleting a check symbol. Moreover, if $\mathcal{C}^\bullet$ is non-trivial, then for its minimum distance we have $d - 1 \leq d^\bullet \leq d$; in particular, if $x_n = 0$ for all $[x_1, \ldots, x_n] \in \mathcal{C}$, then $d^\bullet = d$.

ii) Extending by adding a check symbol yields $\widehat{\mathcal{C}} := \{[x_1, \ldots, x_n, x_{n+1}] \in \mathbb{F}_q^{n+1}; [x_1, \ldots, x_n] \in \mathcal{C}, \sum_{i=1}^{n+1} x_i = 0\} \leq \mathbb{F}_q^{n+1}$. Hence $\widehat{\mathcal{C}}$ has check matrix $\widehat{H} := \left[\begin{array}{c|c} H & 0_{n-k}^{\mathrm{tr}} \\ \hline 1_n & 1 \end{array}\right] \in \mathbb{F}_q^{(n+1-k)\times(n+1)}$, thus in any case we have $\widehat{k} := \dim_{\mathbb{F}_q}(\widehat{\mathcal{C}}) = k$.

If $\mathcal{C}$ is non-trivial, then $\widehat{\mathcal{C}}$ has minimum distance $d \leq \widehat{d} \leq d + 1$: Since any $(d-1)$-subset of columns of H is $\mathbb{F}_q$-linearly independent, distinguishing the cases whether or not the last column of $\widehat{H}$ is involved, we conclude that any

$(d-1)$-subset of columns of $\widehat{H}$ is $\mathbb{F}_q$-linearly independent as well; and since there is an $\mathbb{F}_q$-linearly dependent d-subset of columns of H, there is an $\mathbb{F}_q$-linearly dependent $(d+1)$-subset of columns of $\widehat{H}$.

In particular, for $q = 2$ the additional condition corresponding to the row 1_{n+1} in $\widehat{H}$ amounts to $\mathrm{wt}(v) \in \mathbb{N}_0$ even, for all $v \in \widehat{\mathcal{C}}$; hence if $\mathcal{C}$ is non-trivial then $\widehat{d}$ is even, implying $\widehat{d} = d + 1$ for d odd, and $\widehat{d} = d$ for d even.

Puncturing the extended code $\widehat{\mathcal{C}} \leq \mathbb{F}_q^{n+1}$ we recover $(\widehat{\mathcal{C}})^\bullet = \{[x_1, \ldots, x_n] \in \mathbb{F}_q^n; [x_1, \ldots, x_n, x_{n+1}] \in \widehat{\mathcal{C}}\} = \{[x_1, \ldots, x_n] \in \mathbb{F}_q^n; [x_1, \ldots, x_n] \in \mathcal{C}\} = \mathcal{C} \leq \mathbb{F}_q^n$.

b) i) Expurgating by throwing away certain codewords yields the code $\mathcal{C}' := \{[x_1, \ldots, x_n] \in \mathcal{C}; \sum_{i=1}^n x_i = 0\} \leq \mathcal{C} \leq \mathbb{F}_q^n$. Hence for the minimum distance of $\mathcal{C}'$ we have $d' \geq d$. Moreover, we have $k - 1 \leq k' := \dim_{\mathbb{F}_q}(\mathcal{C}') \leq k$. If $k' = k - 1$ then $\mathcal{C}'$ has check matrix $H' := \left[\begin{array}{c} H \\ \hline 1_n \end{array}\right] \in \mathbb{F}_q^{(n-k+1)\times n}$; in particular, if $1_n \in \mathcal{C}$ then we have $1_n \notin \mathcal{C}'$ if and only if $\gcd(q, n) = 1$, thus in this case $k' = k - 1$.

In particular, for $q = 2$ we have $\mathcal{C}' := \{v \in \mathcal{C}; \mathrm{wt}(v) \in \mathbb{N}_0 \text{ even}\} \leq \mathcal{C}$, being called the **even-weight subcode** of $\mathcal{C}$; hence if $\mathcal{C}'$ is non-trivial then d' is even, and we have $k' = k - 1$ if and only if $\mathcal{C}$ contains elements of odd weight.

ii) Augmenting by adding certain codewords yields $\widetilde{\mathcal{C}} := \langle \mathcal{C}, 1_n \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^n$. Hence for the minimum distance of $\widetilde{\mathcal{C}}$ we have $\widetilde{d} \leq d$. Moreover, we have $k \leq \widetilde{k} := \dim_{\mathbb{F}_q}(\widetilde{\mathcal{C}}) \leq k + 1$. We have $\widetilde{k} = k + 1$ if and only if $1_n \notin \mathcal{C}$; in this case $\widetilde{\mathcal{C}}$ has generator matrix $\widetilde{G} := \left[\begin{array}{c} G \\ \hline 1_n \end{array}\right] \in \mathbb{F}_q^{(k+1)\times n}$.

In particular, for $q = 2$ we have $\widetilde{\mathcal{C}} := \mathcal{C} \cup (1_n + \mathcal{C}) \leq \mathbb{F}_2^n$, consisting of the elements of $\mathcal{C}$ and their **complements**. If $\mathcal{C}$ is non-trivial, since $\mathrm{wt}(1_n + v) = n - \mathrm{wt}(v)$ for all $v \in \mathbb{F}_2^n$, we get $\widetilde{d} = \min\{d, n - D\}$, where $D := \max\{\mathrm{wt}(v) \in \mathbb{N}_0; 1_n \neq v \in \mathcal{C}\}$; if $1_n \in \mathcal{C}$ we have $1_n + \mathcal{C} = \mathcal{C}$ and thus $D = n - d$ and $\widetilde{d} = d$ anyway.

If $1_n \in \mathcal{C}$, then augmenting the expurgated code $\mathcal{C}' \leq \mathcal{C} \leq \mathbb{F}_q^n$ yields $\mathcal{C}' \leq \widetilde{(\mathcal{C}')} = \langle \mathcal{C}', 1_n \rangle_{\mathbb{F}_q} \leq \mathcal{C}$, hence $\widetilde{(\mathcal{C}')} = \mathcal{C}$ if $\gcd(q, n) = 1$, and $\widetilde{(\mathcal{C}')} = \mathcal{C}'$ if $\gcd(q, n) > 1$. Moreover, if $\gcd(q, n) > 1$, then we have $\langle 1_n \rangle'_{\mathbb{F}_q} = \langle 1_n \rangle_{\mathbb{F}_q}$, and thus expurgating the augmented code $\widetilde{\mathcal{C}} \leq \mathbb{F}_q^n$ yields $(\widetilde{\mathcal{C}})' = \langle \mathcal{C}, 1_n \rangle'_{\mathbb{F}_q} = \langle \mathcal{C}', 1_n \rangle_{\mathbb{F}_q} = \widetilde{(\mathcal{C}')}$.

c) i) Shortening by **taking a cross section**, for $n \geq 2$, is the concatenation of n**-expurgation** and subsequent puncturing, yielding the code $\mathcal{C}^\circ := \{[x_1, \ldots, x_{n-1}] \in \mathbb{F}_q^{n-1}; [x_1, \ldots, x_n] \in \mathcal{C}, x_n = 0\} = (\mathcal{C}^{(n)})^\bullet \leq \mathbb{F}_q^{n-1}$, where $\mathcal{C}^{(n)} := \{[x_1, \ldots, x_n] \in \mathcal{C}; x_n = 0\} \leq \mathbb{F}_q^n$. For the minimum distance of $\mathcal{C}^\circ$ we have $d^\circ \geq d$. Moreover, we have $k - 1 \leq k^\circ := \dim_{\mathbb{F}_q}(\mathcal{C}^\circ) \leq k$. We have $k^\circ = k - 1$ if and only if $\mathcal{C}^{(n)} < \mathcal{C}$, that is if and only if there is $[x_1, \ldots, x_n] \in \mathcal{C}$ such that $x_n \neq 0$; in this case $\mathcal{C}^\circ$ has check matrix $H^\circ \in \mathbb{F}_q^{(n-k)\times(n-1)}$ obtained from H by deleting column n, amounting to deleting an information symbol.

ii) Lengthening is the concatenation of augmentation and subsequent extension, yielding $\widehat{\widetilde{\mathcal{C}}} \leq \mathbb{F}_q^{n+1}$. For the minimum distance of $\widehat{\widetilde{\mathcal{C}}}$ we have $\widehat{\widetilde{d}} \leq d+1$. Moreover, we have $k \leq \widehat{\widetilde{k}} := \dim_{\mathbb{F}_q}(\widehat{\widetilde{\mathcal{C}}}) \leq k+1$. We have $\widehat{\widetilde{k}} = k+1$ if and only if $\mathcal{C} < \widetilde{\mathcal{C}}$, that is $1_n \notin \mathcal{C}$, amounting to adding an information symbol.

Thus, shortening the extended code $\widehat{\mathcal{C}} \leq \mathbb{F}_q^{n+1}$ yields $(\widehat{\mathcal{C}})^\circ = \{[x_1,\ldots,x_n] \in \mathbb{F}_q^n; [x_1,\ldots,x_{n+1}] \in \widehat{\mathcal{C}}, x_{n+1} = 0\} = \{[x_1,\ldots,x_n] \in \mathcal{C}, \sum_{i=1}^n x_i = 0\} = \mathcal{C}'$; in particular, shortening the lengthened code $\widehat{\widetilde{\mathcal{C}}} \leq \mathbb{F}_q^{n+1}$ yields $(\widehat{\widetilde{\mathcal{C}}})^\circ = (\widetilde{\mathcal{C}})'$.

5 Bounds for codes

(5.1) Theorem: Plotkin bound [1960]. Let $\mathcal{C}$ be a non-trivial (n,m,d)-code over an alphabet $\mathcal{X}$ such that $q := |\mathcal{X}|$. Then we have $m \cdot (d - n \cdot \frac{q-1}{q}) \leq d$.

In particular, if we have equality then $d(v,w) = d$ for all $v \neq w \in \mathcal{C}$, that is $\mathcal{C}$ is an **equidistant** code. Note that the above inequality is fulfilled for all $m \in \mathbb{N}$ whenever $\frac{d}{n} \leq \frac{q-1}{q}$, hence giving no obstruction at all in this case.

Proof. We compute two estimates of $\Delta := \sum_{[v,w]\in\mathcal{C}^2,\, v\neq w} d(v,w) \in \mathbb{N}$. Firstly, since $d(v,w) \geq d$ for all $v \neq w \in \mathcal{C}$, we get $\Delta \geq m(m-1)d$.

Secondly, letting $\mathcal{X} = \{x_1,\ldots,x_q\}$, let $m_{ij} \in \mathbb{N}_0$ be the number of occurrences of the symbol x_i, for $i \in \{1,\ldots,q\}$, in position $j \in \{1,\ldots,n\}$ of the various words in $\mathcal{C}$. Hence we have $\sum_{i=1}^q m_{ij} = m$, and the Cauchy-Schwarz inequality, applied to the tuples $[m_{ij}; i \in \{1,\ldots,q\}] \in \mathbb{R}^q$ and $1_q \in \mathbb{R}^q$, yields $m^2 = (\sum_{i=1}^q m_{ij})^2 \leq q \cdot \sum_{i=1}^q m_{ij}^2$, thus $\sum_{i=1}^q m_{ij}^2 \geq \frac{m^2}{q}$. Now, there are m_{ij} words in $\mathcal{C}$ having entry x_i at position j, and $m - m_{ij}$ words having a different entry there. This accounts for all contributions to Δ, hence we get $\Delta = \sum_{j=1}^n \sum_{i=1}^q m_{ij}(m - m_{ij}) = \sum_{j=1}^n (m^2 - \sum_{i=1}^q m_{ij}^2) \leq \sum_{j=1}^n m^2(1 - \frac{1}{q}) = nm^2 \cdot \frac{q-1}{q}$.

Thus we get $m(m-1)d \leq \Delta \leq nm^2 \cdot \frac{q-1}{q}$, entailing $(m-1)d \leq nm \cdot \frac{q-1}{q}$. In particular, equality implies that $\Delta = m(m-1)d$, thus $\mathcal{C}$ is equidistant. ♯

(5.2) Theorem: Griesmer bound [1960]. a) If $\mathcal{C} \leq \mathbb{F}_q^n$ is an $[n,k,d]$-code such that $k \geq 2$, there is an $[n-d, k-1, d^*]$-code $\mathcal{C}^* \leq \mathbb{F}_q^{n-d}$ such that $d^* \geq \lceil \frac{d}{q} \rceil$.
b) If $\mathcal{C} \leq \mathbb{F}_q^n$ is a non-trivial $[n,k,d]$-code, then we have $\sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil \leq n$.

Proof. a) We use the **Helgert-Stinaff construction** [1973]: Let $G \in \mathbb{F}_q^{k\times n}$ be a generator matrix of $\mathcal{C}$. Up to linear equivalence we may assume that $G = \left[\begin{array}{c|c} 1_d & 0_{n-d} \\ \hline G^{**} & G^* \end{array}\right]$, where $G^{**} \in \mathbb{F}_q^{(k-1)\times d}$ and $G^* \in \mathbb{F}_q^{(k-1)\times(n-d)}$; note that the Singleton bound yields $d < n$. We show that the **residual code** $\mathcal{C}^* \leq \mathbb{F}_q^{n-d}$ generated by the rows of G^* is an $[n-d, k-1, d^*]$-code such that $d^* \geq \lceil \frac{d}{q} \rceil$:

We first show that $\mathrm{rk}_{\mathbb{F}_q}(G^*) = k-1$: Assume to the contrary that $\mathrm{rk}_{\mathbb{F}_q}(G^*) \leq k-2$, then up to row operations may assume that $[G^{**} \mid G^*] = \left[\begin{array}{c|c} w & 0_{n-d} \\ \hline * & * \end{array}\right] \in \mathbb{F}_q^{(k-1)\times n}$, for some $w \in \mathbb{F}_q^d$. If $w = x \cdot 1_d \in \mathbb{F}_q^d$ for some $x \in \mathbb{F}_q^d$, then we have $\mathrm{rk}_{\mathbb{F}_q}(G) < k$, a contradiction; otherwise we have $0 \neq [w - x \cdot 1_d, 0_{n-d}] \in \mathcal{C}$ for all $x \in \mathbb{F}_q$, but $\mathrm{wt}(w - x \cdot 1_d) < d$ for some $x \in \mathbb{F}_q$, a contradiction as well.

We show that the minimum weight $\mathrm{wt}(\mathcal{C}^*) \in \mathbb{N}$ is bounded below by $\lceil \frac{d}{q} \rceil$: Let $0 \neq v \in \mathcal{C}^*$, and let $[w \mid v] \in \mathcal{C}$ for some $w \in \mathbb{F}_q^d$. Then for some $x \in \mathbb{F}_q$ there are at least $\lceil \frac{d}{q} \rceil$ entries of w equal to x, hence $\mathrm{wt}(w - x \cdot 1_d) \leq d - \lceil \frac{d}{q} \rceil$. Since $0 \neq [w - x \cdot 1_d \mid v] \in \mathcal{C}$ has weight at least d, we conclude that $\mathrm{wt}(v) \geq \lceil \frac{d}{q} \rceil$.

b) This now follows by induction on k, the case $k = 1$ being trivial: For $k \geq 2$ we have $n - d \geq \sum_{i=0}^{k-2} \lceil \frac{d^*}{q^i} \rceil$, thus $n \geq \lceil \frac{d}{q^0} \rceil + \sum_{i=0}^{k-2} \lceil \frac{d}{q^{i+1}} \rceil = \sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil$. ♯

We have a couple of immediate comments: Firstly, the Griesmer bound for linear codes improves the (non-linear) Plotkin bound: The former entails $n \geq d \cdot \sum_{i=0}^{k-1} \frac{1}{q^i} = d \cdot \frac{q^{-k}-1}{q^{-1}-1} = d \cdot \frac{q}{q^k} \cdot \frac{q^k-1}{q-1}$, or equivalently $nq^k \cdot \frac{q-1}{q} \geq d(q^k - 1)$, that is $d \geq q^k(d - n \cdot \frac{q-1}{q})$, which is the Plotkin bound.

Secondly, if $\mathcal{C}$ is an MDS code such that $k \geq 2$, then the Griesmer bound entails $d \leq q$, showing that the minimum distance of MDS codes is severely restricted: Assume to the contrary that $d > q$, then we have $n \geq \sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil \geq d + 2 + \sum_{i=2}^{k-1} 1 = d + 2 + (k-2) = d + k = n + 1$, a contradiction.

(5.3) Theorem: Gilbert bound [1952]. a) Let $\mathcal{X}$ be an alphabet such that $q := |\mathcal{X}|$, and let $n, m, d \in \mathbb{N}$ such that $2 \leq m \leq q^n$. If $m \cdot |B_{d-1}(\cdot)| = m \cdot \sum_{i=0}^{d-1} \binom{n}{i} \cdot (q-1)^i \leq q^n$, there is an (n, m, d')-code over $\mathcal{X}$ such that $d' \geq d$.
b) Let $n, k, d \in \mathbb{N}$ such that $k \leq n$. If $|B_{d-1}(0_n)| = \sum_{i=0}^{d-1} \binom{n}{i} \cdot (q-1)^i < q^{n-k+1}$, then there is an $[n, k, d']$-code over $\mathbb{F}_q$ such that $d' \geq d$,

Proof. a) We construct a suitable code successively, starting from a singleton set $\mathcal{C} \subseteq \mathcal{X}^n$; recall that $\mathcal{C}$ has infinite minimum distance. As long as $|\mathcal{C}| < m$ we have $|\mathcal{C}| \cdot |B_{d-1}(\cdot)| < q^n = |\mathcal{X}^n|$, hence $\bigcup_{v \in \mathcal{C}} B_{d-1}(v) \subset \mathcal{X}^n$ is a proper subset of $\mathcal{X}^n$. Thus there is $w \in \mathcal{X}^n$ having distance at least d from all elements of $\mathcal{C}$. Hence, since $\mathcal{C}$ has minimum distance at least d, this also holds for $\mathcal{C} \,\dot\cup\, \{w\}$.

b) The assumption implies that $d \leq n$. If $k = 1$ then the repetition $[n, 1, n]$-code is as desired; hence we may assume that $k \geq 2$. Then by induction there is an $[n, k-1, d']$-code $\mathcal{C} \leq \mathbb{F}_q^n$ such that $d' \geq d$. Since $q^{k-1} \cdot |B_{d-1}(0_n)| < q^n$, we conclude that $\bigcup_{v \in \mathcal{C}} B_{d-1}(v) \subset \mathbb{F}_q^n$ is a proper subset of $\mathbb{F}_q^n$. Thus there is $w \in \mathbb{F}_q^n$ having distance at least d from all elements of $\mathcal{C}$. Hence we have $\mathrm{wt}(w) \geq d$, and $aw \in \mathbb{F}_q^n$ has the same distance properties from $\mathcal{C}$, for all $a \in \mathbb{F}_q^*$. Thus we have $\mathcal{C} \cap \langle w \rangle_{\mathbb{F}_q} = \{0\}$, and we let $\mathcal{C}^+ := \mathcal{C} \oplus \langle w \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^n$. Hence we have $\dim_{\mathbb{F}_q}(\mathcal{C}^+) = k$, and $d(aw + v, bw + v') = d((a-b)w, v' - v) \geq d$, for all $a, b \in \mathbb{F}_q$

and $v, v' \in \mathcal{C}$ such that $a \neq b$ or $v \neq v'$, entails that $\mathcal{C}^+$ has minimum distance at least d, saying that $\mathcal{C}^+$ is as desired. $\sharp$

(5.4) Theorem: Gilbert-Varshamov bound [1952, 1957]. Let $n, k, d \in \mathbb{N}$ such that $k \leq n$. If $|B_{d-2}(0_{n-1})| = \sum_{i=0}^{d-2} \binom{n-1}{i} \cdot (q-1)^i < q^{n-k}$, where we let $B_{-1}(0_{n-1}) := \emptyset$, then there is an $[n, k, d']$-code over $\mathbb{F}_q$ such that $d' \geq d$.

Proof. The assumption implies that $d \leq n$, and we may assume that $d \geq 2$. For $k = n$, the above inequality is fulfilled if and only if $d = 1$, hence we may assume that $k < n$. We construct an $\mathbb{F}_q$-generating set $\mathcal{B}_n := \{v_1, \ldots, v_n\} \subseteq \mathbb{F}_q^{n-k}$ such that any $(d-1)$-subset of $\mathcal{B}_n$ is $\mathbb{F}_q$-linearly independent; then $[v_1^{\text{tr}}, \ldots, v_n^{\text{tr}}] \in \mathbb{F}_q^{(n-k)\times n}$ is the check matrix of a code as desired. We proceed by induction to find subsets $\mathcal{B}_{n-k} \subseteq \mathcal{B}_{n-k+1} \subseteq \cdots \subseteq \mathcal{B}_j \subseteq \cdots \subseteq \mathcal{B}_n$ of cardinality $|\mathcal{B}_j| = j$, where $\mathcal{B}_{n-k} := \{v_1, \ldots, v_{n-k}\} \subseteq \mathbb{F}_q^{n-k}$ is an $\mathbb{F}_q$-basis. For $j \in \{n-k, \ldots, n-1\}$, the number of vectors in $\mathbb{F}_q^{n-k}$ being $\mathbb{F}_q$-linear combinations of at most $d-2$ elements of $\mathcal{B}_j$ is at most $\sum_{i=0}^{d-2} \binom{j}{i} \cdot (q-1)^i \leq \sum_{i=0}^{d-2} \binom{n-1}{i} \cdot (q-1)^i < q^{n-k}$, hence there is $v_{j+1} \in \mathbb{F}_q^{n-k}$ such that any $(d-1)$-subset of $\mathcal{B}_{j+1} := \mathcal{B}_j \,\dot{\cup}\, \{v_{j+1}\}$ is $\mathbb{F}_q$-linearly independent. $\sharp$

We again have an immediate comment, saying that the Gilbert-Varshamov bound for linear codes improves the linear Gilbert bound: Given the inequality $\sum_{i=0}^{d-1} \binom{n}{i} \cdot (q-1)^i < q^{n-k+1}$, the latter ensures the existence of an $[n, k, d']$-code such that $d' \geq d$, while the former even ensures the existence of an $[n+1, k, d'']$-code $\mathcal{C}$ such that $d'' \geq d+1$, which indeed is an improvement, inasmuch, since $d'' \geq d+1 \geq 2$, the punctured code $\mathcal{C}^\bullet \leq \mathbb{F}_q^n$ has $\mathbb{F}_q$-dimension k and has minimum distance at least $d'' - 1 \geq d$.

(5.5) Optimal codes. Let $\mathbb{F}_q$ be the field with q elements, being kept fixed. For $n, d \in \mathbb{N}$ such that $d \leq n$ let

$$K_q(n, d) := \max\{k \in \mathbb{N};\ \text{there is an } [n, k, d']\text{-Code over } \mathbb{F}_q \text{ such that } d' \geq d\};$$

note that the existence of repetition $[n, 1, n]$-codes entails that $K_q(n, d) \leq n$ is well-defined. In particular, for $d = 1$ the $[n, n, 1]$-code $\mathbb{F}_q^n$ shows that $K_q(n, 1) = n$; for $d = n$ the Singleton bound implies $k \leq 1$, showing that $K_q(n, n) = 1$.

We have $K_q(n, d) = \max\{k \in \mathbb{N};\ \text{there is an } [n, k, d]\text{-Code over } \mathbb{F}_q\}$, as well as $K_q(n, d+1) \leq K_q(n, d)$: Let $\mathcal{C}$ be an $[n, k, d+1]$-code, where we may assume that there is $[x_1, \ldots, x_n] \in \mathcal{C}$ such that $x_n \neq 0$ and having minimal weight $d+1 \geq 2$. Then the punctured code $\mathcal{C}^\bullet \leq \mathbb{F}_q^{n-1}$ is an $[n-1, k, d]$-code. Hence adding an n-th component consisting of zeroes only we get the $[n, k, d]$-code $\{[x_1, \ldots, x_{n-1}, 0] \in \mathbb{F}_q^n; [x_1, \ldots, x_{n-1}] \in \mathcal{C}^\bullet\} \leq \mathbb{F}_q^n$, proving both assertions.

Upper bounds on $k \geq 1$ for an $[n, k, d]$-code over $\mathbb{F}_q$ to possibly exist, thus upper bounds on $K_q(n, d)$ are, where (i)–(iii) also hold for non-linear codes:

i) Singleton bound $k \leq n-d+1$, **ii)** Hamming bound $\sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i} \cdot (q-1)^i \leq q^{n-k}$,

iii) Plotkin bound $q^k(d - n \cdot \frac{q-1}{q}) \leq d$, **iv)** Griesmer bound $\sum_{i=0}^{k-1} \lceil \frac{d}{q^i} \rceil \leq n$.

Lower bounds on k for an $[n, k, d']$-code over $\mathbb{F}_q$ such that $d' \geq d$ to exist, that is lower bounds on $K_q(n, d)$ are, where (v) also holds for non-linear codes:

v) Gilbert bound $\sum_{i=0}^{d-1} \binom{n}{i} \cdot (q-1)^i \leq q^{n-k}$, **vi)** linear Gilbert bound $\sum_{i=0}^{d-1} \binom{n}{i} \cdot (q-1)^i < q^{n+1-k}$, **vii)** Gilbert-Varshamov bound $\sum_{i=0}^{d-2} \binom{n-1}{i} \cdot (q-1)^i < q^{n-k}$.

A code $\mathcal{C} \leq \mathbb{F}_q^n$ such that $d(\mathcal{C}) = d$ and $\dim_{\mathbb{F}_q}(\mathcal{C}) = K_q(n, d)$ is called **optimal**. Thus any linear code fulfilling one of the upper bounds mentioned is optimal; for example MDS codes and perfect codes are so, but for these classes of codes d is severely restricted. Actually, $K_q(n, d)$ is not at all known precisely, and improving the bounds for $K_q(n, d)$, aiming at determining $K_q(n, d)$ affirmatively, and finding and classifying optimal codes continue to be major problems of combinatorial coding theory.

Example. For $q = 2$ and $n = 13$ and $d = 5$ we get the following:

i) The Singleton bound yields $k \leq 9$; **ii)** the Hamming bound yields $2^{13-k} \geq \sum_{i=0}^{2} \binom{13}{i} = 92$, that is $k \leq 6$; **iv)** the Griesmer bound yields $\sum_{i=0}^{k-1} \lceil \frac{5}{2^i} \rceil = 5 + 3 + 2 + 1 + 1 + 1 + \cdots \leq 13$, that is $k \leq 6$.

iii) Since $5 \leq \frac{13}{2}$, the Plotkin bound does not yield immediately. But, noting that we may assume that $k \geq 4$ by the lower bounds to follow, if we first extend a binary $[13, k, 5]$-code to a $[14, k, 6]$-code, and then shorten it three times to obtain a $[11, k-3, 6]$-code (and possibly puncturing and adding zero components), then the Plotkin bound yields $2^{k-3}(6 - \frac{11}{2}) \leq 6$, that is $k \leq 6$.

Assume that there is a binary $[13, 6, 5]$-code. Then by the Helgert-Stinaff construction there is a residual $[8, 5, d]$-code, where $d \geq 3$. Now the Hamming bound $9 = \sum_{i=0}^{1} \binom{8}{i} \leq \sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{8}{i} \leq 2^{8-5} = 8$ yields a contradiction. Hence there is no binary $[13, 6, 5]$-code, so that $k \leq 5$.

Conversely, **v)** the Gilbert bound yields $2^{13-k} \geq \sum_{i=0}^{4} \binom{13}{i} = 1093$, that is $k \geq 2$; **vi)** the linear Gilbert bound yields $2^{14-k} > 1093$, that is $k \geq 3$; **vii)** the Gilbert-Varshamov bound yields $2^{13-k} > \sum_{i=0}^{3} \binom{12}{i} = 299$, that is $k \geq 4$.

We finally show that a binary $[13, 5, 5]$-code exists: In view of the Helgert-Stinaff construction we begin with a binary $[8, 4, 4]$-code, namely the (self-dual) **extended Hamming code** $\mathcal{C}^* := \widehat{\mathcal{H}}_3 \leq \mathbb{F}_2^8$, which has generator and check matrices as follows; see also (6.2), and (4.2) where $(\widehat{\mathcal{H}}_3)^\bullet = \mathcal{H}_3 \leq \mathbb{F}_2^7$ is given:

$$G^* := \left[\begin{array}{ccccccc|c} . & . & . & 1 & 1 & 1 & 1 & . \\ . & 1 & 1 & 1 & 1 & . & . & . \\ 1 & 1 & . & . & 1 & 1 & . & . \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right] \qquad H^* := \left[\begin{array}{cccccccc} . & . & . & 1 & 1 & 1 & 1 & . \\ . & 1 & 1 & . & . & 1 & 1 & . \\ 1 & . & 1 & . & 1 & . & 1 & . \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right]$$

As any 3-subset of columns of H^* is $\mathbb{F}_2$-linearly independent, we conclude that indeed $d(\mathcal{C}^*) = 4$. Now let $\mathcal{C} \leq \mathbb{F}_2^{13}$ be the code generated by

$$G := \left[\begin{array}{ccccc|cccccccc} 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . & . & . \\ \hline . & . & 1 & . & . & . & . & . & 1 & 1 & 1 & 1 & . \\ . & . & . & 1 & . & . & 1 & 1 & 1 & 1 & . & . & . \\ . & . & . & . & 1 & 1 & 1 & . & . & 1 & 1 & . & . \\ . & . & . & . & . & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right] \in \mathbb{F}_2^{4 \times 13}.$$

We show that $d(\mathcal{C}) = 5$: Let $0 \neq v = [v' \mid v''] \in \mathcal{C}$, where $v' \in \mathbb{F}_2^5$ and $v'' \in \mathbb{F}_2^8$. If $v' = 0$, then we have $v'' = 1_8$, hence $\mathrm{wt}(v) = 8$; if $v'' = 0$, then we have $v' = 1_5$, hence $\mathrm{wt}(v) = 5$; if both $v' \neq 0$ and $v'' \neq 0$, then we have $\mathrm{wt}(v) = \mathrm{wt}(v') + \mathrm{wt}(v'') \geq 1 + 4 = 5$.

Hence we conclude that $K_2(13, 5) = 5$. We note that there is a (unique optimal non-linear) binary $(13, 2^6, 5)$-code, which via shortening is related to the (optimal non-linear) binary **Nadler** $(12, 2^5, 5)$-code. ♯

(5.6) Asymptotic bounds. We consider the question how good codes might be asymptotically for $n \gg 0$. Since the error correction capabilities of a non-trivial $[n, k, d]$-code $\mathcal{C} \leq \mathbb{F}_q^n$, which are governed by its minimum distance d, should grow proportionally with respect to its length n, we let $\delta(\mathcal{C}) := \frac{d}{n} \leq 1$ be the **relative minimum distance** of $\mathcal{C}$; recall that $\rho(\mathcal{C}) = \frac{k}{n} \leq 1$ is the information rate of $\mathcal{C}$.

For $0 \leq \delta \leq 1$ we let $\kappa_q(\delta) := \limsup_{n \to \infty} \frac{1}{n} \cdot K_q(n, \lceil \delta n \rceil)$, that is

$$\kappa_q(\delta) = \limsup_{n \to \infty} \Big(\max \Big\{ \frac{k}{n} \in \mathbb{R};\ \text{there is an } [n, k, d]\text{-code such that } \frac{d}{n} \geq \delta \Big\} \Big).$$

Since we may assume that $d = \lceil \delta n \rceil$, this amounts to saying that $0 \leq \kappa_q(\delta) \leq 1$ is largest such that there is a sequence of codes of unbounded length whose relative minimum distance is approaches δ from above, and whose information rate tends towards $\kappa_q(\delta)$.

Hence $\kappa_q(\delta)$ is decreasing, where for $\delta = 0$ from $K_q(n, 1) = n$ we get $\kappa_q(0) = 1$, while for $\delta = 1$ from $K_q(n, n) = 1$ we get $\kappa_q(1) = 0$. Again, the bounds (i)–(iv) above provide upper bounds for $\kappa_q(\delta)$, while (v)–(vii) give lower bounds for $\kappa_q(\delta)$, but in general $\kappa_q(\delta)$ is not known. We proceed to derive the associated asymptotic bounds explicitly; they are depicted for $q = 2$ in Table 4.

(5.7) Linear bounds. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a non-trivial $[n, k, d]$-code, and $0 \leq \delta \leq 1$.

Theorem: Asymptotic Singleton bound. We have $\kappa_q(\delta) \leq 1 - \delta$.

Proof. The Singleton bound says that $\rho(\mathcal{C}) = \frac{k}{n} \leq 1 + \frac{1}{n} - \frac{d}{n} = 1 + \frac{1}{n} - \delta(\mathcal{C})$. Hence, given $0 \leq \delta \leq 1$, whenever $\delta(\mathcal{C}) \geq \delta$ we have $\rho(\mathcal{C}) \leq 1 + \frac{1}{n} - \delta$. Thus for $n \to \infty$ we get $\kappa_q(\delta) \leq 1 - \delta$. ♯

Next, we consider the Plotkin bound, yielding the asymptotic result to follow, which actually supersedes the asymptotic Singleton bound. The Griesmer bound, although for specific cases often being better than the Plotkin bound, yields the same asymptotic bound; indeed, asymptotically there is no loss in applying the estimate used to show in the first place that the Griesmer bound implies the Plotkin bound.

Theorem: Asymptotic Plotkin bound. We have $\kappa_q(\delta) = 0$ for $\frac{q-1}{q} \leq \delta \leq 1$, and $\kappa_q(\delta) \leq 1 - \frac{q}{q-1} \cdot \delta$ for $0 \leq \delta \leq \frac{q-1}{q}$.

Proof. We may assume that $\delta \not\in \{0, \frac{q-1}{q}\}$. Let first $\frac{q-1}{q} < \delta \leq \delta(\mathcal{C})$. The Plotkin bound $q^k(d - n \cdot \frac{q-1}{q}) \leq d$ can be written as $q^k(\frac{d}{n} - \frac{q-1}{q}) \leq \frac{d}{n}$, in other words $q^k \leq \frac{\delta(\mathcal{C})}{\delta(\mathcal{C}) - \frac{q-1}{q}}$. This entails $q^k \leq \frac{\delta}{\delta - \frac{q-1}{q}}$, thus k is bounded above, implying that $\rho(\mathcal{C}) = \frac{k}{n} \to 0$, for $n \to \infty$, hence $\kappa_q(\delta) = 0$.

Let now $0 < \delta \leq \delta(\mathcal{C}) < \frac{q-1}{q}$, and we may assume that $d \geq 2$. Letting $n' := \lfloor \frac{q(d-1)}{q-1} \rfloor$, we get $1 \leq n' \leq \frac{q(d-1)}{q-1} \leq \frac{n(d-1)}{d} < n$ and $d - n' \cdot \frac{q-1}{q} = d - \frac{q-1}{q} \cdot \lfloor \frac{q(d-1)}{q-1} \rfloor \geq d - \frac{q-1}{q} \cdot \frac{q(d-1)}{q-1} = 1$. Shortening $n - n'$ times we get an $[n', k', d']$-code where $k' \geq k - (n - n')$ and $d' \geq d$. Hence there also is an $[n', k', d]$-code, which by the Plotkin bound fulfills $q^{k'} \leq \frac{d}{d - n' \cdot \frac{q-1}{q}} \leq d$. Thus we have $q^k \leq q^{k'} q^{n-n'} \leq dq^{n-n'}$, hence $k \leq n - n' + \log_q(d)$, thus $\rho(\mathcal{C}) = \frac{k}{n} \leq 1 - \frac{n'}{n} + \frac{1}{n} \cdot \log_q(d)$. We have $\lim_{n\to\infty} \frac{n'}{n} = \lim_{n\to\infty} \frac{1}{n} \cdot \lfloor \frac{q(d-1)}{q-1} \rfloor = \frac{q}{q-1} \cdot \lim_{n\to\infty} \frac{d-1}{n} = \frac{q}{q-1} \cdot \lim_{n\to\infty} \delta(\mathcal{C}) = \frac{q}{q-1} \cdot \delta$, recall that we may assume that $\delta(\mathcal{C}) \to \delta$. Since $\frac{1}{n} \cdot \log_q(d) \to 0$ anyway, from this we infer that $\kappa_q(\delta) \leq 1 - \frac{q}{q-1} \cdot \delta$. ♯

(5.8) Bounds based on sphere packing. a) In order to proceed, we need a few preparations: Let $q \in \mathbb{R}$ such that $q > 1$. For $0 < \alpha \leq \frac{q-1}{q} = 1 - \frac{1}{q}$ let $H_q(\alpha) := \alpha \log_q(q-1) - \alpha \log_q(\alpha) - (1-\alpha)\log_q(1-\alpha)$; since $\lim_{\alpha\to 0^+} H_q(\alpha) = 0$ we extend $H_q(\alpha)$ continuously for $\alpha = 0$ by letting $H_q(0) := 0$. We have $H_q(\frac{q-1}{q}) = \frac{q-1}{q} \cdot \log_q(q-1) - \frac{q-1}{q} \cdot \log_q(\frac{q-1}{q}) - \frac{1}{q}\log_q(\frac{1}{q}) = 1$.

Differentiating yields $\frac{\partial}{\partial\alpha} H_q(\alpha) = \log_q(q-1) + \log_q(\frac{1-\alpha}{\alpha})$, for $0 < \alpha \leq \frac{q-1}{q}$. Hence we have $\frac{\partial}{\partial\alpha} H_q(\alpha) = 0$ if and only if $\frac{\alpha}{1-\alpha} = q - 1$, or equivalently $\alpha = \frac{q-1}{q}$. Since $\frac{\partial}{\partial\alpha} H_q(\alpha)$ is strictly decreasing, such that $\frac{\partial}{\partial\alpha} H_q(\alpha) \to \infty$, for $\alpha \to 0^+$, we conclude that $\frac{\partial}{\partial\alpha} H_q(\alpha) > 0$ for all $0 < \alpha < \frac{q-1}{q}$, implying that $H_q(\alpha)$ is strictly increasing and strictly concave for $0 \leq \alpha \leq \frac{q-1}{q}$.

Note that for $q = 2$ we recover the conditional entropy $H(\mathcal{Y}|\mathcal{X}) = -p\log_2(p) - (1-p)\log_2(1-p) = H_2(p)$ of a symmetric binary channel with error probability $0 \leq p \leq \frac{1}{2}$. Actually, $H_q(p)$ is the conditional entropy of a **symmetric q-ary** channel with error probability $0 \leq p \leq \frac{q-1}{q}$, for which reason H_q is also called the associated **entropy function**.

Lemma. Let $0 < \delta < \frac{q-1}{q}$, and let $[d_n \in \mathbb{N}; n \in \mathbb{N}]$ be a sequence of integers such that $\frac{d_n}{n} \to \delta$, for $n \to \infty$. Then we have

$$\frac{1}{n} \cdot \log_q(|B_{d_n}(0_n)|) = \frac{1}{n} \cdot \log_q \left(\sum_{i=0}^{d_n} \binom{n}{i} \cdot (q-1)^i \right) \to H_q(\delta), \quad \text{for} \quad n \to \infty.$$

Proof. Let $0 \leq i \leq j \leq n \cdot \frac{q-1}{q}$. Then we have $\binom{n}{i} \cdot (q-1)^i \leq \binom{n}{j} \cdot (q-1)^j$: The assertion is equivalent to $\prod_{s=i+1}^{j} \frac{s}{n+1-s} = \frac{j!}{i!} \cdot \frac{(n-j)!}{(n-i)!} \leq (q-1)^{j-i}$. The factors of the product are increasing with s increasing, hence the left hand side is bounded above by $(\frac{j}{n+1-j})^{j-i}$. Now, from $j \leq n \cdot \frac{q-1}{q}$ we get $j + j(q-1) = jq \leq n(q-1) \leq (n+1)(q-1)$, implying $j \leq (n+1-j)(q-1)$, thus $\frac{j}{n+1-j} \leq q-1$. Hence we conclude that the left hand side indeed is bounded above by $(q-1)^{j-i}$.

This yields $\binom{n}{j} \cdot (q-1)^j \leq \sum_{i=0}^{j} \binom{n}{i} \cdot (q-1)^i \leq (j+1) \cdot \binom{n}{j} \cdot (q-1)^j$ for $j \leq n \cdot \frac{q-1}{q}$. In particular, this applies to $j := d_n$ for $n \gg 0$. Since $\frac{1}{n} \cdot \log_q(n \cdot \frac{q-1}{q}) \to 0$, for $n \to \infty$, it suffices to show that $L_n := \frac{1}{n} \cdot \log_q(\binom{n}{j} \cdot (q-1)^j) \to H_q(\delta)$.

To this end, note that Stirling's formula $\lim_{n \to \infty} \frac{n! \cdot e^n}{n^n \cdot \sqrt{2\pi n}} = 1$ implies that $n! = (\frac{n}{e})^n \cdot \sqrt{2\pi n} \cdot (1 + o(1))$, for $n \to \infty$, where $o(1) \colon \mathbb{N} \to \mathbb{R}$ fulfills $o(1) \to 0$, for $n \to \infty$. Thus we get $\log_q(n!) = (n + \frac{1}{2}) \log_q(n) - n \log_q(e) + \log_q(\sqrt{2\pi}) + o(1)$.

Recalling that $j = d_n$, we have $j \to \infty$, for $n \to \infty$; and since $n - j \geq n \cdot \frac{1}{q}$ we also have $n - j \to \infty$, for $n \to \infty$. Thus we obtain $\log_q(\binom{n}{j} \cdot (q-1)^j) = (n + \frac{1}{2}) \log_q(n) - (j + \frac{1}{2}) \log_q(j) - (n - j + \frac{1}{2}) \log_q(n-j) + j \log_q(q-1) - \log_q(\sqrt{2\pi}) + o(1)$. This entails $L_n = \log_q(n) - \frac{j}{n} \cdot \log_q(\frac{j}{n} \cdot n) - \frac{n-j}{n} \cdot \log_q(\frac{n-j}{n} \cdot n) + \frac{j}{n} \cdot \log_q(q-1) + o(1) = -\frac{j}{n} \cdot \log_q(\frac{j}{n}) - \frac{n-j}{n} \cdot \log_q(\frac{n-j}{n}) + \frac{j}{n} \cdot \log_q(q-1) + o(1)$. Since $\frac{j}{n} \to \delta$, for $n \to \infty$, this yields $L_n \to \delta \log_q(q-1) - \delta \log_q(\delta) - (1-\delta) \log_q(1-\delta) = H_q(\delta)$. ♯

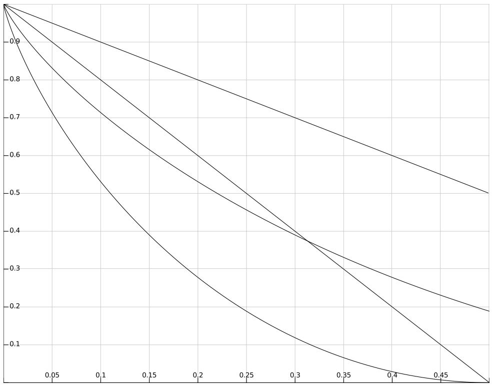
b) Now let $\mathcal{C} \leq \mathbb{F}_q^n$ be a non-trivial $[n, k, d]$-code, and $0 \leq \delta \leq 1$.

Theorem: Asymptotic Hamming bound. We have $\kappa_q(\delta) \leq 1 - H_q(\frac{\delta}{2})$.

Proof. We may assume that $0 < \delta < 1$. Since for $a \geq 0$ we have $2 \cdot \lceil \frac{a}{2} \rceil \leq \lceil a \rceil + 1$, we may weaken the condition $d \geq \lceil \delta n \rceil$ by just assuming that $d \geq 2 \cdot \lceil \frac{\delta}{2} \cdot n \rceil - 1$. The Hamming bound $q^k \cdot |B_{\lfloor \frac{d-1}{2} \rfloor}(0_n)| \leq q^n$ yields $k \leq n - \log_q(|B_{\lfloor \frac{d-1}{2} \rfloor}(0_n)|) \leq n - \log_q(|B_{\lceil \frac{\delta}{2} \cdot n \rceil - 1}(0_n)|)$, that is $\rho(\mathcal{C}) = \frac{k}{n} \leq 1 - \frac{1}{n} \cdot \log_q(|B_{\lceil \frac{\delta}{2} \cdot n \rceil - 1}(0_n)|)$. Since $\frac{1}{n} \cdot (\lceil \frac{\delta}{2} \cdot n \rceil - 1) \to \frac{\delta}{2}$, for $n \to \infty$, we get $\kappa_q(\delta) \leq 1 - H_q(\frac{\delta}{2})$. ♯

Finally, we provide an asymptotic lower bound, which is based on the Gilbert-Varshamov bound. The weaker estimates given by the Gilbert and linear Gilbert bounds yield the same asymptotic bound; indeed, the estimates used in the proof given below show that essentially the Gilbert bound is used.

Table 4: Asymptotic bounds for $q = 2$.



Theorem: Asymptotic Gilbert-Varshamov bound. For $0 \leq \delta \leq \frac{q-1}{q}$ we have $\kappa_q(\delta) \geq 1 - H_q(\delta)$.

Proof. We may assume that $0 < \delta < \frac{q-1}{q}$ and that $d := \lceil \delta n \rceil \geq 2$. Let $k \in \mathbb{N}$ be maximal such that $|B_{d-2}(0_{n-1})| < q^{n-k}$, then the Gilbert-Varshamov bound there exists an $[n, k, d]$-code $\mathcal{C}$. By maximality we have $q^{k+1} \geq \frac{q^n}{|B_{d-2}(0_{n-1})|}$, and thus $\rho(\mathcal{C}) = \frac{k+1}{n} - \frac{1}{n} = -\frac{1}{n} + \frac{1}{n} \cdot \log_q(q^{k+1}) \geq \frac{n-1}{n} - \frac{1}{n} \cdot \log_q(|B_{d-2}(0_{n-1})|) \geq \frac{n-1}{n} - \frac{1}{n} \cdot \log_q(|B_d(0_n)|)$. Since $\frac{d}{n} \to \delta$, for $n \to \infty$, we get $\kappa_q(\delta) \geq 1 - H_q(\delta)$. $\sharp$

(5.9) Remark. a) We mention a few further, better asymptotic bounds:

i) The **asymptotic Elias-Bassalygo bound** [1967], being based on an improvement of the strategy used to prove the Plotkin bound, says that for $0 \leq \delta \leq \frac{q-1}{q}$ we have $\kappa_q(\delta) \leq 1 - H_q(\frac{q-1}{q} - \sqrt{\frac{q-1}{q} \cdot (\frac{q-1}{q} - \delta)})$; for $q = 2$ we get $\kappa_2(\delta) \leq 1 - H_2(\frac{1}{2}(1 - \sqrt{1 - 2\delta}))$. This improves the Hamming and Plotkin bounds, and was the best asymptotic upper bound at that time.

For $q = 2$ (and $0.15 \sim \delta_0 < \delta$) the latter is superseded by the **asymptotic McEliece-Rodemich-Rumsey-Welch bound** [1977], based on the **linear programming bound** [Delsarte, 1973], saying that for $0 \leq \delta \leq \frac{1}{2}$ we have $\kappa_q(\delta) \leq H_2(\frac{1}{2} - \sqrt{\delta(1-\delta)})$, and being the best asymptotic upper bound known.

ii) On the other side, the asymptotic Gilbert-Varshamov bound was long considered to be the best possible asymptotic lower bound. But using **algebraic-geometric codes**, which in turn are based on the idea of **Goppa codes** [1981], Tsfasman-Vladut-Zink [1982] provided a series of linear codes over $\mathbb{F}_{p^2}$, where $p \geq 7$, which exceed the asymptotic Gilbert-Varshamov bound. These are still the asymptotically best codes known.

b) Finally, a similar approach works for non-linear codes: Let $\mathcal{X}$ be a fixed alphabet such that $q := |\mathcal{X}|$. For $n, d \in \mathbb{N}$ such that $d \leq n$ let

$$M(n,d) := \max\{m \in \mathbb{N}; \text{ there is an } (n,m,d')\text{-Code over } \mathcal{X} \text{ such that } d' \geq d\},$$

and for $0 \leq \delta \leq 1$ let $\mu(\delta) := \limsup_{n\to\infty} \frac{1}{n} \cdot \log_q(M(n, \lceil \delta n \rceil))$, that is

$$\mu(\delta) = \limsup_{n\to\infty} \Big(\max\Big\{ \frac{\log_q(m)}{n} \in \mathbb{R}; \text{ there is } (n,m,d)\text{-code such that } \frac{d}{n} \geq \delta \Big\} \Big).$$

Since the Singleton, Hamming, Plotkin and Gilbert bounds are all non-linear bounds, the asymptotic bounds given above also hold for non-linear codes. Similarly, the Elias-Bassalygo and McEliece-Rodemich-Rumsey-Welch bounds hold for non-linear codes.

6 Hamming codes

(6.1) Hamming codes [1950]. Let $\mathbf{P}^{k-1}(\mathbb{F}_q) := \{\langle v \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^k; 0 \neq v \in \mathbb{F}_q^k\}$ be the $(k-1)$-dimensional **projective space** over $\mathbb{F}_q$, where $k \geq 2$, and let $n := |\mathbf{P}^{k-1}(\mathbb{F}_q)| = \frac{q^k-1}{q-1} \geq 3$. Let $H_k \in \mathbb{F}_q^{k\times n}$ having columns being in bijection with $\mathbf{P}^{k-1}(\mathbb{F}_q)$; note that H_k is unique up to linear equivalence. Thus we have $\mathrm{rk}_{\mathbb{F}_q}(H_k) = k$, and any 2-set of columns of H_k is $\mathbb{F}_q$-linearly independent, while there are $\mathbb{F}_q$-linearly dependent 3-sets of columns.

Let the **Hamming code** $\mathcal{H}_k \leq \mathbb{F}_q^n$ be defined by having check matrix H_k, hence being unique up to linear equivalence. Thus $\mathcal{H}_k$ is an $[n, n-k, 3]$-code, which since $q^{n-k} \cdot \sum_{i=0}^{1} \binom{n}{i} \cdot (q-1)^i = q^{n-k}(1 + n(q-1)) = q^{n-k}(1 + (q-1) \cdot \frac{q^k-1}{q-1}) = q^{n-k}q^k = q^n$ is perfect; see also (3.5).

Conversely, any $[n, n-k, 3]$-code $\mathcal{C}$ is linearly equivalent to $\mathcal{H}_k$: Let $H \in \mathbb{F}_q^{k\times n}$ be a check matrix of $\mathcal{C}$, then any 2-set of columns of H is $\mathbb{F}_q$-linearly independent, that is the columns of H generate pairwise distinct 1-dimensional subspaces of $\mathbb{F}_q^k$, and hence the $n = |\mathbf{P}^{k-1}(\mathbb{F}_q)|$ columns are in bijection with $\mathbf{P}^{k-1}(\mathbb{F}_q)$.

We may choose the columns of $H_k \in \mathbb{F}_q^{k\times n}$ according to the q-ary representation of the integers in $\{1, \ldots, q^k - 1\}$ having 1 as their highest digit. This allows for a fast decoding algorithm: Since $\mathcal{H}_k$ has minimum distance 3, the $(q-1)\cdot n = q^k - 1$ vectors in $\mathbb{F}_q^n$ of weight 1 belong to pairwise distinct non-trivial cosets of $\mathcal{H}_k$. Since there are $q^{n-(n-k)} - 1 = q^k - 1$ such cosets, this induces a bijection between the vectors of weight 1 and the non-trivial cosets. Thus the non-trivial coset leaders are precisely the vectors $xe_1, \ldots, xe_n \in \mathbb{F}_q^n$, where $x \in \mathbb{F}_q^*$ and $e_i \in \mathbb{F}_q^n$ is the i-th unit vector, for $i \in \{1, \ldots, n\}$. Now given $v = w + xe_i \in \mathbb{F}_q^n \setminus \mathcal{H}_k$, where $w \in \mathcal{H}_k$, the associated syndrome is $vH_k^{\mathrm{tr}} = xe_iH_k^{\mathrm{tr}} \in \mathbb{F}_q^k$, that is the transpose of the x-fold of the i-th column of H_k, which can be translated into the q-ary representation of the i-th integer with highest digit 1 and the scalar x, revealing both the position of the error and saying how to correct it.

(6.2) Binary Hamming codes. Keeping the notation of (6.1), let $q := 2$. Ordering the columns of $H_k \in \mathbb{F}_2^{k \times n}$, where $k \geq 2$ and $n = 2^k - 1$, according to the binary representation of $i \in \{1, \ldots, n\}$, letting $H_1 := [1] \in \mathbb{F}_2^{1 \times 1}$, we recursively get

$$[0_k^{\mathrm{tr}} \mid H_k] = \left[\begin{array}{c|c|c|c} \multicolumn{2}{c|}{0_{2^{k-1}}} & \multicolumn{2}{c}{1_{2^{k-1}}} \\ \hline 0_{k-1}^{\mathrm{tr}} & H_{k-1} & 0_{k-1}^{\mathrm{tr}} & H_{k-1} \end{array}\right] \in \mathbb{F}_2^{k \times 2^k};$$

for example, we get $H_2 = \begin{bmatrix} . & 1 & 1 \\ 1 & . & 1 \end{bmatrix}$ and $H_3 = \begin{bmatrix} . & . & . & 1 & 1 & 1 & 1 \\ . & 1 & 1 & . & . & 1 & 1 \\ 1 & . & 1 & . & 1 & . & 1 \end{bmatrix}$.

Note that $\mathcal{H}_2$ is a $[3, 1, 3]$-code, hence is the repetition code, and that the $[7, 4, 3]$-code $\mathcal{H}_3$ has been considered in (4.2); moreover, the $[6, 3, 3]$-code obtained by shortening $\mathcal{H}_3$ with respect to the 4-th component has been considered in (3.4).

For $k \geq 2$, all rows of H_k have weight 2^{k-1}, which is even, hence we have $1_n \in \mathcal{H}_k$; in particular we have $\widetilde{\mathcal{H}}_k = \mathcal{H}_k$. For $k \geq 3$ we get $H_k H_k^{\mathrm{tr}} = 0 \in \mathbb{F}_2^{k \times k}$, that is $\mathcal{H}_k^\perp \leq \mathcal{H}_k$: Letting $H_k = [w_1, \ldots, w_k] \in \mathbb{F}_2^{k \times n}$, using the standard $\mathbb{F}_2$-bilinear form we get $\langle w_i, w_i \rangle = \mathrm{wt}(w_i) = 2^{k-1} = 0 \in \mathbb{F}_2$ for all $i \in \{1, \ldots, k\}$, as well as $\langle w_1, w_i \rangle = 2^{k-2} = 0 \in \mathbb{F}_2$ and $\langle w_i, w_j \rangle = 0 \in \mathbb{F}_2$ for $j > i \geq 2$.

We apply the modifications described in (4.5), see Table 5:

i) Extending $\mathcal{H}_k \leq \mathbb{F}_2^n$ yields the **extended Hamming code** $\widehat{\mathcal{H}}_k \leq \mathbb{F}_2^{n+1}$, an $[n+1, n-k, 4]$-code, and puncturing $\widehat{\mathcal{H}}_k$ yields $(\widehat{\mathcal{H}}_k)^\bullet = \mathcal{H}_k$ again; note that $\widehat{\mathcal{H}}_2$ is a $[4, 1, 4]$-code, hence is the repetition code. For $k \geq 3$, since $\langle 1_{n+1}, 1_{n+1} \rangle = 0 \in \mathbb{F}_2$ and $\langle 1_{n+1}, [w_i \mid 0] \rangle = \mathrm{wt}(w_i) = 0 \in \mathbb{F}_2$ for all $i \in \{1, \ldots, k\}$, the associated check matrix $\widehat{H}_k \in \mathbb{F}_2^{(k+1) \times (n+1)}$ fulfills $\widehat{H}_k \widehat{H}_k^{\mathrm{tr}} = 0 \in \mathbb{F}_2^{(k+1) \times (k+1)}$, that is $(\widehat{\mathcal{H}}_k)^\perp \leq \widehat{\mathcal{H}}_k$. In particular, since $\dim((\widehat{\mathcal{H}}_3)^\perp) = 4 = \dim(\widehat{\mathcal{H}}_3)$, we conclude that $(\widehat{\mathcal{H}}_3)^\perp = \widehat{\mathcal{H}}_3$ is a self-dual $[8, 4, 4]$-code.

ii) Expurgating $\mathcal{H}_k \leq \mathbb{F}_2^n$ yields the **even-weight Hamming code** $\mathcal{H}'_k \leq \mathbb{F}_2^n$. Since $1_n \in \mathcal{H}_k$ and n is odd, we conclude that $\mathcal{H}'_k$ is an $[n, n-k-1, d']$-code, where $d' \geq 3$, and augmenting $\mathcal{H}'_k$ yields $\widetilde{(\mathcal{H}'_k)} = \mathcal{H}_k$ again. While $\mathcal{H}'_2 \leq \mathbb{F}_2^3$ is the trivial code, for $k \geq 3$ we have $d' = 4$:

Since $d' \geq 3$ is even, it suffices to show that $d' \leq 4$: We choose an $\mathbb{F}_2$-linearly dependent 4-subset of columns of H_k, such that all its 3-subsets are $\mathbb{F}_2$-linearly independent; for example we may take the following and extend by zeroes below:

$$\begin{bmatrix} . & . & 1 & 1 \\ . & 1 & . & 1 \\ 1 & . & . & 1 \end{bmatrix}.$$

Summing up these columns yields $0 \in \mathbb{F}_2^{k \times 1}$. Hence summing up the associated columns of H'_k, whose last row consists of 1_n, yields $0 \in \mathbb{F}_2^{(k+1) \times 1}$, saying that $\mathcal{H}'_k$ has an $\mathbb{F}_2$-linearly dependent 4-subset of columns. ♯

Table 5: Modified binary Hamming and simplex codes.

$$\begin{tikzcd}
& \widehat{\mathcal{H}}_k \arrow[dl, shift left, "\text{puncture}"] \arrow[dr, shift left, "\text{shorten}"] \arrow[rrr, leftrightarrow, "\text{dualise}"] & & & \mathcal{R}_k \arrow[dl, shift left, "\text{puncture}"] \arrow[dr, shift left, "\text{shorten}"] & \\
\mathcal{H}_k \arrow[ur, shift left, "\text{extend}"] \arrow[rr, shift right, "\text{expurgate}"'] & & \mathcal{H}'_k \arrow[ul, shift left, "\text{lengthen}"] \arrow[ll, shift right, "\text{augment}"'] & \mathcal{R}^\bullet_k \arrow[ur, shift left, "\text{extend}"] \arrow[rr, shift right, "\text{expurgate}"'] & & \mathcal{S}_k \arrow[ul, shift left, "\text{lengthen}"] \arrow[ll, shift right, "\text{augment}"']
\end{tikzcd}$$

iii) Shortening $\widehat{\mathcal{H}}_k \leq \mathbb{F}_2^{n+1}$ yields $(\widehat{\mathcal{H}}_k)^\circ = \mathcal{H}'_k \leq \mathbb{F}_2^n$ again, and lengthening $\mathcal{H}'_k \leq \mathbb{F}_2^n$ yields $\widehat{(\mathcal{H}'_k)} = \widehat{\mathcal{H}}_k \leq \mathbb{F}_2^{n+1}$ again.

(6.3) Simplex codes. a) For $k \geq 2$ and $n := \frac{q^k-1}{q-1}$ let $H_k \in \mathbb{F}_q^{k \times n}$ be as in (6.1). Then the code $\mathcal{S}_k := \mathcal{H}_k^\perp \leq \mathbb{F}_q^n$ having H_k as a generator matrix is called the associated **simplex code**. We show that all $0 \neq v \in \mathcal{S}_k$ have weight $\mathrm{wt}(v) = q^{k-1}$; in particular $\mathcal{S}_k$ is an equidistant $[n, k, q^{k-1}]$-code:

Let $\{v_1, \ldots, v_k\} \subseteq \mathcal{S}_k$ be the $\mathbb{F}_q$-basis given by the rows of H_k. Then there is $0 \neq [x_1, \ldots, x_k] \in \mathbb{F}_q^k$ such that $v = \sum_{i=1}^k x_i v_i \in \mathbb{F}_q^n$. The j-th entry of v, for $j \in \{1, \ldots, n\}$, is zero if and only if $0 = \langle v, e_j \rangle = \sum_{i=1}^k x_i \langle v_i, e_j \rangle$. This amounts to saying that $[\langle v_1, e_j \rangle, \ldots, \langle v_k, e_j \rangle] \in \mathbb{F}_q^k$ is an element of $U := \langle [x_1, \ldots, x_k] \rangle_{\mathbb{F}_q}^\perp \leq \mathbb{F}_q^k$. Now $[\langle v_1, e_j \rangle, \ldots, \langle v_k, e_j \rangle]^{\mathrm{tr}}$ coincides with the j-th column of H_k, hence $\dim_{\mathbb{F}_q}(U) = k-1$ shows that there are precisely $\frac{q^{k-1}-1}{q-1}$ such columns. Thus we have $\mathrm{wt}(v) = n - \frac{q^{k-1}-1}{q-1} = \frac{q^k-1}{q-1} - \frac{q^{k-1}-1}{q-1} = q^{k-1}$. ♯

Note that $\sum_{i=0}^{k-1} \lceil \frac{q^{k-1}}{q^i} \rceil = \sum_{i=0}^{k-1} q^{k-i-1} = \sum_{i=0}^{k-1} q^i = \frac{q^k-1}{q-1}$ and $q^k(q^{k-1} - \frac{q^k-1}{q-1} \cdot \frac{q-1}{q}) = q^{k-1}$ show that $\mathcal{S}_k$ fulfills the Griesmer bound and the Plotkin bound, respectively; recall that the latter also implies that $\mathcal{S}_k$ is an equidistant code.

b) Conversely, any $[n, k, q^{k-1}]$-code $\mathcal{C} \leq \mathbb{F}_q^n$ is linearly equivalent to $\mathcal{S}_k$: Let $d^\perp \in \mathbb{N}$ be the minimum distance of $\mathcal{C}^\perp \leq \mathbb{F}_q^n$. We show that $d^\perp \geq 3$; then since $\dim_{\mathbb{F}_q}(\mathcal{C}^\perp) = n - k$ and any $[n, n-k, 3]$-code is perfect, that is fulfills the Hamming bound, we conclude that $d^\perp = 3$, and hence $\mathcal{C}^\perp$ is linearly equivalent to the Hamming code $\mathcal{H}_k = \mathcal{S}_k^\perp$:

To this end, in turn, let $H \in \mathbb{F}_q^{(n-k) \times n}$ be a check matrix of $\mathcal{C}$, hence H is a generator matrix of $\mathcal{C}^\perp \leq \mathbb{F}_q^n$. Firstly, assume that $d^\perp = 1$, then we may assume that $[0, \ldots, 0, 1] \in \mathcal{C}^\perp \leq \mathbb{F}_q^n$. Thus any word in $\mathcal{C}$ has zero n-th entry, implying that the shortened code $\mathcal{C}^\circ$ is an $[n-1, k, q^{k-1}]$-code. Hence the Griesmer bound implies $n - 1 \geq \sum_{i=0}^{k-1} \lceil \frac{q^{k-1}}{q^i} \rceil = \frac{q^k-1}{q-1} = n$, a contradiction.

Secondly, assume that $d^\perp = 2$, then we may assume that $[0, \ldots, 0, 1, 1] \in \mathcal{C}^\perp \leq \mathbb{F}_q^n$. Thus any word in $\mathcal{C}$ is of the form $[*, \ldots, *, -x, x] \in \mathbb{F}_q^n$, for some $x \in \mathbb{F}_q$, implying that any word in the shortened code $\mathcal{C}^\circ$ has zero $(n-1)$-st entry, thus the doubly shortened code $\mathcal{C}^{\circ\circ}$ is an $[n-2, k^{\circ\circ}, d^{\circ\circ}]$-code, where $k-1 \leq k^{\circ\circ} \leq k$ and $d^{\circ\circ} \geq q^{k-1}$. Hence the Griesmer bound implies $n-2 \geq \sum_{i=0}^{k^{\circ\circ}-1} \lceil \frac{d^{\circ\circ}}{q^i} \rceil \geq \sum_{i=0}^{k^{\circ\circ}-1} \lceil \frac{q^{k-1}}{q^i} \rceil = \frac{q^k - q^{k-k^{\circ\circ}}}{q-1} = n - \frac{q^{k-k^{\circ\circ}}-1}{q-1} \geq n-1$, a contradiction again. ♯

(6.4) Binary simplex codes and Reed-Muller codes. a) Keeping the notation of (6.3), let $q := 2$, hence for $k \geq 2$ we have $n = 2^k - 1$, thus $\mathcal{S}_k$ is a $[2^k - 1, k, 2^{k-1}]$-code. For example, the elements of $\mathcal{S}_2 = \langle [0,1,1], [1,0,1] \rangle_{\mathbb{F}_2} = \{0_3, [0,1,1], [1,0,1], [1,1,0]\} \leq \mathbb{F}_2^3$ can be seen as the vertices of a tetrahedron inscribed into a cube; note that $\mathcal{S}_2$ is the even-weight $[3,2,2]$-code.

We apply the modifications described in (4.5), see Table 5: Since all elements of $\mathcal{S}_k$ have even weight, we get $\mathcal{S}_k' = \mathcal{S}_k$ and $\widehat{\mathcal{S}}_k = \{[v, 0] \in \mathbb{F}_2^{n+1}; v \in \mathcal{S}_k\}$.

i) Dualizing the even-weight Hamming code $\mathcal{H}_k'$ yields a code which has $H_k' := \left[\begin{array}{c} H_k \\ \hline 1_n \end{array}\right] \in \mathbb{F}_2^{(k+1) \times n}$ as a generator matrix, thus $(\mathcal{H}_k')^\perp = \widetilde{\mathcal{S}}_k \leq \mathbb{F}_2^n$ is the augmented simplex code. Since for all $0 \neq v \in \mathcal{S}_k$ we have $\mathrm{wt}(v) = 2^{k-1}$ and $\mathrm{wt}(1_n + v) = n - 2^{k-1} = 2^{k-1} - 1$, we conclude that $\widetilde{\mathcal{S}}_k$ has minimum distance $2^{k-1} - 1$, thus is a $[2^k - 1, k+1, 2^{k-1} - 1]$-code.

Moreover, expurgating $\widetilde{\mathcal{S}}_k$ yields $(\widetilde{\mathcal{S}}_k)' = (\mathcal{S}_k \mathbin{\dot\cup} (1_n + \mathcal{S}_k))' = \mathcal{S}_k$ again. In particular, for $k = 2$ we infer that $\widetilde{\mathcal{S}}_2$ is a $[3,3,1]$-code, thus $\widetilde{\mathcal{S}}_2 = \mathbb{F}_2^3$; for $k = 3$ we get the Hamming $[7,4,3]$-code $\widetilde{\mathcal{S}}_3 = \mathcal{H}_3$.

Note that $\sum_{i=0}^{k} \lceil \frac{2^{k-1}-1}{2^i} \rceil = (2^{k-1} - 1) + 1 + \sum_{i=1}^{k-1} \lceil 2^{k-1-i} - \frac{1}{2^i} \rceil = 2^{k-1} + \sum_{i=0}^{k-2} 2^i = 2^{k-1} + (2^{k-1} - 1) = 2^k - 1$ shows that $\widetilde{\mathcal{S}}_k$ fulfills the Griesmer bound; but since $(2^{k-1} - 1) - (2^k - 1) \cdot \frac{1}{2} = -\frac{1}{2}$ the Plotkin bound does not yield.

ii) Dualizing the extended Hamming code $\widehat{\mathcal{H}}_k$ yields the **Reed-Muller code** $\mathcal{R}_k := (\widehat{\mathcal{H}}_k)^\perp \leq \mathbb{F}_2^{n+1}$ with generator matrix $\widehat{H}_k = \left[\begin{array}{c|c} H_k & 0_k^{\mathrm{tr}} \\ \hline 1_n & 1 \end{array}\right] \in \mathbb{F}_2^{(k+1) \times (n+1)}$. Hence $\mathcal{R}_k = \widehat{\widetilde{\mathcal{S}}}_k$ is obtained by extending $\widetilde{\mathcal{S}}_k$, that is by lengthening $\mathcal{S}_k$; note that in particular $1_{2^k} \in \mathcal{R}_k$. Since $\mathcal{R}_k$ extends $\widetilde{\mathcal{S}}_k$, it has minimum distance 2^{k-1}, thus $\mathcal{R}_k$ is a $[2^k, k+1, 2^{k-1}]$-code.

Moreover, shortening $\mathcal{R}_k$ yields $\mathcal{R}_k^\circ = (\widehat{\widetilde{\mathcal{S}}}_k)^\circ = (\widetilde{\mathcal{S}}_k)' = \mathcal{S}_k$ again. In particular, for $k = 2$ we infer that $\mathcal{R}_2$ is the even-weight $[4,3,2]$-code, while for $k = 3$ we get the extended Hamming $[8,4,4]$-code $\mathcal{R}_3 = \widehat{\mathcal{H}}_3^\perp = \widehat{\mathcal{H}}_3$.

Note that $\sum_{i=0}^{k} \lceil \frac{2^{k-1}}{2^i} \rceil = 1 + \sum_{i=0}^{k-1} 2^i = 2^k$ shows that $\mathcal{R}_k$ fulfills the Griesmer bound; but since $2^{k-1} - 2^k \cdot \frac{1}{2} = 0$ the Plotkin bound does not yield.

iii) The **punctured Reed-Muller code** is $\mathcal{R}_k^\bullet = (\widehat{\widetilde{\mathcal{S}}}_k)^\bullet = \widetilde{\mathcal{S}}_k$ again, and extending $\mathcal{R}_k^\bullet$ yields $\widehat{(\mathcal{R}_k^\bullet)} = \widehat{\widehat{\mathcal{S}}}_k = \mathcal{R}_k$ again.

b) Conversely, any binary $[2^k, k+1, 2^{k-1}]$-code $\mathcal{C}$ is linearly equivalent to $\mathcal{R}_k$: We proceed by induction on $k \geq 1$, letting $\mathcal{R}_1 := \mathbb{F}_2^2$ be the unique $[2,2,1]$-code.

Let $k \geq 2$. Since $\mathcal{C}$ fulfills the Griesmer bound, it cannot possibly possess a zero component, hence the shortened code $\mathcal{C}^\circ$ is a $[2^k - 1, k, d^\circ]$-code, where $d^\circ \geq 2^{k-1}$. Since any $[2^k - 1, k, 2^{k-1}]$-code fulfills the Griesmer bound, we conclude that $d^\circ = 2^{k-1}$, thus we may assume that $\mathcal{C}^\circ = \mathcal{S}_k$. Note that this also shows that shortening with respect to any component yields a code linearly equivalent to the simplex code, implying that any word in $\mathcal{C} \setminus \{0_{2^k}, 1_{2^k}\}$ has weight 2^{k-1}. We show that $1_{2^k} \in \mathcal{C}$; then we have $\mathcal{C} = \langle \widehat{(\mathcal{C}^\circ)}, 1_{2^k} \rangle_{\mathbb{F}_2} = \widehat{\mathcal{S}}_k + \langle \widehat{1_{2^k-1}} \rangle_{\mathbb{F}_2} = \widehat{\widehat{\mathcal{S}}}_k = \mathcal{R}_k$:

Let $G = \left[\begin{array}{c|c|c} \multicolumn{2}{c|}{1_{2^{k-1}}} & 0_{2^{k-1}} \\ \hline * & 0_k^{\mathrm{tr}} & G^* \end{array}\right] \in \mathbb{F}_q^{(k+1) \times 2^k}$ be a generator matrix of $\mathcal{C}$, and let $\mathcal{C}^* \leq \mathbb{F}_q^{2^{k-1}}$ be the residual $[2^{k-1}, k, d^*]$-code generated by $G^* \in \mathbb{F}_q^{k \times 2^{k-1}}$, where $d^* \geq \frac{2^{k-1}}{2} = 2^{k-2}$. Since any $[2^{k-1}, k, 2^{k-2}]$-code fulfills the Griesmer bound, we conclude that $d^* = 2^{k-2}$, and thus we may assume by induction that $\mathcal{C}^* = \mathcal{R}_{k-1}$; note that for $k = 2$ we indeed see that $\mathcal{C}^*$ is a $[2,2,1]$-code. Hence we have $1_{2^{k-1}} \in \mathcal{C}^*$, and thus $\mathcal{C}$ contains a word of the form $[* \mid 0 \mid 1_{2^{k-1}}] + [1_{2^{k-1}} \mid 0_{2^{k-1}}] = [* \mid 1 \mid 1_{2^{k-1}}]$, which has weight at least $2^{k-1} + 1$, hence equals 1_{2^k}. $\sharp$

The Reed-Muller codes $\mathcal{R}_k$ are **Hadamard** codes, being defined by **Hadamard matrices of Sylvester type**, see Exercise (15.28), and thus have a particularly fast decoding algorithm (outperforming the general one for higher order Reed-Muller codes, which are discussed below). Together with their large relative minimum distance $\delta(\mathcal{R}_k) = \frac{2^{k-1}}{2^k} = \frac{1}{2}$ this outweighs their low information rate $\rho(\mathcal{R}_k) = \frac{k+1}{2^k}$, making them suitable for very noisy channels.

For example, the $[32, 6, 16]$-code $\mathcal{R}_5$ was used in the 'Mariner' expeditions to planet Mars [1969–1976]: The 6 information symbols are used to encode picture data based on dots on a grey-scale with $2^6 = 64$ steps, where $\mathcal{R}_5$ has a low information rate of $\rho(\mathcal{R}_5) = \frac{6}{32} \sim 0.2$, but is able to correct $\lfloor \frac{16-1}{2} \rfloor = 7$ errors.

(6.5) Higher order Reed-Muller codes. Reed-Muller codes are merely the first ones in the series of binary **higher order Reed-Muller codes** [1954], which in turn belong to the class of **geometric codes**, being based on finite geometries, having a rich algebraic structure, and having a fast decoding algorithm, being called **multistep majority decoding**. Moreover, higher order Reed-Muller codes have been generalized to codes over arbitrary finite fields.

a) Let first $\mathbb{F}_q$ be the field with q elements, let $\mathcal{C}'$ be an $[n, k', d']$-code and $\mathcal{C}''$ be an $[n, k'', d'']$-code over $\mathbb{F}_q$, and let $\mathcal{C} := \mathcal{C}' \ltimes \mathcal{C}'' := \{[v \mid v+w] \in \mathbb{F}_q^{2n}; v \in \mathcal{C}', w \in \mathcal{C}''\} \leq \mathbb{F}_q^{2n}$ be their **Plotkin sum**. Then $\mathcal{C}$ is a $[2n, k' + k'', \min\{2d', d''\}]$-code:

The $\mathbb{F}_q$-linear map $\mathcal{C}' \oplus \mathcal{C}'' \to \mathcal{C}$: $[v, w] \mapsto [v \mid v + w]$ being injective, we get $\dim_{\mathbb{F}_q}(\mathcal{C}) = k' + k''$. We turn to the minimum distance $d = d(\mathcal{C})$: If both $\mathcal{C}'$ and $\mathcal{C}''$ are trivial then $\mathcal{C}$ is trivial as well, and we have $\min\{\infty, \infty\} = \infty = d$;

if $\mathcal{C}'$ is non-trivial and $\mathcal{C}''$ is trivial, then we have $\mathcal{C} = \{[v \mid v] \in \mathbb{F}_q^{2n}; v \in \mathcal{C}'\}$ and $\min\{2d', \infty\} = 2d' = d$; if $\mathcal{C}'$ is trivial and $\mathcal{C}''$ is non-trivial, then we have $\mathcal{C} = \{[0 \mid w] \in \mathbb{F}_q^{2n}; w \in \mathcal{C}''\}$ and $\min\{\infty, d''\} = d'' = d$.

Thus let both $\mathcal{C}'$ and $\mathcal{C}''$ be non-trivial, and let $0 \neq u := [v \mid v+w] \in \mathcal{C}$, where $v \in \mathcal{C}'$ and $w \in \mathcal{C}''$. If $w = 0$ then $\mathrm{wt}(u) = 2 \cdot \mathrm{wt}(v) \geq 2d'$, and equality is attained for $v \neq 0$ of minimum weight; if $v = 0$ then $\mathrm{wt}(u) = \mathrm{wt}(w) \geq d''$, and equality is attained for $w \neq 0$ of minimum weight. Hence letting both $v \neq 0$ and $w \neq 0$, then using $\mathrm{wt}(v) \geq |\mathrm{supp}(v) \cap \mathrm{supp}(w)|$ we get $\mathrm{wt}(u) = \mathrm{wt}(v) + \mathrm{wt}(v+w) \geq \mathrm{wt}(v) + (\mathrm{wt}(v) + \mathrm{wt}(w) - 2 \cdot |\mathrm{supp}(v) \cap \mathrm{supp}(w)|) \geq \mathrm{wt}(w) \geq d''$. $\sharp$

b) We are now prepared to define the **Reed-Muller code** $\mathcal{R}_k^{(r)} \leq \mathbb{F}_2^{2^k}$ of **order** $r \in \mathbb{N}_0$, where $k \in \mathbb{N}_0$ such that $k \geq r$, recursively as follows:

For $k \in \mathbb{N}_0$ let $\mathcal{R}_k^{(0)} = \{0_{2^k}, 1_{2^k}\} \leq \mathbb{F}_2^{2^k}$ be the repetition $[2^k, 1, 2^k]$-code, and let $\mathcal{R}_k^{(k)} := \mathbb{F}_2^{2^k}$ be the $[2^k, 2^k, 1]$-code; in particular we have $\mathcal{R}_0^{(0)} = \mathbb{F}_2$, while $\mathcal{R}_1^{(0)} = \{0_2, 1_2\} \leq \mathbb{F}_2^2$ and $\mathcal{R}_1^{(1)} = \mathbb{F}_2^2$. Then, recursing over $k \geq 2$, for $r \in \{1, \ldots, k-1\}$ let $\mathcal{R}_k^{(r)} := \mathcal{R}_{k-1}^{(r)} \ltimes \mathcal{R}_{k-1}^{(r-1)} \leq \mathbb{F}_2^{2^{k-1}} \oplus \mathbb{F}_2^{2^{k-1}} \cong \mathbb{F}_2^{2^k}$.

Then $\mathcal{R}_k^{(r)}$ is a $[2^k, \sum_{i=0}^r \binom{k}{i}, 2^{k-r}]$-code: We have $\dim_{\mathbb{F}_2}(\mathcal{R}_k^{(0)}) = 1 = \binom{k}{0}$ and $\dim_{\mathbb{F}_2}(\mathcal{R}_k^{(k)}) = 2^k = \sum_{i=0}^k \binom{k}{i}$, and for $k \geq 2$ and $r \in \{1, \ldots, k-1\}$ we get $\dim_{\mathbb{F}_2}(\mathcal{R}_k^{(r)}) = \dim_{\mathbb{F}_2}(\mathcal{R}_{k-1}^{(r)}) + \dim_{\mathbb{F}_2}(\mathcal{R}_{k-1}^{(r-1)}) = \sum_{i=0}^r \binom{k-1}{i} + \sum_{i=0}^{r-1} \binom{k-1}{i} = 1 + \sum_{i=0}^{r-1}(\binom{k-1}{i+1} + \binom{k-1}{i}) = 1 + \sum_{i=0}^{r-1} \binom{k}{i+1} = \sum_{i=0}^r \binom{k}{i}$. Moreover, we have $d(\mathcal{R}_k^{(0)}) = 2^k$ and $d(\mathcal{R}_k^{(k)}) = 1$, and for $k \geq 2$ and $r \in \{1, \ldots, k-1\}$ we obtain $d(\mathcal{R}_k^{(r)}) = \min\{2 \cdot d(\mathcal{R}_{k-1}^{(r)}), d(\mathcal{R}_{k-1}^{(r-1)})\} = \min\{2 \cdot 2^{k-r-1}, 2^{k-r}\} = 2^{k-r}$. $\sharp$

The Reed-Muller codes considered in (6.4) are indeed linearly equivalent to the first order Reed-Muller codes: We have $\mathcal{R}_1^{(1)} = \mathbb{F}_2^2 = \mathcal{R}_1$, and $\mathcal{R}_k^{(1)}$ is a $[2^k, k+1, 2^{k-1}]$-code, thus is linearly equivalent to $\mathcal{R}_k$, for $k \geq 2$.

(6.6) Boolean functions. a) We present an alternative construction of higher order Reed-Muller codes: A function $p \colon \mathbb{F}_2^k \to \mathbb{F}_2$ is called a **Boolean function** in $k \in \mathbb{N}_0$ variables. Identifying $x = [x_1, \ldots, x_k] \in \mathbb{F}_2^k$ with the integer $\sum_{i=1}^k x_i \cdot 2^{i-1} \in \{0, \ldots, 2^k - 1\}$, where we silently lift the elements of $\mathbb{F}_2$ to $\mathbb{Z}_2 \subseteq \mathbb{Z}$, and ordering the elements of $\mathbb{F}_2^k$ accordingly, p can be identified with an element of $\mathbb{F}_2^{2^k}$ by listing the values it assumes. Moreover, identifying the values 0 and 1 with the Boolean values false and true, the Boolean operations exor, and, or and not can be translated into the operations $p+q$, pq, $p+q+pq$ and $1_{2^k} + p$, respectively, where $p, q \in \mathbb{F}_2^{2^k}$ and products are taken pointwise.

For $i \in \{1, \ldots, k\}$ let $p_i \colon \mathbb{F}_2^k \to \mathbb{F}_2$ be the projection onto the i-th component, thus we have $p_1 = [0, 1, \ldots, 0, 1] \in \mathbb{F}_2^{2^k}$, $p_2 = [0, 0, 1, 1, \ldots, 0, 0, 1, 1] \in \mathbb{F}_2^{2^k}$, and so forth up to $p_k = [0, \ldots, 0, 1, \ldots, 1] \in \mathbb{F}_2^{2^k}$. For $\mathcal{I} \subseteq \{1, \ldots, k\}$ let $p_{\mathcal{I}} := \prod_{i \in \mathcal{I}} p_i \in \mathbb{F}_2^{2^k}$, where $p_\emptyset := 1_{2^k}$, and $|\mathcal{I}| \in \{0, \ldots, k\}$ is called the **degree** of $p_{\mathcal{I}}$. Since the function $p_{\mathcal{I}} \colon \mathbb{F}_2^k \to \mathbb{F}_2$ only depends on $|\mathcal{I}|$ variables, for $\mathcal{I} \neq \{1, \ldots, k\}$

the vector $p_{\mathcal{I}} \in \mathbb{F}_2^{2^k}$ has even weight, while $p_{\{1,\ldots,k\}} = [0,\ldots,0,1] \in \mathbb{F}_2^{2^k}$.

Then, using the disjunctive normal form of Boolean logic, any Boolean function can be written as a sum $\sum_{\mathcal{I}\subseteq\{1,\ldots,k\}} a_{\mathcal{I}} p_{\mathcal{I}}$, where $a_{\mathcal{I}} \in \mathbb{F}_2$ and $p_{\mathcal{I}} := \prod_{i\in\mathcal{I}} p_i \in \mathbb{F}_2^{2^k}$, letting $p_\emptyset := 1_{2^k}$. Since there are precisely $2^{2^k} = |\mathbb{F}_2^{2^k}|$ sums of this shape, we deduce that these are pairwise distinct, thus $\{p_{\mathcal{I}} \in \mathbb{F}_2^{2^k}; \mathcal{I} \subseteq \{1,\ldots,k\}\}$ is an $\mathbb{F}_2$-basis of the space of Boolean functions in k variables.

For example, for $k = 3$ let $p := [0,0,0,1,1,0,0,0] \in \mathbb{F}_2^8$, having value 1 at positions $\{3,4\} \subseteq \{0,\ldots,7\}$, which hence corresponds to the Boolean function assuming the value true if and only if the variables assume the values $[\mathsf{true},\mathsf{true},\mathsf{false}]$ or $[\mathsf{false},\mathsf{false},\mathsf{true}]$. This translates into $p = p_1p_2(1+p_3) + (1+p_1)(1+p_2)p_3 = p_{\{1,2\}} + p_{\{1,2,3\}} + p_3 + p_{\{1,3\}} + p_{\{2,3\}} + p_{\{1,2,3\}} = p_3 + p_{\{1,2\}} + p_{\{1,3\}} + p_{\{2,3\}}$. Indeed, we have $p_3 = [0,0,0,0,1,1,1,1]$ and $p_{\{1,2\}} = [0,0,0,1,0,0,0,1]$ and $p_{\{1,3\}} = [0,0,0,0,0,1,0,1]$ and $p_{\{2,3\}} = [0,0,0,0,0,0,1,1]$, thus we get $p_3 + p_{\{1,2\}} + p_{\{1,3\}} + p_{\{2,3\}} = [0,0,0,1,1,0,0,0] = p$.

b) Now let $r \in \{0,\ldots,k\}$, and let $\mathcal{C}_k^{(r)} := \langle p_{\mathcal{I}} \in \mathbb{F}_2^{2^k}; |\mathcal{I}| \leq r\rangle_{\mathbb{F}_2} \leq \mathbb{F}_2^{2^k}$ be the linear code spanned by the Boolean functions in k variables of degree at most r. In particular, we immediately see that $\dim_{\mathbb{F}_2}(\mathcal{C}_k^{(r)}) = \sum_{i=0}^r \binom{k}{i}$, but we have no clue about the minimum distance of $\mathcal{C}_k^{(r)}$. Anyway, we show that $\mathcal{C}_k^{(r)} = \mathcal{R}_k^{(r)}$:

We have $\mathcal{C}_k^{(0)} = \langle 1_{2^k}\rangle_{\mathbb{F}_2} = \mathcal{R}_k^{(0)}$, and $\dim_{\mathbb{F}_2}(\mathcal{C}_k^{(r)}) = \sum_{i=0}^k \binom{k}{i} = 2^k$ implies that $\mathcal{C}_k^{(k)} = \mathbb{F}_2^{2^k} = \mathcal{R}_k^{(k)}$. For $k \geq 2$ and $r \in \{1,\ldots,k-1\}$, any $p \in \mathcal{C}_k^{(r)}$ can be written as $p = \sum_{\mathcal{I}\subseteq\{1,\ldots,k\},|\mathcal{I}|\leq r} a_{\mathcal{I}} p_{\mathcal{I}} = \sum_{k\notin\mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I}} + p_k \cdot \sum_{k\in\mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I}\setminus\{k\}}$. Then the first sum $\sum_{k\notin\mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I}}$ can be identified with $v \in \mathcal{R}_{k-1}^{(r)} \leq \mathbb{F}_2^{2^{k-1}}$, which embeds into $\mathbb{F}_2^{2^k}$ as $[v \mid v]$. The second sum $\sum_{k\in\mathcal{I}} a_{\mathcal{I}} p_{\mathcal{I}\setminus\{k\}}$ can be identified with $w \in \mathcal{R}_{k-1}^{(r-1)} \leq \mathbb{F}_2^{2^{k-1}}$, so that $p_k \cdot w$ embeds into $\mathbb{F}_2^{2^k}$ as $[0_{2^{k-1}} \mid w]$. Thus we conclude that indeed $\mathcal{C}_k^{(r)} = \mathcal{R}_{k-1}^{(r)} \ltimes \mathcal{R}_{k-1}^{(r-1)} = \mathcal{R}_k^{(r)}$.

c) This allows to read off further properties of $\mathcal{R}_k^{(r)}$: By construction we have $\langle 1_{2^k}\rangle_{\mathbb{F}_2} = \mathcal{R}_k^{(0)} \leq \mathcal{R}_k^{(1)} \leq \cdots \leq \mathcal{R}_k^{(k-1)} \leq \mathcal{R}_k^{(k)} = \mathbb{F}_2^{2^k}$, for $k \in \mathbb{N}_0$; in particular we observe that $1_{2^k} \in \mathcal{R}_k^{(r)}$ for all $k \geq r \geq 0$. Moreover, since $\mathcal{R}_k^{(k-1)}$ has an $\mathbb{F}_2$-basis consisting of vectors of even weight, and $\sum_{i=0}^{k-1} \binom{k}{i} = 2^k - 1$, we conclude that $\mathcal{R}_k^{(k-1)}$ is the even-weight $[2^k, 2^k-1, 2]$-code, for $k \geq 1$. Finally, noting that $(\mathcal{R}_k^{(k)})^\perp = (\mathbb{F}_2^{2^k})^\perp = \{0\}$, for $k > r \geq 0$ we have $(\mathcal{R}_k^{(r)})^\perp = \mathcal{R}_k^{(k-r-1)}$:

Let $p \in \mathcal{R}_k^{(r)}$ and $q \in \mathcal{R}_k^{(k-r-1)}$. Then, since $p_i^2 = p_i$ for $i \in \{1,\ldots,k\}$, we conclude that $pq \in \mathbb{F}_2^{2^k}$ has degree at most $r+(k-r-1) = k-1$, thus $pq \in \mathcal{R}_k^{(k-1)}$. Since the latter is the even-weight code, we infer $\langle p,q\rangle \equiv \mathrm{wt}(pq) \equiv 0 \pmod 2$. This shows that $\mathcal{R}_k^{(k-r-1)} \leq (\mathcal{R}_k^{(r)})^\perp$. Since $\dim_{\mathbb{F}_2}(\mathcal{R}_k^{(r)}) + \dim_{\mathbb{F}_2}(\mathcal{R}_k^{(k-r-1)}) = \sum_{i=0}^r \binom{k}{i} + \sum_{i=0}^{k-r-1} \binom{k}{i} = \sum_{i=0}^r \binom{k}{i} + \sum_{i=0}^{k-r-1} \binom{k}{k-i} = \sum_{i=0}^r \binom{k}{i} + \sum_{i=r+1}^k \binom{k}{i} = \sum_{i=0}^k \binom{k}{i} = 2^k = \dim_{\mathbb{F}_2}(\mathbb{F}_2^{2^k})$ we conclude that $\mathcal{R}_k^{(k-r-1)} = (\mathcal{R}_k^{(r)})^\perp$. ♯

In particular, from $(\mathcal{R}_k^{(k-1)})^\perp = \mathcal{R}_k^{(0)} = \langle 1_{2^k}\rangle_{\mathbb{F}_2}$ for $k \geq 1$, we recover the fact that $\mathcal{R}_k^{(k-1)} = (\langle 1_{2^k}\rangle_{\mathbb{F}_2})^\perp \leq \mathbb{F}_2^{2^k}$ is the even-weight code. More interestingly, from $(\mathcal{R}_k^{(k-2)})^\perp = \mathcal{R}_k^{(1)}$, the latter being linearly equivalent to $\mathcal{R}_k$, we conclude that $\mathcal{R}_k^{(k-2)}$ is linearly equivalent to $(\mathcal{R}_k)^\perp = \widehat{\mathcal{H}}_k$, the extended Hamming code; recall that $\sum_{i=0}^{k-2}\binom{k}{i} = 2^k - k - 1$ entails that $\mathcal{R}_k^{(k-2)}$ is a $[2^k, 2^k-k-1, 4]$-code.

7 Weight enumerators

(7.1) Weight enumerators. Let $\mathcal{X}$ be an alphabet and let $\mathcal{C} \subseteq \mathcal{X}^n$ be a code. For $i \in \mathbb{N}_0$ let $w_i = w_i(\mathcal{C}) := |\{v \in \mathcal{C}; \operatorname{wt}(v) = i\}| \in \mathbb{N}_0$. Hence we have $w_0 \leq 1$, and $w_i = 0$ for $i \in \{1, \ldots, \operatorname{wt}(\mathcal{C}) - 1\}$, as well as $w_{\operatorname{wt}(\mathcal{C})} \geq 1$, and $w_i = 0$ for $i \geq n+1$, and $\sum_{i=0}^n w_i = |\mathcal{C}|$.

Let $\{X, Y\}$ be indeterminates. The **homogeneous generating function** $A_\mathcal{C} := \sum_{i=0}^n w_i X^i Y^{n-i} = \sum_{v \in \mathcal{C}} X^{\operatorname{wt}(v)} Y^{n-\operatorname{wt}(v)} \in \mathbb{C}[X, Y]$ associated with the sequence $[w_0, w_1, \ldots, w_n]$ is called the **(Hamming) weight enumerator** of $\mathcal{C}$. Hence $A_\mathcal{C}$ is homogeneous of total degree n and has non-negative rational integers as its coefficients. By **dehomogenizing**, that is specializing $X \mapsto X$ and $Y \mapsto 1$, we obtain the **(ordinary) generating function** $A_\mathcal{C}(X, 1) = \sum_{i=0}^n w_i X^i = \sum_{v \in \mathcal{C}} X^{\operatorname{wt}(v)} \in \mathbb{C}[X]$.

Example. For the trivial code $\mathcal{C} := \{0_n\} \leq \mathbb{F}_q^n$ we get $A_\mathcal{C} = Y^n$. For the code $\mathcal{C}^\perp = \mathbb{F}_q^n$ we get $A_{\mathcal{C}^\perp} = \sum_{i=0}^n \binom{n}{i}(q-1)^i X^i Y^{n-i} = (Y + (q-1)X)^n$. Thus we have $A_{\mathcal{C}^\perp}(X, Y) = A_\mathcal{C}(Y - X, Y + (q-1)X)$, in accordance with (7.2) below.

For the binary repetition code $\mathcal{C} := \{0_n, 1_n\} \leq \mathbb{F}_2^n$ we get $A_\mathcal{C} = Y^n + X^n$. For the binary even-weight code $\mathcal{C}^\perp = (\mathbb{F}_2^n)' \leq \mathbb{F}_2^n$ we get $A_{\mathcal{C}^\perp} = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2i} X^{2i} Y^{n-2i} = \frac{1}{2} \cdot \sum_{i=0}^n \binom{n}{i} X^i Y^{n-i} + \frac{1}{2} \cdot \sum_{i=0}^n (-1)^i \binom{n}{i} X^i Y^{n-i} = \frac{1}{2} \cdot \big((Y+X)^n + (Y-X)^n\big)$, thus $A_{\mathcal{C}^\perp}(X, Y) = \frac{1}{2} \cdot A_\mathcal{C}(Y - X, Y + X)]$, in accordance with (7.2) below.

(7.2) Theorem: MacWilliams [1963]. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a linear code such that $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \in \mathbb{N}_0$, and let $\mathcal{C}^\perp \leq \mathbb{F}_q^n$ be its dual. Then for the associated weight enumerators we have $q^k \cdot A_{\mathcal{C}^\perp}(X, Y) = A_\mathcal{C}(Y - X, Y + (q-1)X) \in \mathbb{C}[X, Y]$.

In particular, if $\mathcal{C} = \mathcal{C}^\perp$ is self-dual, then for the weight enumerators we have $A_\mathcal{C}(X, Y) = A_\mathcal{C}(\frac{Y-X}{\sqrt{q}}, \frac{Y+(q-1)X}{\sqrt{q}}) \in \mathbb{C}[X, Y]$; recall that in this case $k = \frac{n}{2}$.

Proof. i) Let $\chi \colon \mathbb{F}_q \to \mathbb{C}^* := \mathbb{C} \setminus \{0\}$ be a **character** of $\mathbb{F}_q$, that is a group homomorphism from the additive group $(\mathbb{F}_q, +)$ to the multiplicative group $(\mathbb{C}^*, \cdot)$; note that there always is the **trivial character** $\underline{1} \colon \mathbb{F}_q \to \mathbb{C}^* \colon a \mapsto 1$. Let V be a $\mathbb{C}$-vector space, and let $\omega \colon \mathbb{F}_q^n \to V$ be a map. Then the map $\chi_\omega \colon \mathbb{F}_q^n \to V \colon v \mapsto \sum_{w \in \mathbb{F}_q^n} \chi(\langle v, w \rangle)\omega(w)$ is called the **discrete Fourier transform** or **Hadamard transform** of ω with respect to χ. We show that for any character $\chi \neq \underline{1}$ we have $\sum_{v \in \mathcal{C}} \chi_\omega(v) = q^k \cdot \sum_{w \in \mathcal{C}^\perp} \omega(w)$:

For the left hand side we obtain $\sum_{v\in\mathcal{C}}\chi_\omega(v)=\sum_{v\in\mathcal{C}}\sum_{w\in\mathbb{F}_q^n}\chi(\langle v,w\rangle)\omega(w)=\sum_{w\in\mathcal{C}^\perp}\sum_{v\in\mathcal{C}}\chi(\langle v,w\rangle)\omega(w)+\sum_{w\in\mathbb{F}_q^n\setminus\mathcal{C}^\perp}\sum_{v\in\mathcal{C}}\chi(\langle v,w\rangle)\omega(w)\in V$, where since $\chi(\langle v,w\rangle)=\chi(0)=1\in\mathbb{C}^*$ the first summand evaluates to $|\mathcal{C}|\cdot\sum_{w\in\mathcal{C}^\perp}\omega(w)\in V$, which coincides with the right hand side of the above equation. Hence we have to show that the second summand vanishes:

Given $w\in\mathbb{F}_q^n\setminus\mathcal{C}^\perp$, the map $\mathcal{C}\to\mathbb{F}_q\colon v\mapsto\langle v,w\rangle$ is $\mathbb{F}_q$-linear and non-zero, hence surjective, and thus for any $a\in\mathbb{F}_q$ we have $|\{v\in\mathcal{C};\langle v,w\rangle=a\}|=\frac{|\mathcal{C}|}{|\mathbb{F}_q|}=q^{k-1}$. Thus the second summand evaluates to $q^{k-1}\cdot\big(\sum_{a\in\mathbb{F}_q}\chi(a)\big)\cdot\big(\sum_{w\in\mathcal{C}^\perp}\omega(w)\big)\in V$, hence it suffices to show that $\sum_{a\in\mathbb{F}_q}\chi(a)=0\in\mathbb{C}$: Since $\chi\neq\underline{1}$, there is $b\in\mathbb{F}_q$ such that $\chi(b)\neq1$; then we have $\chi(b)\cdot\sum_{a\in\mathbb{F}_q}\chi(a)=\sum_{a\in\mathbb{F}_q}\chi(a+b)=\sum_{a\in\mathbb{F}_q}\chi(a)$, implying $(\chi(b)-1))\cdot\sum_{a\in\mathbb{F}_q}\chi(a)=0\in\mathbb{C}$.

ii) Let now $V:=\mathbb{C}[X,Y]_n\leq\mathbb{C}[X,Y]$ be the $\mathbb{C}$-vector space of all homogeneous polynomials of total degree n, including the zero polynomial, and let $\omega\colon\mathbb{F}_q^n\to\mathbb{C}[X,Y]_n\colon v\mapsto X^{\mathrm{wt}(v)}Y^{n-\mathrm{wt}(v)}$. Moreover, let $\delta\colon\mathbb{F}_q\to\{0,1\}$ be defined by $\delta(0)=0$, and $\delta(a)=1$ for $a\neq0$. Thus for any character $\chi\neq\underline{1}$ the associated discrete Fourier transform is given as, where $v=[x_1,\dots,x_n]\in\mathbb{F}_q^n$,

$$\begin{array}{rcl}\chi_\omega(v)&=&\sum_{w\in\mathbb{F}_q^n}\chi(\langle v,w\rangle)X^{\mathrm{wt}(w)}Y^{n-\mathrm{wt}(w)}\\ &=&\sum_{[y_1,\dots,y_n]\in\mathbb{F}_q^n}\chi(\sum_{i=1}^n x_iy_i)X^{\sum_{i=1}^n\delta(y_i)}Y^{\sum_{i=1}^n(1-\delta(y_i))}\\ &=&\sum_{[y_1,\dots,y_n]\in\mathbb{F}_q^n}\big(\prod_{i=1}^n\chi(x_iy_i)X^{\delta(y_i)}Y^{1-\delta(y_i)}\big)\\ &=&\prod_{i=1}^n\big(\sum_{a\in\mathbb{F}_q}\chi(ax_i)X^{\delta(a)}Y^{1-\delta(a)}\big).\end{array}$$

If $x_i=0$ then $\chi(ax_i)=\chi(0)=1\in\mathbb{C}^*$ shows that the associated factor equals $\sum_{a\in\mathbb{F}_q}X^{\delta(a)}Y^{1-\delta(a)}=Y+(q-1)X\in V$. If $x_i\neq0$ then, using $\sum_{a\in\mathbb{F}_q}\chi(a)=0\in\mathbb{C}$ again, the associated factor evaluates as $\sum_{a\in\mathbb{F}_q}\chi(ax_i)X^{\delta(a)}Y^{1-\delta(a)}=Y+\big(\sum_{a\in\mathbb{F}_q^*}\chi(ax_i)\big)\cdot X=Y+\big(\sum_{a\in\mathbb{F}_q^*}\chi(a)\big)\cdot X=Y-\chi(0)\cdot X=Y-X\in V$. Thus we get $\chi_\omega(v)=(Y-X)^{\mathrm{wt}(v)}(Y+(q-1)X)^{n-\mathrm{wt}(v)}\in V$.

In conclusion we have $q^k\cdot A_{\mathcal{C}^\perp}(X,Y)=q^k\cdot\sum_{w\in\mathcal{C}^\perp}\omega(w)=\sum_{v\in\mathcal{C}}\chi_\omega(v)=\sum_{v\in\mathcal{C}}(Y-X)^{\mathrm{wt}(v)}(Y+(q-1)X)^{n-\mathrm{wt}(v)}=A_{\mathcal{C}}(Y-X,Y+(q-1)X)\in V$. ♯

Example. For $k\geq2$, the simplex code $\mathcal{S}_k\leq\mathbb{F}_q^n$ is an equidistant $[n,k,q^{k-1}]$-code, where $n:=\frac{q^k-1}{q-1}$, hence $A_{\mathcal{S}_k}=Y^{\frac{q^k-1}{q-1}}+(q^k-1)X^{q^{k-1}}Y^{\frac{q^{k-1}-1}{q-1}}\in\mathbb{C}[X,Y]$. Thus for the associated Hamming code $\mathcal{H}_k=\mathcal{S}_k^\perp$ we get $A_{\mathcal{H}_k}(X,Y)=\frac{1}{q^k}\cdot A_{\mathcal{S}_k}(Y-X,Y+(q-1)X)\in\mathbb{C}[X,Y]$.

In particular, for $q=2$ we have $n=2^k-1$, and hence $A_{\mathcal{S}_k}=Y^n+nX^{\frac{n+1}{2}}Y^{\frac{n-1}{2}}\in\mathbb{C}[X,Y]$, which yields $A_{\mathcal{H}_k}=\frac{1}{n+1}\cdot((Y+X)^n+n(Y-X)^{\frac{n+1}{2}}(Y+X)^{\frac{n-1}{2}})$. Dehomogenizing, that is specializing $X\mapsto X$ and $Y\mapsto1$, yields $\sum_{i=0}^n w_i(\mathcal{H}_k)X^i=A_{\mathcal{H}_k}(X,1)=\frac{1}{n+1}\cdot((1+X)^n+n(1-X)^{\frac{n+1}{2}}(1+X)^{\frac{n-1}{2}})\in\mathbb{C}[X]$.

For example we have $A_{\mathcal{H}_2}(X,1) = 1 + X^3$, showing again that $\mathcal{H}_2$ is the binary repetition code, and $A_{\mathcal{H}_3}(X,1) = 1 + 7X^3 + 7X^4 + X^7$, and $A_{\mathcal{H}_4}(X,1) = 1 + 35X^3 + 105X^4 + 168X^5 + 280X^6 + 435X^7 + 435X^8 + 280X^9 + 168X^{10} + 105X^{11} + 35X^{12} + X^{15}$.

III

8 Cyclic codes

(8.1) Cyclic codes. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a linear code of length $n \in \mathbb{N}$ over $\mathbb{F}_q$. If for all $[c_0, \ldots, c_{n-1}] \in \mathcal{C}$ we have $[c_{n-1}, c_0, \ldots, c_{n-2}] \in \mathcal{C}$ as well, that is if $P_{(1,\ldots,n)} \in \mathrm{Aut}_{\mathbb{F}_q}(\mathcal{C})$, then $\mathcal{C}$ is called **cyclic**.

Example. The repetition code $\mathcal{C} := \{[c, \ldots, c] \in \mathbb{F}_q^n; c \in \mathbb{F}_q\}$ and the associated dual code, the parity check code $\mathcal{C}^\perp := \{[c_0, \ldots, c_{n-1}] \in \mathbb{F}_q^n; \sum_{i=0}^{n-1} c_i = 0\}$, are cyclic; note that in both cases the full symmetric group $\mathcal{S}_n$ is a subgroup of $\mathrm{Aut}_{\mathbb{F}_q}(\mathcal{C})$. A generator matrix of $\mathcal{C}$ is given as $G := [1, \ldots, 1] \in \mathbb{F}_q^n$, and a generator matrix of $\mathcal{C}^\perp$, that is a check matrix of $\mathcal{C}$, is given as

$$H := \begin{bmatrix} 1 & -1 & . & . & \ldots & . \\ . & 1 & -1 & . & \ldots & . \\ \vdots & & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \vdots \\ . & . & \ldots & . & 1 & -1 \end{bmatrix} \in \mathbb{F}_q^{(n-1)\times n}.$$

Example. The binary Hamming code $\mathcal{H} \leq \mathbb{F}_2^7$, whose elements are given in (4.2), is not cyclic, but applying $P_{(3,4)(5,7,6)} \in I_n(\mathbb{F}_q)$ yields the linearly equivalent code $\mathcal{C} \leq \mathbb{F}_2^7$, whose elements are given by the rows of the following matrices, which hence is cyclic:

$$\left[\begin{array}{ccccccc} . & . & . & . & . & . & . \\ \hline 1 & 1 & . & 1 & . & . & . \\ . & 1 & 1 & . & 1 & . & . \\ . & . & 1 & 1 & . & 1 & . \\ . & . & . & 1 & 1 & . & 1 \\ 1 & . & . & . & 1 & 1 & . \\ . & 1 & . & . & . & 1 & 1 \\ 1 & . & 1 & . & . & . & 1 \end{array}\right] \qquad \left[\begin{array}{ccccccc} 1 & 1 & 1 & . & . & 1 & . \\ . & 1 & 1 & 1 & . & . & 1 \\ 1 & . & 1 & 1 & 1 & . & . \\ . & 1 & . & 1 & 1 & 1 & . \\ . & . & 1 & . & 1 & 1 & 1 \\ 1 & . & . & 1 & . & 1 & 1 \\ 1 & 1 & . & . & 1 & . & 1 \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right]$$

Moreover, a generator matrix $G \in \mathbb{F}_2^{4\times 7}$ and a check matrix $H \in \mathbb{F}_2^{3\times 7}$ of $\mathcal{C}$ are

$$G := \begin{bmatrix} 1 & 1 & . & 1 & . & . & . \\ . & 1 & 1 & . & 1 & . & . \\ . & . & 1 & 1 & . & 1 & . \\ . & . & . & 1 & 1 & . & 1 \end{bmatrix} \quad \text{and} \quad H := \begin{bmatrix} 1 & . & 1 & 1 & 1 & . & . \\ . & 1 & . & 1 & 1 & 1 & . \\ . & . & 1 & . & 1 & 1 & 1 \end{bmatrix}.$$

(8.2) Cyclic codes as ideals. a) Let $\mathbb{F}_q[X]$ be the polynomial ring over $\mathbb{F}_q$ in the indeterminate X. For $0 \neq f = \sum_{i=0}^{d} c_i X^i \in \mathbb{F}_q[X]$, where $d \in \mathbb{N}_0$ and $c_i \in \mathbb{F}_q$ such that $c_d \neq 0$, let $\deg(f) := d$ denote its degree and let $\mathrm{lc}(f) := c_d$ denote its leading coefficient; if $\mathrm{lc}(f) = 1$ then f is called monic. Moreover, $\mathbb{F}_q[X]$ is an $\mathbb{F}_q$-algebra, and an Euclidean ring with respect to polynomial division, hence in particular is a principal ideal domain.

For $n \in \mathbb{N}$ let $\langle X^n - 1\rangle = \{(X^n-1)\cdot f; f \in \mathbb{F}_q[X]\} \trianglelefteq \mathbb{F}_q[X]$ be the principal ideal generated by $X^n - 1 \in \mathbb{F}_q[X]$, let ${}^{-}\colon \mathbb{F}_q[X] \to \overline{\mathbb{F}_q[X]}\colon f \mapsto \overline{f} := f + \langle X^n - 1\rangle$ be the natural epimorphism of $\mathbb{F}_q$-algebras, where $\overline{\mathbb{F}_q[X]} := \mathbb{F}_q[X]/\langle X^n - 1\rangle$ is the associated quotient $\mathbb{F}_q$-algebra.

Then polynomial division shows that $\mathbb{F}_q[X] = \mathbb{F}_q[X]_{<n} \oplus \langle X^n - 1\rangle$ as $\mathbb{F}_q$-vector spaces, where $\mathbb{F}_q[X]_{<n} := \{f \in \mathbb{F}_q[X] \setminus \{0\}; \deg(f) < n\} \,\dot\cup\, \{0\} \leq \mathbb{F}_q[X]$ is the $\mathbb{F}_q$-subspace of all polynomials of degree less than n, including the zero polynomial. Hence $\mathbb{F}_q[X]_{<n}$ is a set of representatives of the cosets in $\overline{\mathbb{F}_q[X]}$, and ${}^{-}\colon \mathbb{F}_q[X]_{<n} \to \overline{\mathbb{F}_q[X]}$ is an $\mathbb{F}_q$-isomorphism. Since $\{X^0, \ldots, X^{n-1}\} \subseteq \mathbb{F}_q[X]_{<n}$ is an $\mathbb{F}_q$-basis, we conclude that $\{\overline{X}^0, \ldots, \overline{X}^{n-1}\} \subseteq \overline{\mathbb{F}_q[X]}$ is an $\mathbb{F}_q$-basis, in particular we have $\dim_{\mathbb{F}_q}(\mathbb{F}_q[X]_{<n}) = \dim_{\mathbb{F}_q}(\overline{\mathbb{F}_q[X]}) = n$.

Let $\nu\colon \mathbb{F}_q^n \to \mathbb{F}_q[X]_{<n}\colon [c_0, \ldots, c_{n-1}] \mapsto \sum_{i=0}^{n-1} c_i X^i$, that is we consider the elements of $\mathbb{F}_q^n$ as coefficient vectors of polynomials in $\mathbb{F}_q[X]_{<n}$ with respect to the above $\mathbb{F}_q$-basis, and let $\overline{\nu} := {}^{-}\circ\nu\colon \mathbb{F}_q^n \to \overline{\mathbb{F}_q[X]}\colon [c_0, \ldots, c_{n-1}] \mapsto \sum_{i=0}^{n-1} c_i \overline{X}^i$; thus both ν and $\overline{\nu}$ are $\mathbb{F}_q$-isomorphisms.

Multiplication by $\overline{X}$ acts on $\overline{\mathbb{F}_q[X]}$ as follows: Given $v := [c_0, \ldots, c_{n-1}] \in \mathbb{F}_q^n$, we have $\overline{\nu}(v) \cdot \overline{X} = (\sum_{i=0}^{n-1} c_i \overline{X}^i) \cdot \overline{X} = \overline{\sum_{i=0}^{n-1} c_i X^{i+1}} = c_{n-1}\overline{X}^0 + \sum_{i=1}^{n-1} c_{i-1}\overline{X}^i = \overline{\nu}(w) \in \overline{\mathbb{F}_q[X]}$, where $w := [c_{n-1}, c_0, \ldots, c_{n-2}] \in \mathbb{F}_q^n$. Thus the action of $P_{(1,\ldots,n)} \in I_n(\mathbb{F}_q^n)$ on $\mathbb{F}_q^n$ is transported to multiplication with $\overline{X}$ on $\overline{\mathbb{F}_q[X]}$.

b) Hence a code $\mathcal{C} \leq \mathbb{F}_q^n$ can be identified via ν with the $\mathbb{F}_q$-subspace $\nu(\mathcal{C}) \leq \mathbb{F}_q[X]_{<n}$, and via $\overline{\nu}$ with the $\mathbb{F}_q$-subspace $\overline{\nu}(\mathcal{C}) \leq \overline{\mathbb{F}_q[X]}$. Moreover, $\mathcal{C}$ is cyclic if and only if $\overline{\nu}(\mathcal{C}) \leq \overline{\mathbb{F}_q[X]}$ is invariant under multiplying with $\overline{X}$, or equivalently under multiplying with $\overline{\mathbb{F}_q[X]}$, that is if and only if $\overline{\nu}(\mathcal{C}) \trianglelefteq \overline{\mathbb{F}_q[X]}$ is an ideal.

In this case, the preimage $\nu(\mathcal{C}) + \langle X^n - 1\rangle \subseteq \mathbb{F}_q[X]$ of $\overline{\nu}(\mathcal{C}) \subseteq \overline{\mathbb{F}_q[X]}$ with respect to ${}^{-}$ is an ideal of $\mathbb{F}_q[X]$. Since $\mathbb{F}_q[X]$ is a principal ideal domain, there is a **generator polynomial** $g \in \mathbb{F}_q[X]$, unique up to scalar multiples, such that $\langle g\rangle = \nu(\mathcal{C}) + \langle X^n - 1\rangle \trianglelefteq \mathbb{F}_q[X]$, in particular implying $\langle \overline{g}\rangle = \overline{\nu}(\mathcal{C})$. Moreover, from $\langle X^n - 1\rangle \subseteq \langle g\rangle \trianglelefteq \mathbb{F}_q[X]$ we infer that $g \mid X^n - 1$; see Table 6.

Conversely, any polynomial $g \in \mathbb{F}_q[X]$ such that $g \mid X^n - 1$ yields an ideal $\langle X^n - 1\rangle \subseteq \langle g\rangle \trianglelefteq \mathbb{F}_q[X]$, hence via $\overline{\nu}$ we get an ideal $\langle \overline{g}\rangle \trianglelefteq \overline{\mathbb{F}_q[X]}$, which in turn can be identified with a cyclic code. Thus we conclude that the cyclic codes $\mathcal{C} \leq \mathbb{F}_q^n$ are in bijective correspondence with the monic divisors g of $X^n - 1 \in \mathbb{F}_q[X]$.

Thus, if $\mathcal{C} \leq \mathbb{F}_q^n$ is cyclic with generator polynomial $g \in \mathbb{F}_q[X]$, then for $v \in \mathbb{F}_q^n$

Table 6: Cyclic codes.

$$\begin{tikzcd}
 & & \mathbb{F}_q[X] \arrow[r, "\overline{}"] \arrow[d, dash] & \overline{\mathbb{F}_q[X]} \arrow[d, dash] \\
 & & \langle g\rangle \arrow[r] \arrow[d, dash] & \langle\overline{g}\rangle \arrow[d, dash] \\
\mathbb{F}_q^n \arrow[r, "\nu"] \arrow[d, dash] & \mathbb{F}_q[X]_{<n} \arrow[uur, dash] \arrow[d, dash] & \langle X^n-1\rangle \arrow[r] & \{0\} \\
\mathcal{C} \arrow[r] \arrow[d, dash] & \nu(\mathcal{C}) \arrow[uur, dash] \arrow[d, dash] & & \\
\{0\} \arrow[r] & \{0\} \arrow[uur, dash] & &
\end{tikzcd}$$

we have $v \in \mathcal{C}$ if and only if $g \mid \nu(v) \in \mathbb{F}_q[X]$. Moreover, if $\mathcal{C}$ is non-trivial then we have $g \in \nu(\mathcal{C}) \leq \mathbb{F}_q[X]_{<n}$, thus g is given as $\gcd(\nu(\mathcal{C}))$, or likewise as a non-zero polynomial of smallest degree in $\nu(\mathcal{C})$. Hence if $\mathcal{C}' \leq \mathbb{F}_q^n$ is cyclic with generator polynomial $g' \in \mathbb{F}_q[X]$, then we have $\mathcal{C}' \leq \mathcal{C}$ if and only if $g \mid g'$; in particular $\mathcal{C}+\mathcal{C}' \in \mathbb{F}_q^n$ and $\mathcal{C}\cap\mathcal{C}' \in \mathbb{F}_q^n$ have generator polynomial $\gcd(g,g') \in \mathbb{F}_q[X]$ and $\operatorname{lcm}(g,g') \in \mathbb{F}_q[X]$, respectively.

Example. For $g = 1 \in \mathbb{F}_q[X]$ we get $\langle\overline{1}\rangle = \overline{\mathbb{F}_q[X]}$, thus $\mathcal{C} = \mathbb{F}_q^n$; while $g = X^n - 1 \in \mathbb{F}_q[X]$ yields $\langle\overline{X^n-1}\rangle = \{0\} \trianglelefteq \overline{\mathbb{F}_q[X]}$, thus $\mathcal{C} = \{0\} \leq \mathbb{F}_q^n$.

The repetition code $\mathcal{C} := \langle[1,\ldots,1]\rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^n$ corresponds to $\langle\overline{g}\rangle = \overline{\langle g\rangle_{\mathbb{F}_q}} \trianglelefteq \overline{\mathbb{F}_q[X]}$, where $g := \nu([1,\ldots,1]) = \sum_{i=0}^{n-1} X^i = \frac{X^n-1}{X-1} \in \mathbb{F}_q[X]$ is the associated monic generator polynomial.

We consider the parity check code $\mathcal{C}^\perp = \{[c_0,\ldots,c_{n-1}] \in \mathbb{F}_q^n; \sum_{i=0}^{n-1} c_i = 0\} \leq \mathbb{F}_q^n$: For $f := \sum_{i=0}^{n-1} c_i X^i \in \mathbb{F}_q[X]$ we have $\sum_{i=0}^{n-1} c_i = 0$ if and only if $f(1) = 0$, which holds if and only if $X - 1 \mid f$. Hence $\mathcal{C}^\perp$ corresponds to $\langle\overline{h}\rangle \trianglelefteq \overline{\mathbb{F}_q[X]}$, where $h := \nu([-1,1,0,\ldots,0]) = X-1 \in \mathbb{F}_q[X]$ is the associated monic generator polynomial; note that we have $gh = X^n - 1$ in accordance with (8.3) below.

Example. The non-zero elements of the Hamming code $\mathcal{H} \leq \mathbb{F}_2^7$, up to the linear equivalence applied in (8.1), consist of the cyclic shifts of $[1,1,0,1,0,0,0]$ and of $[1,0,1,1,1,0,0]$, together with 1_7. Hence $\nu([1,1,0,1,0,0,0]) = X^3 + X + 1 \in \mathbb{F}_2[X]$ is the non-zero polynomial of smallest degree in $\nu(\mathcal{H})$. Thus $\mathcal{H}$ corresponds to $\overline{\nu}(\mathcal{H}) = \langle\overline{g}\rangle \trianglelefteq \mathbb{F}_2[X]$ with generator polynomial $g := X^3+X+1 \in \mathbb{F}_2[X]$; note that $X^7 + 1 = (X+1)(X^3+X+1)(X^3+X^2+1) \in \mathbb{F}_2[X]$.

In accordance with (8.3) below this can also be read off from the generator matrix of $\mathcal{H}$ given above; see also (9.4). Moreover, we have the check polynomial $h := \frac{X^7+1}{g} = (X+1)(X^3+X^2+1) = X^4+X^2+X+1 \in \mathbb{F}_2[X]$, hence $\mathcal{H}^\perp$ has generator polynomial $h^* := X^4+X^3+X^2+1 \in \mathbb{F}_2[X]$, as can also been seen from the check matrix of $\mathcal{H}$ given above.

(8.3) Theorem. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a cyclic code with generator polynomial $g = \sum_{i=0}^k g_i X^i \in \mathbb{F}_q[X]$ of degree $k := \deg(g) \in \{0, \ldots, n\}$. Let $h = \sum_{i=0}^{n-k} h_i X^i \in \mathbb{F}_q[X]$ such that $X^n - 1 = gh \in \mathbb{F}_q[X]$; hence we have $\deg(h) = n-k$.

a) Then we have $\dim_{\mathbb{F}_q}(\mathcal{C}) = n-k$, and a generator matrix of $\mathcal{C}$ is given as

$$G := \begin{bmatrix} g_0 & g_1 & \cdots & g_k & \cdot & \cdot & \cdots & \cdot \\ \cdot & g_0 & \cdots & g_{k-1} & g_k & \cdot & \cdots & \cdot \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot & g_0 & \cdots & g_{k-1} & g_k \end{bmatrix} \in \mathbb{F}_q^{(n-k)\times n}.$$

b) The dual code $\mathcal{C}^\perp \leq \mathbb{F}_q^n$ is cyclic as well, and is generated by the **reversed polynomial** $h^* := X^{\deg(h)} \cdot h(X^{-1}) = \sum_{i=0}^{n-k} h_{n-k-i} X^i \in \mathbb{F}_q[X]$ associated with h; hence h is called a **check polynomial** of $\mathcal{C}$. Thus a generator matrix of $\mathcal{C}^\perp$, that is a check matrix of $\mathcal{C}$, is given as

$$H := \begin{bmatrix} h_{n-k} & h_{n-k-1} & \cdots & h_0 & \cdot & \cdot & \cdots & \cdot \\ \cdot & h_{n-k} & \cdots & h_1 & h_0 & \cdot & \cdots & \cdot \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \cdot & \cdot & \cdots & \cdot & h_{n-k} & \cdots & h_1 & h_0 \end{bmatrix} \in \mathbb{F}_q^{k\times n}.$$

Note that, by reversing the order of the columns, the cyclic codes generated by h and h^* are linearly equivalent.

Proof. a) For any $v \in \mathcal{C}$ we have $\overline{\nu}(v) = \overline{gf} \in \overline{\mathbb{F}_q[X]}$ for some $f \in \mathbb{F}_q[X]$. Let $f = qh + r \in \mathbb{F}_q[X]$, where $q, r \in \mathbb{F}_q[X]$ such that $r = 0$ or $\deg(r) < \deg(h) = n-k$. Then we have $\nu(v) - gf = \nu(v) - g(qh+r) = \nu(v) - gr - (X^n-1)q \in \langle X^n - 1\rangle \trianglelefteq \mathbb{F}_q[X]$, which implies $\overline{\nu}(v) = \overline{gr} \in \overline{\mathbb{F}_q[X]}$. Thus since $\dim_{\mathbb{F}_q}(\mathcal{C}) = \dim_{\mathbb{F}_q}(\langle\overline{g}\rangle) = \dim_{\mathbb{F}_q}(\overline{\mathbb{F}_q[X]}) - \dim_{\mathbb{F}_q}(\mathbb{F}_q[X]/\langle g\rangle) = n-k$ we conclude that $\{\overline{g}, \overline{g}\overline{X}, \ldots, \overline{g}\overline{X}^{n-k-1}\} \subseteq \langle\overline{g}\rangle = \overline{\nu}(\mathcal{C})$ is an $\mathbb{F}_q$-basis, which consists of the images under $\overline{\nu}$ of the rows of G.

b) Note first that, since evaluation of polynomials is a monoid homomorphism, and by additivity of polynomial degrees, we have $a^*b^* = (ab)^* \in \mathbb{F}_q[X]$, for any $0 \neq a, b \in \mathbb{F}_q[X]$. Now, from $gh = X^n - 1$ we conclude that $h_0 \neq 0$, hence we have $\deg(h^*) = n-k$, and from $g^*h^* = (gh)^* = (X^n-1)^* = 1 - X^n \in \mathbb{F}_q[X]$, we infer that $h^* \mid X^n - 1 \in \mathbb{F}_q[X]$. Hence H is a generator matrix of a cyclic code with generator polynomial h^*, having dimension $\mathrm{rk}_{\mathbb{F}_q}(H) = k = \dim_{\mathbb{F}_q}(\mathcal{C}^\perp)$. Thus it suffices to show that the rows of H are orthogonal to the rows of G:

For $i \in \{1, \ldots, n-k\}$ the i-th row of G is $v_i := [0, \ldots, 0, g_0, \ldots, g_k, 0, \ldots, 0] \in \mathbb{F}_q^n$, where g_0 is the i-th entry, and for $j \in \{1, \ldots, k\}$ the j-th row if H is $w_j := [0, \ldots, 0, h_{n-k}, \ldots, h_0, 0, \ldots, 0] \in \mathbb{F}_q^n$, where h_{n-k} is the j-th entry. Thus letting $g_l := 0$ for $l > k$ and $l < 0$, and $h_l := 0$ for $l > n-k$ and $l < 0$, we have $\langle v_i, w_j \rangle = \sum_{l=1}^n g_{l-i} h_{n-k+j-l} \in \mathbb{F}_q$. Since $1 - i \leq 0$ and $n - i \geq k$, and $(n-k+j) - 1 \geq n-k$ and $(n-k+j) - n \leq 0$, the latter sum equals the coefficient of $X^{(n-k+j)-i}$ in $gh = X^n - 1 \in \mathbb{F}_q[X]$. Since $1 \leq n-k+j-i \leq n-1$, from that we conclude $\langle v_i, w_j \rangle = 0$. ♯

(8.4) Cyclic redundancy check (CRC) codes [Peterson, 1961]. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be cyclic with generator polynomial $g = \sum_{i=0}^k g_i X^i \in \mathbb{F}_q[X]$ of degree $\deg(g) = k$ and associated generator matrix $G \in \mathbb{F}_q^{(n-k)\times n}$. Since $g_k \neq 0$ the matrix G can be transformed by Gaussian row elimination to $[A \mid E_{n-k}] \in \mathbb{F}_q^{(n-k)\times n}$, for some $A \in \mathbb{F}_q^{(n-k)\times k}$; this does not affect the cyclicity of $\mathcal{C}$.

Using the generator matrix $[A \mid E_{n-k}]$, a word $v = [a_0, \ldots, a_{n-k-1}] \in \mathbb{F}_q^{n-k}$ is encoded into $w = [b_0, \ldots, b_{k-1}; a_0, \ldots, a_{n-k-1}] \in \mathbb{F}_q^n$, where we have to find $[b_0, \ldots, b_{k-1}] \in \mathbb{F}_q^k$: We have $\nu(w) = \nu([b_0, \ldots, b_{k-1}]) + X^k \cdot \nu(v) \in \mathbb{F}_q[X]$. Polynomial division yields $X^k \cdot \nu(v) = \sum_{i=0}^{n-k-1} a_i X^{i+k} = qg + r$, for $q, r \in \mathbb{F}_q[X]$ such that $r = 0$ or $\deg(r) < \deg(g) = k$. Since $w \in \mathcal{C}$ we have $g \mid \nu(w) = (qg + r) + (\sum_{i=0}^{k-1} b_i X^i)$, thus $r + \sum_{i=0}^{k-1} b_i X^i = 0$. Hence $\nu([b_0, \ldots, b_{k-1}])$ is the remainder of the shifted polynomial $-X^k \cdot \nu(v)$ upon polynomial division by g.

Error detection, which is the typical application, runs as follows: Given $w \in \mathbb{F}_q^n$, we have $w \in \mathcal{C}$ if and only if $g \mid \nu(w) \in \mathbb{F}_q[X]$. Again polynomial division yields $\nu(w) = qg + r$, for $q, r \in \mathbb{F}_q[X]$ such that $r = 0$ or $\deg(r) < \deg(g) = k$. Hence we have $w \in \mathcal{C}$ if and only if $r = 0$, and in this case $w = [b_0, \ldots, b_{k-1}; b_k, \ldots, b_{n-1}] \in \mathcal{C}$ is decoded to $[b_k, \ldots, b_{n-1}] \in \mathbb{F}_q^{n-k}$. We discuss a few types of errors:

i) A **burst error** of length $l \in \{0, \ldots, n\}$ is an error vector $u = [c_0, \ldots, c_{n-1}] \in \mathbb{F}_q^n$, such that $c_i \neq 0$ only if $i \in \{j, \ldots, j+l-1\} \subseteq \mathbb{Z}_n$, for some $j \in \mathbb{Z}_n$.

Then $\mathcal{C}$ detects all burst errors of length $l \leq k$; in particular, if $k \geq 1$ all single errors are detected: We may assume that $u = [0, \ldots, 0, c_0, \ldots, c_{l-1}, 0, \ldots, 0] \in \mathbb{F}_q^n$, where c_0 occurs at position $j \in \{1, \ldots, n\}$, then $\nu(u) = X^{j-1} \cdot \sum_{i=0}^{l-1} c_i X^i \in \mathbb{F}_q[X]$, hence from $\gcd(g, X) = 1$ and $\deg(g) = k$ we infer $g \nmid \nu(u)$, thus $u \notin \mathcal{C}$.

ii) We show that we have $\mathcal{C} = \mathcal{C}' \leq \mathbb{F}_q^n$, the latter denoting the expurgated code, if and only if $X - 1 \mid g \in \mathbb{F}_q[X]$: We have $\mathcal{C} = \mathcal{C}'$ if and only if for all $w = [b_0, \ldots, b_{n-1}] \in \mathcal{C}$ we have $\nu(w)(1) = (\sum_{i=0}^{n-1} b_i X^i)(1) = \sum_{i=0}^{n-1} b_i = 0$, that is $X - 1 \mid \nu(w)$ for all $w \in \mathcal{C}$, in other words $\langle g \rangle = \nu(\mathcal{C}) \subseteq \langle X - 1 \rangle \trianglelefteq \mathbb{F}_q[X]$.

In particular, for $q = 2$ we have $X + 1 \mid g \in \mathbb{F}_2[X]$ if and only if $\mathcal{C}$ is an even-weight code. In this case, if $u = [c_0, \ldots, c_{n-1}] \in \mathbb{F}_2^n$ is an error vector of odd weight, then we have $\nu(u)(1) = \sum_{i=0}^{n-1} c_i \equiv \mathrm{wt}(u) \not\equiv 0 \pmod 2$, hence $g \nmid \nu(u)$, saying that $\mathcal{C}$ detects all errors of odd weight.

iii) Letting $q = 2$, for a double error occurring in positions $i < j \in \{1, \ldots, n\}$, we have the error vector $u = e_i + e_j \in \mathbb{F}_2^n$, hence $\nu(u) = X^{i-1}(X^{j-i}+1) \in \mathbb{F}_2[X]$. Thus, since $\gcd(g, X) = 1$ and $1 \leq j-i \leq n-1$, all double errors are detected if and only if $g \nmid X^m + 1$ for all $m \in \{1, \ldots, n-1\}$, that is if and only if g has an n-primitive divisor; recall that an irreducible polynomial $f \in \mathbb{F}_q[X]$ such that $f \mid X^n - 1$, but $f \nmid X^m - 1$ for all $m \in \{1, \ldots, n-1\}$, is called ***n*-primitive**.

Example. Actually, CRC codes over $\mathbb{F}_2$ are used throughout information technology. In particular, polynomial division over $\mathbb{F}_2$ is extremely fast, on a machine level just consisting of bit shifts and `exor` commands.

For example, this is used in the **Universal Serial Bus (USB)** [≥1996] data transmission standard: The 'CRC-5-USB' polynomial $X^5 + X^2 + 1 \in \mathbb{F}_2[X]$ is used to add $k = 5$ check bits to 'token' packets consisting of 11 information bits, making up a code of length 16, that is 2 Bytes. The polynomial $X^5 + X^2 + 1 \in \mathbb{F}_2[X]$ is irreducible, hence splits in $\mathbb{F}_{2^5} = \mathbb{F}_{32}$ and thus divides $X^{31} + 1 \in \mathbb{F}_2[X]$, and since 31 is a prime it is 31-primitive, entailing $n = 31$. Thus the code used actually used is a 15-fold shortened cyclic code; note that the encoding and decoding algorithms are not affected by shortening.

Similarly, for 'data' packets, having length up to 1023 Bytes, that is up to 8184 bits, the 'CRC-16-USB' polynomial $X^{16}+X^{15}+X^2+1 = (X+1)(X^{15}+X+1) \in \mathbb{F}_2[X]$ is used. The polynomial $X^{15} + X + 1 \in \mathbb{F}_2[X]$ is the lexicographically smallest irreducible polynomial of degree 15, hence splits in $\mathbb{F}_{2^{15}} = \mathbb{F}_{32768}$ and thus divides $X^{32767} + 1 \in \mathbb{F}_2[X]$, actually it is 32767-primitive, where $2^{15} - 1 = 32767 = 7 \cdot 31 \cdot 151$, entailing $n = 32767$. Thus the code used actually used is a 24583-fold shortened cyclic (even-weight) code.

(8.5) Example: The RWTH-ID [Bunsen, J.M., 2007]. Identity management is a task which all large organizations dealing with many customers are faced with. The aim is to associate an identity number with any customer, in order to uniquely identify them. It should have the following properties: The set of available numbers should be large enough; the number should not convey any further information about the customer in question; the number should be easy to remember to human beings; and it should be possible to detect simple transmission errors.

To create identity numbers, an alphabet $\mathcal{X}$ consisting of 32 alpha-numerical symbols, decimal digits and capital Latin letters, is used; in order to avoid mixing up symbols, the letters `I`, `J`, `O` and `V`, resembling `1`, `0` and `U`, respectively, are not allowed. Thus using 5 information symbols, we obtain a set of $|\mathcal{X}|^5 = 32^5 = 33\,554\,432 \sim 3 \cdot 10^7$ words over $\mathcal{X}$, to which we add a single check symbol, yielding identity numbers being words of length 6. To ease remembering identity numbers, these are written as two words of length three each, connected by a hyphen, for example `SL8-BRX`.

By source coding, $\mathcal{X}$ is encoded into the elements of $\mathbb{F}_2^5$ as given in Table 7. Thus we get a linear binary code $\mathcal{D} \leq \mathbb{F}_2^{30}$ of length $6 \cdot 5 = 30$ and dimension $\dim_{\mathbb{F}_2}(\mathcal{D}) = 5 \cdot 5 = 25$. To ease practical implementation, and to achieve the desired error detection properties, namely to be able to detect single errors and adjacent transposition errors, we aim at choosing $\mathcal{D}$ related to a cyclic code.

To this end, we look for a suitable cyclic code $\mathcal{C} \leq \mathbb{F}_2^n$ of length $n \geq 30$ and dimension $\dim_{\mathbb{F}_2}(\mathcal{C}) = n-5$, then $\mathcal{D} \leq \mathbb{F}_2^{30}$ such that $\dim_{\mathbb{F}_2}(\mathcal{D}) = 25$ is obtained by $(n-30)$-fold shortening; recall that the encoding and decoding algorithms are not affected by shortening. Thus we look for a suitable generator polynomial $g \in \mathbb{F}_2[X]$ of degree $k := \deg(g) = 5$, dividing the polynomial $X^n + 1 \in \mathbb{F}_2[X]$. We consider the relevant error types:

A single error yields a burst error of length 5, hence any such error is detected by any cyclic code with the above parameters. Moreover, an adjacent transposition error yields an error vector $u = [0, \ldots, 0; c_0, \ldots, c_4; c_0, \ldots, c_4; 0, \ldots, 0] \in \mathbb{F}_2^n$, where $[c_0, \ldots, c_4] \in \mathbb{F}_2^5$. Hence we have $\nu(u) = X^j(X^5+1) \cdot \sum_{i=0}^4 c_i X^i \in \mathbb{F}_2[X]$, where the leftmost c_0 occurs in entry $j \in \{0, \ldots, n-1\}$. Hence all adjacent transposition errors are detected if and only if $g \nmid \nu(u)$ for all error vectors as above. This in turn holds if and only if $\gcd(g, X^5+1) = 1$; note that we have the factorization $X^5 + 1 = (X+1)(X^4+X^3+X^2+X+1) \in \mathbb{F}_2[X]$.

Since $\gcd(g, X) = 1 = \gcd(g, X+1)$ we conclude that g cannot possibly have a linear factor, thus either g is the product of two irreducible polynomials of degree 2 and 3, respectively, or g is irreducible of degree 5. Now $X^2+X+1 \in \mathbb{F}_2[X]$ is the unique irreducible polynomial of degree 2, and $X^3+X+1 \in \mathbb{F}_2[X]$ and $X^3+X^2+1 \in \mathbb{F}_2[X]$ are those of degree 3, hence leading to the candidate polynomials $(X^2+X+1)(X^3+X+1) = X^5+X^4+1$ and $(X^2+X+1)(X^3+X^2+1) = X^5+X+1$, which both split in $\mathbb{F}_{2^{\mathrm{lcm}(2,3)}} = \mathbb{F}_{2^6} = \mathbb{F}_{64}$.

Moreover, as further candidates there are the six irreducible polynomials of degree 5, which split in $\mathbb{F}_{2^5} = \mathbb{F}_{32}$. Indeed, for $n := 31 = 2^5 - 1$ we find that

$$\begin{array}{rcl} X^{31}+1 & = & (X+1) \cdot (X^5+X^2+1) \cdot (X^5+X^3+1) \\ & & \cdot\, (X^5+X^3+X^2+X+1) \cdot (X^5+X^4+X^2+X+1) \\ & & \cdot\, (X^5+X^4+X^3+X+1) \cdot (X^5+X^4+X^3+X^2+1) \in \mathbb{F}_2[X]. \end{array}$$

Thus either of the irreducible polynomials of degree 5 is a suitable generator polynomial; since 31 is a prime all of them are 31-primitive. For the RWTH-ID the 'CRC-5-USB' polynomial $g := X^5+X^2+1 \in \mathbb{F}_2[X]$ is used; let $\mathcal{C} \leq \mathbb{F}_2^{31}$ be the associated cyclic code and $\mathcal{D} := \mathcal{C}^\circ \leq \mathbb{F}_2^{30}$.

For example, for `L8BRX` from Table 7 we get

$$\begin{array}{rrcl} h & = & & (1+X^3+X^4) \\ & + & X^5 \cdot & (X) \\ & + & X^{10} \cdot & (X+X^3+X^4) \\ & + & X^{15} \cdot & (1+X) \\ & + & X^{20} \cdot & (1+X+X^2+X^4) \in \mathbb{F}_2[X], \end{array}$$

Table 7: The alphabet of the RWTH-ID.

0	00000	8	01000	G	10000	R	11000
1	00001	9	01001	H	10001	S	11001
2	00010	A	01010	K	10010	T	11010
3	00011	B	01011	L	10011	U	11011
4	00100	C	01100	M	10100	W	11100
5	00101	D	01101	N	10101	X	11101
6	00110	E	01110	P	10110	Y	11110
7	00111	F	01111	Q	10111	Z	11111

where polynomial division of $X^5 \cdot h$ by g yields the remainder $1+X+X^4 \in \mathbb{F}_2[X]$, which belongs to the symbol S, saying that SL8-BRX is a valid identity number.

9 BCH codes

(9.1) Roots of unity. Let $\mathbb{F}_q$ be the field with q elements, and let $n \in \mathbb{N}$ such that $\gcd(q,n) = 1$. We consider the polynomial $X^n - 1 \in \mathbb{F}_q[X]$. Since $n \neq 0 \in \mathbb{F}_q$ we have $\gcd(\frac{\partial}{\partial X}(X^n-1), X^n-1) = \gcd(nX^{n-1}, X^n-1) = 1 \in \mathbb{F}_q[X]$, implying that $X^n - 1 \in \mathbb{F}_q[X]$ is square-free, that is a product of pairwise distinct monic irreducible polynomials.

We aim at describing its factorization more precisely: Let $\mathbb{F}_q \subseteq \overline{\mathbb{F}}$ be the algebraic closure of $\mathbb{F}_q$. Since $X^n - 1 \in \overline{\mathbb{F}}[X]$ still is square-free, we conclude that it splits into pairwise distinct linear factors, thus letting $\mathcal{V}_n := \mathcal{V}(X^n-1) \subseteq \overline{\mathbb{F}}^*$ be its **set of zeroes**, we have $|\mathcal{V}_n| = n$ and $X^n - 1 = \prod_{\zeta \in \mathcal{V}_n}(X - \zeta) \in \overline{\mathbb{F}}[X]$.

Since whenever $\zeta, \zeta' \in \mathcal{V}_n$ we also have $\zeta^{-1}\zeta' \in \mathcal{V}_n$, we conclude that $\mathcal{V}_n$ is a finite subgroup of $\overline{\mathbb{F}}^*$, hence by Artin's Theorem is cyclic. Thus there is a **primitive n-th root of unity** $\zeta_n \in \mathcal{V}_n$, that is an element of multiplicative order n, so that $\mathcal{V}_n = \{\zeta_n^i \in \overline{\mathbb{F}}^*; i \in \mathbb{Z}_n\}$. Hence we have a group isomorphism $\mathbb{Z}_n \to \mathcal{V}_n \colon i \mapsto \zeta_n^i$, the left hand side being an additive group. Moreover, $\zeta_n^i \in \overline{\mathbb{F}}^*$ has order $\min\{j \in \mathbb{N}; \zeta_n^{ij} = 1 \in \overline{\mathbb{F}}^*\} = \min\{j \in \mathbb{N}; n \mid ij\} = \frac{n}{\gcd(i,n)}$; in particular $\zeta_n^i \in \mathcal{V}_n$ is a primitive n-th root of unity if and only if $i \in \mathbb{Z}_n^*$.

Let $\mathbb{F}_q(\zeta_n) \subseteq \overline{\mathbb{F}}$ be the field generated by ζ_n over $\mathbb{F}_q$. Then $\mathbb{F}_q(\zeta_n)$ is the splitting field of $X^n - 1$, hence $\mathbb{F}_q \subseteq \mathbb{F}_q(\zeta_n)$ is a Galois extension. Thus its automorphism group $\Gamma := \mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F}_q(\zeta_n))$ has finite order $|\Gamma| = [\mathbb{F}_q(\zeta_n) \colon \mathbb{F}_q] := \dim_{\mathbb{F}_q}(\mathbb{F}_q(\zeta_n))$, hence $|\mathbb{F}_q(\zeta_n)| = |\mathbb{F}_q|^{[\mathbb{F}_q(\zeta_n)\colon \mathbb{F}_q]} = q^{|\Gamma|}$, that is $\mathbb{F}_q(\zeta_n) = \mathbb{F}_{q^{|\Gamma|}} \subseteq \overline{\mathbb{F}}$.

The group Γ is cyclic, generated by the Frobenius automorphism $\varphi_q \colon \mathbb{F}_q(\zeta_n) \to \mathbb{F}_q(\zeta_n) \colon a \mapsto a^q$, having order $|\varphi_q| = \min\{i \in \mathbb{N}; \varphi_q^i = \mathrm{id}_{\mathbb{F}_q(\zeta_n)}\} = \min\{i \in \mathbb{N}; \zeta_n^{q^i} = \zeta_n\} = \min\{i \in \mathbb{N}; q^i = 1 \in \mathbb{Z}_n\}$. Thus $|\Gamma| = |\varphi_q|$ coincides with the order of $q \in \mathbb{Z}_n^*$. Identifying $\mathcal{V}_n$ with $\mathbb{Z}_n$, the group Γ acts by the multiplication map $\varphi_q\colon \mathbb{Z}_n \to \mathbb{Z}_n\colon i \mapsto iq$, the Γ-orbits in $\mathbb{Z}_n$ being called **cyclotomic cosets**.

The monic divisors $g \mid X^n - 1 \in \mathbb{F}_q(\zeta_n)[X]$ are uniquely determined by their sets of zeroes $\mathcal{V}(g) \subseteq \mathcal{V}_n$ as $g = \prod_{\zeta\in\mathcal{V}(g)}(X-\zeta) \in \mathbb{F}_q(\zeta_n)[X]$. Since $\mathrm{Fix}_{\mathbb{F}_q(\zeta_n)}(\Gamma) = \mathbb{F}_q$, we have $g \in \mathbb{F}_q[X]$ if and only if $\mathcal{V}(g)$ is a union of Γ-orbits. Thus the monic irreducible divisors of $X^n - 1 \in \mathbb{F}_q[X]$, being called **cyclotomic polynomials** over $\mathbb{F}_q$, are given by the Γ-orbits on $\mathcal{V}_n$ as $\mu_i := \prod_{\zeta\in(\zeta_n^i)^\Gamma}(X-\zeta) \in \mathbb{F}_q[X]$.

The polynomial $\mu_i \in \mathbb{F}_q[X]$ is the minimum polynomial of $\zeta_n^i \in \mathbb{F}_q(\zeta_n)$ over $\mathbb{F}_q$, hence we have $[\mathbb{F}_q(\zeta_n^i)\colon \mathbb{F}_q] = \deg(\mu_i) = |(\zeta_n^i)^\Gamma| \mid |\Gamma| = [\mathbb{F}_q(\zeta_n)\colon \mathbb{F}_q]$. Thus we have equality $\deg(\mu_i) = |\Gamma|$ if and only if $\mathbb{F}_q(\zeta_n^i) = \mathbb{F}_q(\zeta_n)$; in particular this holds whenever ζ_n^i is a primitive n-th root of unity, that is $i \in \mathbb{Z}_n^*$. Note that Γ acts semi-regularly on the set of primitive n-th roots of unity, but not necessarily regularly, that is not necessarily transitively.

Example. For $q := 2$ and $n := 7$ we find that $2 \in \mathbb{Z}_7^*$ has order 3, thus $\mathbb{F}_2(\zeta_7) = \mathbb{F}_8$ and $\varphi_2 \in \mathrm{Aut}_{\mathbb{F}_2}(\mathbb{F}_8)$ has order 3. The Γ-orbits on $\mathcal{V}_7$ are $\mathcal{V}_7 = \{1\} \,\dot\cup\, \{\zeta_7^i; i \in \mathcal{O}'\} \,\dot\cup\, \{\zeta_7^i; i \in \mathcal{O}''\} \subseteq \mathbb{F}_8$, where the associated cyclotomic cosets are $\mathcal{O}' := \{1,2,4\}$ and $\mathcal{O}' := \{3,5,6\}$, thus $X^7 + 1 = (X+1)\cdot\prod_{i\in\mathcal{O}'}(X+\zeta_7^i)\cdot\prod_{i\in\mathcal{O}''}(X+\zeta_7^i) = \mu_0\mu_1\mu_3 = (X+1)\cdot(X^3+X+1)\cdot(X^3+X^2+1) \in \mathbb{F}_8[X]$, where the latter are irreducible in $\mathbb{F}_2[X]$; note that here we do not specify which of the factors has the chosen primitive 7-th root of unity ζ_7 as a zero.

(9.2) Zeroes of cyclic codes. a) Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q,n) = 1$, be the cyclic code with monic generator polynomial $g \in \mathbb{F}_q[X]$ of degree $k \in \{0,\ldots,n\}$. Let $\mathcal{V}(\mathcal{C}) := \mathcal{V}(g) \subseteq \mathcal{V}_n$ be the **set of zeroes** of $\mathcal{C}$; hence $|\mathcal{V}(\mathcal{C})| = \deg(g) = k$. Recall that the monic divisors of $X^n - 1$ are in bijection with the Γ-invariant subsets of $\mathcal{V}_n$, where $\Gamma := \mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F}_q(\zeta_n))$.

Hence any subset $\mathcal{V} \subseteq \mathcal{V}_n$ whose smallest Γ-invariant superset equals $\mathcal{V}(\mathcal{C})$, that is $\mathcal{V}(\mathcal{C}) = \bigcup_{\zeta\in\mathcal{V}}\zeta^\Gamma \subseteq \mathcal{V}_n$, is called a **defining set** of $\mathcal{C}$, and $\mathcal{C}$ is called the code **associated** with $\mathcal{V}$; in particular, $\mathcal{V}(\mathcal{C})$ is the unique maximal defining set of $\mathcal{C}$.

For $v \in \mathbb{F}_q^n$ we have $v \in \mathcal{C}$ if and only if $g \mid \nu(v) \in \mathbb{F}_q[X]$. Since $g \in \mathbb{F}_q(\zeta_n)[X]$ is square-free, this is equivalent to $\mathcal{V}(\mathcal{C}) \subseteq \mathcal{V}(\nu(v)) \subseteq \overline{\mathbb{F}}$, where $\overline{\mathbb{F}}$ is the algebraic closure of $\mathbb{F}_q$; note that $\mathcal{V}(\nu(v)) \cap \mathcal{V}_n$ is Γ-invariant. By taking Γ-orbits, this in turn is equivalent to $\mathcal{V} \subseteq \mathcal{V}(\nu(v))$ for any defining set $\mathcal{V}$ of $\mathcal{C}$; that is we have $\mathcal{C} = \{[c_0,\ldots,c_{n-1}] \in \mathbb{F}_q^n; \sum_{i=0}^{n-1} c_i\zeta^i = 0 \in \mathbb{F}_q(\zeta_n) \text{ for all } \zeta \in \mathcal{V}\}$. Moreover, we recover $\mathcal{V}(\mathcal{C})$ as $\mathcal{V}(\mathcal{C}) = \bigcap_{v\in\mathcal{C}}\mathcal{V}(\nu(v)) \subseteq \overline{\mathbb{F}}$.

b) We determine $\mathcal{V}(\mathcal{C}^\perp) \subseteq \mathcal{V}_n$ for the dual code $\mathcal{C}^\perp \leq \mathbb{F}_q^n$: Letting $h \in \mathbb{F}_q[X]$ such that $gh = X^n - 1 \in \mathbb{F}_q[X]$, we have $h = \prod_{\zeta\in\mathcal{V}_n\setminus\mathcal{V}(\mathcal{C})}(X-\zeta) \in \mathbb{F}_q(\zeta_n)[X]$, thus we get $h^* = \prod_{\zeta\in\mathcal{V}_n\setminus\mathcal{V}(\mathcal{C})}(X-\zeta)^* = \prod_{\zeta\in\mathcal{V}_n\setminus\mathcal{V}(\mathcal{C})}((-\zeta)\cdot(X-\zeta^{-1})) \in \mathbb{F}_q(\zeta_n)[X]$, from which we infer $\mathcal{V}(\mathcal{C}^\perp) = \mathcal{V}(h^*) = \{\zeta \in \mathcal{V}_n; \zeta^{-1} \notin \mathcal{V}(\mathcal{C})\}$.

(9.3) Theorem. a) Let $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q,n) = 1$, be the cyclic code with set of zeroes $\mathcal{V}(\mathcal{C}) = \{\zeta_n^{a_1},\ldots,\zeta_n^{a_k}\} \subseteq \mathcal{V}_n$, where $\{a_1,\ldots,a_k\} \subseteq \mathbb{Z}_n$ and

$|\mathcal{V}(\mathcal{C})| = k \in \{0, \ldots, n\}$. Then the **Delsarte matrix** [1975]

$$H = H(\mathcal{V}(\mathcal{C})) := [\zeta_n^{(j-1)a_i}]_{ij} = \begin{bmatrix} 1 & \zeta_n^{a_1} & \zeta_n^{2a_1} & \ldots & \zeta_n^{(n-1)a_1} \\ 1 & \zeta_n^{a_2} & \zeta_n^{2a_2} & \ldots & \zeta_n^{(n-1)a_2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \zeta_n^{a_k} & \zeta_n^{2a_k} & \ldots & \zeta_n^{(n-1)a_k} \end{bmatrix} \in \mathbb{F}_q(\zeta_n)^{k \times n}$$

is a generalized check matrix of $\mathcal{C}$, that is we have $\mathcal{C} = \ker(H^{\mathrm{tr}}) \cap \mathbb{F}_q^n$. Moreover, if $\mathcal{V} \subseteq \mathcal{V}(\mathcal{C})$ is a defining set of $\mathcal{C}$, then we have $\mathcal{C} = \ker(H(\mathcal{V})^{\mathrm{tr}}) \cap \mathbb{F}_q^n$.

b) If $\mathcal{V}(\mathcal{C})$ contains a **consecutive set** $\{\zeta_n^a, \zeta_n^{a+b}, \ldots, \zeta_n^{a+(\delta-2)b}\}$, where $a \in \mathbb{Z}_n$ and $b \in \mathbb{Z}_n^*$ and $\delta \in \{1, \ldots, n+1\}$, then $\mathcal{C}$ has minimum distance $d(\mathcal{C}) \geq \delta$.

Note that for $\delta = 1$ the consecutive set is empty, and we trivially have $d(\mathcal{C}) \geq \delta$; and that for $\delta = n+1$ we have $\mathcal{V}(\mathcal{C}) = \mathcal{V}_n$, implying $\mathcal{C} = \{0\}$, thus $d(\mathcal{C}) = \infty \geq \delta$.

Proof. a) The submatrix of H consisting of the first k columns is a Vandermonde matrix associated with the pairwise distinct roots of unity $\{\zeta_n^{a_1}, \ldots, \zeta_n^{a_k}\}$, hence is invertible. Thus we infer $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H) = k$, hence $\dim_{\mathbb{F}_q(\zeta_n)}(\ker(H^{\mathrm{tr}})) = n-k$, which implies $\dim_{\mathbb{F}_q}(\ker(H^{\mathrm{tr}}) \cap \mathbb{F}_q^n) \leq n-k$; note that whenever $\mathcal{S} \subseteq \mathbb{F}_q^n$ then Gaussian elimination shows that $\dim_{\mathbb{F}_q}(\langle \mathcal{S} \rangle_{\mathbb{F}_q}) = \dim_{\mathbb{F}_q(\zeta_n)}(\langle \mathcal{S} \rangle_{\mathbb{F}_q(\zeta_n)})$.

Since $\dim_{\mathbb{F}_q}(\mathcal{C}) = n-k$ it suffices to show that $\mathcal{C} \leq \ker(H^{\mathrm{tr}})$: Let $g = \sum_{i=0}^k g_i X^i := \prod_{\zeta \in \mathcal{V}(\mathcal{C})}(X - \zeta) \in \mathbb{F}_q(\zeta_n)[X]$ be the monic generator polynomial of $\mathcal{C}$. For the i-th row $v_i = [0, \ldots, 0, g_0, \ldots, g_k, 0, \ldots, 0] \in \mathbb{F}_q^n$ of the generator matrix of $\mathcal{C}$ associated with g, where $i \in \{1, \ldots, n-k\}$, and the j-th row $w_j \in \mathbb{F}_q(\zeta_n)^n$ of H, where $j \in \{1, \ldots, k\}$, we get $\langle v_i, w_j \rangle = \sum_{l=0}^k g_l \zeta_n^{(i+l-1)a_j} = \zeta_n^{(i-1)a_j} \cdot \sum_{l=0}^k g_l (\zeta_n^{a_j})^l = \zeta_n^{(i-1)a_j} \cdot g(\zeta_n^{a_j}) = 0 \in \mathbb{F}_q(\zeta_n)$, hence $v_i \in \ker(H^{\mathrm{tr}})$.

Finally, for any row $w \in \mathbb{F}_q(\zeta_n)^n$ of H there is a row $u \in \mathbb{F}_q(\zeta_n)^n$ of $H(\mathcal{V})$ such that $w = u^{q^i} = \varphi_q^i(u)$, for some $i \in \mathbb{N}_0$, where $\varphi_q \in \Gamma$ is applied componentwise. Since for $v \in \mathbb{F}_q^n$ we have $v^q = v$, we from $\langle v, w \rangle = \langle v, u^{q^i} \rangle = \langle v^{q^i}, u^{q^i} \rangle = \langle v, w \rangle^{q^i} \in \mathbb{F}_q(\zeta_n)$ infer that $v \in \ker(H^{\mathrm{tr}}) \cap \mathbb{F}_q^n$ if and only if $v \in \ker(H(\mathcal{V})^{\mathrm{tr}}) \cap \mathbb{F}_q^n$.

b) Since $b \in \mathbb{Z}_n^*$ we conclude that $\zeta_n^b \in \mathcal{V}_n$ is a primitive n-th root of unity as well. Letting $c := ab^{-1} \in \mathbb{Z}_n$, we observe that $\{\zeta_n^a, \zeta_n^{a+b}, \ldots, \zeta_n^{a+(\delta-2)b}\} = \{(\zeta_n^b)^c, (\zeta_n^b)^{c+1}, \ldots, (\zeta_n^b)^{c+\delta-2}\}$. Hence we may assume that $b = 1$.

We consider the rows of H corresponding to the consecutive set, that is the matrix $H(\{\zeta_n^a, \ldots, \zeta_n^{a+\delta-2}\}) \in \mathbb{F}_q(\zeta_n)^{(\delta-1) \times n}$, and show that any $(\delta-1)$-subset of its columns is $\mathbb{F}_q(\zeta_n)$-linearly independent. Note that, although the argument given below is valid for all $\delta \in \{1, \ldots, n+1\}$, we may assume that $\delta \notin \{1, n+1\}$, so that in order to apply (4.3) we may moreover assume that $\mathcal{C} \neq \{0\}$, where it would be sufficient to show $\mathbb{F}_q$-linear independence only. Now:

Picking columns $\{j_1, \ldots, j_{\delta-1}\} \subseteq \{1, \ldots, n\}$ yields the square matrix

$$\widetilde{H} := \begin{bmatrix} \zeta_n^{(j_1-1)a} & \ldots & \zeta_n^{(j_{\delta-1}-1)a} \\ \zeta_n^{(j_1-1)(a+1)} & \ldots & \zeta_n^{(j_{\delta-1}-1)(a+1)} \\ \vdots & & \vdots \\ \zeta_n^{(j_1-1)(a+\delta-2)} & \ldots & \zeta_n^{(j_{\delta-1}-1)(a+\delta-2)} \end{bmatrix} \in \mathbb{F}_q(\zeta_n)^{(\delta-1)\times(\delta-1)}.$$

Hence we have

$$\widetilde{H} = \begin{bmatrix} 1 & \ldots & 1 \\ \zeta_n^{j_1-1} & \ldots & \zeta_n^{j_{\delta-1}-1} \\ \vdots & & \vdots \\ \zeta_n^{(j_1-1)(\delta-2)} & \ldots & \zeta_n^{(j_{\delta-1}-1)(\delta-2)} \end{bmatrix} \cdot \mathrm{diag}[\zeta_n^{(j_1-1)a}, \ldots, \zeta_n^{(j_{\delta-1}-1)a}],$$

where the left hand factor is a Vandermonde matrix associated with the pairwise distinct roots of unity $\{\zeta_n^{j_1-1}, \ldots, \zeta_n^{j_{\delta-1}-1}\}$, thus is invertible. $\sharp$

(9.4) Example: Hamming codes. We show that, generically, Hamming codes are linearly equivalent to cyclic codes: Let $\mathbb{F}_q$ be the field with q elements, let $k \geq 2$ such that $\gcd(k, q-1) = 1$, and let $n := \frac{q^k-1}{q-1}$; note that $\gcd(q, n) = 1$, and that the condition on k always holds for $q = 2$. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be the cyclic code with defining set $\{\zeta_n\}$, that is $\mathcal{V}(\mathcal{C}) = (\zeta_n)^\Gamma \subseteq \mathcal{V}_n$, in other words having $g = \mu_1 \in \mathbb{F}_q[X]$ as a generator polynomial. Then $\mathcal{C}$ is an $[n, n-k, 3]$-code, thus is linearly equivalent to the Hamming code $\mathcal{H}_k \leq \mathbb{F}_q^n$:

We show that the order of $q \in \mathbb{Z}_n^*$ equals k: We have $n \mid q^k - 1$, hence the order in question divides k; and assuming that $\frac{q^k-1}{q-1} = n \mid q^l - 1$ for some $l \in \{1, \ldots, k-1\}$, then we have $q^k - 1 \mid (q-1)(q^l-1) = q^{l+1} - q^l - q + 1$, thus $q^k \leq q^k - 2q + 2 \leq q^k - 2$, a contradiction. This implies $(\zeta_n)^\Gamma = \{\zeta_n, \zeta_n^q, \ldots, \zeta_n^{q^{k-1}}\}$, or equivalently $\deg(\mu_1) = k$, implying that $\dim_{\mathbb{F}_q}(\mathcal{C}) = n - k$.

We show that $\mathcal{C}$ has minimum distance $d(\mathcal{C}) = 3$: Since any $[n, n-k, 3]$-code is perfect, that is fulfills the Hamming bound, we have $d(\mathcal{C}) \leq 3$. Moreover, $\mathcal{V}(\mathcal{C})$ contains the consecutive set $\{\zeta_n, \zeta_n^q\}$ of length 2 and step size $q - 1$. Now $n = \frac{q^k-1}{q-1} = \sum_{i=0}^{k-1} q^i \equiv k \pmod{q-1}$ implies $\gcd(n, q-1) = \gcd(k, q-1) = 1$, hence we have $q - 1 \in \mathbb{Z}_n^*$, entailing $d(\mathcal{C}) \geq 3$. $\sharp$

(9.5) BCH-Codes [Bose, Ray-Chaudhuri, 1960; Hocquenghem, 1959]. **a)** A cyclic code $\mathcal{C} \leq \mathbb{F}_q^n$, where $\gcd(q, n) = 1$, associated with a (genuinely) consecutive set $\{\zeta_n^a, \ldots, \zeta_n^{a+\delta-2}\} \subseteq \mathcal{V}_n$ of length $\delta - 1$, where $a \in \mathbb{Z}_n$ and $\delta \in \{1, \ldots, n+1\}$, is called a **BCH code** of **designed distance** δ. Note that $\mathcal{C}$ might be a BCH code with respect to consecutive sets of varying lengths, or varying step sizes amounting to changing the chosen primitive n-th root of unity; the largest designed distance occurring is called the **Bose distance**. Hence for the minimum distance of $\mathcal{C}$ we have the **BCH bound** $d(\mathcal{C}) \geq \delta$.

If $n = q^{|\Gamma|} - 1$, that is the multiplicative group $\mathbb{F}_q(\zeta_n)^* = \mathbb{F}_{q^{|\Gamma|}}^*$ is generated by the **primitive element** $\zeta_n = \zeta_{q^{|\Gamma|}-1}$, then $\mathcal{C}$ is called **primitive**. If $a = 1$, that is the consecutive set considered is $\{\zeta_n, \ldots, \zeta_n^{\delta-1}\}$, then $\mathcal{C}$ is called a **narrow sense** BCH code. In particular, in the narrow sense, for $\delta = 1$ we get $\mathcal{V}(\mathcal{C}) = \emptyset$, thus $\mathcal{C} = \mathbb{F}_q^n$; for $\delta = n + 1$ we get $\mathcal{V}(\mathcal{C}) = \mathcal{V}_n$, thus $\mathcal{C} = \{0\}$; and for $\delta = n$ we get $\mathcal{V}(\mathcal{C}) = \{\zeta_n, \ldots, \zeta_n^{n-1}\} = \mathcal{V}_n \setminus \{1\} = \mathcal{V}(\frac{X^n-1}{X-1})$, thus $\mathcal{C}$ is the repetition code.

b) We comment on the parameters of BCH codes. Firstly, we consider the dimension of BCH codes: We have $n - \dim_{\mathbb{F}_q}(\mathcal{C}) = |\mathcal{V}(\mathcal{C})| = |\bigcup_{i=0}^{\delta-2}(\zeta_n^{a+i})^\Gamma| \leq (\delta-1)\cdot|\Gamma|$, hence in general we have $\dim_{\mathbb{F}_q}(\mathcal{C}) \geq n-(\delta-1)\cdot|\Gamma|$. More specially, if $q = 2$ and $\mathcal{C}$ is a narrow sense BCH code, then from $\mu_i = \mu_{2i} \in \mathbb{F}_2[X]$, for all $i \in \mathbb{Z}_n$, we get $\mathcal{V}(\mathcal{C}) = \bigcup_{i=1}^{\delta-1}(\zeta_n^i)^\Gamma = \bigcup_{i=1}^{\lfloor\frac{\delta}{2}\rfloor}(\zeta_n^{2i-1})^\Gamma \subseteq \mathcal{V}_n$, which yields the better estimate $\dim_{\mathbb{F}_2}(\mathcal{C}) \geq n - \lfloor\frac{\delta}{2}\rfloor \cdot |\Gamma|$.

Secondly, we consider the minimum distance of BCH codes, which in general might be strictly larger than the Bose distance. But we have **Peterson's Theorem [1967]**, saying that if $\mathcal{C}$ is a narrow sense BCH code of designed distance $\delta \mid n$, then $\mathcal{C}$ actually has minimum distance δ:

Let $n = l\delta$, where $l \in \mathbb{N}$. Then we have $X^n - 1 = (X^l - 1)\cdot\sum_{i=0}^{\delta-1} X^{il} \in \mathbb{F}_q[X]$. Since $\zeta_n^{il} \neq 1 \in \mathbb{F}_q(\zeta_n)$, for all $i \in \{1, \ldots, \delta-1\}$, we conclude that $\{\zeta_n, \ldots, \zeta_n^{\delta-1}\} \cap \mathcal{V}(X^l - 1) = \emptyset$. Thus we have $\{\zeta_n, \ldots, \zeta_n^{\delta-1}\} \subseteq \mathcal{V}(\sum_{i=0}^{\delta-1} X^{il})$, which implies that $\mathcal{V}(\mathcal{C}) = \bigcup_{i=1}^{\delta-1}(\zeta_n^i)^\Gamma \subseteq \mathcal{V}(\sum_{i=0}^{\delta-1} X^{il}) = \mathcal{V}(\nu(v))$, where $v := \sum_{i=0}^{\delta-1} e_{il} = [1, 0, \ldots, 0; \ldots; 1, 0, \ldots, 0] \in \mathbb{F}_q^n$; hence $v \in \mathcal{C}$, having weight δ. ♯

Example. For $q = 2$ and $n = 2^k - 1$, where $k \in \{2, 3, 4\}$, we get the following: We have $X^3+1 = \mu_0\mu_1 = (X+1)(X^2+X+1) \in \mathbb{F}_2[X]$ and $X^7+1 = \mu_0\cdot\mu_1\mu_3 = (X+1)\cdot(X^3+X+1)(X^3+X^2+1) \in \mathbb{F}_2[X]$; see (9.1). Moreover, we have $X^{15}+1 = \mu_0\mu_5\mu_3\cdot\mu_1\mu_7 = (X+1)(X^2+X+1)(X^4+X^3+X^2+X+1)\cdot(X^4+X+1)(X^4+X^3+1) \in \mathbb{F}_2[X]$, where $\mu_5 = \prod_{i\in\{5,10\}}(X-\zeta_{15}^i) = \prod_{i\in\{1,2\}}(X-\zeta_3^i)$ and $\mu_3 = \prod_{i\in\{3,6,9,12\}}(X-\zeta_{15}^i) = \prod_{i\in\{1,2,3,4\}}(X-\zeta_5^i)$ and $\mu_1 = \prod_{i\in\{1,2,4,8\}}(X-\zeta_{15}^i)$ and $\mu_7 = \prod_{i\in\{7,11,13,14\}}(X-\zeta_{15}^i) = \prod_{i\in\{1,2,4,8\}}(X-\zeta_{15}^{-i})$.

Hence we have the associated narrow sense primitive binary BCH codes $\mathcal{C}$ as given in Table 8, where we indicate the Bose distance δ, the monic generator polynomial $g \in \mathbb{F}_2[X]$, the union $\mathcal{O}$ of cyclotomic cosets associated with $\mathcal{V}(\mathcal{C})$, the dimension $\dim_{\mathbb{F}_2}(\mathcal{C}) = n - \deg(g) = n - |\mathcal{O}|$, and the minimum distance d. Note that in all cases given we observe that $\delta = d$: Except for $[k, \delta] \in \{[3, 3], [4, 7]\}$ this is explained by Peterson's Theorem, while the named cases are covered by (10.3) below; alternatively, for $[k, \delta] = [3, 3]$ we recall that any binary $[7, 4, 3]$-code already fulfills the Hamming bound; see (9.4).

(9.6) Reed-Solomon codes [1954]. a) We consider the first case of primitive BCH codes, that is $n := q - 1$. We have $\mathbb{F}_q^* = \langle\zeta_{q-1}\rangle$, hence $[\mathbb{F}_q(\zeta_{q-1}) \colon \mathbb{F}_q] = 1$, thus Γ is trivial, and $X^{q-1} - 1 = \prod_{i=0}^{q-2}(X - \zeta_{q-1}^i) \in \mathbb{F}_q[X]$. A primitive BCH code $\mathcal{C} \leq \mathbb{F}_q^{q-1}$ is called a **Reed-Solomon code**. Thus $\mathcal{V}(\mathcal{C})$ coincides with the

Table 8: Narrow sense primitive binary BCH codes.

δ	g	$\mathcal{O}$	dim	d
1	1	$\emptyset$	3	1
3	μ_1	$\{1,2\}$	1	3
4	$\mu_1\mu_0$	$\mathbb{Z}_3$	0	∞

δ	g	$\mathcal{O}$	dim	d
1	1	$\emptyset$	7	1
3	μ_1	$\{1,2,4\}$	4	3
7	$\mu_1\mu_3$	$\{1,\ldots,6\}$	1	7
8	$\mu_1\mu_3\mu_0$	$\mathbb{Z}_7$	0	∞

δ	g	$\mathcal{O}$	dim	d
1	1	$\emptyset$	15	1
3	μ_1	$\{1,2,4,8\}$	11	3
5	$\mu_1\mu_3$	$\{1,2,3,4,6,8,9,12\}$	7	5
7	$\mu_1\mu_3\mu_5$	$\{1,2,3,4,5,6,8,9,10,12\}$	5	7
15	$\mu_1\mu_3\mu_5\mu_7$	$\{1,\ldots,14\}$	1	15
16	$\mu_1\mu_3\mu_5\mu_7\mu_0$	$\mathbb{Z}_{15}$	0	∞

defining consecutive set $\mathcal{V} := \{\zeta_{q-1}^a, \ldots, \zeta_{q-1}^{a+\delta-2}\} = \zeta_{q-1}^{a-1} \cdot \{\zeta_{q-1}, \ldots, \zeta_{q-1}^{\delta-1}\} \subseteq \mathbb{F}_q^*$, where $a \in \mathbb{Z}_{q-1}$ and $\delta \in \{1, \ldots, q\}$ is the designed distance of $\mathcal{C}$.

Thus we have $k := \dim_{\mathbb{F}_q}(\mathcal{C}) = (q-1) - (\delta - 1) = q - \delta$. Hence if $k \geq 1$, that is $\delta < q$, then from the Singleton and BCH bounds we get $\delta - 1 = (q-1) - k \geq d - 1 \geq \delta - 1$, where $d := d(\mathcal{C}) \in \mathbb{N}$ is the minimum distance of $\mathcal{C}$, showing that $d = \delta$, implying that $\mathcal{C}$ is an MDS $[q-1, q-\delta, \delta]$-code.

We describe a fast encoding procedure for $\mathcal{C}$: We observe first that $\mathcal{V}^\perp := \mathcal{V}(\mathcal{C}^\perp) = \mathbb{F}_q^* \setminus \mathcal{V}^{-1} = \{\zeta_{q-1}^i \in \mathbb{F}_q^*; i \in \mathbb{Z}_{q-1} \setminus (-a - \{0, \ldots, \delta-2\})\} = \{\zeta_{q-1}^i \in \mathbb{F}_q^*; i \in (-a + \{1, \ldots, q-\delta\}) \subseteq \mathbb{Z}_{q-1}\}$; in particular $\mathcal{C}^\perp$ is a Reed-Solomon code again. Hence the (conventional) check matrix $H(\mathcal{V}^\perp) = [\zeta_{q-1}^{(i-a)(j-1)}]_{ij} \in \mathbb{F}_q^{(q-\delta)\times(q-1)}$ of $\mathcal{C}^\perp \leq \mathbb{F}_q^{q-1}$ is a generator matrix of $\mathcal{C}$:

$$H(\mathcal{V}^\perp) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \zeta_{q-1} & \zeta_{q-1}^2 & \cdots & \zeta_{q-1}^{q-2} \\ 1 & \zeta_{q-1}^2 & \zeta_{q-1}^4 & \cdots & \zeta_{q-1}^{2(q-2)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \zeta_{q-1}^{q-\delta-1} & \zeta_{q-1}^{2(q-\delta-1)} & \cdots & \zeta_{q-1}^{(q-2)(q-\delta-1)} \end{bmatrix} \cdot \operatorname{diag}([\zeta_{q-1}^{(q-a)(j-1)}]_j)$$

Thus for $v := [a_0, \ldots, a_{q-\delta-1}] \in \mathbb{F}_q^{q-\delta}$ the associated codeword $v \cdot H(\mathcal{V}^\perp) \in \mathbb{F}_q^{q-1}$ is given as follows: Letting $w_j := [\zeta_{q-1}^{(i-a)(j-1)}]_i \in \mathbb{F}_q^{q-\delta}$ be the transpose of the j-th column of $H(\mathcal{V}^\perp)$, for $j \in \{1, \ldots, q-1\}$, we get $\sum_{i=0}^{q-\delta-1} a_i w_{j,i+1} = \sum_{i=0}^{q-\delta-1} a_i (\zeta_{q-1}^{j-1})^{i-(a-1)} = \sum_{i=0}^{q-\delta-1} a_i (\zeta_{q-1}^{j-1})^{i+(q-a)} = (X^{q-a}\nu(v))(\zeta_{q-1}^{j-1})$. Hence we have $v \cdot H(\mathcal{V}^\perp) = [(X^{q-a}\nu(v))(\zeta_{q-1}^{j-1})]_j \in \mathbb{F}_q^{q-1}$, where the latter is obtained by evaluating the polynomial $X^{q-a}\nu(v) \in \mathbb{F}_q[X]$ at all places in $\mathbb{F}_q^*$.

In particular, for narrow sense Reed-Solomon codes, that is $a = 1$, we obtain $\mathcal{C} = \{[\nu(v)(1), \nu(v)(\zeta_{q-1}), \ldots, \nu(v)(\zeta_{q-1}^{q-2})] \in \mathbb{F}_q^{q-1}; v \in \mathbb{F}_q^{q-\delta}\} \leq \mathbb{F}_q^{q-1}$, being obtained by evaluating all polynomials in $\mathbb{F}_q[X]_{<q-\delta}$ at all places in $\mathbb{F}_q^*$.

b) Generalizing this idea leads to **generalized Reed-Solomon codes**, which in turn belong to the class of **algebro-geometric codes**, as follows:

Letting n be arbitrary and $k \in \{1, \ldots, n\}$, we choose pairwise distinct elements $\alpha := [\alpha_1, \ldots, \alpha_n] \subseteq \mathbb{F}_q$, and a vector $v := [v_1, \ldots, v_n] \in \mathbb{F}_q^n$ such that $\mathrm{wt}(v) = n$, and then let $\mathrm{GRS}_k(\alpha, v) := \{[v_1 f(\alpha_1), \ldots, v_n f(\alpha_n)] \in \mathbb{F}_q^n; f \in \mathbb{F}_q[X]_{<k}\} \leq \mathbb{F}_q^n$. In particular, narrow sense Reed-Solomon codes are given by $n := q - 1$ and $k := q - \delta$ and $\alpha = [1, \zeta_{q-1}, \ldots, \zeta_{q-1}^{q-2}]$ and $v := 1_{q-1}$.

Hence for $\mathcal{C} := \mathrm{GRS}_k(\alpha, v)$ we have $\dim_{\mathbb{F}_q}(\mathcal{C}) \leq k$. Moreover, since any $0 \neq f \in \mathbb{F}_q[X]_{<k}$ has at most $k-1$ zeroes in $\mathbb{F}_q$, we infer that $d(\mathcal{C}) \geq n-k+1$. Thus by the Singleton bound we have $n-k \geq n-\dim_{\mathbb{F}_q}(\mathcal{C}) \geq d(\mathcal{C})-1 \geq n-k$, implying equality throughout, so that $\mathcal{C}$ is an MDS $[n, k, n-k+1]$-code.

c) Finally, in order to prepare the application below, we observe the following: Let $n \in \mathbb{N}$ be arbitrary, and let $\mathcal{C} \leq \mathbb{F}_q^n$ be an MDS $[n, k, d]$-code, where $k \geq 2$. Then for the associated shortened $[n-1, k^\circ, d^\circ]$-code $\mathcal{C}^\circ \leq \mathbb{F}_q^{n-1}$ we have $k-1 \leq k^\circ \leq k$ and $d \leq d^\circ$. The Singleton bound for $\mathcal{C}^\circ$ yields $n-k = (n-1)-(k-1) \geq (n-1)-k^\circ \geq d^\circ - 1 \geq d-1$. Hence the Singleton bound $n-k = d-1$ being sharp for $\mathcal{C}$ yields equality throughout, thus $k^\circ = k-1$ and $d^\circ = d$, so that $\mathcal{C}^\circ$ is an MDS $[n-1, k-1, d]$-code as well.

Thus, starting with a Reed-Solomon $[q-1, q-\delta, \delta]$-code, successive shortening yields MDS $[q-i, (q-i)-(\delta-1), \delta]$-codes, for $i \in \{1, \ldots, q-\delta\}$. For example, for $q := 2^8 = 256$ and $n = 255$ and designed distance $\delta = 5$, starting with the narrow sense Reed-Solomon $[255, 251, 5]$-code, thus having generator polynomial $\prod_{i=1}^4 (X - \zeta_{255}^i) \in \mathbb{F}_{256}[X]$, we get the 2-error correcting $[32, 28, 5]$- and $[28, 24, 5]$-codes over $\mathbb{F}_{256}$ being used in the following application:

(9.7) Example: The Audio Compact Disc [1982]. The **Red Book Standard**, nowadays called **DIN EN 60908**, for the **compact disc digital audio (CD-DA) system** has been developed by the companies ‘Sony’ and ‘Philips’.

The amplitude of the analog audio data is sampled at a frequency of 44.1 kHz. By the **Nyquist-Shannon Theorem** frequencies up to half of the sampling frequency can be encoded and decoded, thus here up to ∼ 22 kHz. To prevent producing **moire** artifacts, the analog signal has to run through a **low pass (anti-aliasing) filter** before digitalization.

The analog signal is encoded using 16-bit **pulse code modulation (PCM)**. Hence using $2^8 = 256$ symbols instead of only the symbols 0 and 1, that is Bytes instead of bits, a stereo audio signal sample needs 4 Byte. Thus digitalization produces $4 \cdot 44100 \ \frac{\text{Byte}}{\text{s}} = 176400 \ \frac{\text{Byte}}{\text{s}} = 1411200 \ \frac{\text{bit}}{\text{s}}$. Given the running time of 74min, this yields a total of $74 \cdot 60 \cdot 176400$ Byte $= 783216000$ Byte ~ 783 MB.

Now 6 samples form a word of 24 Byte = 192 bit, being called a **frame**. These are encoded using a **cross-interleaved Reed-Solomon code (CIRC)**, which essentially works as follows: First, using an **outer** $[28, 24, 5]$-code $\mathcal{C}_2$, which is shortened from the narrow sense Reed-Solomon $[255, 251, 5]$-code over $\mathbb{F}_{256}$, words of length 24 are encoded into words of length 28. Then an **interleaver** with **offset** four is applied: Codewords $[x_{i1}, \ldots, x_{in}] \in \mathbb{F}_q^n$, for $i \in \mathbb{Z}$, are written diagonally into a matrix and are read out column-wise, as the following scheme with offset one shows:

$$\begin{bmatrix} \cdots & x_{i1} & x_{i+1,1} & \cdots & & & \\ & \cdots & x_{i2} & x_{i+1,2} & \cdots & & \\ & & & \ddots & \ddots & & \\ & & & \cdots & x_{in} & x_{i+1,n} & \cdots \end{bmatrix}$$

Next, an **inner** $[32, 28, 5]$-code $\mathcal{C}_1$, again shortened from the narrow sense Reed-Solomon $[255, 251, 5]$-code over $\mathbb{F}_{256}$, encodes words of length 28 into words of length 32. Finally, a further Byte is added containing **subchannel information**, yielding words of total length 33.

The idea of this encoding scheme is as follows: The code $\mathcal{C}_1$ has minimum distance 5, hence is 2-error correcting, where single **C1** errors are corrected, while words with two errors (typically) are marked as erasures. The resulting words of length 28 are de-interleaved, leading to a distribution of erasures, called **C2** errors. The code $\mathcal{C}_2$ has minimum distance 5 as well, thus is able to correct four erased positions in any word. Hence, given $g \in \mathbb{N}$ consecutive erasures, that is columns of the above scheme, due to offset four any diagonally written word is affected in at most $\lceil \frac{g}{4} \rceil$ known positions. Thus burst errors, which for example result from surface scratches, with a loss of up to 16 words can be corrected this way. Still remaining **CU errors** are treated by **interpolation**, and finally **oversampling** is applied against aliasing.

The data is stored as a spiral track of **pits** moulded into a polycarbonate layer. The pits are 100nm deep, 500nm wide, and at least 850nm long; the regions between pits are called **lands**. The data is read by a 780nm solid state laser, where a pit-land or a land-pit change is read as a 1, and 0 otherwise.

This technique requires that between two read 1's there must be at least two and at most ten read 0's. This is achieved by **eight-to-fourteen modulation (EFM)**, where each Byte, that is each 8-bit word, is replaced by a 14-bit word, using table lookup. Then a suitable 3-bit **merging** word is added between two 14-bit words. Finally, a 3 Byte **synchronization** word is added, together with another 3-bit merging word. The synchronization word does not occur elsewhere in the bit stream, hence can be used to detect the beginning of a frame.

Hence a frame consists of $\big(33 \cdot (14+3) + (24+3)\big)$ bit $=$ 588 bit, which amounts to an information rate of $\frac{192}{588} = \frac{16}{49} \sim 0.33$, hence a bit rate of $\frac{588}{192} \cdot 1411200 \, \frac{\text{bit}}{\text{s}} = 4321800 \, \frac{\text{bit}}{\text{s}} = 540225 \, \frac{\text{Byte}}{\text{s}}$, and a total of $74 \cdot 60 \cdot 540225$ Byte $=$ 2398599000 Byte $\sim$ 2.4 GB. Moreover, a burst error of 16 words of length

32 Byte is contained in $16 \cdot 588$ bit $= 9408$ bit, since a bit needs some 300nm of track length, this amounts to some $9408 \cdot 300\text{nm} = 2822400\text{nm} \sim 2.8\text{mm}$.

10 Minimum distance of BCH codes

We have already seen that the BCH bound for the minimum distance of a BCH code is not necessarily sharp. We now proceed into two opposite directions: Firstly, we improve on the idea behind the BCH bound in order to obtain better bounds. Secondly, in the narrow sense primitive case, we provide sufficient criteria ensuring that the BCH bound is actually sharp.

(10.1) Van-Lint-Wilson bound [1986]. We need a definition first: Letting $A = [a_{ij}]_{ij} \in \mathbb{F}_q^{r \times n}$ and $B = [b_{i'j}]_{i'j} \in \mathbb{F}_q^{s \times n}$, where $n \in \mathbb{N}$ and $r, s \in \mathbb{N}_0$, let $A * B := [a_{ij} b_{i'j}]_{(i-1)s+i',j} \in \mathbb{F}_q^{rs \times n}$, where $i \in \{1, \ldots, r\}$ and $i' \in \{1, \ldots, s\}$. Note that $A * B$ in general does not have full rank, even if A and B have.

Theorem. Let $v \in \mathcal{C} := \ker((A * B)^{\text{tr}}) \leq \mathbb{F}_q^n$, and let $A_{\mathcal{J}} \in \mathbb{F}_q^{r \times |\mathcal{J}|}$ and $B_{\mathcal{J}} \in \mathbb{F}_q^{s \times |\mathcal{J}|}$ be the submatrices of A and B, respectively, consisting of the columns in $\mathcal{J} := \text{supp}(v) \subseteq \{1, \ldots, n\}$. Then we have $\text{rk}_{\mathbb{F}_q}(A_{\mathcal{J}}) + \text{rk}_{\mathbb{F}_q}(B_{\mathcal{J}}) \leq |\mathcal{J}| = \text{wt}(v)$.

Proof. Let $v = [x_1, \ldots, x_n]$, where we may assume that $\mathcal{J} = \{1, \ldots, n\}$, that is $x_j \neq 0$ for all $j \in \{1, \ldots, n\}$. Letting $B' := B \cdot \text{diag}[x_1, \ldots, x_n] = [b_{i'j} x_j]_{i'j} \in \mathbb{F}_q^{s \times n}$, we have $\text{rk}_{\mathbb{F}_q}(B) = \text{rk}_{\mathbb{F}_q}(B') \in \mathbb{N}_0$. Moreover, the condition $v \cdot (A * B)^{\text{tr}} = [\sum_{j=1}^n a_{ij} b_{i'j} x_j]_{(i-1)s+i'} = 0 \in \mathbb{F}_q^{rs}$ can be rewritten as $A \cdot B'^{\text{tr}} = 0 \in \mathbb{F}_q^{r \times s}$. In other words, the row space of A is contained in the orthogonal space of the row space of B', hence we have $\text{rk}_{\mathbb{F}_q}(A) \leq n - \text{rk}_{\mathbb{F}_q}(B') = n - \text{rk}_{\mathbb{F}_q}(B)$. ♯

Corollary. If for all $\emptyset \neq \mathcal{I} \subseteq \{1, \ldots, n\}$ such that $|\mathcal{I}| \leq d - 1$, for some $d \in \mathbb{N}$, we have $\text{rk}_{\mathbb{F}_q}(A_{\mathcal{I}}) + \text{rk}_{\mathbb{F}_q}(B_{\mathcal{I}}) > |\mathcal{I}|$, then $\mathcal{C}$ has minimum distance at least d.

(10.2) Theorem: Roos bound [1983]. For $n \in \mathbb{N}$ let $\mathcal{V}' \subseteq \mathcal{V}_n \subseteq \mathbb{F}_q(\zeta_n)$, where $\gcd(q, n) = 1$, be a consecutive set of length $\delta - 1$, where $\delta \in \{2, \ldots, n+1\}$. Moreover, let $\emptyset \neq \mathcal{V}'' \subseteq \mathcal{V}_n$ be any subset such that there is a consecutive subset of $\mathcal{V}_n$ containing $\mathcal{V}''$ and having length $|\mathcal{V}''| + \delta - 2$. Then the cyclic code $\mathcal{C} \leq \mathbb{F}_q^n$ associated with $\mathcal{V} := \mathcal{V}' \cdot \mathcal{V}'' \subseteq \mathcal{V}_n$ has minimum distance at least $\delta - 1 + |\mathcal{V}''|$; note that we recover the BCH bound from $\mathcal{V}'' = \{1\}$.

Proof. Since the matrices $H(\mathcal{V}') * H(\mathcal{V}'') \in \mathbb{F}_q(\zeta_n)^{(|\mathcal{V}'| \cdot |\mathcal{V}''|) \times n}$ and $H(\mathcal{V}) \in \mathbb{F}_q(\zeta_n)^{|\mathcal{V}| \times n}$ have the same set of rows, we have $\mathcal{C} = \ker(H(\mathcal{V})^{\text{tr}}) \cap \mathbb{F}_q^n = \ker((H(\mathcal{V}') * H(\mathcal{V}''))^{\text{tr}}) \cap \mathbb{F}_q^n$. We aim at applying the van-Lint-Wilson bound:

For $\emptyset \neq \mathcal{I} \subseteq \{1, \ldots, n\}$, by the BCH bound we have $\text{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}')_{\mathcal{I}}) = |\mathcal{I}|$ if $|\mathcal{I}| \leq \delta - 1$, and thus $\text{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}')_{\mathcal{I}}) \geq \delta - 1$ if $|\mathcal{I}| \geq \delta$. Since we always have $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}'')_{\mathcal{I}}) \geq 1$, we get $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}')_{\mathcal{I}}) + \mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}'')_{\mathcal{I}}) > |\mathcal{I}|$ if $|\mathcal{I}| \leq \delta - 1$. This settles the case $|\mathcal{V}''| = 1$. Hence let now $|\mathcal{V}''| \geq 2$ and $|\mathcal{I}| \geq \delta$:

Let $\mathcal{V}'' \subseteq \mathcal{W} \subseteq \mathcal{V}_n$ be a consecutive set, again by the BCH bound yielding $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{W})_{\mathcal{I}}) \geq |\mathcal{I}|$ if $|\mathcal{I}| \leq |\mathcal{W}|$. Since deleting the rows of $H(\mathcal{W})$ corresponding to $\mathcal{W} \setminus \mathcal{V}''$ yields $H(\mathcal{V}'')$, we infer $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}'')_{\mathcal{I}}) \geq |\mathcal{I}| - |\mathcal{W}| + |\mathcal{V}''|$ if $|\mathcal{I}| \leq |\mathcal{W}|$. This yields $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}')_{\mathcal{I}}) + \mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{V}'')_{\mathcal{I}}) \geq \delta - 1 + |\mathcal{I}| - |\mathcal{W}| + |\mathcal{V}''|$ if $\delta \leq |\mathcal{I}| \leq |\mathcal{W}|$. Here, the right hand side exceeds $|\mathcal{I}|$ if and only if $\delta \leq |\mathcal{I}| \leq |\mathcal{W}| \leq |\mathcal{V}''| + \delta - 2$, in which case $\mathcal{C}$ has minimum distance at least $|\mathcal{I}| + 1$. Now the assertion follows from choosing $|\mathcal{I}| = |\mathcal{W}| = |\mathcal{V}''| + \delta - 2$. ♯

Corollary: Hartmann, Tzeng [1972]. Let $\mathcal{V}' \subseteq \mathcal{V}_n$ be a consecutive set of length $\delta - 1$, where $\delta \in \{2, \ldots, n+1\}$, and let $\emptyset \neq \mathcal{V}'' \subseteq \mathcal{V}_n$ be a generalized consecutive set of some step size in $\mathbb{Z}_n^*$ such that $|\mathcal{V}''| \leq n+2-\delta$. Then the cyclic code associated with $\mathcal{V} := \mathcal{V}' \cdot \mathcal{V}''$ has minimum distance at least $\delta - 1 + |\mathcal{V}''|$.

Proof. The set $\mathcal{V}''$ can be extended to a consecutive set with the same stpe size and having length $|\mathcal{V}''| + \delta - 2$. Recall that the rank estimates for Delsarte matrices also hold for generalized consecutive sets. ♯

Example. Let $q := 2$ and $n := 35$. Then $2 \in \mathbb{Z}_{35}^*$ has order 12, and the cyclotomic cosets are given as

$$\begin{aligned} \{0\} \; & \dot\cup \{5,10,20\} \dot\cup \{15,25,30\} \dot\cup \{7,14,21,28\} \\ & \dot\cup \{1,2,4,8,9,11,16,18,22,23,29,32\} \\ & \dot\cup \{3,6,12,13,17,19,24,26,27,31,33,34\}. \end{aligned}$$

Let $\mathcal{C} \leq \mathbb{F}_2^{35}$ be the cyclic code associated with $g := \mu_1\mu_5\mu_7 \in \mathbb{F}_2[X]$. Hence $\mathcal{V}(\mathcal{C})$ is given by $\{1,2,4,5,7,8,9,10,11,14,16,18,20,21,22,23,28,29,32\} \subseteq \mathbb{Z}_{35}$, thus for the minimum distance of $\mathcal{C}$ the BCH bound yields $d(\mathcal{C}) \geq 6$.

But $\mathcal{C}$ is also associated with $\mathcal{O} := \{7,8,9,10,11\} \dot\cup \{20,21,22,23\}$, which letting $\mathcal{O}' := \{7,8,9,10\}$ and $\mathcal{O}'' := \{0,1,13\}$ can be written as $\mathcal{O} = \mathcal{O}' + \mathcal{O}'' \subseteq \mathbb{Z}_{35}$. In order to apply the Roos bound with $\delta = 5$, entailing $d(\mathcal{C}) \geq \delta - 1 + |\mathcal{O}''| = 7$, we have to embed $\mathcal{O}''$ in a consecutive set of length $|\mathcal{O}''| + \delta - 2 = 6$: Since $3 \in \mathbb{Z}_{35}^*$ we conclude that $\zeta_{35}^3 \in \mathcal{V}_{35}$ is a primitive 35-th root of unity as well; recall that the rank estimates of check matrices associated with consecutive sets do not depend on a particular choice of a primitive root of unity. Thus we indeed get $3 \cdot \mathcal{O}'' = \{0,3,4\} \subseteq \{0,\ldots,5\} \subseteq \mathbb{Z}_{35}$.

To show conversely that $d(\mathcal{C}) \leq 7$, we choose $\mu_1 := X^{12} + X^{11} + X^{10} + X^8 + X^5 + X^4 + X^3 + X^2 + 1 \in \mathbb{F}_2[X]$, which entails $\mu_3 = X^{12} + X^{10} + X^9 + X^8 + X^7 + X^4 + X^2 + X + 1$, as well as $\mu_5 = X^3 + X + 1$ and $\mu_{15} = X^3 + X^2 + 1$, while $\mu_7 = X^4 + X^3 + X^2 + X + 1$ anyway. This yields $g = \mu_1\mu_5\mu_7 = X^{19} + X^{15} + X^{14} + X^{13} + X^{12} + X^{10} + X^9 + X^7 + X^6 + X^2 + 1$. It turns out that $g \mid f := X^{28} + X^{16} + X^{14} + X^{11} + X^7 + X + 1$, or equivalently $f(\zeta_{35}) = f(\zeta_{35}^5) = f(\zeta_{35}^7) = 0$. Hence $f \in \mathbb{F}_2[X]$ corresponds to a codeword of weight 7. ♯

Example. Let $q := 2$ and $n := 127 = 2^7 - 1$. Then $2 \in \mathbb{Z}_{127}^*$ has order 7, and the cyclotomic cosets are given as

$$\begin{aligned}
\{0\} \;&\dot\cup\; \{1,2,4,8,16,32,64\} \,\dot\cup\, \{3,6,12,24,48,65,96\} \\
&\dot\cup\; \{5,10,20,33,40,66,80\} \,\dot\cup\, \{7,14,28,56,67,97,112\} \\
&\dot\cup\; \{9,17,18,34,36,68,72\} \,\dot\cup\, \{11,22,44,49,69,88,98\} \\
&\dot\cup\; \{13,26,35,52,70,81,104\} \,\dot\cup\, \{15,30,60,71,99,113,120\} \\
&\dot\cup\; \{19,25,38,50,73,76,100\} \,\dot\cup\, \{21,37,41,42,74,82,84\} \\
&\dot\cup\; \{23,46,57,75,92,101,114\} \,\dot\cup\, \{27,51,54,77,89,102,108\} \\
&\dot\cup\; \{29,39,58,78,83,105,116\} \,\dot\cup\, \{31,62,79,103,115,121,124\} \\
&\dot\cup\; \{43,45,53,85,86,90,106\} \,\dot\cup\, \{47,61,87,94,107,117,122\} \\
&\dot\cup\; \{55,59,91,93,109,110,118\} \,\dot\cup\, \{63,95,111,119,123,125,126\}.
\end{aligned}$$

Let $\mathcal{C}^\perp \leq \mathbb{F}_2^{127}$ be the narrow sense primitive BCH code with designed distance 11. Hence $\mathcal{C}^\perp$ is associated both with $\{1,3,5,7,9\} \subseteq \{1,\ldots,10\}$, thus $\mathcal{V}(\mathcal{C}^\perp)$ has cardinality 35 and is given by

$$\begin{aligned}
\mathcal{O}^\perp := \{ \quad & 1,2,3,4,5,6,7,8,9,10,12,14,16,17,18,20,24,28,32, \\
& 33,34,36,40,48,56,64,65,66,67,68,72,80,96,97,112 \quad \} \subseteq \mathbb{Z}_{127}.
\end{aligned}$$

Let $\mathcal{C} := (\mathcal{C}^\perp)^\perp \leq \mathbb{F}_2^{127}$. Hence $\mathcal{V}(\mathcal{C})$ is given by $\mathcal{O} := \mathbb{Z}_{127} \setminus (-\mathcal{O}^\perp)$, that is

$$\begin{aligned}
\mathcal{O} = \{ \quad & 0,1,2,3,4,5,6,7,8,9,10,11,12,13,14; \\
& 16,17,18,19,20,21,22,23,24,25,26,27,28,29; \\
& 32,33,34,35,36,37,38,39,40,41,42,43,44,45,46; \\
& 48,49,50,51,52,53,54;\ 56,57,58;\ 64,65,66,67,68,69,70; \\
& 72,73,74,75,76,77,78;\ 80,81,82,83,84,85,86;\ 88,89,90; \\
& 92;\ 96,97,98;\ 100,101,102;\ 104,105,106;\ 108;\ 112;\ 114;\ 116 \quad \}.
\end{aligned}$$

Thus $\mathcal{C}$ has minimal defining set $\mathcal{Q} := \{0,1,3,5,7,9,11,13,19,21,23,27,29,43\}$, where $\mathcal{V}(\mathcal{C})$ has cardinality 92. This implies $\mathcal{O}^\perp \subseteq \mathcal{O}$, thus $\mathcal{C} \leq \mathcal{C}^\perp$ is weakly self-dual. We proceed to determine the minimum distance d of $\mathcal{C}$:

i) The BCH bound yields $d \geq 16$ and the Singleton bound yields $d \leq 127 - 92 + 1 = 36$. Moreover, we show that d is divisible by 4:

To this end, we consider the localization $\mathbb{F}_2[X^{\pm 1}] := \mathbb{F}_2[X]_X \subseteq \mathbb{F}_2(X)$, and let $\bar{\ }\colon \mathbb{F}_2[X^{\pm 1}] \rightarrow \overline{\mathbb{F}_2[X]} := \mathbb{F}_2[X]/\langle X^{127} - 1\rangle \cong \bigoplus_{i \in \mathbb{Z}_{127}} \mathbb{F}_2[X]/\langle X - \zeta_{127}^i\rangle$ be the extension of the natural epimorphism, where the latter isomorphism results from the Chinese Remainder Theorem. Now, from $\mathcal{O}^\perp \subseteq \mathcal{O}$ we get $\mathcal{O} \cup (-\mathcal{O}) = \mathcal{O} \cup (\mathbb{Z}_{127} \setminus \mathcal{O}^\perp) = \mathbb{Z}_{127}$, thus for any $i \in \mathbb{Z}_{127}$ we have $g(\zeta_{127}^i)g(\zeta_{127}^{-i}) = 0$, entailing that $\overline{g(X)g(X^{-1})} = 0 \in \overline{\mathbb{F}_2[X]}$.

It suffices to show that for any $v = [a_0, \ldots, a_{126}] \in \mathcal{C}$ we have $4 \mid \mathrm{wt}(v) =: s$: Let $f := \nu(v) = \sum_{i=0}^{126} a_i X^i \in \mathbb{F}_2[X]$, then $f(X^{-1}) = \sum_{j=0}^{126} a_j X^{-j} \in \mathbb{F}_2[X^{\pm 1}]$. From $g \mid f \in \mathbb{F}_2[X]$ we infer that $\overline{f(X)f(X^{-1})} = \sum_{k=0}^{126} (\sum_{j \in \mathbb{Z}_{127}} a_{k+j} a_j) \overline{X}^k = 0 \in \overline{\mathbb{F}_2[X]}$, Thus for all $k \in \mathbb{Z}_{127}$ we have $\sum_{j \in \mathbb{Z}_{127}} a_{k+j} a_j = 0 \in \mathbb{F}_2$. In other words, letting $\mathcal{J} := \mathrm{supp}(v) \subseteq \mathbb{Z}_{127}$, this says that $|\{[i,j] \in \mathcal{J}^2; i = j + k\}|$ is even. For $k = 0$ this yields that $|\mathcal{J}| = s$ is even. Since 127 is odd, for $k \neq 0$ we have $-k \neq k \in \mathbb{Z}_{127}$, hence $4 \mid |\{[i,j] \in \mathcal{J}^2; i \in \{j \pm k\}\}|$, and thus $4 \mid |\{[i,j] \in \mathcal{J}^2; i \neq j\}| = s^2 - s = s(s-1)$, implying that $4 \mid s$.

ii) We aim at applying the Roos bound with $\delta = 15$: Letting $\mathcal{O}' := \{0, \ldots, 13\}$ and $\mathcal{O}'' := \{0, 1, 16, 32, 33\}$ we get $\mathcal{O}' + \mathcal{O}'' = \{0, \ldots, 14\} \,\dot\cup\, \{16, \ldots, 29\} \,\dot\cup\, \{32, \ldots, 46\} \subseteq \mathcal{O}$, hence $\mathcal{C}$ is associated with $\mathcal{O}' + \mathcal{O}''$. Moreover, since $8 \in \mathbb{Z}_{127}^*$ we conclude that $\zeta_{127}^8 \in \mathcal{V}_{127}$ is a primitive 127-th root of unity as well. Thus from $8 \cdot \mathcal{O}'' = \{0, 1, 2, 8, 10\} \subseteq \{0, \ldots, 17\} \subseteq \mathbb{Z}_{127}$, where $18 = |\mathcal{O}''| + \delta - 2$, the Roos bound yields $d \geq \delta - 1 + |\mathcal{O}''| = 19$. Hence we get $d \geq 20$.

iii) We aim at applying the van-Lint-Wilson bound: Let $\mathcal{P}' := \{16, \ldots, 29\} \,\dot\cup\, \{32, \ldots, 44\} \subseteq \mathcal{O}$ and $\mathcal{P}'' := \{0, -16, -15\}$. Then we get $\mathcal{P}' + \mathcal{P}'' = \{0, \ldots, 14\} \,\dot\cup\, \{16, \ldots, 29\} \,\dot\cup\, \{32, \ldots, 44\} \subseteq \mathcal{O}' + \mathcal{O}'' \subseteq \mathcal{O}$, hence $\mathcal{C}$ is associated with $\mathcal{P}' + \mathcal{P}''$.

We consider the check matrix $H(\mathcal{P}') \in \mathbb{F}_2^{27 \times 127}$: Since the set $\{16, \ldots, 29\}$ and $\{16, \ldots, 44\}$ are consecutive of length 14 and 29, respectively, for $\emptyset \neq \mathcal{I} \subseteq \{1, \ldots, 127\}$ we get $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}')_{\mathcal{I}}) \geq |\mathcal{I}|$ if $|\mathcal{I}| \leq 14$, and $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}')_{\mathcal{I}}) \geq 14$ if $|\mathcal{I}| = 15$, and $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}')_{\mathcal{I}}) \geq |\mathcal{I}| - 2$ if $16 \leq |\mathcal{I}| \leq 29$.

We consider the check matrix $H(\mathcal{P}'') \in \mathbb{F}_2^{3 \times 127}$: From $8 \cdot \mathcal{P}'' = \{-1, 0, 7\} \subseteq \{-1, \ldots, 7\}$, we get $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}'')_{\mathcal{I}}) \geq |\mathcal{I}|$ if $|\mathcal{I}| \leq 2$, and $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}'')_{\mathcal{I}}) \geq 2$ if $3 \leq |\mathcal{I}| \leq 7$, and $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}'')_{\mathcal{I}}) \geq |\mathcal{I}| - 6$ if $8 \leq |\mathcal{I}| \leq 9$, and $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}'')_{\mathcal{I}}) = 3$ if $|\mathcal{I}| \geq 10$; note that $|\{-1, \ldots, 7\} \setminus \{-1, 0, 7\}| = 6$.

Thus we have $\mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}')_{\mathcal{I}}) + \mathrm{rk}_{\mathbb{F}_q(\zeta_n)}(H(\mathcal{P}'')_{\mathcal{I}}) > |\mathcal{I}|$ whenever $|\mathcal{I}| \leq 29$, hence the van-Lint-Wilson bound yields $d \geq 30$. Hence we get $d \geq 32$.

iv) Finally, to show conversely that $d \leq 32$, we choose $\mu_1 := X^7 + X + 1 \in \mathbb{F}_2[X]$, and thus fix all the polynomials $\mu_i \in \mathbb{F}_2[X]$ for $i \in \mathbb{Z}_{127}$. Then the generator polynomial $g := \prod_{i \in \mathcal{Q}} \mu_i \in \mathbb{F}_2[X]$ of $\mathcal{C}$ turns out to be

$$\begin{array}{rcl} g & = & X^{92} + X^{91} + X^{89} + X^{86} + X^{85} + X^{84} + X^{83} + X^{80} + \\ & & X^{77} + X^{76} + X^{74} + X^{73} + X^{71} + X^{67} + X^{65} + X^{62} + \\ & & X^{60} + X^{58} + X^{56} + X^{53} + X^{52} + X^{51} + X^{49} + X^{48} + \\ & & X^{47} + X^{46} + X^{45} + X^{43} + X^{39} + X^{38} + X^{36} + X^{35} + \\ & & X^{34} + X^{32} + X^{28} + X^{24} + X^{23} + X^{19} + X^{18} + X^{17} + \\ & & X^{16} + X^{11} + X^{10} + X^{6} + X^{4} + X^{3} + X + 1. \end{array}$$

Letting $f \in \mathbb{F}_2[X]$ to be given as

$$\begin{array}{rcl} f & := & X^{99} + X^{98} + X^{97} + X^{96} + X^{94} + X^{87} + X^{83} + X^{79} + \\ & & X^{78} + X^{75} + X^{72} + X^{69} + X^{62} + X^{61} + X^{59} + X^{57} + \\ & & X^{56} + X^{54} + X^{51} + X^{49} + X^{48} + X^{45} + X^{42} + X^{28} + \\ & & X^{26} + X^{25} + X^{22} + X^{17} + X^{13} + X^{10} + X^{3} + 1, \end{array}$$

it turns out that $g \mid f$, or equivalently $f(\zeta_{127}^i) = 0$ for all $i \in \mathcal{Q}$. Hence f corresponds to a codeword of weight 32. ♯

(10.3) Theorem. Let $n := q^s - 1$, for some $s \in \mathbb{N}$, and $\mathcal{C} \leq \mathbb{F}_q^n$ be a narrow sense primitive BCH code of designed distance $\delta = q^t - 1$, for some $t \in \{1, \ldots, s\}$. Then $\mathcal{C}$ has minimum distance δ.

Proof. We show that $\mathcal{C}$ contains an element of weight δ:

i) Since $q \in \mathbb{Z}_n^*$ has order s, we have $|\Gamma| = s$, from which we infer that $\zeta_n^\Gamma = \{\zeta_n, \zeta_n^q, \ldots, \zeta_n^{q^{s-1}}\} \subseteq \mathcal{V}_n$. Hence the elements $\{\zeta_n, \zeta_n^q, \ldots, \zeta_n^{q^{t-1}}\}$ are pairwise distinct, implying the invertibility of the associated Vandermonde matrix

$$A := [\zeta_n^{(j-1)q^{i-1}}]_{ij} = \begin{bmatrix} 1 & \zeta_n & \cdots & \zeta_n^{t-1} \\ 1 & \zeta_n^q & \cdots & \zeta_n^{(t-1)q} \\ \vdots & \vdots & & \vdots \\ 1 & \zeta_n^{q^{t-1}} & \cdots & \zeta_n^{(t-1)q^{t-1}} \end{bmatrix} \in \mathbb{F}_q(\zeta_n)^{t \times t}.$$

Hence the system of linear equations $[X_0, \ldots, X_{t-1}] \cdot A = -[1, \zeta_n^{q^t}, \ldots, \zeta_n^{(t-1)q^t}]$ has a unique solution $[a_0, \ldots, a_{t-1}] \in \mathbb{F}_q(\zeta_n)^t$. Let $f := X^{q^t} + \sum_{i=0}^{t-1} a_i X^{q^i} \in \mathbb{F}_q(\zeta_n)[X]$ be the associated **q-linearized polynomial** of degree q^t; note that we have $f(ax+y) = af(x) + f(y) \in \mathbb{F}_q(\zeta_n)$, for all $x, y \in \mathbb{F}_q(\zeta_n)$ and $a \in \mathbb{F}_q$.

By construction we have $(\zeta_n^j)^{q^t} + \sum_{i=0}^{t-1} a_i (\zeta_n^j)^{q^i} = 0$, for all $j \in \{0, \ldots, t-1\}$, hence $\mathcal{V} := \{1, \zeta_n, \ldots, \zeta_n^{t-1}\} \subseteq \mathcal{V}(f) \subseteq \overline{\mathbb{F}}$ consists of zeroes of f. Moreover, $\mathcal{V} \subseteq \mathbb{F}_q(\zeta_n)$ is $\mathbb{F}_q$-linearly independent: Let $[b_0, \ldots, b_{t-1}] \in \mathbb{F}_q^t$ such that $\sum_{j=0}^{t-1} b_j \zeta_n^j = 0 \in \mathbb{F}_q(\zeta_n)$, thus from $b_j^q = b_j$ we get $\sum_{j=0}^{t-1} b_j \zeta_n^{jq^i} = 0$, for all $i \in \{0, \ldots, t-1\}$, that is $A \cdot [b_0, \ldots, b_{t-1}]^{\mathrm{tr}} = 0 \in \mathbb{F}_q(\zeta_n)^{t \times 1}$, implying that $[b_0, \ldots, b_{t-1}] = 0$.

Thus letting $V := \langle \mathcal{V} \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q(\zeta_n)$ we have $|V| = q^t$, and from f being $\mathbb{F}_q$-linear we infer that $V \subseteq \mathcal{V}(f)$ consists of zeroes of f, implying that $V = \mathcal{V}(f)$, hence f splits over $\mathbb{F}_q(\zeta_n)$ as $f = \prod_{c \in V}(X - c) \in \mathbb{F}_q(\zeta_n)[X]$.

ii) We need an auxiliary construction: Let $\mathcal{X} := \{X_1, \ldots, X_m\}$ be indeterminates, where $m \in \mathbb{N}$. For $k \in \{0, \ldots, m\}$ let $e_{m,k} \in \mathbb{F}_q[\mathcal{X}]$ be the associated elementary symmetric polynomial of degree k, for example $e_{m,0} = 1$ and $e_{m,1} = \sum_{i=1}^m X_i$ and $e_{m,2} = \sum_{1 \leq iyj \leq m} X_i X_j$ and $e_{m,m} = X_1 \cdots X_m$, and for $k \in \mathbb{N}$ let $p_{m,k} := \sum_{i=1}^m X_i^k \in \mathbb{F}_q[\mathcal{X}]$ be the associated power sum polynomial.

Letting $h := \prod_{i=1}^m (1 - X_i X) = \sum_{j=0}^m (-1)^j e_{m,j} X^j \in \mathbb{F}_q[\mathcal{X}][X] \subseteq \mathbb{F}_q((\mathcal{X}, X))$, we have $\frac{\partial}{\partial X} h = \sum_{j=1}^m (-1)^j j e_{m,j} X^{j-1}$, and the product rule yields

$$\begin{aligned} \tfrac{\partial}{\partial X} h &= -\textstyle\sum_{j=1}^m \Big(X_j \cdot \prod_{i \in \{1, \ldots, m\} \setminus \{j\}} (1 - X_i X) \Big) \\ &= -h \cdot \textstyle\sum_{j=1}^m \frac{X_j}{1 - X_j X} \\ &= -h \cdot \textstyle\sum_{j=1}^m \Big(\sum_{k \geq 0} X_j^{k+1} X^k \Big) \\ &= -h \cdot \textstyle\sum_{k \geq 0} p_{m,k+1} X^k, \end{aligned}$$

implying $\sum_{i=1}^m (-1)^{i-1} i e_{m,i} X^{i-1} = \big(\sum_{j=0}^m (-1)^j e_{m,j} X^j \big) \cdot \big(\sum_{k \geq 1} p_{m,k} X^{k-1} \big) \in \mathbb{F}_q[\mathcal{X}][[X]]$. Thus we get the **Newton identities** $\sum_{j=1}^i (-1)^{j-1} e_{m,i-j} p_{m,j} =$

$ie_{m,i}$ for $i \in \{1,\ldots,m\}$, as well as $\sum_{j=i-m}^{i}(-1)^{j-1}e_{m,i-j}p_{m,j} = 0$ for $i \geq m+1$; the former can also be written as $(-1)^{i-1}p_{m,i} = ie_{m,i} + \sum_{j=1}^{i-1}(-1)^j e_{m,i-j}p_{m,j}$.

iii) Now let $m := q^t$. Evaluating $e_{m,i} \in \mathbb{F}_q[\mathcal{X}]$, for $i \in \{0,\ldots,m\}$, at the elements of $V \subseteq \mathbb{F}_q(\zeta_n)$, we get $f = X^{q^t} + \sum_{i=0}^{t-1} a_i X^{q^i} = \prod_{c \in V}(X-c) = \sum_{j=0}^{m}(-1)^{m-j}e_{m,m-j}(V)X^j \in \mathbb{F}_q(\zeta_n)[X]$. This implies that $e_{m,i}(V) \neq 0$ possibly only for $i = m - q^j$, where $j \in \{0,\ldots,t\}$. Since $q \mid m$, we thus have $ie_{m,i}(V) = 0$ for all $i \in \{0,\ldots,m-2\}$. Evaluating $p_{m,i} \in \mathbb{F}_q[\mathcal{X}]$ at the elements of V as well, from the Newton identities we by induction on $i \in \mathbb{N}$ get $\sum_{c \in V \setminus \{0\}} c^i = \sum_{c \in V} c^i = p_{m,i}(V) = 0$, for all $i \in \{1,\ldots,m-2\}$.

Since $|\mathbb{F}_q(\zeta_n)^*| = q^s - 1 = n$, the map $\mathbb{Z}_n \to \mathbb{F}_q(\zeta_n)^* \colon i \mapsto \zeta_n^i$ is a bijection. Let $v = [x_0,\ldots,x_{n-1}] \in \{0,1\}^n \subseteq \mathbb{F}_q^n$ be such that $V \setminus \{0\} = \{\zeta_n^i \in \mathbb{F}_q(\zeta_n); i \in \mathbb{Z}_n, x_i = 1\}$. Then we have $\mathrm{wt}(v) = |V| - 1 = m - 1 = \delta$, and we get $\nu(v)(\zeta_n^j) = \sum_{i \in \mathbb{Z}_n, x_i = 1}(\zeta_n^j)^i = \sum_{i \in \mathbb{Z}_n, x_i=1}(\zeta_n^i)^j = \sum_{c \in V \setminus \{0\}} c^j = 0$, for $j \in \{1,\ldots,\delta-1\}$, which since $\mathcal{C}$ is associated with $\{\zeta_n,\ldots,\zeta_n^{\delta-1}\} \subseteq \mathcal{V}_n$ implies that $v \in \mathcal{C}$. ♯

(10.4) Corollary. Let $\mathcal{C} \leq \mathbb{F}_q^n$ be a narrow sense primitive BCH code of designed distance $\delta \in \{1,\ldots,n\}$. Then $\mathcal{C}$ has minimum distance at most $q\delta - 1$.

Proof. Let $n = q^s - 1$, and let $t \in \{1,\ldots,s\}$ such that $q^{t-1} \leq \delta \leq q^t - 1$. Since $\mathcal{C}$ is associated with $\{\zeta_n,\ldots,\zeta_n^{\delta-1}\}$, it contains the code associated with $\{\zeta_n,\ldots,\zeta_n^{q^t-2}\}$. Since the latter has minimum distance $q^t - 1$, the minimum distance of $\mathcal{C}$ is bounded above by $q^t - 1 = q \cdot q^{t-1} - 1 \leq q\delta - 1$. ♯

Finally, we just mention the following theorem, whose proof requires further tools we do not have at our disposal here:

(10.5) Theorem. Let $\mathcal{C}$ be a non-trivial narrow sense primitive binary BCH code. Then the minimum distance of $\mathcal{C}$ is odd. ♯

11 Quadratic residue codes

(11.1) Quadratic residues. We collect a few number theoretic facts.

a) Let p be an odd prime. Then by Artin's Theorem $\mathbb{Z}_p^*$ is a cyclic group of even order $p-1$. Hence the set of **squares** $\mathcal{Q}_p := \{i^2 \in \mathbb{Z}_p^*; i \in \mathbb{Z}_p^*\} \leq \mathbb{Z}_p^*$ is the unique subgroup of $\mathbb{Z}_p^*$ of index 2, and consists of the elements of $\mathbb{Z}_p^*$ of order dividing $\frac{p-1}{2}$. Let $\mathcal{N}_p := \mathbb{Z}_p^* \setminus \mathcal{Q}_p$ be the set of **non-squares** in $\mathbb{Z}_p^*$; hence we have $|\mathcal{Q}_p| = |\mathcal{N}_p| = \frac{p-1}{2}$.

For $i \in \mathbb{Z}_p^*$ let the **Legendre symbol** be defined as $\left(\frac{i}{p}\right) := 1$ if $i \in \mathcal{Q}_p$, and $\left(\frac{i}{p}\right) := -1$ if $i \in \mathcal{N}_p$. Hence for $i, j \in \mathbb{Z}_p^*$ we have $\left(\frac{ij}{p}\right) = \left(\frac{i}{p}\right)\left(\frac{j}{p}\right)$, thus

$\left(\frac{\cdot}{p}\right)\colon \mathbb{Z}_p^* \to \{\pm 1\}$ is a group homomorphism with kernel $\mathcal{Q}_p$. Moreover, we extend $\left(\frac{\cdot}{p}\right)$ to $\mathbb{Z} \setminus \mathbb{Z}_p$ via the natural epimorphism $\mathbb{Z} \to \mathbb{Z}_p$.

Lemma. $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$, that is $\left(\frac{-1}{p}\right) = 1$ if and only if $p \equiv 1 \pmod 4$.

Proof. From $X^2 - 1 = (X-1)(X+1) \in \mathbb{Z}_p[X]$ we infer that $-1 \in \mathbb{Z}_p^*$ is the unique primitive second root of unity. Hence $-1 \in \mathbb{Z}_p^*$ is a square if and only if $\mathbb{Z}_p^*$ has an element of order 4, which by Artin's Theorem is equivalent to $p \equiv 1 \pmod 4$, the latter in turn being equivalent to $(-1)^{\frac{p-1}{2}} = 1$. ♯

b) Now let $q \neq p$ be a prime. Then, by **Dirichlet's Theorem** on primes in an arithmetic progression, given p there are infinitely many q such that $\left(\frac{q}{p}\right) = 1$.

The **Quadratic Reciprocity Law** [Gauss, 1796], says that **i)** $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}$ whenever q is odd, that is $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$ except $p, q \equiv -1 \pmod 4$, and **ii)** $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$, that is $\left(\frac{2}{p}\right) = 1$ if and only if $p \equiv \pm 1 \pmod 8$.

Consequently, given q there are infinitely many p such that $\left(\frac{q}{p}\right) = 1$.

c) Let still p be an odd prime, and let $q \neq p$ be a prime. Then for the associated **Gaussian sum**, being defined as $\gamma_p := \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^i \in \mathbb{F}_q(\zeta_p)$, we have:

Lemma. i) We have $\gamma_p^2 = \left(\frac{-1}{p}\right) \cdot p \in \mathbb{F}_q$, in particular $\gamma_p \neq 0$.
ii) If $\left(\frac{q}{p}\right) = 1$, then we have $\gamma_p \in \mathbb{F}_q$.

Proof. i) We have $\gamma_p^2 = \sum_{i,j \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right)\left(\frac{-j}{p}\right) \zeta_p^{i-j} = \left(\frac{-1}{p}\right) \cdot \sum_{i,j \in \mathbb{Z}_p^*} \left(\frac{ij}{p}\right) \zeta_p^{i-j}$. Since multiplication with $j \in \mathbb{Z}_p^*$ induces the bijection $\mathbb{Z}_p^* \to \mathbb{Z}_p^* \colon i \mapsto ij$ we get $\gamma_p^2 = \left(\frac{-1}{p}\right) \cdot \sum_{i,j \in \mathbb{Z}_p^*} \left(\frac{ij^2}{p}\right) \zeta_p^{(i-1)j} = \left(\frac{-1}{p}\right) \cdot \sum_{i,j \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^{(i-1)j}$. Using $\sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) = 0$ we get $\gamma_p^2 = \left(\frac{-1}{p}\right) \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\left(\frac{i}{p}\right) \cdot \sum_{j \in \mathbb{Z}_p} (\zeta_p^{i-1})^j \right)$. From $X^p - 1 = (X-1) \cdot \sum_{i=0}^{p-1} X^i \in \mathbb{F}_q[X]$ we conclude that $\sum_{j \in \mathbb{Z}_p} (\zeta_p^i)^j = 0$ whenever $i \in \mathbb{Z}_p^*$, hence in the above outer sum only the case $i = 1$ remains, yielding $\gamma_p^2 = \left(\frac{-1}{p}\right) \cdot p$.

ii) Recalling that $\mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F}_q(\zeta_p)) = \langle \varphi_q \rangle$, we show that $\gamma_p^q = \gamma_p \in \mathbb{F}_q(\zeta_p)$: We have $\gamma_p^q = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right)^q \zeta_p^{iq} = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^{iq}$; note that for $q = 2$ we have $\left(\frac{i}{p}\right) = 1 \in \mathbb{F}_2$. From $\left(\frac{iq}{p}\right) = \left(\frac{i}{p}\right)$ we get $\gamma_p^q = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{iq}{p}\right) \zeta_p^{iq} = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^i = \gamma_p$. ♯

Note that in the latter case it depends on the chosen minimum polynomial $\mu_1 \in \mathbb{F}_q[X]$ of ζ_p which square root of $\left(\frac{-1}{p}\right) \cdot p \in \mathbb{F}_q$ equals $\gamma_p \in \mathbb{F}_q$.

(11.2) Quadratic residue codes. a) Let p be an odd prime, and let $q \neq p$ be a prime such that $\left(\frac{q}{p}\right) = 1$. Since $\left(\frac{iq}{p}\right) = \left(\frac{i}{p}\right)$ for all $i \in \mathbb{Z}_p^*$, we conclude that both $\{\zeta_p^i \in \mathbb{F}_q(\zeta_p); i \in \mathcal{Q}_p\}$ and $\{\zeta_p^i \in \mathbb{F}_q(\zeta_p); i \in \mathcal{N}_p\}$ are Γ-invariant, where $\Gamma := \mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F}_q(\zeta_p))$. Hence letting $\rho_p := \prod_{i \in \mathcal{Q}_p}(X - \zeta_p^i) \in \mathbb{F}_q(\zeta_p)[X]$ and $\eta_p := \prod_{i \in \mathcal{N}_p}(X - \zeta_p^i) \in \mathbb{F}_q(\zeta_p)[X]$, we infer that both ρ_p and η_p have coefficients in $\mathbb{F}_q$. Thus we get $\rho_p\eta_p = \prod_{i \in \mathbb{Z}_p^*}(X - \zeta_p^i) = \frac{X^p-1}{X-1} = \sum_{i=0}^{p-1} X^i \in \mathbb{F}_q[X]$, hence $(X-1) \cdot \rho_p\eta_p = X^p - 1$.

This gives rise to the following **quadratic residue (QR) codes**: Let $\mathcal{Q}^p \leq \mathbb{F}_q^p$ and $\mathcal{N}^p \leq \mathbb{F}_q^p$ be the cyclic codes having generator polynomial ρ_p and η_p, respectively. Hence we have $\dim_{\mathbb{F}_q}(\mathcal{Q}^p) = \dim_{\mathbb{F}_q}(\mathcal{N}^p) = p - \frac{p-1}{2} = \frac{p+1}{2}$.

Moreover, let $(\mathcal{Q}^p)' \leq \mathbb{F}_q^p$ and $(\mathcal{N}^p)' \leq \mathbb{F}_q^p$ be the associated expurgated codes, that is having generator polynomials $(X-1) \cdot \rho_p$ and $(X-1) \cdot \eta_p$, respectively; recall that for $v = [a_0, \ldots, a_{p-1}] \in \mathbb{F}_q^p$ the condition $\sum_{i=0}^{p-1} a_i = 0 \in \mathbb{F}_q$ is equivalent to $\nu(v)(1) = (\sum_{i=0}^{p-1} a_i X^i)(1) = 0$, that is $X - 1 \mid \nu(v) \in \mathbb{F}_q[X]$. Hence we have $\dim_{\mathbb{F}_q}((\mathcal{Q}^p)') = \dim_{\mathbb{F}_q}((\mathcal{N}^p)') = p - \frac{p+1}{2} = \frac{p-1}{2}$.

b) For $\mathcal{C}^p \in \{\mathcal{Q}^p, \mathcal{N}^p\}$, the associated **extended quadratic residue code** is defined as $\widehat{\mathcal{C}}^p := \{[a_0, \ldots, a_{p-1}, a_\infty] \in \mathbb{F}_q^{p+1}; [a_0, \ldots, a_{p-1}] \in \mathcal{C}^p, a_\infty = \frac{\epsilon\gamma_p}{p} \cdot \sum_{i=0}^{p-1} a_i\}$, for $\epsilon \in \{\pm 1\}$; note that the choices yield linearly equivalent codes.

The reason for twisting the original definition of an extended code will become clear in (11.5) below. In particular, for $q = 2$ we have $\widehat{\mathcal{C}}^p := \{[a_0, \ldots, a_{p-1}, a_\infty] \in \mathbb{F}_2^{p+1}; [a_0, \ldots, a_{p-1}] \in \mathcal{C}^p, a_\infty + \sum_{i=0}^{p-1} a_i = 0\}$ anyway, for $q = 3$ we still may choose ϵ such that $\widehat{\mathcal{C}}^p := \{[a_0, \ldots, a_{p-1}, a_\infty] \in \mathbb{F}_3^{p+1}; [a_0, \ldots, a_{p-1}] \in \mathcal{C}^p, a_\infty + \sum_{i=0}^{p-1} a_i = 0\}$, so that in both cases we recover the conventional extended code.

(11.3) Example. For $q := 2$ and $p := 7$ we find that $2 \in \mathbb{Z}_7^*$ has order $3 = \frac{7-1}{2}$, thus $\mathbb{F}_2(\zeta_7) = \mathbb{F}_8$ and $\varphi_2 \in \Gamma := \mathrm{Aut}_{\mathbb{F}_2}(\mathbb{F}_8)$ has order 3. Moreover, we conclude that $2 \in \mathcal{Q}_7$, that is $\left(\frac{2}{7}\right) = 1$. Hence the Γ-orbits on $\mathcal{V}_7$ are $\mathcal{V}_7 = \{1\} \mathbin{\dot\cup} \{\zeta_7^i; i \in \mathcal{Q}_7\} \mathbin{\dot\cup} \{\zeta_7^i; i \in \mathcal{N}_7\}$, where $\mathcal{Q}_7 := \{1, 2, 4\}$ and $\mathcal{N}_7 := \{3, 5, 6\}$.

Thus we have $X^7 + 1 = (X+1) \cdot \prod_{i \in \mathcal{Q}_7}(X + \zeta_7^i) \cdot \prod_{i \in \mathcal{N}_7}(X + \zeta_7^i) = \mu_0\mu_1\mu_3 \in \mathbb{F}_8[X]$. Actually, we have $X^7 + 1 = (X+1) \cdot g'g'' \in \mathbb{F}_2[X]$, where $g' := X^3 + X + 1 \in \mathbb{F}_2[X]$ and $g'' := X^3 + X^2 + 1 \in \mathbb{F}_2[X]$, hence the latter are both irreducible; see (9.1).

Let $\mathcal{C} \leq \mathbb{F}_2^7$ be the cyclic code generated by g'; since $(g')^* = g''$ the code generated by g'' is linearly equivalent to $\mathcal{C}$. Hence the even-weight subcode $\mathcal{C}' \leq \mathcal{C}$ has generator polynomial $(X+1) \cdot g' = X^4 + X^3 + X^2 + 1$. Moreover, $\mathcal{C}$ has check polynomial $h := (X+1) \cdot g'' = X^4 + X^2 + X + 1$, thus $\mathcal{C}^\perp$ has generator polynomial $h^* = (X+1)^* \cdot (g'')^* = (X+1) \cdot g'$, showing that $\mathcal{C}^\perp = \mathcal{C}' \leq \mathcal{C}$.

Choosing a primitive 7-th root of unity $\zeta_7 \in \mathbb{F}_8^*$ having minimum polynomial g', we conclude that $\mathcal{C}$ is a QR code of type $\mathcal{Q}$; the code generated by g'' then is the associated QR code of type $\mathcal{N}$. Moreover, extending yields $\widehat{\mathcal{C}} \leq \mathbb{F}_2^8$.

The code $\mathcal{C}$ has defining set $\{\zeta_7\}$, so that by (9.4) taking $q := 2$ and $k := 3$ there, we conclude that $\mathcal{C}$ is linearly equivalent to the Hamming $[7,4,3]$-code $\mathcal{H}_3$. Thus $\mathcal{C}'$ is linearly equivalent to the even-weight Hamming $[7,3,4]$-code $\mathcal{H}_3'$, and $\widehat{\mathcal{C}}$ is linearly equivalent to the self-dual extended Hamming $[8,4,4]$-code $\widehat{\mathcal{H}}_3$. (In view of (11.5) below the duality properties of these codes are not surprising.)

(11.4) Theorem. Let p be an odd prime, and $q \neq p$ a prime such that $\left(\frac{q}{p}\right) = 1$.
a) Then $\mathcal{Q}^p$ and $\mathcal{N}^p$ are linearly equivalent, and so are $(\mathcal{Q}^p)'$ and $(\mathcal{N}^p)'$, as well as are $\widehat{\mathcal{Q}}^p$ and $\widehat{\mathcal{N}}^p$ with either choice of ϵ.
b) Let $v \in \mathcal{Q}^p \setminus (\mathcal{Q}^p)'$. Then for $d := \mathrm{wt}(v) \in \mathbb{N}$ we have the **square root bound** $d^2 \geq p$. Moreover, if $p \equiv -1 \pmod 4$ then we have $d^2 - d + 1 \geq p$; and if $p \equiv -1 \pmod 8$ and $q = 2$ then we have $d \equiv 3 \pmod 4$.

Proof. a) Let $j \in \mathcal{N}_p$. Then from $\left(\frac{ij}{p}\right) = \left(\frac{i}{p}\right)\left(\frac{j}{p}\right) = -\left(\frac{i}{p}\right)$, for all $i \in \mathbb{Z}_p^*$, we conclude that the bijection $\pi \colon \mathbb{Z}_p \to \mathbb{Z}_p \colon i \mapsto ij$ interchanges the sets $\mathcal{Q}_p$ and $\mathcal{N}_p$, while 0 is kept fixed. We consider the linear isometry induced by letting π permute the components of $\mathbb{F}_q^p$:

For $v := [a_0, \ldots, a_{p-1}] \in \mathbb{F}_q^p$ we get $v^\pi = [a_{i\pi^{-1}}; i \in \mathbb{Z}_p] \in \mathbb{F}_q^p$, that is $\nu(v^\pi) = \sum_{i \in \mathbb{Z}_p} a_{i\pi^{-1}} X^i = \sum_{i \in \mathbb{Z}_p} a_i X^{i\pi} = \sum_{i \in \mathbb{Z}_p} a_i X^{ij}$, where exponents are taken in $\mathbb{Z}_p$, that is we are computing in $\overline{\mathbb{F}_q[X]} = \mathbb{F}_q[X]/\langle X^p - 1\rangle$. Since evaluation at a p-th root of unity factors through $\overline{\mathbb{F}_q[X]}$, for $k \in \mathbb{Z}_p$ we get $\nu(v^\pi)(\zeta_p^k) = \sum_{i \in \mathbb{Z}_p} a_i \zeta_p^{ijk} = \sum_{i \in \mathbb{Z}_p} a_i (\zeta_p^{kj})^i = \nu(v)(\zeta_p^{kj}) \in \mathbb{F}_q(\zeta_p)$, thus $\zeta_p^k \in \mathcal{V}(\nu(v^\pi))$ if and only if $\zeta_p^{kj} \in \mathcal{V}(\nu(v))$. Hence we have $v \in \mathcal{Q}^p$ if and only if $v^\pi \in \mathcal{N}^p$.

Finally, since the linear equivalence between $\mathcal{Q}^p$ and $\mathcal{N}^p$ is induced by a permutation of components, it induces a linear equivalence between $(\mathcal{Q}^p)'$ and $(\mathcal{N}^p)'$ and a linear equivalence between $\widehat{\mathcal{Q}}^p$ and $\widehat{\mathcal{N}}^p$.

b) We have $\rho_p \mid \nu(v) \in \mathbb{F}_q[X]$, but $(X-1) \nmid \nu(v)$. Recalling the equality $\nu(v^\pi)(\zeta_p^k) = \nu(v)(\zeta_p^{kj})$ for all $k \in \mathbb{Z}_p$, we get $\eta_p \mid \nu(v^\pi)$, but $(X-1) \nmid \nu(v^\pi)$. Hence we have $\sum_{i=0}^{p-1} X^i = \rho_p\eta_p \mid \nu(v)\nu(v^\pi)$, but $X^p - 1 = (X-1)\cdot\rho_p\eta_p \nmid \nu(v)\nu(v^\pi)$. Let $w \in \mathbb{F}_q^p$ be the vector associated with $\nu(v)\nu(v^\pi) \in \mathbb{F}_q[X]$. Then we have $w \neq 0$, and since $\sum_{i=0}^{p-1} X^i$ generates the repetition subcode of $\mathbb{F}_q^p$ we conclude that $w = [a, \ldots, a]$ for some $0 \neq a \in \mathbb{F}_q$, hence $\mathrm{wt}(w) = p$.

Now for $\nu(v)\nu(v^\pi)$ we get d^2 products of non-zero coefficients of $\nu(v)$ and $\nu(v^\pi)$, respectively. Hence $\nu(v)\nu(v^\pi)$ has at most d^2 non-zero coefficients, thus $d^2 \geq p$.

If $p \equiv -1 \pmod 4$, that is $\left(\frac{-1}{p}\right) = -1$, then we may take $j = -1$, thus $\pi \colon \mathbb{Z}_p \to \mathbb{Z}_p \colon i \mapsto -i$. Then d of the above products belong to the constant coefficient of $\nu(v)\nu(v^\pi)$, hence the latter has at most $d^2 - d + 1$ non-zero coefficients.

Finally, if addionally $q = 2$ then d is odd. Hence $d-1$ of the products belonging to the constant, that is 0-th, coefficient cancel. Moreover, if two products belonging to the i-th coefficient of $\nu(v)\nu(v^\pi)$ cancel, where $i \in \mathbb{Z}_p^*$, then there are two products belonging to the $(-i)$-th coefficient canceling as well. Hence cancellation for these coefficients occurs in quadruples. Thus we get $d^2-d+1 \equiv p \equiv -1 \pmod 4$, thus $d(d-1) \equiv 2 \pmod 4$, which implies $d \equiv 3 \pmod 4$. $\sharp$

(11.5) Theorem. Let p be an odd prime, and $q \neq p$ a prime such that $\left(\frac{q}{p}\right) = 1$.
a) If $p \equiv -1 \pmod 4$, then we have $(\mathcal{Q}^p)^\perp = (\mathcal{Q}^p)'$, as well as $(\widehat{\mathcal{Q}}^p)^\perp = \widehat{\mathcal{Q}}^p$ with either choice of ϵ.
b) If $p \equiv 1 \pmod 4$, then we have $(\mathcal{Q}^p)^\perp = (\mathcal{N}^p)'$, as well as $(\widehat{\mathcal{Q}}^p)^\perp = \widehat{\mathcal{N}}^p$ with opposite choices of ϵ.

Proof. i) We first consider $(\mathcal{Q}^p)^\perp$: Recall that $\mathcal{Q}^p$ has generator polynomial ρ_p. Now we have $\rho_p^* = \prod_{i \in \mathcal{Q}_p}(X-\zeta_p^i)^* = \prod_{i \in \mathcal{Q}_p}(-\zeta_p^i)(X-\zeta_p^{-i}) \in \mathbb{F}_q(\zeta_p)[X]$. Moreover, we have $\left(\frac{-i}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{i}{p}\right)$ for all $i \in \mathbb{Z}_p^*$.

Hence if $p \equiv -1 \pmod 4$, then $\left(\frac{-i}{p}\right) = -\left(\frac{i}{p}\right)$ implies $\prod_{i \in \mathcal{Q}_p}(X-\zeta_p^{-i}) = \prod_{i \in \mathcal{N}_p}(X-\zeta_p^i) = \eta_p$, thus $\rho_p^* \sim \eta_p \in \mathbb{F}_q[X]$. Now $(\mathcal{Q}^p)^\perp$ has generator polynomial $(X-1)^* \cdot \eta_p^* \sim (X-1)\cdot\rho_p$, so that $(\mathcal{Q}^p)^\perp = (\mathcal{Q}^p)'$.

If $p \equiv 1 \pmod 4$, then $\left(\frac{-i}{p}\right) = \left(\frac{i}{p}\right)$ implies $\prod_{i \in \mathcal{Q}_p}(X-\zeta_p^{-i}) = \prod_{i \in \mathcal{Q}_p}(X-\zeta_p^i) = \rho_p$, thus $\rho_p^* \sim \rho_p$ and hence $\eta_p^* \sim \eta_p$. Now $(\mathcal{Q}^p)^\perp$ has generator polynomial $(X-1)^* \cdot \eta_p^* \sim (X-1)\cdot\eta_p$, so that $(\mathcal{Q}^p)^\perp = (\mathcal{N}^p)'$.

ii) We now consider $(\widehat{\mathcal{C}}^p)^\perp$, where $\mathcal{C}^p \in \{\mathcal{Q}^p, \mathcal{N}^p\}$: Let $G' \in \mathbb{F}_q^{\frac{p-1}{2}\times p}$ be a generator matrix of $(\mathcal{C}^p)'$. For the vector $1_p \in \mathbb{F}_q^p$ we have $\nu(1_p) = \sum_{i=0}^{p-1} X^i = \rho_p\eta_p \in \mathbb{F}_q[X]$, entailing that $1_p \in \mathcal{C}^p$. But since $X-1 \nmid \nu(1_p)$ we have $1_p \notin (\mathcal{C}^p)'$. Hence we recover $\mathcal{C}^p$ by augmenting $(\mathcal{C}^p)'$ again, thus $G := \left[\begin{array}{c} G' \\ \hline 1_p \end{array}\right] \in \mathbb{F}_q^{\frac{p+1}{2}\times p}$ is a generator matrix of $\mathcal{C}^p$. Now, for $[a_0, \ldots, a_{p-1}] \in (\mathcal{C}^p)'$ we have $\sum_{i=0}^{p-1} a_i = 0$, hence we get $[a_0, \ldots, a_{p-1}, 0] \in \widehat{\mathcal{C}}^p$, and for $1_p \in \mathcal{C}^p$ we get $v := [1, \ldots, 1, \epsilon\gamma_p] \in \widehat{\mathcal{C}}^p$. Hence $\widehat{G} := \left[\begin{array}{c|c} G' & 0^{\text{tr}}_{\frac{p-1}{2}} \\ \hline 1_p & \epsilon\gamma_p \end{array}\right] \in \mathbb{F}_q^{\frac{p+1}{2}\times(p+1)}$ is a generator matrix of $\widehat{\mathcal{C}}^p$.

If $p \equiv -1 \pmod 4$, then we have $\langle \mathcal{Q}^p, (\mathcal{Q}^p)' \rangle = \{0\}$ and $\langle v, v \rangle = p + \gamma_p^2 = \left(1 + \left(\frac{-1}{p}\right)\right) \cdot p = 0 \in \mathbb{F}_q$. This shows that $\widehat{\mathcal{Q}}^p \leq (\widehat{\mathcal{Q}}^p)^\perp$, hence $\dim_{\mathbb{F}_q}(\widehat{\mathcal{Q}}^p) = \frac{p+1}{2} = \dim_{\mathbb{F}_q}((\widehat{\mathcal{Q}}^p)^\perp)$ entails equality.

If $p \equiv 1 \pmod 4$, then we still have $\langle \mathcal{Q}^p, (\mathcal{N}^p)' \rangle = \{0\}$, but now we get $\langle [1, \ldots, 1, \epsilon\gamma_p], [1, \ldots, 1, -\epsilon\gamma_p] \rangle = p - \gamma_p^2 = \left(1 - \left(\frac{-1}{p}\right)\right) \cdot p = 0 \in \mathbb{F}_q$. This shows that $\widehat{\mathcal{Q}}^p \leq (\widehat{\mathcal{N}}^p)^\perp$ with opposite choices of ϵ, hence $\dim_{\mathbb{F}_q}(\widehat{\mathcal{Q}}^p) = \frac{p+1}{2} =$

$\dim_{\mathbb{F}_q}((\widehat{\mathcal{N}}^p)^\perp)$ entails equality. ♯

For the next result we already need the Gleason-Prange Theorem (12.4) to be proven below, which in particular says that the linear automorphism group of an extended QR code induces a transitive group of component permutations.

(11.6) Proposition. Let p be an odd prime, and $q \neq p$ be a prime such that $\left(\frac{q}{p}\right) = 1$. Then we have $d((\mathcal{Q}^p)') = d(\mathcal{Q}^p)+1$; in particular, the assertions of the square root bound (11.4) hold for $d(\mathcal{Q}^p)$. Moreover, we have $d(\widehat{\mathcal{Q}}^p) = d(\mathcal{Q}^p)+1$.

Proof. Let $v = [a_0, \ldots, a_{p-1}] \in (\mathcal{Q}^p)'$ such that $\mathrm{wt}(v) = d((\mathcal{Q}^p)')$. Then we have $\widehat{v} := [a_0, \ldots, a_{p-1}, 0] \in \widehat{\mathcal{Q}}^p$, hence by (12.4) there is $w = [b_0, \ldots, b_{p-1}] \in \mathcal{Q}^p \setminus (\mathcal{Q}^p)'$ such that $\widehat{w} := [b_0, \ldots, b_{p-1}, b_\infty] \in \widehat{\mathcal{Q}}^p$, where $\mathrm{wt}(\widehat{w}) = \mathrm{wt}(\widehat{v})$ and $b_\infty \neq 0$. Since $\mathrm{wt}(w) = \mathrm{wt}(\widehat{w}) - 1 = \mathrm{wt}(\widehat{v}) - 1 = \mathrm{wt}(v) - 1 = d((\mathcal{Q}^p)') - 1$ we conclude that $d(\mathcal{Q}^p) \leq d((\mathcal{Q}^p)') - 1 \leq p - 1$.

Conversely, let $v = [a_0, \ldots, a_{p-1}] \in \mathcal{Q}^p$ such that $\mathrm{wt}(v) = d(\mathcal{Q}^p) \leq p-1$, and let $\widehat{v} := [a_0, \ldots, a_{p-1}, a_\infty] \in \widehat{\mathcal{Q}}^p$. Again by (12.4) there is $w = [b_0, \ldots, b_{p-1}] \in (\mathcal{Q}^p)'$ such that $\widehat{w} := [b_0, \ldots, b_{p-1}, 0] \in \widehat{\mathcal{Q}}^p$, where $\mathrm{wt}(\widehat{v}) = \mathrm{wt}(\widehat{w})$. Since $\mathrm{wt}(w) = \mathrm{wt}(\widehat{w}) = \mathrm{wt}(\widehat{v}) \leq \mathrm{wt}(v)+1 = d(\mathcal{Q}^p)+1$ we conclude that $d((\mathcal{Q}^p)') \leq d(\mathcal{Q}^p)+1$.

The second assertion follows from recalling that for $v = [a_0, \ldots, a_{p-1}] \in \mathcal{Q}^p$ and $[a_0, \ldots, a_{p-1}, a_\infty] \in \widehat{\mathcal{Q}}^p$ we have $a_\infty = 0$ if and only if $v \in (\mathcal{Q}^p)'$. ♯

12 Automorphisms of quadratic residue codes

(12.1) Automorphisms of codes. a) We need an additional general piece of notation: Given a linear code $\mathcal{C} \leq \mathbb{F}_q^n$, let $A(\mathcal{C}) := \mathrm{Aut}_{\mathbb{F}_q}(\mathcal{C}) \leq I_n(\mathbb{F}_q) \cong (\mathbb{F}_q^*)^n \rtimes \mathcal{S}_n$ be its linear automorphism group. Then let $P(\mathcal{C}) := A(\mathcal{C})/(A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n) \cong A(\mathcal{C})(\mathbb{F}_q^*)^n/(\mathbb{F}_q^*)^n \leq I_n(\mathbb{F}_q)/(\mathbb{F}_q^*)^n \cong \mathcal{S}_n$ be the group of component permutations induced by $A(\mathcal{C})$, and let $\overline{}\colon A(\mathcal{C}) \to P(\mathcal{C})$ be the natural epimorphism.

By linearity we always have $\mathbb{F}_q^* \cdot E_n \leq A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n$, for the trivial code we have $A(\{0\}) \cap (\mathbb{F}_q^*)^n = (\mathbb{F}_q^*)^n$, and for $q = 2$ we have $A(\mathcal{C}) \cap (\mathbb{F}_2^*)^n \cong \{1\}$ anyway. The question arises how $A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n \trianglelefteq A(\mathcal{C})$ looks like in general.

b) More can be said if $\mathcal{C} \leq \mathbb{F}_q^n$ is a non-trivial cyclic code such that $\gcd(q, n) = 1$: (The argument to follow was indicated to me by my student C. Kirch [2022].)

Let $g = \sum_{i=0}^k g_i X^i \in \mathbb{F}_q[X]$ be a generator polynomial of $\mathcal{C}$, having degree $k \in \{0, \ldots, n-1\}$, and let $\mathrm{supp}(g) := \mathrm{supp}(\nu^{-1}(g)) \subseteq \mathbb{Z}_n$ be the **support** of g; note that $0 \in \mathrm{supp}(g)$. Moreover, let $\langle \mathrm{supp}(g) \rangle = \langle \gamma \rangle \leq \mathbb{Z}_n$, where $\gamma := \gcd(\mathrm{supp}(g) \,\dot\cup\, \{n\}) \in \mathbb{Z}_n$, greatest common divisors being taken in $\mathbb{Z}$.

Let $D := \mathrm{diag}[a_0, \ldots, a_{n-1}] \in A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n$. Then, conjugating with the permutation $(0, \ldots, n-1) \in A(\mathcal{C})$, we have $\mathrm{diag}[a_{i+1}, \ldots, a_{n-1}, a_0, \ldots, a_i] \in A(\mathcal{C})$ as well, for $i \in \mathbb{Z}_n$. Now let $i, i' \in \mathbb{Z}_n$ such that $i - i' \in \mathrm{supp}(g) \subseteq \mathbb{Z}_n$. In order to show that $a_i = a_{i'}$, by cyclicity we may assume that $i' = 0$, hence $i \in \operatorname{supp}(g)$; transporting the action of D to $\mathbb{F}_q[X]_{<n}$, we have $gD = \sum_{i=0}^{k} a_i g_i X^i \in \nu(\mathcal{C})$, that is $g \mid gD$, entailing $gD \sim g \in \mathbb{F}_q[X]$, implying that $a_0 = a_i$. This shows that the diagonal entries of D are constant along the cosets of $\gamma\mathbb{Z}_n$ in $\mathbb{Z}_n$.

Conversely, if $D := \operatorname{diag}[a_0, \ldots, a_{n-1}] \in (\mathbb{F}_q^*)^n$ having diagonal entries being constant along the cosets of $\gamma\mathbb{Z}_n$ in $\mathbb{Z}_n$. We consider the $\mathbb{F}_q$-basis $\{X^j g \in \mathbb{F}_q[X]_{<n}; j \in \{0, \ldots, n-k-1\}\} \subseteq \nu(\mathcal{C})$. Then we have $(X^j g)D \in \nu(\mathcal{C})$, that is $g \mid (X^j g)D$, and since $X^j \mid (X^j g)D$ we infer that $(X^j g)D \sim X^j g \in \mathbb{F}_q[X]$, entailing $D \in A(\mathcal{C})$. Hence we have $A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n \cong (\mathbb{F}_q^*)^{[\mathbb{Z}_n \colon \gamma\mathbb{Z}_n]} \cong (C_{q-1})^\gamma$; in particular, we have $A(\mathcal{C}) \cap (\mathbb{F}_q^*)^n = \mathbb{F}_q^* \cdot E_n$ if and only if $\gamma = 1$.

(12.2) Automorphisms of quadratic residue codes. a) Let p be an odd prime, and let $q \neq p$ be a prime such that $\left(\frac{q}{p}\right) = 1$. Since the cyclic QR codes $\mathcal{Q}^p \leq \mathbb{F}_q^p$ and $(\mathcal{Q}^p)' \leq \mathbb{F}_q^p$ have prime length p and non-constant generator polynomials of degree $\frac{p-1}{2}$ and $\frac{p+1}{2}$, respectively, in both cases we have $\gamma = 1$, entailing $A(\mathcal{Q}^p) \cap (\mathbb{F}_q^*)^p = A((\mathcal{Q}^p)') \cap (\mathbb{F}_q^*)^p = \mathbb{F}_q^* \cdot E_p$, thus consisting of the non-zero scalar matrices only.

For the extended QR code $\widehat{\mathcal{Q}}^p \leq \mathbb{F}_q^{p+1}$ we get: If $D := \operatorname{diag}[a_0, \ldots, a_{p-1}, a_\infty] \in A(\widehat{\mathcal{Q}}^p) \cap (\mathbb{F}_q^*)^{p+1}$, then since $(\widehat{\mathcal{Q}}^p)^\bullet = \mathcal{Q}^p$ we have $\operatorname{diag}[a_0, \ldots, a_{p-1}] \in A(\mathcal{Q}^p)$, thus the latter is a scalar matrix, which by the extension condition implies that D is a scalar matrix as well; in other words we have $A(\widehat{\mathcal{Q}}^p) \cap (\mathbb{F}_q^*)^{p+1} = \mathbb{F}_q^* \cdot E_{p+1}$.

b) We show that any automorphism $\alpha \in A(\mathcal{Q}^p) \leq I_p(\mathbb{F}_q)$ extends to an automorphism $\widehat{\alpha} \in A(\widehat{\mathcal{Q}}^p) \leq I_{p+1}(\mathbb{F}_q)$; in other words we have $A(\mathcal{Q}^p) = \operatorname{Stab}_{A(\widehat{\mathcal{Q}}^p)}(\langle [0_p \mid 1] \rangle_{\mathbb{F}_q})$ and $P(\mathcal{Q}^p) = \operatorname{Stab}_{P(\widehat{\mathcal{Q}}^p)}(\infty)$:

Let firstly $p \equiv -1 \pmod 4$. Then we have $(\mathcal{Q}^p)^\perp = (\mathcal{Q}^p)'$, so that α restricts to an automorphism of $(\mathcal{Q}^p)'$. Moreover, we have $\mathcal{Q}^p = (\mathcal{Q}^p)' \oplus \langle 1_p \rangle_{\mathbb{F}_q}$, thus $\alpha(1_p) = a \cdot 1_p + w$ for some $a \in \mathbb{F}_q^*$ and $w \in (\mathcal{Q}^p)'$. Now we have $\widehat{\mathcal{Q}}^p = \{[v \mid 0] \in \mathbb{F}_q^{p+1}; v \in (\mathcal{Q}^p)'\} \oplus \langle [1_p \mid \epsilon\gamma_p] \rangle_{\mathbb{F}_q}$. Hence letting $\widehat{\alpha} \in I_{p+1}(\mathbb{F}_q)$ be the monomial map given as $\widehat{\alpha}([v \mid c]) := [\alpha(v) \mid ac] \in \mathbb{F}_q^{p+1}$ for $v \in \mathbb{F}_q^p$ and $c \in \mathbb{F}_q$, we have $\widehat{\alpha}([v \mid 0]) = [\alpha(v) \mid 0] \in \widehat{\mathcal{Q}}^p$ for $v \in (\mathcal{Q}^p)'$, and $\widehat{\alpha}([1_p \mid \epsilon\gamma_p]) = [\alpha(1_p) \mid a\epsilon\gamma_p] = a \cdot [1_p \mid \epsilon\gamma_p] + [w \mid 0] \in \widehat{\mathcal{Q}}^p$. Thus $\widehat{\alpha} \in A(\widehat{\mathcal{Q}}^p)$ is an automorphism extending α.

Let secondly $p \equiv 1 \pmod 4$. Then we have $(\mathcal{Q}^p)^\perp = (\mathcal{N}^p)'$. so that α yields an automorphism of $(\mathcal{N}^p)'$. Moreover, we have $\mathcal{N}^p = (\mathcal{N}^p)' \oplus \langle 1_p \rangle_{\mathbb{F}_q}$, thus $\alpha(1_p) = a \cdot 1_p + w$ for some $a \in \mathbb{F}_q^*$ and $w \in (\mathcal{N}^p)'$. Now we have $\widehat{\mathcal{N}}^p = \{[v \mid 0] \in \mathbb{F}_q^{p+1}; v \in (\mathcal{N}^p)'\} \oplus \langle [1_p \mid -\epsilon\gamma_p] \rangle_{\mathbb{F}_q}$, where we use the opposite choice of $\epsilon \in \{\pm 1\}$. Hence letting $\widehat{\alpha} \in I_{p+1}(\mathbb{F}_q)$ be the monomial map given as $\widehat{\alpha}([v \mid c]) := [\alpha(v) \mid ac] \in \mathbb{F}_q^{p+1}$ for $v \in \mathbb{F}_q^p$ and $c \in \mathbb{F}_q$, we have $\widehat{\alpha}([v \mid 0]) = [\alpha(v) \mid 0] \in \widehat{\mathcal{N}}^p$ for $v \in (\mathcal{N}^p)'$, and $\widehat{\alpha}([1_p \mid -\epsilon\gamma_p]) = [\alpha(1_p) \mid -a\epsilon\gamma_p] = a \cdot [1_p \mid -\epsilon\gamma_p] + [w \mid 0] \in \widehat{\mathcal{N}}^p$. Thus $\widehat{\alpha} \in A(\widehat{\mathcal{N}}^p)$ is an automorphism extending α, which hence yields an an automorphism of $\widehat{\mathcal{Q}}^p = (\widehat{\mathcal{N}}^p)^\perp$. ♯

(12.3) Lemma. Let p be an odd prime and $q \neq p$ be a prime such that $\left(\frac{q}{p}\right) = 1$.
a) The map $\sigma \colon \mathbb{F}_q^{p+1} \to \mathbb{F}_q^{p+1} \colon [a_0, \ldots, a_{p-1}, a_\infty] \mapsto [b_0, \ldots, b_{p-1}, b_\infty]$ given by $b_i := \left(\frac{-i^{-1}}{p}\right) a_{-i^{-1}}$ for $i \in \mathbb{Z}_p^*$, while $b_0 := -\epsilon \left(\frac{-1}{p}\right) a_\infty$ and $b_\infty := -\epsilon a_0$, induces a linear automorphism of $\widehat{\mathcal{Q}}^p$, that is we have $\sigma \in A(\widehat{\mathcal{Q}}^p)$.
b) We have $\sigma \in \mathrm{SL}_{p+1}(\mathbb{F}_q)$ such that $\sigma^2 = \left(\frac{-1}{p}\right) \cdot \mathrm{id}_{\mathbb{F}_q^{p+1}}$. Moreover, we have $\mathrm{tr}(\sigma) = 0$ if $p \equiv -1 \pmod 4$, and $\mathrm{tr}(\sigma) = 2\left(\frac{\zeta_4}{p}\right)$ if $p \equiv 1 \pmod 4$, where $\zeta_r \in \mathbb{F}_p^*$ denotes a primitive 4-th root of unity.

Proof. a) For any vector $v := [a_0, \ldots, a_{p-1}, a_\infty] \in \mathbb{F}_q^{p+1}$ we write $w := \sigma(v) = [b_0, \ldots, b_{p-1}, b_\infty] \in \mathbb{F}_q^{p+1}$. Let $\alpha_j := \sum_{i \in \mathbb{Z}_p} a_i \zeta_p^{ij}$, for $j \in \mathbb{Z}_p$, be the **discrete Fourier transform** of v, leaving out a_∞. Using $\sum_{k \in \mathbb{Z}_p} \zeta_p^k = 0$, for $i \in \mathbb{Z}_p$ we get the **inverse transform** $\sum_{j \in \mathbb{Z}_p} \alpha_j \zeta_p^{-ij} = \sum_{k \in \mathbb{Z}_p} (a_k \cdot \sum_{j \in \mathbb{Z}_p} \zeta_p^{(k-i)j}) = pa_i$. Since σ is a monomial $\mathbb{F}_q$-linear map, we have $\sigma \in I_{p+1}(\mathbb{F}_q)$. Hence we have to show that if $v \in \widehat{\mathcal{Q}}^p$ then $w \in \widehat{\mathcal{Q}}^p$ as well. Being an element of $\widehat{\mathcal{Q}}^p$ is equivalent to saying that $\alpha_j = 0$ for all $j \in \mathcal{Q}_p$, and $a_\infty = \frac{\epsilon\gamma_p}{p} \cdot \sum_{i \in \mathbb{Z}_p} a_i = \frac{\epsilon\gamma_p\alpha_0}{p}$.

i) We show that $\sum_{i \in \mathbb{Z}_p} b_i = \frac{\epsilon p}{\gamma_p} \cdot b_\infty$: For the right hand side, using $\gamma_p^2 = \left(\frac{-1}{p}\right) p$, we get $\frac{\epsilon p}{\gamma_p} \cdot b_\infty = -\gamma_p \left(\frac{-1}{p}\right) a_0$. We turn to the left hand side:

Let $\beta := \sum_{i \in \mathbb{Z}_p^*} b_i = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-i^{-1}}{p}\right) a_{-i^{-1}} = \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) a_i = \frac{1}{p} \cdot \sum_{j \in \mathbb{Z}_p} (\alpha_j \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^{-ij})$. Since for $j = 0$ we get $\sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) = 0$ in the inner sum, we infer $\beta = \frac{1}{p} \left(\frac{-1}{p}\right) \cdot \sum_{j \in \mathbb{Z}_p^*} (\alpha_j \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-i}{p}\right) \zeta_p^{-ij})$. Since $\alpha_j = 0$ for all $j \in \mathcal{Q}_p$ we get $\beta = \frac{-1}{p} \left(\frac{-1}{p}\right) \cdot \sum_{j \in \mathbb{Z}_p^*} (\alpha_j \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-ij}{p}\right) \zeta_p^{-ij}) = \frac{-\gamma_p}{p} \left(\frac{-1}{p}\right) \cdot \sum_{j \in \mathbb{Z}_p^*} \alpha_j$. Since $\sum_{j \in \mathbb{Z}_p} \alpha_j = pa_0$ and $\epsilon\gamma_p\alpha_0 = pa_\infty$ we obtain $\beta = \frac{-\gamma_p}{p} \left(\frac{-1}{p}\right) \cdot (pa_0 - \alpha_0) = \left(\frac{-1}{p}\right) \cdot (-\gamma_p a_0 + \epsilon a_\infty)$. From this we finally get $\sum_{i \in \mathbb{Z}_p} b_i = b_0 + \beta = \left(\frac{-1}{p}\right) (-\epsilon a_\infty - \gamma_p a_0 + \epsilon a_\infty) = -\gamma_p \left(\frac{-1}{p}\right) a_0$, as desired.

ii) We show that the discrete Fourier transform of w fulfills $\beta_j = 0$ for $j \in \mathcal{Q}_p$: For $i \in \mathbb{Z}_p$ the inverse transform yields $pa_i = \alpha_0 + \sum_{k \in \mathbb{Z}_p^*} \alpha_k \zeta_p^{-ik} = \frac{\epsilon p}{\gamma_p} \cdot a_\infty + \sum_{k \in \mathbb{Z}_p^*} \alpha_k \zeta_p^{-ik}$. This yields $\beta_j = b_0 + \sum_{i \in \mathbb{Z}_p^*} b_i \zeta_p^{ij} = -\epsilon \left(\frac{-1}{p}\right) a_\infty + \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-i^{-1}}{p}\right) a_{-i^{-1}} \zeta_p^{ij} = \epsilon x a_\infty + \frac{y}{p}$, where $y := \sum_{i,k \in \mathbb{Z}_p^*} \left(\frac{-i^{-1}}{p}\right) \alpha_k \zeta_p^{ki^{-1}+ij}$ and $x := -\left(\frac{-1}{p}\right) + \frac{1}{\gamma_p} \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{-i^{-1}}{p}\right) \zeta_p^{ij}$. Since $j \in \mathcal{Q}_p$ we obtain $x \left(\frac{-1}{p}\right) = -1 + \frac{1}{\gamma_p} \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{i}{p}\right) \zeta_p^{ij} = -1 + \frac{1}{\gamma_p} \cdot \sum_{i \in \mathbb{Z}_p^*} \left(\frac{ij}{p}\right) \zeta_p^{ij} = -1 + \frac{\gamma_p}{\gamma_p} = 0$.

We have $y = \left(\frac{-1}{p}\right) \cdot \sum_{k \in \mathbb{Z}_p^*} (\sum_{i \in \mathbb{Z}_p^*} \left(\frac{ki^{-1}}{p}\right) \zeta_p^{ki^{-1}+ij}) \left(\frac{k}{p}\right) \alpha_k$. For $k \in \mathcal{Q}_p$ we have $\alpha_k = 0$. We proceed to show that the inner sum $\sum_{i\in\mathbb{Z}_p^*}\left(\frac{ki^{-1}}{p}\right)\zeta_p^{ki^{-1}+ij}$ vanishes whenever $k \in \mathcal{N}_p$, entailing $y = 0$: To this end, let $\pi\colon \mathbb{Z}_p^* \to \mathbb{Z}_p^*\colon i \mapsto \frac{jk}{i}$. Hence we have $\pi^2 = \mathrm{id}_{\mathbb{Z}_p^*}$. Assume that π has a fixed point, i say, then we have $i = \frac{jk}{i} \in \mathbb{Z}_p^*$, thus $jk = i^2 \in \mathcal{Q}_p$, a contradiction. Hence the permutation π consists of 2-cycles only, where we have $\left(\frac{\pi(ij)}{p}\right) = \left(\frac{ki^{-1}}{p}\right) = \left(\frac{k}{p}\right)\left(\frac{i}{p}\right) = -\left(\frac{i}{p}\right) = -\left(\frac{i}{p}\right)\left(\frac{j}{p}\right) = -\left(\frac{ij}{p}\right)$. From this, running through the various cycles of π, we obtain $\sum_{i\in\mathbb{Z}_p^*}\left(\frac{ki^{-1}}{p}\right)\zeta_p^{ki^{-1}+ij} = \sum_{i\in\mathbb{Z}_p^*}\left(\frac{\pi(ij)}{p}\right)\zeta_p^{ij+\pi(ij)} = 0$.

b) Recall that $-i^{-1} = i \in \mathbb{Z}_p^*$ if and only if $i \in \{\pm\zeta_4\} \subseteq \mathbb{Z}_p^*$. Hence, if $p \equiv -1 \pmod 4$, that is $\left(\frac{-1}{p}\right) = -1$, then $\det(\sigma) = \det(\pm\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix})^{\frac{p+1}{2}} = 1$ and $\mathrm{tr}(\sigma) = 0$ and $\sigma^2 = -\mathrm{id}_{\mathbb{F}_q^{p+1}}$. If $p \equiv 1 \pmod 4$, that is $\left(\frac{-1}{p}\right) = 1$, then $\det(\sigma) = \left(\frac{\zeta_4}{p}\right)^2 \cdot \det(\pm\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix})^{\frac{p-1}{2}} = 1$ and $\mathrm{tr}(\sigma) = 2\left(\frac{\zeta_4}{p}\right)$ and $\sigma^2 = \mathrm{id}_{\mathbb{F}_q^{p+1}}$. ♯

(12.4) Theorem: Gleason, Prange. Let p be an odd prime, and let $q \neq p$ be a prime such that $\left(\frac{q}{p}\right) = 1$. Then the group $P(\widehat{\mathcal{Q}}^p) \leq \mathcal{S}_{p+1}$ contains a subgroup isomorphic to $\mathrm{PSL}_2(\mathbb{F}_p)$, with respect to its natural action on $\mathbf{P}^1(\mathbb{F}_p)$.

Proof. i) We first exhibit a certain subgroup of $\mathcal{S}_{p+1}$: Let $S := \mathrm{SL}_2(\mathbb{F}_p)$ be the special linear group of degree 2 over $\mathbb{F}_p$. We have $S = \langle s, t, r\rangle$ (as will be shown below), where $s := \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$ and $t := \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$, and $r := \mathrm{diag}[c, c^{-1}]$, where $\mathbb{F}_p^* = \langle c\rangle$; hence $\langle c^2\rangle = \mathcal{Q}_p$.

We have $Z(S) = \{\pm E_2\}$, giving rise to the natural epimorphism $^-\colon S \to S/Z(S) = \overline{S} := \mathrm{PSL}_2(\mathbb{F}_p)$, where the latter is the associated projective special linear group, having order $\frac{1}{2}p(p-1)(p+1)$. Now, S acts naturally on the projective space $\mathbf{P}^1(\mathbb{F}_p) = \{x_0, \ldots, x_{p-1}, x_\infty\}$, where $x_i := \langle[1, i]\rangle_{\mathbb{F}_q}$ for $i \in \mathbb{Z}_p$, and $x_\infty := \langle[0, 1]\rangle_{\mathbb{F}_q}$. This induces a faithful action of $\overline{S}$ of degree $p + 1$, hence an embedding of $\overline{S}$ into $\mathcal{S}_{p+1}$.

More precisely, we have $x_i t = \langle[1, i]\rangle_{\mathbb{F}_q} \cdot t = \langle[1, i] \cdot t\rangle_{\mathbb{F}_q} = \langle[1, i+1]\rangle_{\mathbb{F}_q} = x_{i+1}$ for $i \in \mathbb{Z}_p$, and $x_\infty t = \langle[0, 1]\rangle_{\mathbb{F}_q} \cdot t = \langle[0, 1] \cdot t\rangle_{\mathbb{F}_q} = \langle[0, 1]\rangle_{\mathbb{F}_q} = x_\infty$; in other words $\overline{t} \in \overline{S}$ induces the p-cycle $(0, \ldots, p-1) \in \mathcal{S}_{p+1}$.

We have $x_i r = \langle[1, i]\rangle_{\mathbb{F}_q} \cdot r = \langle[1, i] \cdot r\rangle_{\mathbb{F}_q} = \langle[c, ic^{-1}]\rangle_{\mathbb{F}_q} = \langle[1, ic^{-2}]\rangle_{\mathbb{F}_q} = x_{ic^{-2}}$ for $i \in \mathbb{Z}_p$, and $x_\infty r = \langle[0, 1]\rangle_{\mathbb{F}_q} \cdot r = \langle[0, 1] \cdot r\rangle_{\mathbb{F}_q} = \langle[0, c^{-1}]\rangle_{\mathbb{F}_q} = \langle[0, 1]\rangle_{\mathbb{F}_q} = x_\infty$; in other words $\overline{r} \in \overline{S}$ induces the permutation $(1, c^{-2}, \ldots, c^2)(c^{-1}, c^{-3}, \ldots, c) \in \mathcal{S}_{p+1}$ of order $\frac{p-1}{2}$, permuting the square and non-squares in $\mathbb{F}_p^*$, respectively.

We have $x_i s = \langle[1, i]\rangle_{\mathbb{F}_q} \cdot s = \langle[1, i] \cdot s\rangle_{\mathbb{F}_q} = \langle[-i, 1]\rangle_{\mathbb{F}_q} = \langle[1, -i^{-1}]\rangle_{\mathbb{F}_q} = x_{-i^{-1}}$ for $i \in \mathbb{Z}_p*$, while $x_0 s = \langle[1, 0]\rangle_{\mathbb{F}_q} \cdot s = \langle[1, 0] \cdot s\rangle_{\mathbb{F}_q} = \langle[0, 1]\rangle_{\mathbb{F}_q}) = x_\infty$ and $x_\infty s = \langle[0, 1]\rangle_{\mathbb{F}_q} \cdot s = \langle[0, 1] \cdot s\rangle_{\mathbb{F}_q} = \langle[-1, 0]\rangle_{\mathbb{F}_q}) = \langle[1, 0]\rangle_{\mathbb{F}_q}) = x_0$; in other words

$\overline{s} \in \overline{S}$ induces the involution $(0,\infty)(1,-1)(2,\frac{-1}{2})\cdots \in \mathcal{S}_{p+1}$. Note that x_i is a fixed point if and only if $i^2 = -1 \in \mathbb{Z}_p$, saying that $\overline{s}$ has fixed points if and only if $p \equiv 1 \pmod 4$, in which case these are $\{\pm\zeta_4\} \subseteq \mathbb{Z}_p^*$.

Hence S acts transitively on $\mathbf{P}^1(\mathbb{F}_p)$, where $\mathrm{Stab}_S(x_\infty)$ consists of the matrices having $[0,1] \in \mathbb{F}_p^2$ as an eigenvector, thus $\mathrm{Stab}_S(x_\infty) = B := \langle t\rangle \rtimes \langle r\rangle \leq S$, the subgroup of upper tringular matrices. Since $t \in B$ we conclude that B acts transitively on $\{x_0,\ldots,x_{p-1}\}$, saying that S acts 2-fold transitively on $\mathbf{P}^1(\mathbb{F}_p)$. In particular, S acts primitively, so that $B < S$ is maximal subgroup, implying that $S = \langle B, s\rangle = \langle s,t,r\rangle$.

ii) We consider the $\mathbb{F}_q$-linear map $\tau\colon \mathbb{F}_q^{p+1} \to \mathbb{F}_q^{p+1}$ induced by the permutation $(0,\ldots,p-1) \in \mathcal{S}_{p+1}$ of the components. Since for $[a_0,\ldots,a_{p-1}] \in \mathcal{Q}^p$ we have $[a_{p-1},a_0,\ldots,a_{p-2}] \in \mathcal{Q}^p$ as well, and the extension condition is trivially fulfilled, we conclude that τ induces a linear automorphism of $\widehat{\mathcal{Q}}^p$, that is we have $\tau \in A(\widehat{\mathcal{Q}}^p)$, inducing $\overline{\tau} = (0,\ldots,p-1) \in P(\widehat{\mathcal{Q}}^p) \leq \mathcal{S}_{p+1}$, fixing ∞.

Next, we consider the $\mathbb{F}_q$-linear map $\rho\colon \mathbb{F}_q^{p+1} \to \mathbb{F}_q^{p+1}$ induced by the permutation $(1,c^{-2},\ldots,c^2)(c^{-1},c^{-3},\ldots,c) \in \mathcal{S}_{p+1}$. Hence for $[a_0,\ldots,a_{p-1}] \in \mathbb{F}_q^p$ we have $\nu(\rho([a_0,\ldots,a_{p-1}])) = a_0 + \sum_{i\in\mathbb{Z}_p^*} a_{ic^{-2}}X^i = a_0 + \sum_{i\in\mathbb{Z}_p^*} a_i X^{ic^2} \in \mathbb{F}_q[X]$. Thus for $[a_0,\ldots,a_{p-1}] \in \mathcal{Q}^p$ and $k \in \mathcal{Q}_p$, noting that $kc^2 \in \mathcal{Q}_p$ as well, we get $\nu(\rho([a_0,\ldots,a_{p-1}]))(\zeta_p^k) = a_0 + \sum_{i=1}^{p-1} a_i(\zeta_p^{kc^2})^i = 0$, implying that $\rho([a_0,\ldots,a_{p-1}]) \in \mathcal{Q}_p$. Since the extension condition is trivially fulfilled, we conclude that ρ induces a linear automorphism of $\widehat{\mathcal{Q}}^p$, that is we have $\rho \in A(\widehat{\mathcal{Q}}^p)$, inducing $\overline{\rho} = (1,c^{-2},\ldots,c^2)(c^{-1},c^{-3},\ldots,c) \in P(\widehat{\mathcal{Q}}^p) \leq \mathcal{S}_{p+1}$, fixing both $\{0,\infty\}$.

Finally, by (12.3) we have $\sigma \in A(\widehat{\mathcal{Q}}^p)$, inducing the permutation in $\overline{\sigma} \in P(\widehat{\mathcal{Q}}^p) \leq \mathcal{S}_{p+1}$ given by $i \mapsto -i^{-1}$ for $i \in \mathbb{Z}_p^*$, interchanging $\{0,\infty\}$. Thus identifying the points of $\mathbf{P}^1(\mathbb{F}_p)$ with the standard $\mathbb{F}_q$-basis of $\mathbb{F}_q^{p+1}$, and comparing with the natrual action of $S = \langle s,t,r\rangle$ on $\mathbf{P}^1(\mathbb{F}_p)$, we conclude that we get an isomorphism $\overline{S} \cong \langle\overline{\sigma},\overline{\tau},\overline{\rho}\rangle \leq P(\widehat{\mathcal{Q}}^p)$, mapping $\overline{s} \mapsto \overline{\sigma}$, $\overline{t} \mapsto \overline{\tau}$, $\overline{r} \mapsto \overline{\rho}$. ♯

Actually, it turns out that, up to three exceptions, we have $P(\widehat{\mathcal{Q}}^p) = \overline{S}$; in this case we have $P(\mathcal{Q}^p) = \mathrm{Stab}_{\overline{S}}(\infty) = \overline{B}$ of order $\frac{1}{2}p(p-1)$. The exceptions are as follows [Knapp, Schmid, 1980]:
i) $[q,p] = [2,7]$, in which case we have $A(\widehat{\mathcal{Q}}^p) \cong P(\widehat{\mathcal{Q}}^p) \cong \mathrm{AGL}_3(\mathbb{F}_2) \cong C_2^3 \rtimes \mathrm{SL}_3(\mathbb{F}_2) \cong C_2^3 \rtimes \mathrm{PSL}_2(\mathbb{F}_7)$, hence $A(\mathcal{Q}^p) \cong P(\mathcal{Q}^p) \cong C_2^3 \rtimes (C_7 \rtimes C_3)$, see (11.3); **ii)** $[q,p] = [2,23]$, see (13.1); **iii)** $[q,p] = [3,11]$, see (13.2).

(12.5) Remark. We consider the group $A := \langle\sigma,\tau,\rho\rangle \leq A(\widehat{\mathcal{Q}}^p) \cap \mathrm{SL}_{p+1}(\mathbb{F}_q)$: For $q = 2$ we have $A \cong \overline{A} \cong \overline{S}$ anyway, so that we may assume that $q \neq 2$. By the matrices given we conclude that A is a subgroup of the group $\{\pm 1\}^{p+1} \rtimes \mathcal{S}_{p+1}$ of signed permutations. In particular, the representation of A considered lifts to a representation $\Delta\colon A \to \mathrm{SL}_{p+1}(\mathbb{Z})$, we have $A \cong \Delta(A)$ independently of q, and we have $A \cap (\mathbb{F}_q^*)^{p+1} \leq \{\pm 1\}^{p+1}$. Since $A(\widehat{\mathcal{Q}}^p) \cap (\mathbb{F}_q^*)^{p+1} = \mathbb{F}_q^* \cdot E_{p+1}$, we conclude that $A \cap (\mathbb{F}_q^*)^{p+1} \leq \{\pm E_{p+1}\} =: Z$, a cyclic group of order 2.

Since $A/(A \cap (\mathbb{F}_q^*)^{p+1}) \cong \overline{A} \cong \overline{S}$ we conclude that either $A \cong \overline{S}$, or A is a central extension of shape $A \cong Z.\overline{S}$. In the latter case, since the Schur multiplier of $\overline{S}$ is cyclic of order 2, we either have a split extension $A \cong Z \times \overline{S}$ or a non-split extension $A \cong Z.\overline{S}$, where in the non-split case for $p \geq 5$, since then $\overline{S}$ is perfect, we necessarily have $A \cong S$, being the unique Schur representation group.

i) If $p \equiv 1 \pmod 4$, then $\sigma \in A$ has order 2. Since S does not possess non-central involutions, this entails $A \cong \overline{S}$ or $A \cong Z \times \overline{S}$, where $\overline{S} = [A, A]$, and $Z = Z(A)$ in the latter case. To decide which of these cases occurs, we determine the character χ of Δ, using the ordinary character table of S, see [4, Ch.12.5]:

If $Z \leq A$, then we have $\chi|_Z = 1_Z^-$, the non-trivial character of Z, so that in this case we have $\chi = 1_Z^- \otimes \chi|_{\overline{S}}$. We proceed to determine $\chi|_{\overline{S}}$, where $\overline{S}$ has the following irreducible characters, subscripts denoting degrees: the trivial character χ_1, the Steinberg character χ_p, two (algebraically conjugate) characters $\chi^{\pm 1}_{\frac{p+1}{2}}$, as well as $\frac{p-1}{4}$ characters χ^i_{p-1} for certain $i \in \mathbb{Z}_{p+1}$, and $\frac{p-5}{4}$ characters χ^j_{p+1} for certain $j \in \mathbb{Z}_{p-1}$.

We show that χ_1 is not a constituent of $\chi|_{\overline{S}}$: Since τ has odd order p, we have $\tau \in \overline{S}$, and $\mathrm{Fix}_{\mathbb{Q}^{p+1}}(\tau) = \{[a, \ldots, a, b] \in \mathbb{Q}^{p+1}; a, b \in \mathbb{Q}\}$. Moreover, one of the elements $\{\pm\sigma\}$ of order 2 belongs to $\overline{S}$. Now $[a, \ldots, a, b] = (\pm\sigma)([a, \ldots, a, b]) = \pm[-\epsilon b, \ldots, \left(\frac{i}{p}\right) a, \ldots, -\epsilon a]$, where $i \in \mathbb{Z}_p^* = \mathcal{Q}_p \,\dot{\cup}\, \mathcal{N}_p$, implies $a = b = 0$. Thus we have $\mathrm{Fix}_{\mathbb{Q}^{p+1}}(\overline{S}) \leq \mathrm{Fix}_{\mathbb{Q}^{p+1}}(\langle \sigma, \tau \rangle) = \{0\}$.

We show that $\chi|_{\overline{S}}$ is reducible: Assume to the contrary that $\chi|_{\overline{S}} = \chi^j_{p+1}$ for some $j \in \mathbb{Z}_{p-1}$. One of the elements $\{\pm\rho\}$ of order $\frac{p-1}{2}$ belongs to the conjugacy class of $\overline{S}$ containing $\overline{r}$, where $r = \mathrm{diag}[c, c^{-1}] \in S$ has order $p-1$. We have $\chi^j_{p+1}(r) = \zeta^j_{p-1} + \zeta^{-j}_{p-1}$, where $\zeta_{p-1} \in \mathbb{C}$ is a primitive $(p-1)$-st root of unity. Now $\chi(\rho) = 2$ entails that $\zeta^j_{p-1} = \zeta^{-j}_{p-1} \in \{\pm 1\}$, implying $\frac{p-1}{2} \mid j$, which is an invalid parameter, a contradiction.

Thus, taking character degrees into account, and recalling that χ is rational, we conclude that $\chi|_{\overline{S}} = \chi^1_{\frac{p+1}{2}} + \chi^{-1}_{\frac{p+1}{2}}$. Since $\left(\frac{c}{p}\right) = -1$ we have $\chi^{\pm 1}_{\frac{p+1}{2}}(r) = -1$, hence $\chi(\rho) = 2$ says that $-\rho \in \overline{S}$, thus $A \cong Z \times \overline{S}$. We note that one of the elements $\{\pm\sigma\}$ of order 2 belongs to the conjugacy class of $\overline{S}$ containing $\overline{s}$, where $s \in S$ has order 4, so that $\chi^{\pm 1}_{\frac{p+1}{2}}(s) = \left(\frac{\zeta_4}{p}\right)$, hence $\chi(\sigma) = 2\left(\frac{\zeta_4}{p}\right)$ entails $\sigma \in \overline{S}$.

ii) If $p \equiv -1 \pmod 4$, then $\sigma \in A$ has order 4, while $\overline{\sigma} \in \overline{A}$ has order 2. This entails $A \cong Z.\overline{S}$, hence for $p \neq 3$ we get $A \cong S$, while an explicit check shows that $A \cong S$ for $p = 3$ as well.

We determine the character χ of Δ: Since Z acts as $-E_{p+1}$, we only consider the faithful irreducible characters of S, which are as follows, subscripts denoting degrees: two (algebraically conjugate) characters $\chi^{\pm 1}_{\frac{p+1}{2}}$, as well as $\frac{p+1}{4}$ characters χ^i_{p-1} for certain $i \in \mathbb{Z}_{p+1}$, and $\frac{p-3}{4}$ characters χ^j_{p+1} for certain $j \in \mathbb{Z}_{p-1}$.

We show that χ is reducible: Assume to the contrary that $\chi = \chi_{p+1}^{j}$ for some $j \in \mathbb{Z}_{p-1}$. The element ρ of order $\frac{p-1}{2}$ belongs to the conjugacy class of S containing $r^2 = \mathrm{diag}[c^2, c^{-2}]$. We have $\chi_{p+1}^{j}(r^2) = \zeta_{\frac{p-1}{2}}^{j} + \zeta_{\frac{p-1}{2}}^{-j}$, where $\zeta_{\frac{p-1}{2}} \in \mathbb{C}$ is a primitive $\frac{p-1}{2}$-th root of unity. Now $\chi(\rho) = 2$ entails that $\zeta_{\frac{p-1}{2}}^{j} = 1$, implying $\frac{p-1}{2} \mid j$, which is an invalid parameter, a contradiction.

Thus, taking character degrees into account, and recalling that χ is rational, we conclude that $\chi = \chi_{\frac{p+1}{2}}^{1} + \chi_{\frac{p+1}{2}}^{-1}$. We note that the element σ of order 4 belongs to the conjugacy class of S containing s, where indeed $\chi_{\frac{p+1}{2}}^{\pm 1}(s) = \chi(\sigma) = 0$. $\sharp$

13 Golay codes

(13.1) Example: Binary Golay codes [1949]. Let $q := 2$ and $p := 23$. We find that $2 \in \mathbb{Z}_{23}^{*}$ has order $11 = \frac{23-1}{2}$, thus $\mathbb{F}_2(\zeta_{23}) = \mathbb{F}_{2^{11}}$ and $\varphi_2 \in \Gamma := \mathrm{Aut}_{\mathbb{F}_2}(\mathbb{F}_2(\zeta_{23}))$ has order 11. Moreover, we conclude that $2 \in \mathcal{Q}_{23}$, that is $\left(\frac{2}{23}\right) = 1$. Hence the Γ-orbits on $\mathcal{V}_{23}$ are $\mathcal{V}_{23} = \{1\} \,\dot\cup\, \{\zeta_{23}^{i}; i \in \mathcal{Q}_{23}\} \,\dot\cup\, \{\zeta_{23}^{i}; i \in \mathcal{N}_{23}\}$, where $\mathcal{Q}_{23} = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$ and $\mathcal{N}_{23} = \{5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22\}$.

Thus we have $X^{23}+1 = (X+1)\cdot\prod_{i\in\mathcal{Q}_{23}}(X+\zeta_{23}^{i})\cdot\prod_{i\in\mathcal{N}_{23}}(X+\zeta_{23}^{i}) = \mu_0\mu_1\mu_5 \in \mathbb{F}_2(\zeta_{23})[X]$. Actually we have $X^{23}+1 = (X+1)\cdot g'g'' \in \mathbb{F}_2[X]$, where $g' := X^{11}+X^{9}+X^{7}+X^{6}+X^{5}+X+1 \in \mathbb{F}_2[X]$ and $g'' := X^{11}+X^{10}+X^{6}+X^{5}+X^{4}+X^{2}+1 \in \mathbb{F}_2[X]$, hence the latter are both irreducible.

Let the **binary Golay code** $\mathcal{G}_{23} \leq \mathbb{F}_2^{23}$ be the cyclic code generated by g', the associated generator matrix $G \in \mathbb{F}_2^{12\times 23}$ being given in Table 9; since $(g')^{*} = g''$ the code generated by g'' is linearly equivalent to $\mathcal{G}_{23}$. Hence the even-weight subcode $\mathcal{G}_{23}' \leq \mathcal{G}_{23}$ has generator polynomial $(X+1)\cdot g' = X^{12}+X^{11}+X^{10}+X^{9}+X^{8}+X^{5}+X^{2}+1$. Moreover, $\mathcal{G}_{23}$ has check polynomial $h := (X+1)\cdot g'' = X^{12}+X^{10}+X^{7}+X^{4}+X^{3}+X^{2}+X+1$, thus $\mathcal{G}_{23}^{\perp}$ has generator polynomial $h^{*} = (X+1)^{*}\cdot(g'')^{*} = (X+1)\cdot g'$, showing that $\mathcal{G}_{23}^{\perp} = \mathcal{G}_{23}'$; the associated check matrix $H \in \mathbb{F}_2^{11\times 23}$ is also given in Table 9.

Choosing a primitive 23-rd root of unity over $\mathbb{F}_2$ having minimum polynomial g', we conclude that $\mathcal{G}_{23}$ is a QR code of type $\mathcal{Q}$; the code generated by g'' then is the associated QR code of type $\mathcal{N}$. Hence we again conclude that $\mathcal{G}_{23}^{\perp} = \mathcal{G}_{23}' \leq \mathcal{G}_{23}$. Moreover, extending yields the **extended binary Golay code** $\mathcal{G}_{24} := \widehat{\mathcal{G}}_{23} \leq \mathbb{F}_2^{24}$; by puncturing $\mathcal{G}_{24}$ again we recover $\mathcal{G}_{24}^{\bullet} = (\widehat{\mathcal{G}}_{23})^{\bullet} = \mathcal{G}_{23}$. We conclude that $\mathcal{G}_{24}^{\perp} = \mathcal{G}_{24}$, that is $\mathcal{G}_{24}$ is self-dual.

We determine minimum distances: The code $\mathcal{G}_{23}$ can be considered as a narrow sense BCH code with irreducible generator polynomial $\mu_1 = g'$ and associated cyclotomic coset $\mathcal{Q}_{23}$, hence has Bose distance $\delta = 5$, so that the BCH bound yields $d(\mathcal{G}_{23}) \geq 5$. The square root bound yields $d(\mathcal{G}_{23}) \geq \lceil\sqrt{23}\rceil = 5$ as well, but additionally $d(\mathcal{G}_{23}) \equiv 3 \pmod 4$, hence $d(\mathcal{G}_{23}) \geq 7$. Now the Hamming bound $2^{23-11}\cdot\sum_{i=0}^{3}\binom{23}{i} = 2^{12}\cdot(1+23+253+1771) = 2^{12}\cdot 2^{11} = 2^{23}$ is fulfilled for

$e := 3 = \frac{7-1}{2}$, showing that $d(\mathcal{G}_{23}) = 7$. Hence $\mathcal{G}_{23}$ is a perfect $[23, 12, 7]$-code, $\mathcal{G}'_{23}$ is a $[23, 11, 8]$-code, and $\mathcal{G}_{24}$ is a quasi-perfect $[24, 12, 8]$-code. ♯

Using the square root bound can be avoided by making use of the following observation: Let $\mathcal{C} \leq \mathcal{C}^\perp \leq \mathbb{F}_2^n$ be a weakly self-dual code. Since for any $v, w \in \mathbb{F}_2^n$ we have $\operatorname{supp}(v + w) = (\operatorname{supp}(v) \setminus \operatorname{supp}(w)) \,\dot\cup\, (\operatorname{supp}(w) \setminus \operatorname{supp}(v))$, we get $\operatorname{wt}(v + w) = \operatorname{wt}(v) + \operatorname{wt}(w) - 2 \cdot |\operatorname{supp}(v) \cap \operatorname{supp}(w)|$. Moreover, we have $\langle v, w\rangle = |\operatorname{supp}(v) \cap \operatorname{supp}(w)| \in \mathbb{F}_2$. Hence if $v, w \in \mathcal{C}$ such that $4 \mid \operatorname{wt}(v)$ and $4 \mid \operatorname{wt}(w)$, since $|\operatorname{supp}(v) \cap \operatorname{supp}(w)|$ is even we have $4 \mid \operatorname{wt}(v + w)$ as well.

Now the generator polynomial of $\mathcal{G}'_{23}$ shows that $\mathcal{G}'_{23}$ has an $\mathbb{F}_2$-basis consisting of vectors of weight 8. Thus the extended code $\mathcal{G}_{24} = \widehat{\mathcal{G}}_{23}$ also has an $\mathbb{F}_2$-basis consisting of vectors of weight 8. Thus we have $4 \mid \operatorname{wt}(v)$ for all $v \in \mathcal{G}_{24}$, that is $\mathcal{G}_{24}$ is 4**-divisible**, hence we have $4 \mid d(\mathcal{G}_{24})$. Since $d(\mathcal{G}_{24}^\bullet) = d(\mathcal{G}_{23}) \geq 5$ this implies $d(\mathcal{G}_{24}) \geq 8$, thus $d(\mathcal{G}_{23}) \geq 7$. ♯

It can be shown that $\mathcal{G}_{24}$ is the unique binary $[24, 12, 8]$-code, up to linear equivalence; its linear automorphism group $A(\mathcal{G}_{24}) = P(\mathcal{G}_{24}) \leq I_{24}(\mathbb{F}_2) \cong \mathcal{S}_{24}$ is isomorphic to the largest sporadic simple **Mathieu group** M_{24}, having order $244823040 \sim 2.4 \cdot 10^8$.

Moreover, it can be shown that $\mathcal{G}_{23}$ is the unique binary $[23, 12, 7]$-code, up to linear equivalence. Its linear automorphism group $A(\mathcal{G}_{23}) = P(\mathcal{G}_{23}) \leq I_{23}(\mathbb{F}_2) \cong \mathcal{S}_{23}$, coinciding with $\operatorname{Stab}_{P(\mathcal{G}_{24})}(\infty)$ and thus having index 24 in $P(\mathcal{G}_{24})$, is isomorphic to the second largest sporadic simple **Mathieu group** M_{23}, having order $10200960 \sim 1.0 \cdot 10^7$.

Finally, we mention that the binary Golay codes are intimately related to **Steiner systems**, in particular the **Witt systems**, occurring in algebraic combinatorics; see Exercises (16.25) and (16.26), where we give a brief indication.

(13.2) Example: Ternary Golay codes [1949]. Let $q := 3$ and $p := 11$. We find that $3 \in \mathbb{Z}_{11}^*$ has order $5 = \frac{11-1}{2}$, thus $\mathbb{F}_3(\zeta_{11}) = \mathbb{F}_{3^5}$ and $\varphi_3 \in \Gamma := \operatorname{Aut}_{\mathbb{F}_3}(\mathbb{F}_3(\zeta_{11}))$ has order 5. Moreover, we conclude that $3 \in \mathcal{Q}_{11}$, that is $\left(\frac{3}{11}\right) = 1$. Hence the Γ-orbits on $\mathcal{V}_{11}$ are $\mathcal{V}_{11} = \{1\} \,\dot\cup\, \{\zeta_{11}^i; i \in \mathcal{Q}_{11}\} \,\dot\cup\, \{\zeta_{11}^i; i \in \mathcal{N}_{11}\}$, where $\mathcal{Q}_{11} := \{1, 3, 4, 5, 9\}$ and $\mathcal{N}_{11} := \{2, 6, 7, 8, 10\}$,

Thus we have $X^{11} - 1 = (X - 1) \cdot \prod_{i \in \mathcal{Q}'_{11}} (X - \zeta_{11}^i) \cdot \prod_{i \in \mathcal{N}_{11}} (X - \zeta_{11}^i) = \mu_0\mu_1\mu_2 \in \mathbb{F}_3(\zeta_{11})[X]$. Actually we have $X^{11} - 1 = (X - 1) \cdot g'g'' \in \mathbb{F}_3[X]$, where $g' := X^5 - X^3 + X^2 - X - 1 \in \mathbb{F}_3[X]$ and $g'' := X^5 + X^4 - X^3 + X^2 - 1 \in \mathbb{F}_3[X]$, hence the latter are both irreducible.

Let the **ternary Golay code** $\mathcal{G}_{11} \leq \mathbb{F}_3^{11}$ be the cyclic code generated by $g' \in \mathbb{F}_3[X]$, the associated generator matrix $G \in \mathbb{F}_3^{6\times 11}$ being given in Table 10; since $(g')^* = -g''$ the code generated by g'' is linearly equivalent to $\mathcal{G}_{11}$. Hence $\mathcal{G}'_{11} \leq \mathcal{G}_{11}$ has generator polynomial $(X - 1) \cdot g' = X^6 - X^5 - X^4 - X^3 + X^2 + 1$. Moreover, $\mathcal{G}_{11}$ has check polynomial $h := (X-1) \cdot g'' = X^6 + X^4 - X^3 - X^2 - X + 1$, thus $\mathcal{G}_{11}^\perp$ has generator polynomial $h^* = (X - 1)^* \cdot (g'')^* = (X - 1) \cdot g'$, showing

Table 9: Generator and check matrices for $\mathcal{G}_{23}$.

$$\left[\begin{array}{ccccccccccccccccccccccc}
1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . & . & . & . & . & . & . \\
. & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . & . & . & . & . & . \\
. & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . & . & . & . & . \\
. & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . & . & . & . \\
. & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . & . & . \\
. & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . & . \\
. & . & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . & . \\
. & . & . & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . & . \\
. & . & . & . & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . & . \\
. & . & . & . & . & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . & . \\
. & . & . & . & . & . & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1 & . \\
. & . & . & . & . & . & . & . & . & . & . & 1 & 1 & . & . & . & 1 & 1 & 1 & . & 1 & . & 1
\end{array}\right]$$

$$\left[\begin{array}{ccccccccccccccccccccccc}
1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . & . & . & . & . \\
. & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . & . & . & . \\
. & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . & . & . \\
. & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . & . \\
. & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . & . & . \\
. & . & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . & . \\
. & . & . & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . & . \\
. & . & . & . & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . & . \\
. & . & . & . & . & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . & . \\
. & . & . & . & . & . & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1 & . \\
. & . & . & . & . & . & . & . & . & . & 1 & . & 1 & . & . & 1 & . & . & 1 & 1 & 1 & 1 & 1
\end{array}\right]$$

that $\mathcal{G}_{11}^{\perp} = \mathcal{G}'_{11}$; the associated check matrix $H \in \mathbb{F}_3^{5\times 11}$ is given in Table 10.

Choosing a primitive 11-th root of unity over $\mathbb{F}_3$ having minimum polynomial g', we conclude that $\mathcal{G}_{11}$ is a QR code of type $\mathcal{Q}$; the code generated by g'' then is the associated QR code of type $\mathcal{N}$. Hence we again conclude that $\mathcal{G}_{11}^{\perp} = \mathcal{G}'_{11} \leq \mathcal{G}_{11}$. Moreover, extending yields the **extended ternary Golay code** $\mathcal{G}_{12} := \widehat{\mathcal{G}}_{11} \leq \mathbb{F}_3^{12}$; by puncturing $\mathcal{G}_{12}$ again we recover $\mathcal{G}_{12}^{\bullet} = (\widehat{\mathcal{G}}_{11})^{\bullet} = \mathcal{G}_{11}$. We conclude that $\mathcal{G}_{12}^{\perp} = \mathcal{G}_{12}$, that is $\mathcal{G}_{12}$ is self-dual.

We determine minimum distances: The code $\mathcal{G}_{11}$ can be considered as a (wide sense) BCH code with irreducible generator polynomial $\mu_1 = g'$ and associated cyclotomic coset $\mathcal{Q}_{11}$, hence has Bose distance $\delta = 4$, so that the BCH bound yields $d(\mathcal{G}_{11}) \geq 4$. The square root bound yields $d(\mathcal{G}_{11}) \geq \lceil\sqrt{11}\rceil = 4$.

We now make use of the following observation: Let $\mathcal{C} \leq \mathcal{C}^{\perp} \leq \mathbb{F}_3^n$ be a weakly self-dual code. Now for any $v = [a_1, \ldots, a_n] \in \mathbb{F}_3^n$ we have $\langle v, v\rangle = \sum_{i=1}^n a_i^2 = |\mathrm{supp}(v)| = \mathrm{wt}(v) \in \mathbb{F}_3$. Hence for $v \in \mathcal{C}$ we have $3 \mid \mathrm{wt}(v)$.

Thus we have $3 \mid \mathrm{wt}(v)$ for $v \in \mathcal{G}_{12}$, that is $\mathcal{G}_{12}$ is 3**-divisible**, hence $3 \mid d(\mathcal{G}_{12})$. Since $d(\mathcal{G}_{12}^{\bullet}) = d(\mathcal{G}_{11}) \geq 4$ this implies that $d(\mathcal{G}_{12}) \geq 6$, and thus $d(\mathcal{G}_{11}) \geq 5$. Now the Hamming bound $3^{11-5} \cdot \sum_{i=0}^{2} \binom{11}{i} \cdot 2^i = 3^6 \cdot (1+11\cdot 2+55\cdot 4) = 3^6 \cdot 3^5 = 3^{11}$ is fulfilled for $e := 2 = \frac{5-1}{2}$, showing $d(\mathcal{G}_{11}) = 5$. Hence $\mathcal{G}_{11}$ is a perfect $[11, 6, 5]$-code, $\mathcal{G}'_{11}$ is an $[11, 5, 6]$-code, and $\mathcal{G}_{12}$ is a quasi-perfect $[12, 6, 6]$-code. ♯

Table 10: Generator and check matrices for $\mathcal{G}_{11}$.

$$G := \begin{bmatrix} 2 & 2 & 1 & 2 & . & 1 & . & . & . & . & . \\ . & 2 & 2 & 1 & 2 & . & 1 & . & . & . & . \\ . & . & 2 & 2 & 1 & 2 & . & 1 & . & . & . \\ . & . & . & 2 & 2 & 1 & 2 & . & 1 & . & . \\ . & . & . & . & 2 & 2 & 1 & 2 & . & 1 & . \\ . & . & . & . & . & 2 & 2 & 1 & 2 & . & 1 \end{bmatrix} \in \mathbb{F}_3^{6\times 11}$$

$$H := \begin{bmatrix} 1 & . & 1 & 2 & 2 & 2 & 1 & . & . & . & . \\ . & 1 & . & 1 & 2 & 2 & 2 & 1 & . & . & . \\ . & . & 1 & . & 1 & 2 & 2 & 2 & 1 & . & . \\ . & . & . & 1 & . & 1 & 2 & 2 & 2 & 1 & . \\ . & . & . & . & 1 & . & 1 & 2 & 2 & 2 & 1 \end{bmatrix} \in \mathbb{F}_3^{5\times 11}$$

It can be shown that $\mathcal{G}_{12}$ is the unique ternary $[12, 6, 6]$-code, up to linear equivalence; its linear automorphism group $A(\mathcal{G}_{12}) \leq I_{12}(\mathbb{F}_3) \cong \{\pm 1\}^{12} \rtimes \mathcal{S}_{12}$ is isomorphic to the non-split two-fold central extension $2.M_{12}$ of the second smallest sporadic simple **Mathieu group** M_{12}, having order 95040; hence $P(\mathcal{G}_{11}) \cong M_{12}$.

Moreover, it can be shown that $\mathcal{G}_{11}$ is the unique ternary $[11, 6, 5]$-code, up to linear equivalence. Its linear automorphism group $A(\mathcal{G}_{11}) \leq I_{11}(\mathbb{F}_3) \cong \{\pm 1\}^{11} \rtimes \mathcal{S}_{11}$, coinciding with $\mathrm{Stab}_{A(\mathcal{G}_{12})}(\langle [0, \ldots, 0, 1] \rangle_{\mathbb{F}_3})$ and thus having index 12 in $A(\mathcal{G}_{12})$, is isomorphic to the direct product $2 \times M_{11}$, where M_{11} is the smallest sporadic simple **Mathieu group**, having order 7920; hence $P(\mathcal{G}_{11}) \cong M_{11}$.

(13.3) Example: Football pool '13er-Wette'. To describe the outcome of a soccer match, we identify 'home team wins' with 1, 'guest team wins' with 2, and 'draw' with 0. Hence the outcome of $n \in \mathbb{N}$ matches can be considered as an element of $\mathbb{F}_3^n$. Now the task is to bet on the outcome of these matches, and the more guesses are correct the higher the reward is. The football pool currently in use in Germany is based on $n = 13$; in the years 1969–2004 it was based on $n = 11$, and in Austria $n = 12$ is used.

• According to the German 'Lotto' company, it is realistic to assume that 10^6 gamblers participate. Betting on a certain outcome actually costs 0.50€, hence there are 500 000€ at stake. From this 60% are handed back to the winners, who have at least 10 correct guesses, according to the schedule below. Assuming independent and uniformly distributed guesses we get the following winning probabilities and quotas, where the latter are obtained from the total rewards by dividing through the associated expected number of winners; these figures indeed fit nicely to the officially published quotas:

hits	%	reward/€	probability $\cdot 3^{13}$	winners	quota/€
13	21	105 000	1	0.63	167403.91
12	12	60 000	$2\cdot\binom{13}{1} = 26$	16.31	3679.21
11	12	60 000	$2^2\cdot\binom{13}{2} = 312$	195.69	306.60
10	15	75 000	$2^3\cdot\binom{13}{3} = 2288$	1435.09	52.26

To facilitate analysis, we assume that a single bet on a certain outcome costs 1€. This yields the following figures, with reward rates p_i, probabilities μ_i and quotas $q_i := \frac{p_i}{\mu_i}$, entailing an expected reward rate of $\sum_{i=0}^{3}\mu_i q_i = \sum_{i=0}^{3} p_i = \frac{3}{5}$:

i	p_i	$q_i/3^{13}$	$\mu_i\cdot 3^{13}$
0	$\frac{21}{100}$	$\frac{21}{100}$	1
1	$\frac{3}{25}$	$\frac{1}{26}\cdot\frac{3}{25}$	$2\cdot\binom{13}{1} = 26$
2	$\frac{3}{25}$	$\frac{1}{104}\cdot\frac{1}{25}$	$2^2\cdot\binom{13}{2} = 312$
3	$\frac{3}{20}$	$\frac{1}{2288}\cdot\frac{3}{20}$	$2^3\cdot\binom{13}{3} = 2288$

• To launch a systematic attack, we look for codes having not too many elements and small covering radius. Thus the best candidates are perfect e-error correcting codes of length n, for some $e \in \mathbb{N}_0$. In this case, the Hamming bound implies that $|B_e(0_n)| = \sum_{i=0}^{e}|B_i(0_n)\setminus B_{i-1}(0_n)| = \sum_{i=0}^{e}\binom{n}{i}\cdot 2^i$ is a 3-power. For $n = 13$ we get the following cardinalities:

e	0	1	2	3	4	5	6	7
$\lvert B_e(0_{13})\rvert$	1	**27**	339	2627	14067	55251	165075	384723

e	8	9	10	11	12	13
$\lvert B_e(0_{13})\rvert$	714195	1080275	1373139	1532883	1586131	1594323

Hence, not surprisingly, next to the trivial code and all of $\mathbb{F}_3^{13}$, we only find the case $e = 1$, into which the 1-error correcting perfect ternary Hamming $[13, 10, 3]$-code $\mathcal{H}_3$ fits; note that indeed the projective space $\mathbf{P}^2(\mathbb{F}_3)$ has $\frac{3^3-1}{3-1} = 13$ elements. Thus with $3^{10} = 59049$ bets, out of a total of $3^{13} = 1594323$, it is possible to guess at least $13 - 1 = 12$ of the outcomes correctly. More precisely, assuming again that outcomes are independent and uniformly distributed, we get the following winning probabilities:

Given an outcome $v \in \mathbb{F}_3^{13}$, we have to count the codewords $w \in \mathcal{H}_3$ such that $d(v, w) = \mathrm{wt}(v - w) = i$, for $i \in \{0, \ldots, 3\}$, which amounts to counting the elements of the coset $v + \mathcal{H}_3 \in \mathbb{F}_3^{13}/\mathcal{H}_3$ having weight i. Averaging over the $3^{13-10} = 3^3$ cosets, yields an expected reward of $\frac{1}{3^3}\cdot\sum_{v\in\mathbb{F}_3^{13}/\mathcal{H}_3}\big(\sum_{i=0}^{3}|\{w \in v+\mathcal{H}_3; \mathrm{wt}(w) = i\}|\cdot q_i\big) = \frac{1}{3^3}\cdot\sum_{i=0}^{3} q_i\cdot\big(\sum_{v\in\mathbb{F}_3^{13}/\mathcal{H}_3}|\{w \in v+\mathcal{H}_3; \mathrm{wt}(w) = i\}|\big) =$

$\frac{1}{3^3} \cdot \sum_{i=0}^{3} q_i \cdot |\{w \in \mathbb{F}_3^{13}; \mathrm{wt}(w) = i\}| = \frac{1}{3^3} \cdot \sum_{i=0}^{3} q_i \cdot (\mu_i \cdot 3^{13}) = 3^{10} \cdot \sum_{i=0}^{3} p_i$. Since we have to place 3^{10} bets, we get an expected reward rate of $\sum_{i=0}^{3} p_i = \frac{3}{5}$, which is precisely as good as random guessing. Note that this is independent of the choice of the reward rates p_i, and that we have not used that $\mathcal{H}_3$ is perfect.

• Another strategy, using expert knowledge, is as follows: If the outcome of two matches is for sure, then we may puncture with respect to these components, and end up in $\mathbb{F}_3^{11}$. For $n = 11$ we get the following cardinalities:

e	0	1	2	3	4	5	6
$\lvert B_e(0_{11})\rvert$	1	23	**243**	1563	6843	21627	51195

e	7	8	9	10	11
$\lvert B_e(0_{11})\rvert$	93435	135675	163835	175099	177147

Hence, next to the trivial code and all of $\mathbb{F}_3^{11}$, we only find the case $e = 2$, into which the 2-error correcting perfect ternary Golay $[11, 6, 5]$-code $\mathcal{G}_{11}$ fits. Thus with $3^6 = 729$ bets, out of a total of $3^{11} = 177147$ next to the above expert knowledge, it is possible to guess at least $13 - 2 = 11$ of the outcomes correctly.

More precisely, assuming that the unsure outcomes are independent and uniformly distributed, and assuming independent and uniformly distributed guesses we get the following winning probabilities μ'_i:

i	p_i	$q_i/3^{13}$	$\mu'_i \cdot 3^{11}$		
0	$\frac{21}{100}$	$\frac{21}{100}$			1
1	$\frac{3}{25}$	$\frac{1}{26} \cdot \frac{3}{25}$	$2 \cdot \binom{11}{1}$	$=$	22
2	$\frac{3}{25}$	$\frac{1}{104} \cdot \frac{1}{25}$	$2^2 \cdot \binom{11}{2}$	$=$	220
3	$\frac{3}{20}$	$\frac{1}{2288} \cdot \frac{3}{20}$	$2^3 \cdot \binom{11}{3}$	$=$	1320

We get an expected reward rate of $\sum_{i=0}^{3} \mu'_i q_i = \sum_{i=0}^{3} \frac{\mu'_i}{\mu_i} \cdot p_i = 3^2 \cdot \sum_{i=0}^{3} \frac{\binom{11}{i}}{\binom{13}{i}} \cdot p_i = 3^2 \cdot (p_0 + \frac{11}{13} \cdot p_1 + \frac{55}{78} \cdot p_2 + \frac{15}{26} \cdot p_3) = 3^2 \cdot \frac{251}{520} = \frac{2259}{520} \sim 4.34$. Hence this is a sensible winning strategy, only depending on expert knowledge, but not on using a code.

Using the code $\mathcal{G}_{11}$, averaging over the $3^{11-6} = 3^5$ cosets in $\mathbb{F}_3^{11}/\mathcal{G}_{11}$ yields an expected reward of $\frac{1}{3^5} \cdot \sum_{v \in \mathbb{F}_3^{11}/\mathcal{G}_{11}} \left(\sum_{i=0}^{3} |\{w \in v + \mathcal{G}_{11}; \mathrm{wt}(w) = i\}| \cdot q_i \right) = \frac{1}{3^5} \cdot \sum_{i=0}^{3} q_i \cdot |\{w \in \mathbb{F}_3^{11}; \mathrm{wt}(w) = i\}| = \frac{1}{3^5} \cdot \sum_{i=0}^{3} q_i \cdot (\mu'_i \cdot 3^{11}) = 3^6 \cdot \sum_{i=0}^{3} \mu'_i q_i$. Since we have to place 3^6 bets, we get an expected reward rate of $\sum_{i=0}^{3} \mu'_i q_i$, which is precisely as good as random guessing.

IV

14 Exercises to Part I (in German)

(14.1) Aufgabe: Quellencodierung.
Alice kann vier verschiedene Nachrichten senden:

00: 'Die Börse ist sehr fest.'	11: 'Sollen wir verkaufen?'
01: 'Die Kurse fallen.'	10: 'Helft uns gegensteuern!'

Diese werden jeweils durch zwei Bits wie angegeben codiert. Bob empfängt die Dauernachricht ...001100110011... Welche Nachrichten will Alice schicken?

(14.2) Aufgabe: Arithmetischer Prüfzeichencode.
Es sei $\mathcal{C} := \{[x_1, \ldots, x_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} ix_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$ der Code des ISBN-10-Standards. Kann $\mathcal{C}$ Zwillingsfehler und Sprungzwillingsfehler erkennen?

(14.3) Aufgabe: Geometrischer Prüfzeichencode.
Für $a \in \mathbb{Z}_{11}$ sei $\mathcal{C}_a := \{[x_1, \ldots, x_{10}] \in \mathbb{Z}_{11}^{10}; \sum_{i=1}^{10} a^i x_i = 0 \in \mathbb{Z}_{11}\} \leq \mathbb{Z}_{11}^{10}$. Man untersuche, in Abhängigkeit von a, wann $\mathcal{C}_a$ Einzelfehler, Drehfehler, Zwillingsfehler und Sprungzwillingsfehler erkennen kann.

(14.4) Aufgabe: Kontonummern-Code.
a) Für $x \in \mathbb{N}_0$ sei $Q(x) \in \mathbb{N}_0$ die **Quersumme** von x bezüglich der Dezimaldarstellung. Man zeige: Faßt man $\mathbb{Z}_{10}$ als Teilmenge von $\mathbb{N}_0$ auf, so erhält man eine Bijektion $\mathbb{Z}_{10} \rightarrow \mathbb{Z}_{10} \colon x \mapsto Q(2x)$.
b) Ein häufig verwendeter Prüfzeichencode bei der Bildung von Kontonummern ist $\mathcal{C} := \{[x_1, \ldots, x_{2n}] \in \mathbb{Z}_{10}^{2n}; \sum_{i=1}^{n}(Q(2x_{2i-1}) + x_{2i}) = 0 \in \mathbb{Z}_{10}\} \subseteq \mathbb{Z}_{10}^{2n}$, für $n \in \mathbb{N}$. Kann $\mathcal{C}$ Einzelfehler und Drehfehler erkennen?

(14.5) Aufgabe: International Bank Account Number.
Man betrachte die gültige IBAN 'DE68 3905 0000 0123 4567 89'. Angenommen, an der führenden Stelle der Kontonummer wird statt '0' eine '9' eingetippt, so daß die BBAN nun '3905 0000 **9**123 4567 89' lautet. Welche möglichen weiteren Tippfehler gibt es, so daß eine so entstehende doppelt fehlerhafte BBAN nicht anhand der Prüfzeichen 'DE68' erkannt werden kann?

(14.6) Aufgabe: Prüfzeichencodes über Gruppen.
Man betrachte einen Prüfzeichencode über der endlichen Gruppe G, bezüglich der bijektiven Abbildungen $\pi_i \colon G \rightarrow G$, für $i \in \{1, \ldots, n\}$ und $n \in \mathbb{N}$. Welche Bedingungen müssen die Abbildungen π_i jeweils erfüllen, damit Zwillingsfehler und Sprungzwillingsfehler erkannt werden?

(14.7) Aufgabe: Prüfzeichencodes über Diedergruppen.
Für $n \geq 2$ seien $D_{2n} := \langle a, b; a^n = 1, b^2 = 1, b^{-1}ab = a^{-1} \rangle$ die **Diedergruppe** der Ordnung $2n$, und $\tau \colon D_{2n} \to D_{2n}$ gegeben durch $a^i \mapsto a^{-i-1}$ und $a^i b \mapsto a^i b$, für $i \in \{0, \ldots, n-1\}$. Man zeige: Die Abbildung τ ist wohldefiniert und bijektiv. Ist n ungerade, so gilt $gh^\tau \neq hg^\tau$ für alle $g \neq h \in D_{2n}$.

(14.8) Aufgabe: Fast-vollständige Abbildungen.
Es seien $G = \mathbb{Z}_n$, wobei $n := 2^a$ für ein $a \in \mathbb{N}$, und $\sigma \colon G \to G$ gegeben durch

$$\sigma \colon 0 \mapsto m \mapsto 1 \mapsto m+1 \mapsto \cdots \mapsto m-1 \mapsto 2m-1 \mapsto 0,$$

wobei $m := \frac{n}{2}$. (Dann ist σ bijektiv.) Für die Abbildung $\tau := \sigma + \mathrm{id}_G \colon G \to G$ zeige man: Es gelten $\tau(G) = G \setminus \{m-1\}$ und $\tau^{-1}(\{n-1\}) = \{n-1, n-1-\lceil \frac{n}{4} \rceil\}$.

(14.9) Aufgabe: Huffman-Codierung.
Zur (optimalen) binären Quellencodierung kann man folgenden rekursiven Algorithmus benutzen: Es seien $\mathcal{X} = \{x_1, \ldots, x_q\}$ ein Alphabet mit Wahrscheinlichkeitsverteilung μ, und $p_k := \mu(x_k)$ für $k \in \{1, \ldots, q\}$. Weiter seien $j \neq i \in \{1, \ldots, q\}$ mit $p_i = \min\{p_1, \ldots, p_q\}$ und $p_j = \min(\{p_1, \ldots, p_q\} \setminus \{p_i\})$. Dann werden $x_i \mapsto 0$ und $x_j \mapsto 1$ codiert. Nun benutzt man das Alphabet $\mathcal{X}' := (\{x_1, \ldots, x_q\} \setminus \{x_i, x_j\}) \mathbin{\dot\cup} \{x_{ij}\}$, mit unveränderten Wahrscheinlichkeiten für x_k, für $k \in \{1, \ldots, q\} \setminus \{i, j\}$, und Wahrscheinlichkeit $p_{ij} := p_i + p_j$ für x_{ij}, um rekursiv Präfixe zu den bereits gefundenen Codierungen zu bestimmen.

a) Man zeige: Die obige **Huffman-Codierung** ergibt eine injektive Funktion $h \colon \mathcal{X} \to (\mathbb{Z}_2)^* \setminus \{\epsilon\}$, und der Code $h(\mathcal{X}) \subseteq (\mathbb{Z}_2)^*$ ist präfix-frei. Außerdem zeige man: Für die mittlere Wortlänge gilt $\sum_{i=1}^{q} p_i \cdot l(h(x_i)) \leq H(\mathcal{X}) + 1$.

b) Nun trage $\mathcal{X}$ die Gleichverteilung, und es sei $k := \lfloor \log_2(q) \rfloor \in \mathbb{N}$. Man zeige: Die zugehörige Huffman-Codierung besteht aus Wörtern der Länge k und $k+1$, und hat die mittlere Wortlänge $k + 2 - \frac{2^{k+1}}{n}$. Was passiert im Falle $q = 2^k$? Was fällt beim Vergleich mit $H(\mathcal{X})$ auf?

c) Etwa erhält man für das Alphabet $\mathbb{Z}_4$ mit den angegebenen Wahrscheinlichkeiten die folgende Huffman-Codierung:

i	p_i	$(\mathbb{Z}_2)^*$
0	0.4	0
1	0.3	11
2	0.2	101
3	0.1	100

Man bestimme die Entropie $H(\mathbb{Z}_4)$ und die mittlere Wortlänge $\sum_{i \in \mathbb{Z}_4} p_i \cdot l(h(i))$ der zugehörigen Huffman-Codierung. Außerdem betrachte man das Alphabet $\mathbb{Z}_4^2$, wobei die beiden Positionen unabhängig seien, und bestimme ebenso $H(\mathbb{Z}_4^2)$, die zugehörige Huffman-Codierung und ihre mittlere Wortlänge. Was fällt auf?

(14.10) Aufgabe: Huffman-Codierung in GAP.
a) Man schreibe ein GAP-Programm zur Berechnung der Huffman-Codierung des Alphabets $\mathcal{X} := \{\mathtt{a}, \ldots, \mathtt{z}\}$ bezüglich der angegebenen relativen Häufigkeiten der Buchstaben in der englischen Sprache. Man bestimme die Entropie $H(\mathcal{X})$ und die mittlere Wortlänge der zugehörigen Huffman-Codierung.

$\mathcal{X}$	$\mathbb{Z}_{26}$	μ
a	0	0.082
b	1	0.015
c	2	0.028
d	3	0.043
e	4	0.127
f	5	0.022
g	6	0.020
h	7	0.061
i	8	0.070
j	9	0.002
k	10	0.008
l	11	0.040
m	12	0.024

$\mathcal{X}$	$\mathbb{Z}_{26}$	μ
n	13	0.067
o	14	0.075
p	15	0.019
q	16	0.001
r	17	0.060
s	18	0.063
t	19	0.091
u	20	0.028
v	21	0.010
w	22	0.023
x	23	0.001
y	24	0.020
z	25	0.001

b) Man schreibe ein GAP-Programm, mit dem man aus der Huffman-Codierung eines Texts den ursprüngichen Text zurückgewinnt.

c) Man codiere und decodiere damit den folgenden Text aus 314 Buchstaben, wobei Leerzeichen und Satzzeichen ignoriert werden können. Wieviele Bit hat die Huffman-Codierung dieses Texts?

```
the almond tree was in a tentative blossom, the days were
longer, often ending with magnificent evenings of corrugated
pink skies, the hunting season was over, with hounds and guns
put away for six months, the vineyards were busy again,
as the well organized farmers treated their vines, and the
more lackadaisical neighbors hurried to do the pruning they
should have done in november
```

(14.11) Aufgabe: Symmetrischer binärer Kanal.
Man bestimme die maximale Kapazität eines symmetrischen binären Kanals mit Fehlerwahrscheinlichkeit $\frac{1}{2} \leq p \leq 1$.

(14.12) Aufgabe: ML-Decodierung.
Man betrachte einen symmetrischen binären Kanal mit Fehlerwahrscheinlichkeit $0 \leq p < \frac{1}{2}$; typische Werte sind etwa $p = 10^{-e}$ für $e \in \{1, 2, 3\}$. Über diesen Kanal sollen die Wörter in $\mathbb{F}_2^3$ übertragen werden, wobei Gleichverteilung auf $\mathbb{F}_2^3$ angenommen und ML-Decodierung verwendet werde.
a) Wie groß ist die Fehlerwahrscheinlichkeit $\gamma(\mathbb{F}_2^3)$, wenn die Wörter ohne Redundanz, also mit Informationsrate $\rho(\mathbb{F}_2^3) = 1$, gesendet werden?

b) Nun werde der Code $\mathcal{C}_0 \leq \mathbb{F}_2^6$ mit Generatormatrix $G_0 := [E_3 \mid E_3] \in \mathbb{F}_2^{3\times 6}$, also mit Informationsrate $\rho(\mathcal{C}_0) = \frac{1}{2}$, verwendet. Wie groß ist die Fehlerwahrscheinlichkeit $\gamma(\mathcal{C}_0)$?
c) Schließlich werde der Code $\mathcal{C} \leq \mathbb{F}_2^6$ mit Generatormatrix

$$G := \begin{bmatrix} 1 & . & . & . & 1 & 1 \\ . & 1 & . & 1 & . & 1 \\ . & . & 1 & 1 & 1 & . \end{bmatrix} \in \mathbb{F}_2^{3\times 6},$$

also ebenfalls mit Informationsrate $\rho(\mathcal{C}_0) = \frac{1}{2}$, verwendet. Wie groß ist nun die Fehlerwahrscheinlichkeit $\gamma(\mathcal{C})$?

Hinweis. Man betrachte die Kugeln $\mathcal{B}_r(c) \subseteq \mathbb{F}_2^6$, für $c \in \mathcal{C}$ und $r \in \{0, 1, 2\}$.

(14.13) Aufgabe: ML-Decodierung.
Zur Datenübertragung durch einen symmetrischen binären Kanal mit Fehlerwahrscheinlichkeit $0 \leq p < \frac{1}{2}$ werde ein $(7, 2^4, 3)$-Code $\mathcal{C}$ (etwa ein Hamming-Code) verwendet, wobei Gleichverteilung auf $\mathcal{C}$ angenommen und ML-Decodierung verwendet werde. Man bestimme die Fehlerwahrscheinlichkeit $\gamma(\mathcal{C})$.

15 Exercises to Part II (in German)

(15.1) Aufgabe: Hamming-Abstand.
a) Es seien $n \in \mathbb{N}$ und $v, w \in \mathbb{F}_2^n$ mit $d := d(v, w) \in \mathbb{N}_0$. Für $r, s \in \mathbb{N}_0$ betrachte man $\mathcal{A} := \{u \in \mathbb{F}_2^n; d(v, u) = r, d(w, u) = s\}$, und es sei $t := \frac{d+r-s}{2}$. Man zeige: Ist $t \notin \mathbb{Z}$, so ist $\mathcal{A} = \emptyset$; ist $t \in \mathbb{Z}$, so ist $|\mathcal{A}| = \binom{d}{t} \cdot \binom{n-d}{r-t}$.
b) Es seien $n \in \mathbb{N}$ und $v, w, x, y \in \mathbb{F}_2^n$ mit paarweisem Abstand $d \in \mathbb{N}$. Man zeige: Der Abstand d ist gerade; und es gibt genau ein $u \in \mathbb{F}_2^n$ mit $d(v, u) = d(w, u) = d(x, u) = \frac{d}{2}$. Gilt notwendig auch $d(y, u) = \frac{d}{2}$?

(15.2) Aufgabe: Auslöschungen.
Es sei $\mathcal{C}$ ein nicht-trivialer Block-Code. Geht bei der Datenübertragung ein Eintrag eines Wortes verloren, so wird er als **Auslöschung** markiert; dies ist also ein Fehler mit bekannter Position. Für $e, g \in \mathbb{N}_0$ zeige man: Der Code $\mathcal{C}$ kann genau dann gleichzeitig e Fehler und g Auslöschungen korrigieren, wenn $2e + g + 1 \leq d(\mathcal{C})$ gilt.

(15.3) Aufgabe: Minimaldistanz.
Es sei $\mathcal{C}$ ein binärer Code der Länge $n \in \mathbb{N}$ und Minimaldistanz $d \geq 3$. Man zeige: Es gilt $|\mathcal{C}| \leq \frac{2^n}{n+1}$; und ist n gerade, so gilt $|\mathcal{C}| \leq \frac{2^n}{n+2}$. Kann die Schranke auch für n ungerade verbessert werden?

Hinweis. Man zähle $\{[v, w] \in \mathcal{C} \times \mathbb{F}_2^n; d(v, w) = 2\}$ mittels Double-Counting.

(15.4) Aufgabe: Nichtlineare Codes.
a) Aus Singleton-, Hamming- und Plotkin-Schranke folgere man: Für einen binären $(6, 9, d)$-Code gilt $d \leq 3$. Man gebe einen binären $(6, 9, 2)$-Code an.
b) Man zeige Es gibt keinen binären $(6, 9, 3)$-Code.

(15.5) Aufgabe: Parameter für Codes.
Für $n \in \mathbb{N}$ betrachte man einen $(n, q^{n-3}, 3)$-Code über einem Alphabet mit q Elementen. Man zeige: Es gilt $n \leq q^2 + q + 1$.

(15.6) Aufgabe: Perfekte Codes.
Es sei $\mathcal{C} \subseteq \mathbb{F}_2^n$ ein perfekter Code mit Minimaldistanz 7. Man zeige: Es ist $n \in \{7, 23\}$. Man bestimme $\mathcal{C}$ für den Fall $n = 7$.

(15.7) Aufgabe: Hamming-Schranke.
Man zeige: Die Parameter $q := 2$, $n := 90$, $m := 2^{78}$ und $d := 5$ erfüllen die Hamming-Schranke, aber es gibt keinen binären (n, m, d)-Code.

Hinweis. Ist $\mathcal{C} \subseteq \mathbb{F}_2^n$ solch ein Code, so betrachte man $\mathcal{A} := \{v = [x_1, \ldots, x_n] \in \mathbb{F}_2^n; x_1 = x_2 = 1, \mathrm{wt}(v) = 3\}$ und $\mathcal{B} := \{w = [y_1, \ldots, y_n] \in \mathcal{C}; y_1 = y_2 = 1, \mathrm{wt}(w) = 5\}$, und zähle $\{[v, w] \in \mathcal{A} \times \mathcal{B}; vw^{\mathrm{tr}} = 1\}$ mittels Double-Counting.

(15.8) Aufgabe: Selbstduale Codes.
a) Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein Code mit Kontrollmatrix $H = [A \mid E_{n-k}] \in \mathbb{F}_q^{(n-k)\times k}$ in Standardform, wobei $k = \dim_{\mathbb{F}_q}(\mathcal{C})$ und $A \in \mathbb{F}_q^{(n-k)\times k}$. Man zeige: Der Code $\mathcal{C}$ ist genau dann selbstdual, wenn $2k = n$ und $AA^{\mathrm{tr}} = -E_{n-k}$ gelten.
b) Es sei $p \in \mathbb{N}$ eine Primzahl. Man gebe selbstduale $\mathbb{F}_p$-lineare Codes der Längen $n = 4$ und $n = 8$ an.

Hinweis zu b). Man unterscheide die Fälle $p = 2$ sowie $p \equiv \pm 1 \pmod 4$.

(15.9) Aufgabe: Gewichtssumme.
Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein Code mit $k := \dim_{\mathbb{F}_q}(\mathcal{C}) \geq 1$, so daß eine Generatormatrix für $\mathcal{C}$ keine Nullspalte enthalte. Man zeige: Es gilt $\sum_{v\in\mathcal{C}} \mathrm{wt}(v) = n(q-1)q^{k-1}$.

(15.10) Aufgabe: Systematische Codes.
Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein Code mit Generatormatrix $G \in \mathbb{F}_q^{k\times n}$. Man zeige: Eine k-elementige Teilmenge von Spalten von G ist genau dann $\mathbb{F}_q$-linear unabhängig, wenn $\mathcal{C}$ systematisch auf den zugehörigen Komponenten ist.

(15.11) Aufgabe: MDS-Codes und Dualität.
Es sei $\mathcal{C} < \mathbb{F}_q^n$ ein nicht-trivialer $[n, k, d]$-Code. Man zeige die Äquivalenz der folgenden Aussagen: **i)** $\mathcal{C}$ ist ein MDS-Code. **ii)** $\mathcal{C}^\perp$ ist ein MDS-Code.
iii) $\mathcal{C}$ ist systematisch auf allen k-elementigen Teilmengen von Komponenten.
iv) In jeder Generatormatrix für $\mathcal{C}$ sind alle k-elementigen Teilmengen von Spalten $\mathbb{F}_q$-linear unabhängig.

v) In jeder Kontrollmatrix für $\mathcal{C}$ sind alle $(n-k)$-elementigen Teilmengen von Spalten $\mathbb{F}_q$-linear unabhängig.

(15.12) Aufgabe: Syndrom-Decodierung.
Man betrachte den Hamming-Code $\mathcal{H} \leq \mathbb{F}_2^7$, der durch die folgende Generatormatrix definiert wird:

$$G := \begin{bmatrix} 1 & . & . & . & . & 1 & 1 \\ . & 1 & . & . & 1 & . & 1 \\ . & . & 1 & . & 1 & 1 & . \\ . & . & . & 1 & 1 & 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{4 \times 7}.$$

Man bestimme Syndrome, zugehörige Nebenklassenführer und decodiere

i) $[1,1,0,0,1,1,0]$, **ii)** $[1,1,1,0,1,1,0]$, **iii)** $[1,1,1,1,1,1,0]$.

(15.13) Aufgabe: Syndrom-Decodierung.
Es sei $\mathcal{C} \leq \mathbb{F}_2^7$ der durch die folgende Generatormatrix definierte Code:

$$G := \begin{bmatrix} 1 & . & . & . & 1 & . & 1 \\ . & 1 & . & . & 1 & . & 1 \\ . & . & 1 & . & . & 1 & 1 \\ . & . & . & 1 & . & 1 & 1 \end{bmatrix} \in \mathbb{F}_2^{4 \times 7}.$$

Man bestimme Syndrome, zugehörige Nebenklassenführer und decodiere

i) $[1,1,0,1,0,1,1]$, **ii)** $[0,1,1,0,1,1,1]$, **iii)** $[0,1,1,1,0,0,0]$.

(15.14) Aufgabe: Eindeutige Decodierbarkeit.
Es sei $\mathcal{C} \leq \mathbb{F}_2^{10}$ der durch die folgende Generatormatrix definierte Code:

$$G := \begin{bmatrix} 1 & . & . & . & . & . & . & . & 1 & 1 \\ . & 1 & . & . & . & . & 1 & 1 & . & . \\ . & . & 1 & . & . & 1 & . & 1 & . & . \\ . & . & . & 1 & . & 1 & 1 & . & . & . \\ . & . & . & . & 1 & 1 & 1 & 1 & . & . \end{bmatrix} \in \mathbb{F}_2^{5 \times 10}.$$

Man zeige: Alle Vektoren in $\mathbb{F}_2^{10}$ sind eindeutig decodierbar. Man bestimme Minimaldistanz und Überdeckungsradius von $\mathcal{C}$.

(15.15) Aufgabe: Äquidistante Codes.
Es sei $\mathcal{C} \subseteq \mathbb{F}_2^{16}$ ein binärer Code mit $\mathrm{wt}(v) = 6$ und $d(v,w) = 8$ für alle $v \neq w \in \mathcal{C}$. Man zeige: Es gilt $|\mathcal{C}| \leq 16$. Gibt es einen solchen Code mit $|\mathcal{C}| = 16$?

(15.16) Aufgabe: Griesmer-Schranke.
a) Man zeige: Die Parameter $q := 3$, $n := 14$, $k := 4$ und $d := 9$ erfüllen die Griesmer-Schranke, aber es gibt keinen ternären $[14,4,9]$-Code.
b) Man zeige: Die Parameter $q := 2$, $n := 15$, $k := 8$ und $d := 5$ erfüllen die Griesmer-Schranke, aber es gibt keinen binären $[15,8,5]$-Code.

Hinweis. Man betrachte jeweils den residualen Code.

(15.17) Aufgabe: Optimale Codes.
Für $n, d \in \mathbb{N}$ sei $K_2(n,d) := \max\{k \in \mathbb{N};$ es gibt einen binären $[n,k,d]$-Code$\}$. Man zeige, daß $K_2(n, 2d-1) = K_2(n+1, 2d)$ und $K_2(n+1,d) \leq K_2(n,d)+1$.

(15.18) Aufgabe: Best-Code.
Ziel ist es, zu zeigen, daß optimale nicht-lineare Codes besser sein können als optimale lineare Codes, aber daß es schwierig sein kann, solche zu finden:
a) Man zeige: Für einen binären $[10,k,4]$-Code gilt $k \leq 5$. Man gebe einen binären $[10,5,4]$-Code an; also gilt $K_2(10,4) = 5$.
b) Es sei $\mathcal{C}_0 \leq \mathbb{F}_2^{10}$ der von $[G \mid G] \in \mathbb{F}_2^{3\times 10}$ erzeugte Code, wobei

$$G := \begin{bmatrix} 1 & . & . & . & 1 \\ . & 1 & . & 1 & 1 \\ . & . & 1 & 1 & . \end{bmatrix} \in \mathbb{F}_2^{3\times 5}.$$

Es seien $v := [1,0,0,0,0;\, 0,0,1,0,0] \in \mathbb{F}_2^{10}$ und $\pi := (1,2,3,4,5)(6,7,8,9,10) \in \mathcal{S}_{10}$, sowie $\mathcal{C} := (v + \mathcal{C}_0)^{\langle\pi\rangle} \subseteq \mathbb{F}_2^{10}$ der nicht-lineare **Best-Code** [1978]. Man zeige: $\mathcal{C}$ ist ein $(10,40,4)$-Code. (Man kann zeigen, daß $\mathcal{C}$ optimal ist.)

(15.19) Aufgabe: Erweiterte Codes.
Es seien $\mathcal{C}_1$ und $\mathcal{C}_2$ die durch die folgenden Generatormatrizen G_1 bzw. G_2 definierten ternären Codes:

$$G_1 := \begin{bmatrix} 1 & . & 1 & . & . \\ . & 1 & . & 1 & 1 \end{bmatrix} \in \mathbb{F}_3^{2\times 5} \quad \text{und} \quad G_2 := \begin{bmatrix} 1 & . & 2 & . & . \\ . & 1 & . & 2 & 2 \end{bmatrix} \in \mathbb{F}_3^{2\times 5}.$$

Man bestimme die Minimaldistanz von $\mathcal{C}_1$ und $\mathcal{C}_2$ sowie von $\widehat{\mathcal{C}}_1$ und $\widehat{\mathcal{C}}_2$.

(15.20) Aufgabe: Modifikation von Codes.
Man wende die Konstruktionen Punktierung, Erweiterung, Bereinigung, Augmentierung, Verkürzung und Verlängerung auf die binären Prüfzeichen- und Wiederholungscodes an, und bestimme die Parameter der so erhaltenen Codes.

(15.21) Aufgabe: Schwach selbstduale binäre Codes.
Es sei $\mathcal{C} \leq \mathbb{F}_2^n$ ein schwach selbstdualer Code.
a) Man zeige: Es ist $\mathcal{C} = \mathcal{C}'$, wobei $\mathcal{C}' \leq \mathbb{F}_2^n$ den bereinigten Code bezeichne. Es gilt genau dann $4 \mid \mathrm{wt}(v)$ für alle $v \in \mathcal{C}$, wenn dies für eine $\mathbb{F}_2$-Basis von $\mathcal{C}$ gilt. Ist n ungerade, so ist $4 \mid \mathrm{wt}(v)$ für alle $v \in \mathcal{C}$.
b) Es sei $\mathcal{C}$ ein $[n, \frac{n-1}{2}, d]$-Code, für $n \in \mathbb{N}$ ungerade. Man zeige: Es gilt $\mathcal{C} < \mathcal{C}^\perp = \widetilde{\mathcal{C}}$, wobei $\widetilde{\mathcal{C}} \leq \mathbb{F}_2^n$ den augmentierten Code bezeichne.
c) Es sei $\mathcal{C}$ selbstdual. Man zeige: Es gilt $\mathcal{C} = \mathcal{C}^\perp = \widetilde{\mathcal{C}}$.

(15.22) Aufgabe: Gerades Gewicht.
Man zeige: Gibt es einen binären (n,m,d)-Code mit $d \in \mathbb{N}$ gerade, so gibt es auch einen binären (n,m,d)-Code, dessen Elemente alle gerades Gewicht haben.

(15.23) Aufgabe: Summen und Produkte von Codes.
Es seien $\mathcal{C}$ ein nicht-trivialer $[n,k,d]$-Code, sowie $\mathcal{C}'$ ein nicht-trivialer $[n',k',d']$-Code über $\mathbb{F}_q$, mit Generatormatrizen $G \in \mathbb{F}_q^{k\times n}$ bzw. $G' \in \mathbb{F}_q^{k'\times n'}$. Man zeige die folgenden Aussagen, und untersuche, was jeweils passiert, wenn man andere Generatormatrizen wählt:
a) Ist $k = k'$, so erzeugen die Zeilen der Matrix $[G \mid G'] \in \mathbb{F}_q^{k\times(n+n')}$ einen $[n+n',k,d'']$-Code mit $d'' \geq d+d'$; er wird **Verkettung** von $\mathcal{C}$ und $\mathcal{C}'$ genannt.
b) Die Zeilen der Matrix $\left[\begin{array}{c|c} G & . \\ \hline . & G' \end{array}\right] \in \mathbb{F}_q^{(k+k')\times(n+n')}$ erzeugen einen $[n+n',k+k',\min\{d,d'\}]$-Code; er wird **direkte Summe** von $\mathcal{C}$ und $\mathcal{C}'$ genannt.
c) Die Menge der Matrizen in $\mathbb{F}_q^{n\times n'}$, deren Spalten bzw. Zeilen Elemente von $\mathcal{C}$ bzw. $\mathcal{C}'$ sind, ist ein $[nn',kk',dd']$-Code, der $G \otimes G' \in \mathbb{F}_q^{(kk')\times(nn')}$ als Generatormatrix besitzt; er wird **direktes Produkt** von $\mathcal{C}$ und $\mathcal{C}'$ genannt.

(15.24) Aufgabe: Produktcodes.
Es sei $\mathcal{C} \leq \mathbb{F}_2^{8\times 16}$ das direkte Produkt der erweiterten binären Hamming-Codes $\widehat{\mathcal{H}}_3$ und $\widehat{\mathcal{H}}_4$; also ist $\mathcal{C}$ ein $[128,44,16]$-Code und 7-fehlerkorrigierend. In der Tat hat $\mathcal{C}$ aber viel bessere Fehlerkorrektureigenschaften:

Angenommen, bei der Übertragung eines Codeworts geschehen Fehler genau in den 14 (zufällig ausgewählten, dem Empfänger nicht bekannten) Positionen

$$[2,7,19,24,27,32,45,51,53,76,82,86,96,121];$$

dabei werden die Einträge der Matrizen zeilenweise durchnumeriert. Man zeige, daß das empfangene Wort eindeutig decodiert werden kann.

(15.25) Aufgabe: Punktierter Simplex-Code.
Durch Punktieren konstruiere man aus dem binären $[31,5,16]$-Simplex-Code einen $[21,5,10]$-Code. Wie verhält er sich zur Griesmer-Schranke?

(15.26) Aufgabe: Hamming- und Simplex-Codes.
Für $k \geq 2$ seien $\mathcal{H}_k$ und $\mathcal{S}_k$ die zugehörigen Hamming- bzw. Simplex-Codes über $\mathbb{F}_q$. Wie verhalten sich jeweils die Informationsrate und die relative Minimaldistanz dieser Codes für $k \to \infty$?

(15.27) Aufgabe: Erweiterte Hamming-Codes.
Es seien $k \geq 2$ und $\mathcal{C}$ ein binärer $[2^k, 2^k-k-1, 4]$-Code. Man zeige: $\mathcal{C}$ ist linear äquivalent zum erweiterten Hamming-Code $\widehat{\mathcal{H}}_k$.

(15.28) Aufgabe: Hadamard-Codes.
a) Eine Matrix $H \in \mathbb{R}^{n\times n}$, für $n \in \mathbb{N}$, die Einträge in $\{\pm 1\}$ hat und $HH^{\mathrm{tr}} = nE_n$ erfüllt, heißt **Hadamard-Matrix**; falls zudem die Einträge in der ersten Zeile und Spalte sämtlich positiv sind, so heißt H **normalisiert**.

Man zeige: Ist $H \in \mathbb{R}^{n \times n}$ eine Hadamard-Matrix mit $n \geq 3$, so ist $n \equiv 0 \pmod 4$. Außerdem zeige man: Sind $H \in \mathbb{R}^{n \times n}$ und $H' \in \mathbb{R}^{n' \times n'}$ Hadamard-Matrizen, so ist $H \otimes H' \in \mathbb{R}^{(nn') \times (nn')}$ ebenfalls eine Hadamard-Matrix.

b) Ersetzt man in einer normalisierten Hadamard-Matrix $H \in \mathbb{R}^{n \times n}$ den Eintrag 1 bzw. -1 durch $0 \in \mathbb{F}_2$ bzw. $1 \in \mathbb{F}_2$, so erhält man die zugehörige **binäre Hadamard-Matrix**. Die Zeilen der binären Matrizen zu H und $-H$ bilden den zugehörigen binären **Hadamard-Code** $\mathcal{A}$. Verkürzt man $\mathcal{A}$ bezüglich der ersten Komponente, so erhält man den **verkürzten Hadamard-Code** $\mathcal{A}^\circ$.

Man zeige: $\mathcal{A}$ ist ein $(n, 2n, \frac{n}{2})$-Code, und $\mathcal{A}^\circ$ ist ein $(n-1, n, \frac{n}{2})$-Code. Was ist ihr Überdeckungsradius? Wie verhalten sie sich zur Plotkin-Schranke?

c) Nun seien $H_2 = H_2^{\otimes 1} := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ und $H_2^{\otimes (k+1)} := H_2^{\otimes k} \otimes H_2 \in \mathbb{R}^{2^k \times 2^k}$, für $k \in \mathbb{N}$. Man zeige: $H_2^{\otimes k}$ eine normalisierte Hadamard-Matrix; sie wird auch **Sylvester-Matrix** genannt.

Weiter zeige man: Die zugehörigen Codes $\mathcal{A}_k$ und $\mathcal{A}_k^\circ$ sind linear; also ist $\mathcal{A}_k$ ein $[2^k, k+1, 2^{k-1}]$-Code, und $\mathcal{A}_k^\circ$ ist ein $[2^k - 1, k, 2^{k-1}]$-Code. Daraus folgere man: Für $k \geq 2$ ist $\mathcal{A}_k$ linear äquivalent zum Reed-Muller-Code $\mathcal{R}_k$, und $\mathcal{A}_k^\circ$ ist linear äquivalent zum Simplex-Code $\mathcal{S}_k$. Welche Codes erhält man für $k \leq 3$?

d) Man zeige, daß die folgende Matrix eine Hadamard-Matrix ist. Sind die zugehörigen binären Hadamard- und verkürzten Hadamard-Codes linear?

$$\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 \\
1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 \\
1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 \\
1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\
1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\
1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 & 1 & -1
\end{bmatrix} \in \mathbb{R}^{12 \times 12}$$

(15.29) Aufgabe: Gewichtsverteilungen.

Es sei $\mathcal{C} \leq \mathbb{F}_2^n$ mit Gewichtsverteilung $A_\mathcal{C} \in \mathbb{C}[X, Y]$.

a) Man gebe die Gewichtsverteilungen des erweiterten Codes $\widehat{\mathcal{C}}$, des bereinigten Codes $\mathcal{C}'$ und des augmentierten Codes $\widetilde{\mathcal{C}}$ an.

b) Ein $f \in \mathbb{C}[X]$ mit $f^* = f$ heißt **Palindrom**. Man gebe eine notwendige und hinreichende Bedingung dafür an, daß $A_\mathcal{C}(X, 1) \in \mathbb{C}[X]$ ein Palindrom ist.

(15.30) Aufgabe: Selbstduale binäre Codes.

a) Für einen selbstdualen binären $[16, 8, d]$-Code zeige man: Es gilt $d \in \{2, 4, 6\}$.

b) Für $d \in \{2, 4\}$ gebe man jeweils einen selbstdualen binären $[16, 8, d]$-Code zusammen mit seiner Gewichtsverteilung an.
c) Man gebe die Gewichtsverteilung eines selbstdualen binären $[16, 8, 6]$-Codes an. Gibt es solch einen Code?

Hinweis zu a). Man benutze die Griesmer-Schranke.

(15.31) Aufgabe: Gewichtsverteilungen binärer Hamming-Codes.
a) Es seien $k \geq 2$ und $n := 2^k - 1$, und $A_{\mathcal{H}_k} = \sum_{i=0}^{n} w_i X^i Y^{n-i} \in \mathbb{C}[X, Y]$ die Gewichtsverteilung des binären Hamming-Codes $\mathcal{H}_k$. Man zeige: Die Koeffizienten erfüllen für $i \geq 2$ die Rekursion $iw_i + w_{i-1} + (n - i + 2)w_{i-2} = \binom{n}{i-1}$, mit Anfangsbedingung $w_0 = 1$ und $w_1 = 0$.
b) Man bestimme die Gewichtsverteilungen der Codes $\widehat{\mathcal{H}}_k$ und $\mathcal{H}'_k$.
c) Man bestimme die Gewichtsverteilungen der Reed-Muller-Codes $\mathcal{R}_k$ und $\mathcal{R}_k^\bullet$.

(15.32) Aufgabe: Gewichtsverteilungen von MDS-Codes.
Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein MDS-Code der Dimension $k \in \mathbb{N}$ mit Gewichtsverteilung $w_i := |\{v \in \mathcal{C}; \mathrm{wt}(v) = i\}| \in \mathbb{N}_0$, für $i \in \{0, \ldots, n\}$.
a) Man zeige: Es gilt $w_0 = 1$, sowie $w_i = 0$ für $i \in \{1, \ldots, n - k\}$, und für $i \in \{0, \ldots, k - 1\}$ gilt $w_{n-i} = \sum_{j=i}^{k-1} (-1)^{j-i} \binom{j}{i} \binom{n}{j} (q^{k-j} - 1)$.
b) Für $k \geq 2$ folgere man daraus: Es gilt $n \leq q + k - 1$.

Hinweis zu a). Es sei $\mathcal{C}_\mathcal{I} := \{[x_1, \ldots, x_n] \in \mathcal{C}; x_i = 0 \text{ für alle } i \in \mathcal{I}\} \leq \mathcal{C}$, für $\mathcal{I} \subseteq \{1, \ldots, n\}$. Man bestimme $|\{[\mathcal{I}, v]; |\mathcal{I}| = j, v \in \mathcal{C}_\mathcal{I} \setminus \{0\}\}|$ mittels Double-Counting, für $j \in \{0, \ldots, k - 1\}$.

16 Exercises to Part III (in German)

(16.1) Aufgabe: Verallgemeinerte zyklische Codes.
Es seien $n \in \mathbb{N}$ und $0 \neq a \in \mathbb{F}_q$. Ein Code $\mathcal{C} \leq \mathbb{F}_q^n$ heißt ***a*-zyklisch**, falls mit $[x_0, \ldots, x_{n-1}] \in \mathcal{C}$ auch stets $[ax_{n-1}, x_0, \ldots, x_{n-2}] \in \mathcal{C}$ ist.

Man gebe eine Korrespondenz von der Menge der a-zyklischen Codes der Länge n zu den Idealen in einem geeigneten Quotienten des Polynomrings $\mathbb{F}_q[X]$ an. Man zeige weiter: Mit $\mathcal{C}, \mathcal{C}' \leq \mathbb{F}_q^n$ sind stets auch $\mathcal{C} + \mathcal{C}' \in \mathbb{F}_q^n$ und $\mathcal{C} \cap \mathcal{C}' \in \mathbb{F}_q^n$ a-zyklisch; man gebe jeweils ein Generatorpolynom an.

(16.2) Aufgabe: Modifikation zyklischer Codes.
Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein zyklischer Code mit Generatorpolynom $g \in \mathbb{F}_q[X]$. Welche der Konstruktionen Punktierung, Erweiterung, Bereinigung, Augmentierung, Verkürzung und Verlängerung ergeben wieder einen zyklischen Code? In diesen Fällen gebe man jeweils ein Generatorpolynom an.

(16.3) Aufgabe: Konstruktion zyklischer Codes.
Es seien $\mathcal{C} \leq \mathbb{F}_q^n$ und $\mathcal{C}' \leq \mathbb{F}_q^{n'}$ nicht-triviale zyklische Codes mit Generatorpolynomen $g \in \mathbb{F}_q[X]$ bzw. $g' \in \mathbb{F}_q[X]$.

a) Es sei $\mathrm{ggT}(n, n') = 1$. Man zeige: Das direkte Produkt von $\mathcal{C}$ und $\mathcal{C}'$ ist ebenfalls zyklisch; man gebe ein Generatorpolynom an.
b) Es seien $q := 2$ und $n = n'$ ungerade, und es gelte $g \mid g'$. Man zeige: Die Plotkin-Summe von $\mathcal{C}$ und $\mathcal{C}'$ ist linear äquivalent zu einem zyklischen Code mit Generatorpolynom $gg' \in \mathbb{F}_2[X]$.

(16.4) Aufgabe: CRC-Codes.
Man schreibe GAP-Programme zur CRC-Codierung und -Decodierung. Welche Eingabedaten müssen zur Verfügung gestellt werden? Was ist die Ausgabe? Welche Konsistenztests sollten gemacht werden?

Man wende die Programme auf den CRC-Code $\mathcal{H} \leq \mathbb{F}_2^7$ mit Generatorpolynom $X^3 + X + 1 \in \mathbb{F}_2[X]$ an: Man codiere die Vektoren

i) $[0,0,0,1]$, **ii)** $[0,0,1,1]$, **iii)** $[0,1,1,1]$, **iv)** $[1,1,1,1]$,

und bestimme, welche der folgenden Vektoren in $\mathcal{H}$ liegen:

i) $[1,1,0,0,1,1,0]$, **ii)** $[0,1,0,1,1,1,0]$, **iii)** $[1,0,0,0,1,0,1]$.

(16.5) Aufgabe: RWTH-ID.
Man schreibe GAP-Programme zur Erzeugung und Verifikation von RWTH-IDs, unter Verwendung der Programme aus Aufgabe (16.4), des Generatorpolynoms $X^5 + X^2 + 1 \in \mathbb{F}_2[X]$ und der Quellencodierung in Tabelle 7.

(16.6) Aufgabe: Kreisteilungspolynome.
Es seien $\mathbb{F}_q$ der Körper mit q Elementen und $n \in \mathbb{N}$.
a) Es sei $m \in \mathbb{N}$. Man zeige: Es ist genau dann $X^m - 1$ ein Teiler von $X^n - 1 \in \mathbb{F}_q[X]$, wenn m ein Teiler von n ist.
b) Es seien p die Charakteristik von $\mathbb{F}_q$, und $\overline{\mathbb{F}}$ ein algebraischer Abschluß von $\mathbb{F}_q$. Wie hängen die Faktorisierungen von $X^{pn} - 1$ und $X^n - 1 \in \mathbb{F}_q[X]$ zusammen? Wie hängen die Nullstellenmengen $\mathcal{V}(X^{pn} - 1)$ und $\mathcal{V}(X^n - 1) \subseteq \overline{\mathbb{F}}$ zusammen?

(16.7) Aufgabe: Klassifikation zyklischer Codes.
a) Man bestimme alle zyklischen binären Codes der Länge $n \in \{1, \ldots, 32\}$ durch Angabe jeweils eines Generatorpolynoms. Außerdem bestimme man jeweils die zugehörige Nullstellenmenge.
b) Man untersuche diese Codes hinsichtlich Dualität und mengentheoretischer Inklusion. Für $n \leq 10$ gebe man jeweils auch die Minimaldistanz an.

(16.8) Aufgabe: Reversible Codes.
Ein zyklischer Code $\mathcal{C} \leq \mathbb{F}_q^n$ mit Generatorpolynom $g \in \mathbb{F}_q[X]$ heißt **reversibel**, falls mit $[x_0, \ldots, x_{n-1}] \in \mathcal{C}$ stets auch $[x_{n-1}, \ldots, x_0] \in \mathcal{C}$ ist.
a) Man zeige die Äquivalenz der folgenden Aussagen: **i)** $\mathcal{C}$ ist reversibel. **ii)** Es gilt $g^* = ag \in \mathbb{F}_q[X]$ für ein $a \in \mathbb{F}_q$. **iii)** Mit $\zeta \in \mathcal{V}(g)$ ist stets auch $\zeta^{-1} \in \mathcal{V}(g)$.
b) Nun seien $\gcd(q, n) = 1$ und $-1 \in \mathbb{Z}_n^*$ eine q-Potenz. Man zeige: Jeder zyklische Code $\mathcal{C} \leq \mathbb{F}_q^n$ ist reversibel.

(16.9) Aufgabe: Hamming-Codes als zyklische Codes.
Für $k \geq 2$ sei $\mathcal{H}_k \leq \mathbb{F}_q^n$ der zugehörige Hamming-Code, wobei $n := \frac{q^k-1}{q-1}$. Ist $\mathcal{H}_k$ auch für $\mathrm{ggT}(k, q-1) > 1$ linear äquivalent zu einem zyklischen Code?

(16.10) Aufgabe: Zyklische binäre Codes gerader Länge.
Für $k \geq 3$ sei $\mathcal{H}_k \leq \mathbb{F}_2^n$ der zugehörige Hamming-Code, wobei $n := 2^k - 1$. Man zeige: Der bereinigte verkürzte Code $(\mathcal{H}_k^\circ)' \leq \mathbb{F}_2^{n-1}$ ist linear äquivalent zu einem zyklischen $[n-1, n-k-2, 4]$-Code.

Hinweis. Man betrachte die Plotkin-Summe von $\mathcal{H}'_{k-1}$ mit einem Even-Weight-Code.

(16.11) Aufgabe: Simplex-Codes.
Für $k \geq 2$ und $n := 2^k - 1$ sei $\mathcal{S}_k \leq \mathbb{F}_2^n$ der zugheörige binäre Simplex-Code, der definiert ist als der Dualcode des binären Hamming-Codes $\mathcal{H}_k \leq \mathbb{F}_2^n$. Man zeige: Der Code $\mathcal{S}_k$ ist linear equivalent zu einem zyklischen Code. Man bestimme die Nullstellenmenge $\mathcal{V}(\mathcal{S}_k)$, und gebe damit neue Beweise für die Aussagen $\dim_{\mathbb{F}_2}(\mathcal{S}_k) = k$ und $d(\mathcal{S}_k) = 2^{k-1}$ an.

(16.12) Aufgabe: BCH-Codes.
Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein BCH-Code. Ist der duale Code $\mathcal{C}^\perp$ ebenfalls ein BCH-Code?

(16.13) Aufgabe: Bose-Distanz von BCH-Codes.
Man gebe einen nicht-trivialen BCH-Code $\mathcal{C}$ im engeren Sinne mit Bose-Distanz δ und Minimaldistanz $d(\mathcal{C}) > \delta$ an.

Hinweis. Man betrachte nicht-primitive Codes.

(16.14) Aufgabe: Dimension binärer BCH-Codes.
Es sei $\mathcal{C} \leq \mathbb{F}_2^n$ ein primitiver BCH-Code im engeren Sinne der Länge $n := 2^k - 1$, für $k \geq 1$, und Entwurfsdistanz δ mit $\lfloor \frac{\delta}{2} \rfloor \leq 2^{\lceil \frac{k}{2} \rceil - 1}$. Man zeige: Es gilt $\dim_{\mathbb{F}_2}(\mathcal{C}) = n - \lfloor \frac{\delta}{2} \rfloor \cdot k$.

(16.15) Aufgabe: Binäre BCH-Codes.
Man bestimme die Dimensionen der primitiven binären BCH-Codes im engeren Sinne der Länge 31.

(16.16) Aufgabe: Ternäre BCH-Codes.
Man bestimme die maximale Dimension eines primitiven ternären BCH-Codes der Länge 26 und der Entwurfsdistanz 5.

(16.17) Aufgabe: Reed-Solomon-Codes.
a) Es sei $\mathcal{C} \leq \mathbb{F}_q^{q-1}$ ein Reed-Solomon-Code. Man zeige: Ist $1 \notin \mathcal{V}(\mathcal{C})$, so ist der erweiterte Code $\widehat{\mathcal{C}} \leq \mathbb{F}_q^q$ ein MDS-Code.

b) Es sei $\mathcal{C} \leq \mathbb{F}_q^n$ ein primitiver BCH-Code. Man zeige: Es gibt einen endlichen Körper $\mathbb{F}_q \subseteq F$ und einen Reed-Solomon-Code $\mathcal{D} \leq F^n$, so daß $\mathcal{C} = \mathcal{D} \cap \mathbb{F}_q^n$ gilt.

(16.18) Aufgabe: Roos-Schranke.
Es seien $n := 2^k - 1$, für ein $k \geq 3$, und $\mathcal{C} \leq \mathbb{F}_2^n$ ein zyklischer Code. Man zeige:
a) Ist $\{\zeta_n, \zeta_n^5\} \subseteq \mathcal{V}(\mathcal{C})$, so hat $\mathcal{C}$ hat Minimaldistanz $d \geq 4$.
b) Ist $\{\zeta_n, \zeta_n^{-1}\} \subseteq \mathcal{V}(\mathcal{C})$, so hat $\mathcal{C}$ hat Minimaldistanz $d \geq 5$, und der bereinigte Code $\mathcal{C}' \leq \mathbb{F}_2^n$ hat Minimaldistanz $d' \geq 6$.

(16.19) Aufgabe: van-Lint-Wilson-Schranke.
a) Es seien $n := 2^{2k} + 1$, für ein $k \geq 1$, und $\mathcal{C} \leq \mathbb{F}_2^n$ der zyklische Code zu $\{\zeta_n\}$. Man zeige: Der Code $\mathcal{C}$ ist reversibel und hat Minimaldistanz $d \geq 5$.
b) Es sei $\mathcal{C} \leq \mathbb{F}_2^{31}$ der zyklische Code zu $\{\zeta_{31}, \zeta_{31}^5, \zeta_{31}^7\}$. Man zeige: Der Code $\mathcal{C}$ hat Minimaldistanz $d \geq 7$. Gilt Gleichheit?

(16.20) Aufgabe: QR-Codes.
a) Man bestimme die Minimaldistanz des binären QR-Codes der Länge 47.
b) Man bestimme alle perfekten 1-fehler-korrigierenden QR-Codes.

(16.21) Aufgabe: Erweiterte QR-Codes.
Man gebe jeweils einen selbstdualen binären $[32, 16, 8]$-Code und einen selbstdualen binären $[48, 24, 12]$-Code an.

(16.22) Aufgabe: Wiege-Problem.
Gegeben seien $n \geq 3$ Münzen, unter denen sich höchstens eine Fälschung mit zu kleinem oder zu großem Gewicht befinde. Wieviele unabhängige Wägungen mit einer Balkenwaage werden benötigt, um die möglicherweise vorhandene gefälschte Münze zu finden?

Hinweis. Man betrachte die Fälle $n = \frac{3^k - 1}{2}$ für $k \geq 2$.

(16.23) Aufgabe: Golay-Codes.
Für den erweiterten ternären Golay-Code $\mathcal{G}_{12} \leq \mathbb{F}_3^{12}$ und den erweiterten binären Golay-Code $\mathcal{G}_{24} \leq \mathbb{F}_2^{24}$ bestimme man die zugehörigen residualen Codes.

(16.24) Aufgabe: Gewichtsverteilungen der Golay-Codes.
Man bestimme die Gewichtsverteilungen der ternären Golay-Codes $\mathcal{G}_{11}$ und $\mathcal{G}_{12}$, sowie der binären Golay-Codes $\mathcal{G}_{23}$ und $\mathcal{G}_{24}$.

(16.25) Aufgabe: Steiner-Systeme.
Ein **Steiner-System** $S(t, k, v)$, wobei $t, k, v \in \mathbb{N}$ mit $t \leq k \leq v$, ist eine Menge $\mathcal{P}$ der Kardinalität v, die **Punkte** genannt werden, zusammen mit einer Menge $\mathcal{B}$ von k-elementigen Teilmengen von $\mathcal{P}$, die **Blöcke** genannt werden, so daß jede t-elementige Menge von Punkten in genau einem Block enthalten ist.

a) Man zeige: Für $s \in \{0, \ldots, t\}$ ist jede s-elementige Menge von Punkten in genau $\lambda_s \in \mathbb{N}$ Blöcken enthalten, wobei $\lambda_s \cdot \binom{k-s}{t-s} = \binom{v-s}{t-s}$. Daraus folgere man: Die Anzahl $b = |\mathcal{B}| \in \mathbb{N}$ der Blöcke ist gegeben durch $b \cdot \binom{k}{t} = \binom{v}{t}$; jeder Punkt gehört zu genau $r \in \mathbb{N}$ Blöcken, wobei $bk = vr$; und für $t = 2$ gilt $r(k-1) = v-1$.

b) Man zeige: Gibt es ein Steiner-System $S(t, k, v)$ mit $t \geq 2$, so gibt es auch ein Steiner-System $S(t-1, k-1, v-1)$.

Steiner-Systeme mit $t = 1$ oder $t = k$ sind uninteressant (warum?), und solche mit $t = 2$ sind leicht zu finden, wie wir sogleich sehen werden. Das ist viel schwieriger für $t \geq 3$, und vermutlich gibt es gar keine für $k > t \geq 6$. Unten werden wir sehen, daß es sporadische Steiner-Systeme mit $t = 4$ und $t = 5$ gibt.

b) Es sei $\mathbf{A}^2(\mathbb{F}_q) := \mathbb{F}_q^2$ die **affine Ebene** über $\mathbb{F}_q$, dabei heißen die Teilmengen $w + \langle v \rangle_{\mathbb{F}_q} \subseteq \mathbb{F}_q^2$, wobei $v, w \in \mathbb{F}_q^2$ mit $v \neq 0$, **affine Geraden**. Man zeige: $\mathbf{A}^2(\mathbb{F}_q)$ bildet zusammen mit den affinen Geraden ein Steiner-System $S(2, q, q^2)$; man bestimme die Parameter λ_s, b und r.

c) Es sei $\mathbf{P}^2(\mathbb{F}_q) := \{\langle v \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^3; 0 \neq v \in \mathbb{F}_q^3\}$ die **projektive Ebene** über $\mathbb{F}_q$; dabei heißen die Teilräume $\langle v, w \rangle_{\mathbb{F}_q} \leq \mathbb{F}_q^3$, wobei $\langle v \rangle_{\mathbb{F}_q} \neq \langle w \rangle_{\mathbb{F}_q} \in \mathbf{P}^2(\mathbb{F}_q)$, **projektive Geraden**. Man zeige: $\mathbf{P}^2(\mathbb{F}_q)$ bildet zusammen mit den projektiven Geraden ein Steiner-System $S(2, q+1, q^2+q+1)$; man bestimme die Parameter λ_s, b und r. Außerdem stelle man **Fano-Ebene** $\mathbf{P}^2(\mathbb{F}_2)$ graphisch dar.

(16.26) Aufgabe: Codes und Steiner-Systeme.
Der Zusammenhang zwischen binären Codes und Steiner-Systemen wird wie folgt hergestellt: Es sei $\mathcal{C} \leq \mathbb{F}_2^n$ ein $[n, k, d]$-Code mit $d \in \mathbb{N}$, und für $\mathcal{P} := \{1, \ldots, n\}$ sei $\mathcal{B} := \{\mathrm{supp}(v) \subseteq \mathcal{P}; v \in \mathcal{C}, \mathrm{wt}(v) = d\}$ die Menge der Träger der Codewörter minimalen positiven Gewichts.

a) Man zeige: Der Code $\mathcal{C} \leq \mathbb{F}_2^n$ ist genau dann perfekt mit $d = 2e+1$, wenn $\mathcal{B}$ die Menge der Blöcke eines Steiner-Systems $S(e+1, 2e+1, n)$ bildet. In diesem Falle führt der erweiterte Code $\widehat{\mathcal{C}} \leq \mathbb{F}_2^{n+1}$ zu einem Steiner-System $S(e+2, 2e+2, n+1)$. Wie hängen diese Steiner-Systeme zusammen?

b) Daraus folgere man: Der Hamming-Code $\mathcal{H}_k$ und der erweiterte Code $\widehat{\mathcal{H}}_k$, für $k \geq 2$, führen zu Steiner-Systemen $S(2, 3, 2^k-1)$ bzw. $S(3, 4, 2^k)$. Der Golay-Code $\mathcal{G}_{23}$ und der erweiterte Code $\mathcal{G}_{24}$ führen zu Steiner-Systemen $S(4, 7, 23)$ bzw. $S(5, 8, 24)$; diese werden auch **Witt-Systeme** genannt.

c) Daraus bestimme man erneut die Gewichtsverteilung von $\mathcal{G}_{24}$.

17 References

[1] E. Assmus, J. Key: Designs and their codes, Cambridge Tracts in Mathematics 103, Cambridge University Press, 1992.

[2] E. Berlekamp: Algebraic coding theory, Mac-Graw Hill, 1968.

[3] A. Betten, H. Fripertinger, A. Kerber, A. Wassermann, K. Zimmermann: Codierungstheorie, Springer, 1998.

[4] F. Digne, J. Michel: Representations of finite groups of Lie type, London Math. Soc. Student Texts 21, Cambridge University Press, 2012.

[5] D. Jungnickel: Codierungstheorie, Spektrum Verlag, 1995.

[6] F. MacWilliams, N. Sloane: The theory of error-correcting codes, North-Holland Mathematical Library 16, North-Holland, 1986.

[7] R. Schulz: Codierungstheorie, Vieweg, 1991.

[8] D. Stinson: Cryptography, theory and practice, third edition, CRC Press Series on Discrete Mathematics and its Applications 36, 2006.

[9] J. van Lint: Introduction to coding theory, 2nd edition, Graduate Texts in Mathematics 86, Springer, 1991.

[10] W. Willems: Codierungstheorie, de Gruyter, 1999.